DOCUMENT RESUME

ED 338 649 TM 017 351

AUTHOR	Cheung, K. C.
TITLE	On Meaningful Measurement: Concepts, Technology and Examples.
PUB DATE	Sep 91
NOTE	43p.; Paper presented at the Regional Symposium on Educational Testing (Beijing, China, September 16-20, 1991).
PUB TYPE	Reports - Evaluative/Feasibility (142) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC02 Plus Postage.
DESCRIPTORS	*Affective Measures; *Cognitive Measurement; *Educational Assessment; Educational Technology; *Epistemology; Foreign Countries; Heuristics; Item Response Theory; Learning Theories; *Mathematical Concepts; Metaphors; *Problem Solving; Test Use
IDENTIFIERS	Authentic Assessment

ABSTRACT

This paper discusses how concepts and procedural skills in problem-solving tasks, as well as affects and emotions, can be subjected to meaningful measurement (MM), based on a multisource model of learning and a constructivist information-processing theory of knowing. MM refers to the quantitative measurement of conceptual and procedural knowledge with qualitative interpretations that should be rooted in a theory of knowing, model of difficult learning, classroom realities, and educational objectives as intended in the programs of study. Knowledge of the problem and search spaces of learning tasks, cognitive and affective schemes, alternative conceptions, and cognitive and procedural barriers contribute to an informed understanding of measurement of the process of development of pertinent cognitive and affective structures and their linkages. Metaphors of pigtails, bamboo stems, and competence ladders are heuristics used to describe these structures and linkages. Quantitative requirements of measurement are elaborated, and associated methodology is described, including focus group interviews, network and latent trait analyses, and dual scaling. Examples based on empirical data are presented. The progression of competence and affective schemes, affordability of problem tasks, conformity of test responses with conceptual and measurement models, simultaneity of conceptualization of a measure, and requirements of specific objectivity are hallmarks of MM. An agenda for a research program is also presented. Six figures and four tables illustrate the discussion. A 54-item list of references is included. (SLD)

On Meaningful Measurement : Concepts, Technology and Examples

K C Cheung

National Institute of Education
Nanyang Technological University
Republic of Singapore

Paper presented to **Regional Symposium on Educational Testing**, Beijing, People's Republic of China.

September 16-20, 1991.

Abstract

This paper discusses how concepts and procedural skills in problem-solving tasks, as well as affects and emotions, can be meaningfully measured, based on a multisource model of learning and a constructivist, information processing theory of knowing. It is proposed that knowledge of the problem and search spaces of learning tasks, cognitive and affective schemes, alternative conceptions, cognitive and procedural barriers contribute to an informed understanding on measurement of the process of development of pertinent cognitive and affective structures and their linkages. The metaphors of pig-tails, bamboo stems, and competence ladders are heuristics used to describe these structures and linkages which are constructed within the cognising minds.

Quantitative requirements of measurement are elaborated, and the associated technology using focus-group interview, network and latent trait analyses, and Dual Scaling on samples of epistemic subjects are developed. The outcomes of the measurement process are quantitative measures with qualitative interpretations that are firmly rooted in the educational objectives of the instructional programmes. Examples based on empirical data are presented to illustrate these concepts and technology of **meaningful measurement**. Under the rubric of meaningful measurement, the ideas of **progression** of competence and affective schemes, **affordance** of problem tasks, **conformity** of test responses with conceptual and measurement models, and **simultaneity** of conceptualisation of a measure and its measurement, and requirements of **specific objectivity** are its hallmarks. An agenda for a research programme currently in progress is also presented.

On Meaningful Measurement : Concepts, Technology and Examples

INTRODUCTION

This article presents concepts, technology, and examples of **meaningful measurement.** The basic rationale concurs with a comment made by Glaser (1991) regarding the need to reap the benefits of advances in **cognitive science** and **psychometrics** in the assessment of **acquired proficiency.**

> To date, we have been reaping the benefits of psychometrics, employing advances in statistical theory, but have had little opportunity to emphasize the **cognitive and developmental theory** that explicate **human thinking and performance.** Given this state of knowledge, most of the work of testing technology ... has occurred after test items were constructed. In the future, test design will entail extensive attention to theory before and during item design as well. ... we must rely on the emerging picture of the **properties of acquired proficiency** in school subjects to make tests responsive to the **structures** and **processes** that develop as individuals move from beginning to advanced learners. Thus, the assessment of achievement can be more closely tied to our understanding of **progression in learning** and the accrual of **results of effective teaching** (Glaser, 1991, p.28, emphases added).

Cheung et al. (1990b, 1991) have been developing the concepts and technology of meaningful measurement. **Meaningful measurement is quantitative measurement of conceptual and procedural knowledge with qualitative interpretations that should be firmly rooted in a theory of knowing, model of difficult learning, classroom realities, and educational objectives as intended in the programmes of study.** This definition is provided here to serve as an advance organiser for the discussion which follows. This research project is still actively going on at the National Institute of Education, Republic of Singapore.

A THEORY OF KNOWING AND A MODEL OF LEARNING

A. Constructivism as a Theory of Knowing

Constructivism is a set of perspectives that collectively guide contemporary developments in pedagogy of learning. It maintains that learning is a **meaning construction** process by which one seeks to organise and make sense of one's experiential world. Knowledge is not transmitted nor imparted as objective reality from teachers to learners. Instead, we construct **viable knowledge schemes** from our everyday experiences for **problem-solving in specific contexts.** This construction process is guided by the specific **habits of the mind** and **prior conceptions** of the learners, allowing them to make **claims** about what they value, believe or feel, and also about how they **perceive** the way the world works. These specificities result in a proliferation of **alternative frameworks** or conceptions despite the following of **objective** ways of enquiry. Furthermore, in order that the personal knowledge that is constructed should have any status, it has to be socially-acknowledged and publicly-mediated. These lead Cheung and Taylor (1991) to assert that **social justification** and **specific objectivity** are the hallmarks of constructivism.

B. Cheung and Taylor's (1991) Humanistic and Constructivist Model of Knowing

This is an epistemological model of knowing, with specific reference made to school science (Cheung & Taylor, 1991). It is also a valuable tool for task analysis, as well as for designing activities for meaningful learning purposes (Qualter et al., 1990). The model extends **Piaget's Spiral of Knowing** to a Double-Spiral, which describes the duality of two basic processes of knowing within an **experienced environment** or a **perceived task context** (see Figure 1). In essence, the process of knowing, a hybrid resultant of maturation and socio-experiential learning, involves an interplay of two basic kinds of knowledge, viz. **conceptual knowledge** with understanding and **procedural knowledge** with proficiency. The Double-Spiral -- an inverted, open-ended, and ever-widening cone -- represents the continuous endogenous cognitive process of **reflexive abstraction**, whereas a hypothesised peripheral envelope, known as the mind-environment interface, represents the process of **empirical abstraction** of the learning experiences from the environment. The process of construction of conceptual and procedural knowledge schemes are self-regulatory and its mechanisms are assimilation and accommodation. Within this framework, **progression** is defined in terms of one's ability to cope with tasks of different degrees of conceptual and procedural complexity, to handle tasks requiring different problem-solving approaches, and to construct more differentiated, integrated knowledge schemes so as to attain higher levels of conceptual understanding and procedural proficiency.

Insert Figure 1 about here

C. Cheung's (1989) Model of School Learning

This model explicates the context elements and process variables which help to bring learners to attain task mastery. The model depicts the educational enterprise as multilevel in structure, ie. learners are clustered within classrooms, classrooms within schools, schools within communities, and so forth. The nested structure implies that the wider community context influences the school environment and its processes, which in turn affect the classroom climate and its events, and these processes ultimately bear upon the learners' various lines of development. One notable feature of this multilevel model of school learning (Cheung 1989, p.45-73) is that it is firmly rooted in constructivism. For the purpose of meaningful measurement, only the sub-model targeted at the learners' level is shown in Figure 2. There is another level focused at the institution where learning is to take place.

Insert Figure 2 about here

The model at the learner's level views task engagement as essential for learning to take place, and it sees comprehension as a meaning construction process. Motivation and attention are the necessary antecedents for deep cognitive processing to take place. **Learning becomes difficult** when students cannot construct meanings by linking and revising their knowledge schemes. The concept of readiness is extended not only to include domain-specific aptitudes and prior learning, but also to the performance goals and structured learning conditions. The domain-specific learning context is set within the wider learning context under the continuous influences of heredity, the home and society. Personality and value orientations are considered as important in attitudinal and motivational style development. They are also important in the pursuance of learning goals. In the model, the notion of comprehension is interpreted with reference to the humanistic, constructivist model of knowing already discussed. Furthermore, it is viewed as a summary mediating process involving task perception, cognitive and memory processes in order to achieve task mastery. This model has been used successfully to model the within-classroom science learning processes at the sixth form level in Hong Kong (Cheung, 1991b).

OTHER USEFUL CONTEMPORARY PERSPECTIVES OF KNOWING AND MODELS OF LEARNING

There are some other contemporary perspectives of knowing and models of learning which are equally legitimate for guiding a principle of meaningful measurement. However, these are mostly not incompatible with Figures 1 and 2. Actually, the ones proposed here follow Einstein's advice that "Everything should be made as simple as possible, but not simpler". Altogether, there are five perspectives/models to be considered for comparison purposes.

A. Greeno's (1991) Model of Knowing as Living in Environments

Greeno views the process of knowing in a conceptual domain as living in an environment within which activities are conducted, resources are made available and utilised, physical situations are perceived and represented, mental models are constructed providing **affordances** for reasoning, and cognitive expertise progressively develops. In his view, flexible mental computation, numerical estimation, and quantitative judgement and inferences are examples of "number sense" which may not be learned through direct instruction because cognitive expertise of "number sense" develops gradually as a result of exploring and relating numbers in a range of contexts, particularly the social collaborative ones, involving holistic and configural understanding rather than based on deliberate rule-based procedures.

B. Sternberg's (1985) Triarchic Model of Human Intelligence

The Model comprises the componential, experiential, and contextual subtheories. Firstly, the **componential subtheory** contains three process components linking intelligent thinking to one's inner worlds. There are executive processes used to plan, monitor, and evaluate one's problem solving (metacomponents); nonexecutive processes used to implement the instructions of the metacomponents (performance components); nonexecutive processes used to learn how to solve problems in the first place (knowledge-acquisition components).

Secondly, the **experiential subtheory** relates intelligent thinking to experiences. It accounts for problem-solving behaviours for relatively novel tasks, and those in the process of becoming automatized. Differential intellectual functioning is to be differentiated from differential past practices when accounting for how expertise develops. Thirdly, the **contextual subtheory** deals with the interactions between intelligent thinking and external world. This may include adaptation to and the shaping of the existing environment, and the selection of new environments. According to Sternberg (1991), attention should be paid on recognising and defining problems, resource allocation, reasoning-based performance and its components, ability of knowledge acquisition, synthetic and insightful thinking, rate and asymptote of automatization of information processing, and tacit knowledge of everyday life.

C. Bereiter's (1990) Educationally Relevant Learning Theory

Bereiter views an educationally relevant learning theory as one which is capable of explaining **processes of difficult learning**, particularly on topics that require new or more complex mental structures than that students already possess. Also, this theory should fulfil traditional functions of other learning theories to explain the process of knowing and enculturation of an individual. This raises the issue that learning is a **multisource** phenomenon. Students actually have a variety of resources available that can help in bootstrapping themselves towards higher levels of cognitive functioning. In recent cognitive science research, autonomous cognitive units of chunks and situated models have been used to describe processes of knowing and its associated knowledge structures.

However, in order to address school learning problems theoretically, the notion of a **contextual module**, which captures the properties of human thoughts and actions and treats environmental situatedness as an emergent property resulting from reasonably well-understood cognitive processes and learning contexts, is proposed as a viable unit of description of difficult learning. For Bereiter, a contextual module comprises of knowledge schemes, goal structures, problem models, affect, self concept, and code of conduct. As an analogy, a contextual module stands as an organic whole in relation to knowledge structures as text set within a text-base stands in relation to its sentences. The notion of contextual module agrees well with the finding that when expertise develops, some knowledge and skills acquired begin to take on modular characteristics that can be very adaptive to very specific real-world contexts.

D. Newell and Simon's (1972) Information Processing Problem-Solving Model

Newell and Simon view problem-solving as an information processing activity in which a problem solver searches from the initial to the goal states within a problem space, which is a mental representation of a task environment. **Chunking of information**, stored within the long-term memory, is viewed as a basic process of learning. Backed by a domain-specific knowledge base in the memory store, this cognitive process of recognising a **deep structure** and not being distracted by the **surface features** of problem tasks differentiates between experts and novices for successful problem-solving. Transfer of problem solving skills can be classified as near/distant transfer and specific/ generalised transfer, depending on the extent of the differences in stimulus features of the task environments, as well as the specificity of the information, skills, or heuristics of the problem tasks respectively (Burton & Magliaro, 1987-1988). On the other hand, Salomon and Perkins (1987) distinguished between **high-road/low-road transfer**. According to them, high-road transfer involves mindful abstraction from one problem context and application to another, whereas low-road transfer occurs when a performance practised to near automaticity in one problem context becomes activated spontaneously in another context.

E. Shuell's (1990) Phases of Meaningful Learning

Shuell noted that during long-term learning there are qualitative changes which are the result of progression of knowledge-based competencies within a particular content domain (Shuell, 1990; Chi, 1978). Shuell has identified three phases of such changes. The **initial phase** involves memorising facts and using preexisting schemata to interpret the isolated pieces of data such that the knowledge acquired is concrete and context-bound. Thus, mnemonic strategies as a form of elaborative encoding is expected to have greater effect on learning than chunking as a form of reductive encoding. The **intermediate phase** entails relating conceptually isolated pieces of information into higher order cognitive structures and networks. The knowledge acquired begins to be decontextualisd and generalised as a consequence of clarifying concepts by relevant examples and applying concepts to different relevant situations. During the **terminal phase**, a learner becomes proficient in handling different task requirements since the knowledge structures formed become more integrated and function more autonomously. Shuell admitted that the knowledge of transitions between phases escapes current understanding and the ways emotions and affects play their part are far from being clear.

In summary, the purpose of recapitulating these relevant theories of knowing and models of learning is to illustrate the unity of these ideas when the two models in Figures 1 and 2 are reexamined within these perspectives. In particular, Greeno's concept of affordances, Newell and Simon's problem and search spaces, and Shuell's phases of learning are significant in understanding task demands and types of progression when task analyses are undertaken and problem-solving behaviours are analysed. Sternberg's triarchic model, Bereiter's contextual module, and Salomon and Perkin's ideas of high-road transfer affirm the multisource, context-bound, socio-experiential nature of learning.

One notable feature of these relevant perspectives is that in order to understand learning difficulties, one has to examine simultaneously many aspects of the learning process organised as an organic whole. Thus, one can regard the model of knowing in Figure 1 as a "zoom-in" version of the model of school learning in Figure 2 when the processes of comprehension within a schooling context are the focus of attention. **With these in mind, one can assert that the ultimate aim of meaningful measurement of the processes and products of learning is to overcome difficult learning, rather than to rank student performances for placement only.** In order to achieve this aim, one has to relate cognition and testing, and thence the need for meaningful measurement.

SOME RESEARCH EFFORTS RELATING COGNITION AND TESTING

There have been some initial successful attempts to develop testing methodologies which incorporate into their measurement procedures a theory of knowing and/or a model of difficult learning so that the measures obtained are meaningful in informing the processes and products of learning. Altogether, three perspectives are discussed here because of their potential contribution for a technology of meaningful measurement.

A. Willett's (1989) Individual Growth Trajectory

Willett noted that learning implies growth and this should be considered as a process of continuous development over time. Unfortunately, this process is often measured by a gain score following an increment model by comparing an individual's status before and after a learning occasion, not knowing that the measures may not remain construct valid nor equatable over time. Moreover, the notion of initial status is not specifically clear and hence its empirical relationship with growth measures is often confusing. According to Willett, what is more desirable is an explicit model of growth with a longitudinal perspective and a multiple-wave data collection design that can be used as bases for describing **individual growth trajectories** and explaining **interindividual heterogeneity in growth.**

A desirable growth measure is one that provides a clear rationale for initial status, maximises discrimination amongst individual growth trajectories, and minimises errors of measurement. The use of **Hierarchical Linear Modelling** is known to be one elegant way to handle quantitative measurement of growth (Raudenbush, 1988; Cheung, Keeves, Sellin & Tsoi, 1990a). However, it should be noted that in order to describe a growth trajectory and explain interindividual heterogeneity in growth, models of knowing and learning such as those in Figures 1 and 2 are urgently needed. Likewise, Shuell's proposed phases of learning are valuable as an alternative perspective informing paths of learning.

B. Biggs and Collis's (1982) SOLO Taxonomy

Gagne (1962) formulated the notion of a learning hierarchy some thirty years ago. Since then, amongst others, Biggs and Collis (1982 & 1989) developed the SOLO Taxonomy (Structure of Observed Learning Outcomes) so that responses to tasks are assessed for **levels of information processing** on their relevant use of pieces of information in a related manner for integrated understanding. Recently, errors made at different SOLO levels of a learning hierarchy are found to concur with their designed levels of information demand, thus substantiating the theoretical underpinnings of the SOLO Taxonomy (Yip & Cheung, 1991).

According to the SOLO Taxonomy, five basic levels of a **learning cycle** can be distinguished, namely, prestructural, unistructural, multistructural, relational, and extended abstract. The beginning and finishing levels mark the progression from one **mode of cognitive functioning** to its next. In sequence of progression, there are five basic modes of cognitive functioning : sensori-motor, iconic, concrete symbolic, formal and post-formal. These modes are responsible for constructing the tacit, intuitive, declarative, and theoretical human knowledge systems respectively. Parallel to these conceptualisations are **curriculum goals** which comprise the development of motor skills, values and aesthetics, symbol systems, knowledge of subject disciplines as well as alternative theories. Biggs and Collis (1989) argued that multi- and inter-modal cognitive functioning in problem-solving and for integrated understanding are desirable educational goals.

C. The van Hieles's Model of Geometrical Thinking

The van Hieles Model (van Hiele, 1957 & 1986; van Hiele-Geldof, 1957) identifies five progressive levels of geometrical thinking, ranging from recognising and visualising geometrical figures by global appearances, analyzing properties of figures and shapes, relating figures and ordering of properties identified without resort to a mathematical system, deducing relationships through the use of proofs within a mathematical system, and working in a variety of axiomatic systems with rigour and understanding. To help from progressing from one level to its next are **five phases of instruction**, namely, inquiry, directed orientation, expliciting, free orientation, and integration, showing a gradual move from very direct instruction to students' independence of the teacher (Hoffer, 1983). By applying the Category Theory and considering the five phases of instruction as a functor (Maclane, 1971), the van Hieles Model can be extended to other topic areas by defining the base elements and their properties, relationships of these properties, partial orderings of these relationships, and properties of these partial orderings.

In summary, although there have been concerted efforts in relating cognition and testing and venturing into the measurement of growth trajectories, the two quantitative and qualitative research camps have not combined their efforts that are sufficient for a comprehensive technology of meaningful measurement. It is noted that the types of progression that may be modelled and the requirements of quantitative measurement have not been explicitly recognised. Consequently, the following two sections suggest the problem sets meaningful measurement are intended to handle and the requirements of quantitative measurement in the light of contemporary advances in cognitive sciences and psychometrics.

PROBLEM SETS OF MEANINGFUL MEASUREMENT

Two topics in Mathematics are provided here to illustrate what type of processes and products of learning that meaningful measurement is intended to handle : (1) the cognitive processes of addition of numbers; and (2) the discontinuity in cognitive processes between arithmetic and algebra. It should be noted that both topics are already very well-researched. Nevertheless, the quantitative measurement of levels of understanding and degrees of competencies is far from adequate for a viable pedagogy in the classroom.

A. Addition and Subtraction of Numbers

Bergeron and Herscovics (1990) viewed the conceptions of number teleologically, both as a measure of plurality (number as size) and a measure of position (number as rank). They clarified Piaget's concept of Conservation of Number by drawing a distinction between Conservation of Plurality (number of objects in a given set) and Quotity (rank of an object in an ordered set). As early as 1976, Klahr and Wallace postulated three distinct but **intertwined quantification processes** in the acquisition of number concepts in early arithmetic. These processes are : subsitizing (instant recognition of a number associated with a configuration), counting and estimating. Subsitizing skills are argued to develop first, followed concurrently by counting and estimating. However, competence in estimating normally reaches maturity later because of the need for other componential skills and resources in learning.

Cognitive science research shows that children around 3-years-old start to learn the number-word sequence first. This competence becomes more elaborate as children grow indicating their progressive levels of development (Fuson et al., 1982). Then, counting is developed as a one-to-one correspondence between the objects to be counted and the counting words. This correspondence is done by means of three distinct procedures : visual-counting, touch-counting, and physical partitioning (Herscovics, Bergeron, & Bergeron, 1986). Bergeron and Herscovics (1990) noted that spatial-physical transformations of objects may affect the procedures and accuracy of counting, which in turn have influences upon the thinking processes of early addition and subtraction problems.

That task context has an effect on cognition is reflected in children's different uses of direct modelling with physical objects, verbal counting, and mental strategies in handling simple additive structures. Children's knowledge of positional notation of a 2-digit number also manifests different levels of understanding, from juxtaposition (position of two digits viewed as not a matter of concern), chronological (the order of the production of the 2-digit number is a matter of concern), and onwards to conventional (the rule of order in representing the 2-digit number is respected). **One notable finding is that a child's performance is often moderated by both the semantic and syntactic features of the assessment tasks.**

B. Discontinuities between Arithmetic and Algebra

Algebra can be learnt by explicitly extracting the operations and relations from given problem situations and expressing these relations using algebraic symbolism. Unfortunately, children find school algebra difficult because there are **discontinuities in cognitive functioning** between their earlier acquired informal methods and symbols in arithmetic and the more systematic procedures used in the symbolic representation of objects and solution of algebraic expressions (Booth, 1984). Understanding of algebraic notions may proceed in stages (Kieran, 1981). Examples of these stages of conceptual understanding are : (1) the meaning of the equal sign, which reveals children's confusion in viewing it as a "do-something" sign rather than a relational symbol indicating equivalence; and (2) treating letters in algebraic expressions as objects rather than as variables or specific unknowns.

There are also three main types of problem-solving approaches when used in certain combinations show students' differential progression in solving algebraic expressions : (1) intuitive (such as use of number facts, counting techniques, and cover methods), (2) trial-and-error substitution, and (3) formal (Kieran, 1990). Again, a notable finding is that a change in task context often leads to a different but inconsistent structuring of the algebraic expression. Even within the formal approaches, there are two different types of cognitive functioning : (1) transposing signs and sides techniques, and (2) performing the same operations on both sides of the equations. In this regard, one teaching experiment shows greatest effect on learning when substitution solving method are accompanied by balancing operations between both sides (Kieran, 1983). Furthermore, Sfard (1987) provides some evidence of the predominance of procedural rather than conceptual understanding amongst children. This is evident when conceptions of an algebraic function as an object than a procedure is to be preferred. **This points to the necessity in understanding the intertwining relationships between procedural and conceptual knowledge.**

QUANTITATIVE REQUIREMENTS OF MEANINGFUL MEASUREMENT

The challenge of meaningful measurement is to specify requirements of quantitative measurement so that the resulting measures can enjoy the same status as those in scientific measurement. The following two quotes from Cheung (1990) explain some of these important requirements:

> 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 often are the inseparable parts of the trait to be constructed ... 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 measures (emphasis added, p.5).
>
> 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 (emphases added, p.6).

One caveat is mentioned here. The use of the Rasch family item response models does not in principle render the measurement process a meaningful one. A theory of knowing and a model of difficult learning within a specific schooling context is required. Notwithstanding this, it is useful to understand the simple Rasch logistic model by considering the responding process when a respondent is confronted with a problem task which affords certain sets of competencies for its mastery. The product term of the demonstrated level of competencies of a respondent and the designed level of affordances of a problem task can then be equated to the odds (ratio of success to failure) of accomplishing the problem task. By taking logarithms on both measures of competencies and affordances in order to transform these scales of progression into a **mutually conformable** measuring scale, the resulting probability model governing a respondent's mastery of the problem task becomes the famous Rasch item response model (Rasch, 1960). **This revised understanding of the Rasch Model in terms of both terminology and conceptual framework acknowledges that not only the knowledge construction process is specifically objective, but also a similar situation occurs for the calibration and measurement of levels of understanding and proficiency of this acquired knowledge.**

In practice, achieving conformity between measurement models and response data is often the decisive step of the Rasch analyses. Rasch (1977) asserted that latent trait modelling calls for clear lawful relationships in which the three aspects -- (1) the definition of each variable involved; (2) the relationships amongst variables; and (3) their measurement -- are simultaneous. Consequently, that the item writing process is the key process in the construction of the measurement scale is not only compatible with the requirement of meaningful measurement to be firmly rooted in the classroom realities and educational objectives of the programmes of study, but also with this hallmark of Rasch modelling -- the **simultaneity** between conceptualisation of measures and its measurement.

ISSUES FOR A TECHNOLOGY OF MEANINGFUL MEASUREMENT

A number of key issues are pertinent for a technology of meaningful measurement. Firstly, characteristics of children's thinking such as those described in the previous two problem sets need to be understood. Concept mapping is considered as a valuable tool for the diagnosis of children's alternative conceptions. These characteristics and cognitive barriers are useful in summarising learning difficulties which may be specifically built into the measurement instrument so as to diagnose difficult learning.

Secondly, the concepts of progression and affordance within an experienced environment or perceived task context (see Figure 1) need to be exemplified to serve as a working model for the purposes of test design and construction. This is illustrated by two metaphors -- the "bamboo stem/pig-tail" and "competence ladder" -- which have been proved viable for an application of psychometrics such as the Rasch, Network and Dual analyses.

Thirdly, a test construction procedure, analogous to those based on a test of specification spanned by the content areas and Bloom's Taxonomy, is needed. In this regard, Vergnaud's (1990) theory of conceptual field, which is compatible with the two models in Figures 1 and 2, is an extremely useful heuristic. For those selected-response type questions such as the multiple-choice, it is argued that the test items, which collectively constitute the conceptual field, should weave and carve out the competence schemes such that the progressive forms of knowing are clearly articulated. The distractors of the multiple choice items should comprise of the alternative conceptions viable to the respondents.

Lastly, for those constructed-response questions, there is a need for an analytical tool which explicates the deep structure(s) of a set of problems and summarises the problem and search spaces of the respondents. Cheung and Mooi's (1991a) network analysis for problem-solver's action has been developed for this purpose. The ensuing sections discuss these four issues with illustrative examples.

A. Alternative Conceptions, Conceptual and Procedural Barriers

The characteristics of the children's thinking need to be understood. Contrary to the deductive and inductive modes of logical thinking, children's reasoning, based on salient perceptual features, very often move from particular to particular, which is a form of transduction. On one hand, children's thoughts suffer from a globality in both thoughts and perceptions, and on the other, a temporal linkage of individual thoughts and perceptions that never add up to a coherent whole. Epistemological commitment to consistency is not appreciated. These pose problems when problem-solving activities are assessed because of the detours and different solution routes in order to arrive at the goal state. Furthermore, children when faced with contradictions often hold on to some ideas and readily change others. As a result, children's explanations often lack transferability across problem contexts. This lack of transferability demands a need for children to distinguish amongst the contexts within which particular conceptions are understood. Otherwise, the possibility of high-road transfer is not favourable.

In meaningful measurement, learners' alternative conceptions, conceptual and procedural barriers within meaningful learning contexts are valuable signposts and resources for test construction. For example, the alternative conceptions which are viable conceptions to the weaker learners can serve as distractors to multiple choice questions, not as options that may be guessed when the keys cannot be secured. **The use of concept mapping, a form of semantic network, is potentially useful to probe children's conceptual frameworks by verbal means** (eg. Novak, Gowin, & Johansen, 1983). A study of these maps generally show that children's alternative conceptions can be very convoluted and elaborated. Those produced by advanced learners show more hierarchical, cross-linking and superordinate conceptual structures than those of the novices. In deploying these concept maps, children's own references to phenomenological experiences and language use should be understood. Nowadays, concept mapping is one of the most popular method in diagnosing alternative conceptions and is an indispensable tool for the purposes of meaningful measurement.

Cheung (1991a) gave a cautionary note regarding the alternative conceptions and errors made by learners of different levels and types of abilities and these should be borne in mind when grading children's test responses.

> Pupils of different abilities commit different types of errors. Very often, errors committed by the high-ability pupils may not be observed in the low-ability pupils. This is because the low-ability pupils may not even understand or represent the problem at all (p.8).

If the ultimate aim of meaningful measurement is to overcome difficult learning, then diagnosis of the conceptual and procedural barriers should be incorporated into the measuring instrument. A literature review of learning difficulties or interviews with classroom teachers would be useful in this regard. The following example from Koh and Cheung (1990) presents a summary of some difficulties in understanding the part-whole concept of fractions. A test based on these difficulties have been constructed by them.

<u>An Example: Difficulties in Learning Fractions</u>

1. Pupils do not understand the notion of equi-partition when parts are related to the whole. For those who do understand, some still lack the necessary skills to do the partitioning adequately.

2. The language used by teachers in naming and denoting a fraction, such as "a third", "one over three", or "1/3", has created cognitive barriers especially for lower primary pupils.

3. Children have perceptual difficulties in interpreting diagrams to represent fractions. Perceptual cues and distractors can both facilitate and hinder children's thinking.

4. The acquired concepts of fractions lack transfer across problem situations because the competence schemes have not fully developed yet.

5. Children have difficulties translating between different models of representation such as between the continuous and the discrete models.

B. Two Metaphors for a Model of Progression and Affordance

As guided by Figure 1, test design requires an attention paid to the measurement of levels of conceptual understanding and procedural proficiency, as well as the intertwining relationships between the conceptual and procedural knowledge. The two metaphors that have proved useful for a conceptualisation of progression and affordance are : (1) the bamboo stem/pig-tail metaphor of progression (Cheung, 1990); and (2) the competence ladder metaphor of cognitive structures (Cheung, 1991a). The following two explanatory quotes are useful for an understanding of these two metaphors:

1. The Bamboo Stem/Pig-tail Metaphor of Progression

> 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 linear 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 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. (Cheung, 1990, p.4-5, emphasis original).

2. The Competence Ladder Metaphor of Cognitive Structures

> Pupils of different abilities commit different types of errors ... Errors can be loosely regarded as "inabilities" of the pupils ... If meaningful measurement is to be pursued, the issue of how these "inabilities" are structured can inform us of the conceptual and procedural barriers towards mastery of problem solving. Errors made by the pupils cannot normally be regarded as unidimensional. Nevertheless, some of these may be hierarchical in nature. Latent trait theories cannot be applied to model the structure of errors. Instead, key types of errors of the questions in the test can be tabulated against the band levels of the problem-solving proficiency continuum, showing the types of errors made by pupils of different levels of problem-solving ability ... The discussions lead to a very important concept on the meaningful assessment of problem-solving activities: The Competence Ladder. **On the left arm of the ladder is a "pig-tail fashioned as a bamboo stem", which represents the progression of conceptual and procedural knowing along the problem-solving ability continuum. The progressive, qualitative bands of this "pig-tail" mark the positions of the rungs of the ladder. The rungs symbolically link the left arm of the proficiency continuum to the right arm of the key structural dimension(s) of errors. From this competence ladder, we know not only how pupils of different abilities perform,**

but also what types of errors they make. (Cheung, 1991a, p.9-10, emphasis original).

C. Conceptual Fields and Competence Schemes

Vergnaud (1990) introduces the theory of conceptual fields in order to explain how children construct mathematical concepts. The theory treats concepts as multi-faceted and learning as multi-source in nature. A concept consists of a **triplet of three sets** -- the **referent**, the **signified**, and the **signifier** -- showing how sets of situations utilise several kinds of concepts, procedures, and symbolic representations that are connected to one another (Narode, 1987). Specifically, the referent set consists of situations that make the concept meaningful. The signified set are invariants that constitute the concept, whereas the signifier set are iconic or symbolic representations that can be used to represent these properties and situations.

Within this framework, competencies are analyzed as organised conceptual and procedural knowledge schemes in the form of implicit concepts-in-action or theorems-in-action whose invariant properties should correspond to those of the situations under consideration. Transformation rules of these implicit invariants, and the ways linguistic signifiers and iconic/symbolic representations have taken place in these knowledge schemes are studied to learn how the mathematical objects in a conceptual field may be better identified and become progressively as real as things in physical reality. **In short, a competence scheme is an evolving network of relevant conceptual and procedural knowledge structures -- ie. progressive forms of knowing -- together with all the problem situations that may be afforded by them.**

1. An Example of a Conceptual Field

This is an example of the measurement of the part-whole concept of fraction (Koh & Cheung, 1990 & 1991). The referent set contains concrete instances of fractions that may be experienced, such as one-half of a pie, three-quarter of a square, one-third of a set of pencils and so forth. The signified set consists of invariants constituting the part-whole concept, such as notions of equipartition and a whole, as well as notions of relating parts of equal-sized units to the whole through the deployment of counting and partitioning schemes. The signifier set comprises the different models of representation that are used to represent the situations and signify the invariant concepts, such as the continuous (eg. a pie/square) and discrete (eg. pencils/stars) model of representation. Figure 3 shows some items that are meant to span the conceptual field of the part-whole concept of fraction.

Insert Figure 3 about here

2. An Example of a Competence Scheme

The "bamboo stem/pig-tail" metaphor of progression is valuable in serving as a working model for conceptualising the competence schemes of the part-whole concept of fractions. Progressive forms of knowing can be identified, ranging from the lower level of recognising names and symbolic notations of fractions, interpreting diagrammatical representations (continuous versus discrete) of fractions and using diagrams to represent fractions, to the higher level of applying the part-whole concept of fractions to more complex situations. Within each level, the item difficulties reflect the conceptual and procedural demands the questions afford and they may be moderated by the different models of representation, types of fractions, and types of perceptual cues or distractors. Counting schemes may be used alone, or in combination with the partitioning scheme, in order to relate the equi-partitioned parts to the whole. Table 1 details a set of competence schemes with suggested levels of understanding of the part-whole concept of fractions.

Insert Table 1 about here

D. Analyses for Problem-Solver's Actions

A theory of problem-solving hinges upon what constitutes a problem which can be delineated on a continuum ranging from routine exercises to non-routine problems. Cheung and Mooi (1991a) set the foundation for the meaningful assessment of problem-solving tasks. They proposed and explicated a theory of perception logic based on connectionism to account for how perceptual cues brings in past experiences for the construction of human thoughts, a theory of information processing that is responsible for generative learning, and a version of problem-solving cycle. There are altogether five key processes which may not proceed in sequence in a problem-solving cycle : (1) from problem-perception to **problem-understanding**; (2) a recognition and explication of relevant schema in **problem-representation**; (3) the use of formalised, goal-oriented rules or heuristic strategies in **problem-execution**; (4) the exercise of meta-level decision-making processes in monitoring the solution paths in **problem-control**; and (5) the formative and summative evaluation of the problem-solving processes in **problem-evaluation**.

This version of problem-solving cycle owes its contribution from Newell and Simon (1972) by focusing on how the critical features of the task environment structure the mental representations of the problem (problem space) in which heuristic strategies may be deployed for the search of viable solution paths (search space). It is noted that a task environment with no recognisable critical features indicative of a **deep structure** and its associated knowledge base may result in an ambiguous problem representation that may be based on the **surface features** only. For the purpose of meaningful measurement, Cheung (1991a) gave the following working definition on what is meant by a deep structure of a problem-solving task :

> In this monograph, questions are said to possess the same "deep structure" if they can be solved by the pupils in particular ways as intended by the teachers, once the pertinent perceptual cues that are built into the questions are recognised. Also, associated with each "deep structure" is a **domain-specific knowledge base** targeted at the ability level of the pupils under examination (emphasis added, p.4).

This working definition allows us to use network analysis for mapping a problem-solver's actions, both for test design and analysis. Cheung and Mooi (1991a) called the resulting maps as **problem-solving networks**. They further likened these maps to the personal constructs obtained in Kelly's (1955) Repertory Grid analysis so as to ascertain the scientific status of this way of understanding how people engage in problem-solving. Since human knowledge is both personally and socially constructed, the network of qualitative data representing such knowledge is necessarily linked to a theory of knowing such as the one in Figure 1. Moreover, the contribution of the emotional system such as motivational and attentional processes depicted in Figure 2 have not been ignored when episodes of problem-solving are analysed.

<u>1. An Example of a Problem-Solving Network</u>

Figure 4 shows an open-ended algebra test for secondary two (grade 8) pupils in Singapore. Figure 5 presents the common designed deep structure of the two core items, item 10 and 11. Loh and Cheung (1991) noted that since the purpose of the test is to establish a progressive, qualitative continuum indicating how competencies may be deployed and barriers may be surmounted, conceptual knowledge and procedural skills leading to the successful solving of items 10 and 11 have to be built into items 1 to 9 of the test. Cheung (1991a) has further explained :

> The rate-determining steps of a question can be anywhere along the solution path. Should the most difficult step be problem-understanding or problem-representation, pupils should find it relatively straight forward to proceed to completion once this step is secured. Otherwise, pupils may be blocked and are denied the opportunity to complete the ensuing steps. If this is the case, the ability levels of the problem-solving steps following the rate-determining steps may not be estimated accurately. Thus, after a set of core questions of the same "deep structure" have been constructed, it is important to include additional questions targeting at the components of the core questions. In this way, conceptual and procedural knowledge, which are needed to solve the core questions, can also be assessed accurately (p.7).

Insert Figures 4 and 5 about here

---

2. An Example of "Bamboo Stem/Pig-tail" and "Competence Ladder"

Based upon the structure of competencies and/or procedural steps of each item in the algebra test, grading students' responses of each item into an ordinal scale of performance levels is then made possible through the use of Partial Credit Modelling (Masters, 1982). One merit of this modelling procedure is that it can tackle procedural steps which are not necessarily ordinal in difficulty. If this happens, some intermediate item performance levels would be less likely to obtain because once the cognitive or procedural barriers have been surmounted the ensuing steps simply follows. Another is that since it is an extension of the Rasch Model, so all the desirable requirements of quantitative measurement such as linearity and specific objectivity are also applicable.

Table 2 shows an item map of the algebra test, showing how the progressive forms of conceptual and procedural knowing relate to the individual item performance levels which are defined by the structure of competencies and procedural steps (for details of these competencies and procedural steps of the 11 algebra items, see Loh and Cheung, 1991). It also gives the conversion of the raw test score, a simple summation of the items' performance levels, with the linearised logit score. A student is thus assigned a problem-solving competence score, which may be interpreted against the four bands of progressive forms of knowing. If required, this can be interpreted against the competencies/procedural steps as represented by the item performance levels. **Consequently, this algebra problem-solving proficiency continuum serves as a realisation of the "bamboo stem/pig-tail" progression discussed earlier.**

Insert Table 2 about here

Five items along this problem-solving proficiency continuum are selected. These items (3, 4, 5, 8, 10) represent a list of concepts, procedural skills, and key types of errors in solving algebra word problems. The dimensional and hierarchical nature of the errors can be ascertained by Dual Scaling (Nishisato, 1980). Table 3 shows a cross-tabulation of key types of errors (Category A to N), including no error or omission, with the four bands of performance levels (Levels 1 to 4) in Table 2. Dual analysis of this table results in a multidimensional decomposition of data with the most informative structural dimension extracted first, then the second most informative dimension, and so forth. The two orthogonal structural dimensions shown in Figure 6 account for a total of 93 percent of the information in the data.

Insert Table 3 and Figure 6 about here

The first structural dimension (72% information) clearly is the problem-solving proficiency continuum upon which key types of errors are structured. By making reference to this structural dimension, the "inabilities" as defined by the key types of errors can be hierarchically structured in the same way as the progression of conceptual and procedural knowing in solving algebra word problems (Cheung & Loh, 1991). **Metaphorically, the first structural dimension constitutes the right arm of the competence ladder, which is to be paired with the left arm of the problem-solving proficiency continuum. The four performance levels act symbolically as the rungs linking the progression of the problem-solving steps on one arm, to the hierarchical structure of key types of error on the other.**

MEANINGFUL MEASUREMENT OF AN EMOTIONAL SYSTEM

The multi-source nature of learning and the need to explain difficult learning such as using Bereiter's contextual module draw our attention to the contribution of the affective/emotional system to cognition. As an initial exploration on the feasibility of meaningful measurement of an emotional system, two issues have been identified for research. The first issue is whether within the emotional system there exists affective schemes that can be modelled in the same way as the conceptual and procedural knowledge schemes. The second issue is whether the respondents are deploying the response scale in the same way when responding to the attitudinal/emotional stimuli. The following two sections report research findings answering these issues and their implications.

A. Affective Schemes and Forms of Affectivity

Mooi and Cheung (1990-91) showed how the measurement of a learner's anxiety towards computer programming can be made meaningful such that the measures are firmly rooted in the classroom realities and educational objectives of the computer programming course. The study shows that both state and trait anxieties can be modelled using latent trait theories. Specifically, they started to examine the nature and manifestations of anxiety towards computer programming, with specific reference to the Junior College students (grade 12) in the Republic of Singapore. State anxiety refers to an unpleasant emotional state of mind such as those that are aroused when confronted with **errors** or difficulties while programming, when in the presence of more capable peers or **significant others**, or when assessed by their teachers. Trait anxiety refers to relatively stable individual differences in anxiety proneness as a personality trait, such as a learner's feelings of inadequacy or lack of **confidence** when doing programming assignments.

Althoug the Pearson correlations of the three anxiety measures -- "Error", "Significant Others", and "Confidence" -- are highly intercorrelated (r ranges from 0.61 to 0.73), it is the progressive **forms of affectivity** of each of the three types of anxiety and the intertwining relationships of the three types of

affective schemes interpreted in the light of classroom realities and course objectives the highlight of the study (see Table 4). Technically, this study employs Rating Scale Analysis on anxiety subscales with a 5-point Likert-format. Consequently, tenability of homogeneity in the use of the response scale and specifically objective anxiety measures are simultaneously established by outcome.

Insert Table 4 about here

B. Deployment of the Response Scale

Despite the success of Mooi and Cheung's (1990/91) study, the quantitative assessment of the emotional system still relies on the more popular scaling procedures, such as ratings, Guttman, Thurstone and Likert scaling. This is because conducting meaningful measurement of the emotional system is not easy. **Conformity** of the implicit item-response model of the measuring instrument with the dimensionality and structural properties of the psychological attributes to be measured is the most critical criterion. Very often, one would find even the powerful latent trait theories fail to account for the observed responses and unveil the affective schemes and forms of affectivity. **Consequently, before one embarks on a modelling of affective behaviour, one key question to ask is whether the respondents deploy the response scale in the same way when the probing stimuli are applied to them.** For example, the number and spacing as well as the position of the neutral point in the Likert response scale need to be clarified. This is because respondents may respond to the stimuli in different manners according to their habits of the mind, instead of based on the psychological attributes which are to be measured.

Cheung and Mooi (1991b) set out to explore this important question concerning the deployment of a Likert-type response scale by assessing students' liking for computer-related activities. There are altogether two subscales, "General" and "Specific". They hoped that through clarifying the tenability of the response scale of the two subscales the latent trait modelling procedures with stronger assumptions and requirements may be applied. By using a response model-free Dual Scaling approach (Nishisato, 1980), the intercorrelated "General" and "Specific" subscales (Pearson r = 0.58) are found to provide a contrast concerning the preference structures of computer-related activities between boys and girls.

Despite Dual Scaling is **response model-free**, seeking to unveil the underlying structural dimensions of a contingency table of response data, the study shows the importance of conformity between respondents and probing stimuli as a fundamental requirement of meaningful measurement. In essence, the Dual Scaling analyses reveal the need for the respondents with a range of affective values to **systematically** discriminate between adjacent categories of a response scale across all probing stimuli which are also with a range of scale values. Consequently, the study shows the **unity** of two powerful scaling methods, the model-free Dual Scaling and model-laden Rating Scale analysis. There is a need for a careful consideration of **targeting** the scale values and response thresholds of the probing stimuli to a distribution of respondents who are of consistent response behaviours. **This would necessitate simultaneous conceptualisation of measures and their measurement, with due attention paid to the development of affective schemes and forms of affectivity, and to the habits of the minds when responding to the sets of probing stimuli.**

MEANINGFUL MEASUREMENT - AN AGENDA FOR RESEARCH

Much of the research work is still actively going on. What has been presented here is just a beginning of a research agenda and the way ahead still needs to be fumbled along. Meaningful measurement is a set of perspectives on how to model human understanding and capabilities, capitalising on advances in cognitive science and psychometrics. If successful, it enjoys not only the same status as scientific measurement, but also the same philosophy of quantification. The multisource nature of learning and that viewing knowing as a meaning construction process will continue to guide the thrusts of meaningful measurement. Some fruitful lines of research have now been identified. These form the next phase of study conducted under the rubric of meaningful measurement. These projects are briefly mentioned below.

1. Modelling of the acquisition of number concepts, especially addition and subtraction of integers, following the example on the part-whole concept of fractions.

2. Examination of the inter-relationships between the conceptual and procedural demands of problem tasks so as to understand the notions of high-road transfer and discontinuity of competence schemes when the acquired understanding of part-whole concept of fractions is to facilitate acquisition of the concept of equivalent fractions.

3. Examination of the inter-relationships between the conceptual and procedural demands of problem tasks so as to understand the notions of high-road transfer and discontinuity of competence schemes when the acquired understanding of arithmetic is to facilitate acquisition of problem-solving of algebra word problems.

3. Examination of metacognitive decision-making behaviours of children of different levels of problem-solving abilities when solving algebra word problems.

4. Examination of the linkages between the cognitive and emotional system by studying how students of different anxiety and interest levels engage in problem-solving of computing problems.

5. Modelling of the van Hieles levels of geometrical thinking using the technology of meaningful measurement and an examination of how these ideas can be applied to other mathematical topic areas.

6. Development of a strategy for the assessment of belief structures, associated with a theory of perception logic.

About the Author:

Dr. K C Cheung obtained his Ph D from the Centre for Educational Studies, King's College, London. His thesis on the modelling of science processes and products won him the Bruce H. Choppin Memorial Award. His research interests cover meaningful measurement and modelling of educational data. He is now serving at the National Institute of Education, Nanyang Technological University, Republic of Singapore.

REFERENCES

Bereiter, C. (1990). Aspects of an Educational Learning Theory. **Review of Educational Research**, 60(4), 603-624.

Bergeron, J.C. & Herscovics N. (1990). Psychological Aspects of Learning Early Arithmetic. In: P. Nesher & J. Kilpatrick (Eds.). **Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education.** New York: Cambridge University Press.

Biggs, J.B. & Collis, K.F. (1982). **Evaluating the Quality of Learning: The SOLO Taxonomy.** New York: Academic Press.

Biggs, J.B. & Collis, K.F. (1989). Towards a School-Based Curriculum Development and Assessment Using the SOLO Taxonomy. **Australian Journal of Education,** 33(2), 151-163.

Booth, L.R. (1984). **Algebra: Children's Strategies and Errors.** Winsor, UK: NFER-Nelson.

Burton, J.K. & Magliaro, S. (1987-1988). Computer Programming and Generalised Problem-Solving Skills: In Search of Direction. **Computers in the Schools,** 4(3/4), 63-90.

Cheung, K.C. (1989). **Product and Process Modelling for Science Achievement at the Sixth Form Level in Hong Kong.** Unpublished Ph.D. Thesis, University of London.

Cheung, K.C. (1990). Bamboo Stems and Pig-tails: Some Thoughts on Meaningful Measurement and Testing in the Classroom. In: K.C. Cheung, W.K. Koh, K.C. Soh, & L.C. Mooi. **Meaningful Measurement in the Classroom Using the Rasch Model: Some Exemplars.** Monograph. Singapore: Institute of Education. ERIC Doc No, TM015767/ED326544.

Cheung, K.C., Keeves, J.P., Sellin, N. & Tsoi, S.C. (1990a). The Analysis of Multilevel Data in Educational Research: Studies of Problems and Their Solutions. **International Journal of Educational Research,** 14(3), 1990.

Cheung, K.C., Koh, W.K., Soh, K.C., & Mooi, L.C. (1990b). **Meaningful Measurement in the Classroom Using the Rasch Model: Some Exemplars.** Monograph. Singapore: Institute of Education. ERIC Doc No, TM015767/ED326544.

Cheung, K.C. (1991a). Climbing up the Competence Ladder : Some Thoughts on Meaningful Assessment of Problem-Solving Tasks in the Classroom. In: K.C. Cheung, L.C. Mooi & W.F. Loh **Meaningful Assessment of Problem-Solving Activities in the Classroom: Some Exemplars.** Monograph. Singapore: National Institute of Education, Nanyang Technological University.

Cheung, K.C. (1991b). PLSPATH Modelling of Within-Classroom Science Learning Processes. **Studies in Educational Evaluation.** (Accepted for publication).

Cheung, K.C. & Loh, W.F. (1991). On Meaningful Measurement: Metacognition and Hierarchical Modelling of Errors in Algebra Word Problems. In: K.C. Cheung, L.C. Mooi & W.F. Loh. **Meaningful Assessment of Problem-Solving Activities in the Classroom: Some Exemplars.** Monograph. Singapore: National Institute of Education, Nanyang Technological University.

Cheung, K.C. & Mooi, L.C. (1991a). A Bane or a Boon ?: Meaningful Assessment of Problem Solving Activities in the Classroom. In: K.C. Cheung, L.C. Mooi & W.F. Loh. **Meaningful Assessment of Problem-Solving Activities in the Classroom: Some Exemplars.** Monograph. Singapore: National Institute of Education, Nanyang Technological University.

Cheung, K.C. & Mooi, L.C. (1991b). **On the Meaningful Assessment of Preference Data: A Dual Scaling Approach to Students' Liking for Computer-Related Activities.** Paper presented to the Regional Symposium on Educational Testing, Beijing, China, 16-20 September, 1991.

Cheung, K.C., Mooi, L.C., & Loh, W.F. (1991). **Meaningful Assessment of Problem-Solving Activities in the Classroom: Some Exemplars.** Monograph. Singapore: National Institute of Education, Nanyang Technological University.

Cheung, K.C. & Taylor R. (1991). Towards a Humanistic Constructivist Model of Science Learning: Changing Perspectives and Research Implications. **Journal of Curriculum Studies**, 23(1), 21-40.

Chi, M.T.H. (1978). Knowledge Structures and Memory Development. In: R.S. Siegler (Ed.). **Children's Thinking: What Develops?** Hillsdale, NJ: Lawrence Erlbaum Associates.

Fuson, K.C., Richards, J., & Briars, D.J. (1982). The Acquisition and Elaboration of the Number Word Sequence. In: C. Brainerd (Ed.). **Progress in Cognitive Development Research: Volume 1. -- Children's Logical and Mathematical Cognition.** New York: Springer-Verlag.

Gagne, R.M. (1962). The Acquisition of Knowledge. **Psychological Review**, 69, 355-65.

Glaser, R. (1991). Expertise and Assessment. In: M.C. Wittrock & E.L. Baker (Eds.). **Testing and Cognition.** New Jersey: Prentice-Hall, Inc.

Greeno, J.G. (1991). Number Sense as Situated Knowing in a Conceptual Domain. **Journal for Research in Mathematics Education**, 22(3), 170-218.

Kelly, G.A. (1955). **The Psychology of Personal Constructs (Volume 1 and 2).** New York: W.W. Norton.

Herscovics, N., Bergeron, J.C., & Bergeron, A. (1986). The Kindergartner's Perception of the Invariance of Number Under Various Transformations. In: G. Lappan & R. Even (Eds.). **Proceedings of the Eighth Annual Meeting of North American Branch of the International Group for the Psychology of Mathematics Education.** East Lansing: Michigan State University.

Hoffer, A. (1983). van Hiele-Based Research. In: R. Lesh & M. Landau (Eds.). **Acquisition of Mathematics Concepts and processes.** New York: Academic Press.

Kieran, C. (1981). Concepts Associated with the Equality Symbol. **Educational Studies in Mathematics,** 12, 317-326.

Kieran, C. (1983). Relationships between Novices' Views of Algebraic Letters and their Use of Symmetric and Asymmetric Equation-Solving Procedures. In: J.C. Bergeron & N. Herscovics (Eds.). **Proceedings of the Fifth Annual Meeting of the North American Branch of the International Group for the Psychology of Mathematics Education.** Montreal, Canada: Universite de Montreal.

Kieran, C. (1990). Cognitive Processes Involved in Learning School Algebra. In: P. Nesher & J. Kilpatrick (Eds.). **Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education.** New York: Cambridge University Press.

Klahr, D. & Wallace, J.G. (1976). **Cognitive Development: An Information Processing View.** Hillsdale, NJ: Erlbaum.

Koh, W.K. & Cheung, K.C. (1990). On Meaningful Measurement: Primary School Pupils' Understanding of Part-Whole Concept of Fractions. In: K.C. Cheung, W.K. Koh, K.C. Soh, & L.C. Mooi. **Meaningful Measurement in the Classroom Using the Rasch Model: Some Exemplars.** Monograph. Singapore: Institute of Education. ERIC Doc No, TM015767/ED326544.

Koh, W.K. & Cheung, K.C. (1991). **On Meaningful Measurement: Progressive Forms of Knowing and Cognitive Barriers of the Part-Whole Concept of Fractions.** Unpublished Manuscript, National Institute of Education, Nanyang Technological University.

Loh, W.F. & Cheung, K.C. (1991). On Meaningful Measurement: Stages of Lower Secondary Pupils' Abilities in Solving Algebraic Word Problems. In K.C. Cheung, L.C. Mooi, & W.F. Loh, (Eds.). **Meaningful Assessment of Problem-Solving Activities in the Classroom: Some Exemplars.** Monograph. Singapore: National Institute of Education, Nanyang Technological University.

MacLane, S. (1971). **Categories for the Working Mathematicians.** New York: Springer-Verlag.

Masters, G.N. (1982). A Rasch Model for Partial Credit Scoring. **Psychometrika**, 47(2), 149-174.

Mooi, L.C. & Cheung, K.C. (1990-91). On Meaningful Measurement: Junior College Pupils' Anxiety Towards Computer Programming. **Journal of Educational Technology Systems**, 19(4), 327-343.

Narode, R.B. (1987). **Constructivism in Math and Science Education.** (EDRS Document Reproduction Service No. ED290616).

Newell, A. & Simon, H.A. (1972). **Human Problem Solving.** Englewood Cliffs, NJ: Prentice Hall.

Nishisato, S. (1980). **Analysis of Categorical Data: Dual Scaling and its Applications.** Toronto, Buffalo, London: University of Toronto Press.

Novak, J.D., Gowin, D.B., & Johansen, G.T. (1983). The Use of Concept Mapping and Vee Mapping with Junior High School Students. **Science Education**, 67, 625-645.

Qualter, A., Strang, J., Swatton, P., & Taylor, R. (1990). **Exploration: A Way of Learning Science.** Oxford: Basil Blackwell.

Rasch, G. (1960). **Probabilistic Models for some Intelligence and Attainment Tests.** Copenhagen: Danmarks Paedogogiske Institut.

Rasch, G. (1977). **On Specific Objectivity: An Attempt at Formalising the Request for Generality and Validity of Scientific Statements.** Danish Yearbook of Philosophy, 14, 58-94.

Raudenbush, S.W. (1988). Educational Applications of Hierarchical Linear Models: A Review. **Journal of Educational Statistics**, 13, 85-116.

Salomon, G. & Perkins, D.N. (1987). Transfer of Cognitive Skills from Programming : When and How? **Journal of Educational Computing Research**, 3, 149-170.

Sfard, A. (1987). Two Conceptions of Mathematical Notions: Operational and Structural. In: J.C. Bergeron, N. Herscovics, & C. Lieran (Eds.). **Proceedings of the 11th International Conference for the Psychology of Mathematics Education.** Montreal, Canada: Universite de Montreal.

Shuell, T.J. (1990). Phases of Meaningful Learning. **Review of Educational Research**, 60(4), 531-547.

Sternberg, R.J. (1985). **Beyond IQ: A Triarchic Theory of Human Intelligence.** New York : Cambridge University Press.

Sternberg, R.J. (1991). Towards Better Intelligence Tests. In : M.C. Wittrock and E.L. Baker (Eds.). **Testing and Cognition.** Englewood Cliffs, New Jersey : Prentice Hall.

van Hiele, P.M. (1957). **De problematiek van het inzicht, gedemonstreerd aan het inzicht van schoolkinderen in meetkunde-leerstof.** Unpublished Doctoral Dissertation, University of Utrecht, the Netherlands.

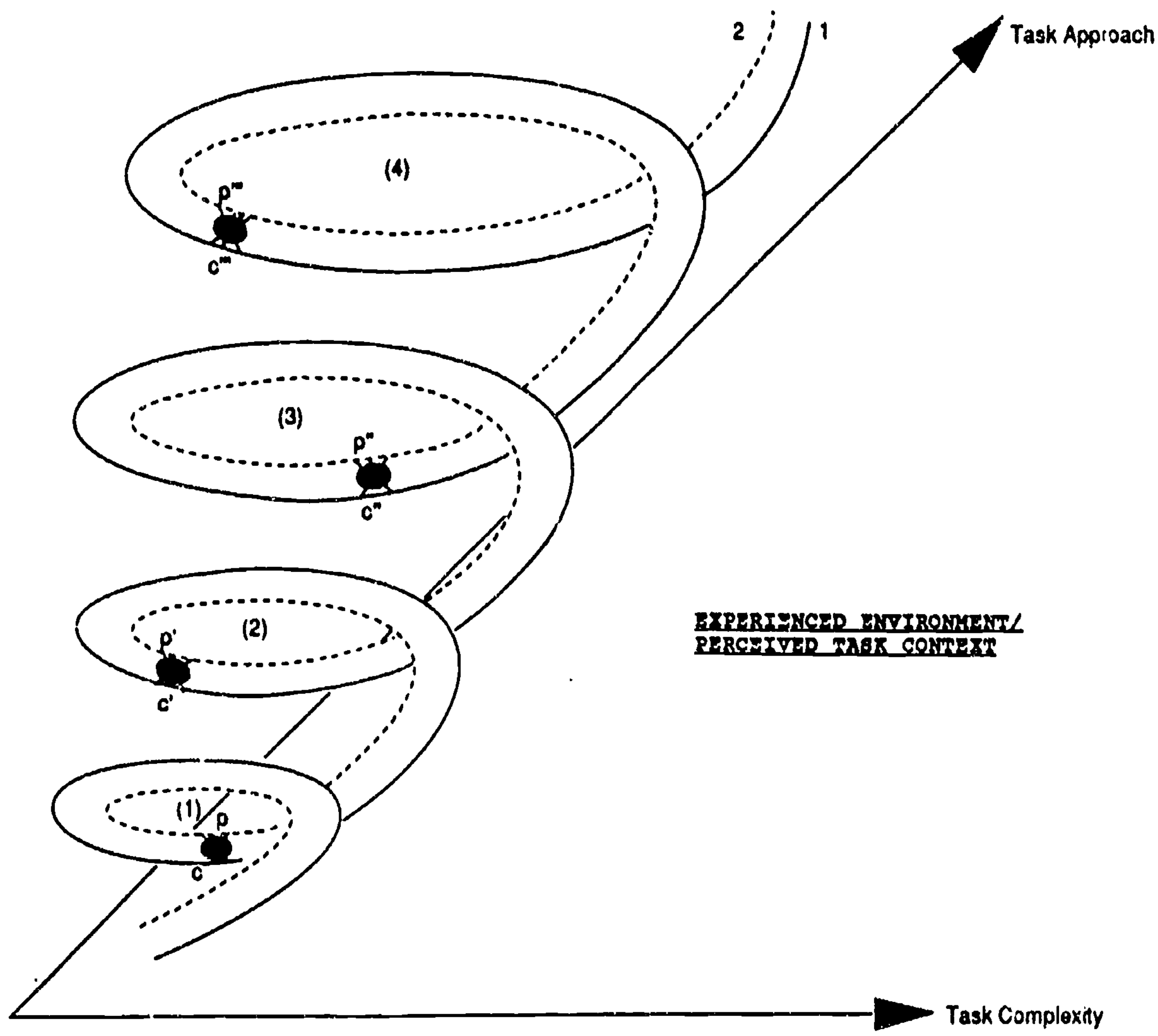
van Hiele, P.M. (1986). **Structure and Insight.** New York : Academic Press.

van Hiele-Geldof, D. (1957). **De didactiek van de meetkunde in de eerste klas van het V.H.M.O.** Unpublished Doctoral dissertation, University of Ultrecht, the Netherlands.

Vergnaud, G. (1990). Epistemology and Psychology of Mathematics Education. In: P. Nesher & J. Kilpatrick (Eds.). **Mathematics and Cognition: A Research Synthesis by the International Group for the Psychology of Mathematics Education.** New York: Cambridge University Press.

Willett, J.B. (1989). Questions and Answers in the Measurement of Change. **Review of Research in Education,** 15, 345-422.

Yip, S.H. & Cheung, K.C. (1991). **The SOLO Approach in the Assessment Cognitive Hierarchies : Analysis of Errors and Information Processing Demands.** Unpublished Manuscript, National Institute of Education, Republic of Singapore.



1. CONCEPTUAL KNOWLEDGE
   = understanding of concepts.

2. PROCEDURAL KNOWLEDGE
   = proficiency of procedures.

(1), (2), (3) and (4) are programmes of study linked to designated key stages of schooling.

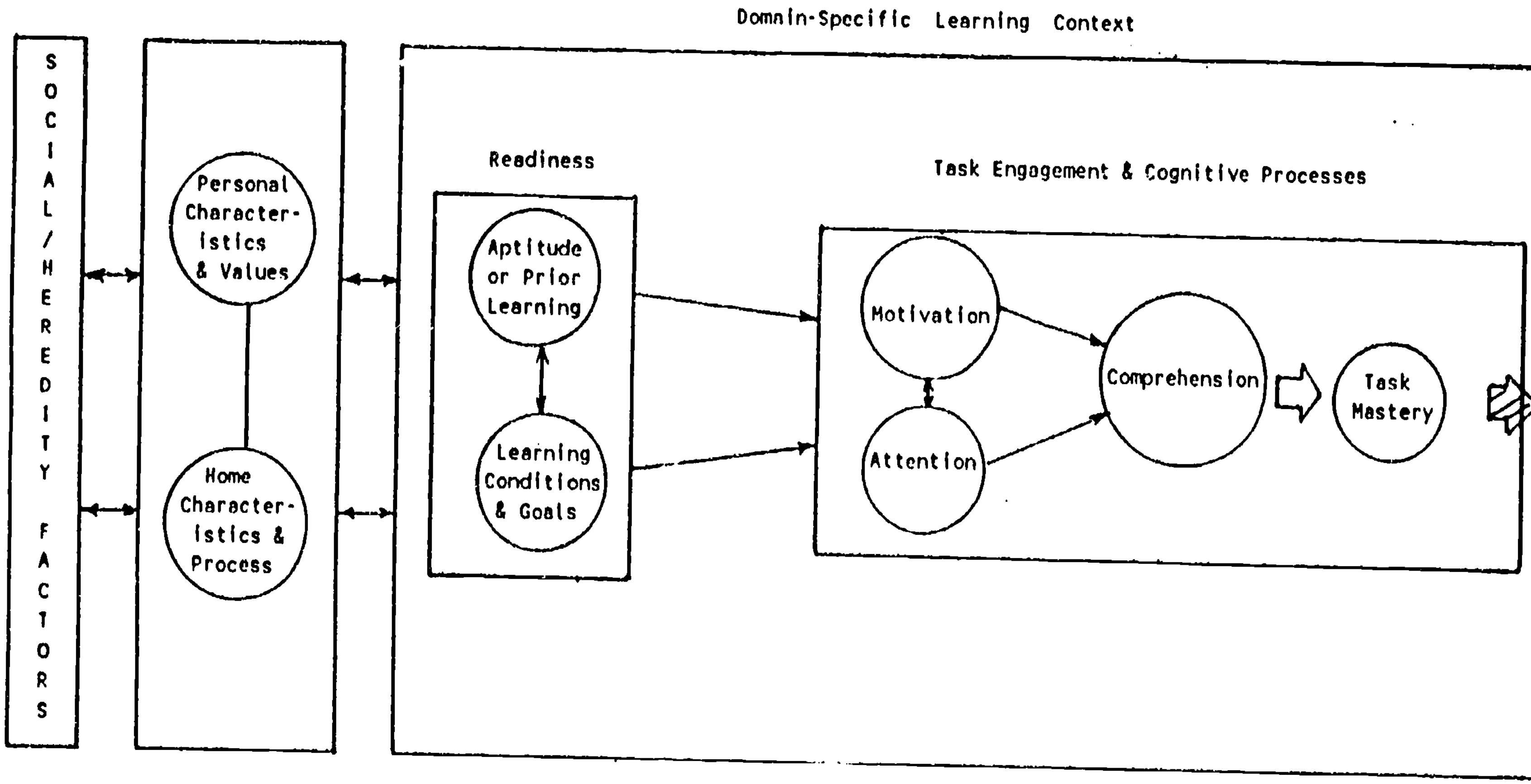
INTERTWINING OF CONCEPTUAL AND PROCEDURAL KNOWLEDGE WITH A DOMAIN-SPECIFIC TASK CONTEXT

conceptual knowledge required

problem tasks

procedural knowledge required

Figure 1. The Double-Spiral Model of Knowing (adapted version; Cheung & Taylor, 1991, p.29)



Note : Comprehension, a summary mediating construct for task perception, cognitive and memory process, is the antecedent of task mastery.

Feedback, tangible or intangible experiences, to the learning tasks and cognitive processing system.

Figure 2. A Model of Difficult Learning (Cheung, 1989, p.72)

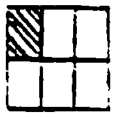
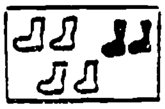
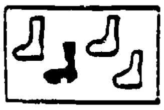
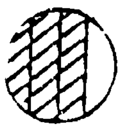
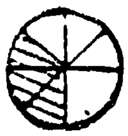
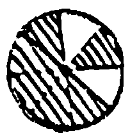
Level	Sample Item
1	Which fractions say $\frac{1}{5}$? (A) one (B) one-sixth (C) one-fifth (D) Not given (E) I don't know
	Which fraction says "three-fifth"? (A) $\frac{3}{5}$ (B) $\frac{3}{8}$ (C) 3 (D) Not given (E) I don't know
2	What fraction of the square is shaded?  (A) $\frac{1}{6}$ (B) $\frac{1}{5}$ (C) 1 (D) Not given (E) I don't know
	Which picture below shows $\frac{1}{2}$ of the set shaded? (A)  (B)  (C)  (D) Not given (E) I don't know
3	Which picture below shows $\frac{1}{3}$ of the set shaded? (A)  (B)  (C)  (D) Not given (E) I don't know
	What fraction of the set is shaded?  (A) three-fifths (B) three-eights (C) six-tenths (D) Not given (E) I don't know
4	What fraction of the set is shaded?  (A) $\frac{2}{6}$ (B) $\frac{1}{2}$ (C) $\frac{1}{4}$ (D) Not given (E) I don't know
	Which circle below shows $\frac{1}{4}$ shaded? (A)  (B)  (C)  (D) Not given (E) I don't know

Figure 3. Sample Fraction Items at Four Progressive Levels of Understanding Spanning a Conceptual Field

ALGEBRA TEST Time:60 minutes

1. Simplify 2a+5b+a.

2. Simplify (8y+6z) - (4y+3z).

3. What is the value of 3a-b+5c if a=2, b=3 and c=5?

4. Given that 4y-8 = 2y-4, find the value of y.

5. Solve the equation 4(x+2) + 2x = 4.

6. One stick is twice as long as the other. If the length of the shorter one is x cm, find the length of the longer one.

7. The price of a desk and a chair is $x. What is the price of the desk if the price of the chair is: i) $20, ii) $y?

8. Mangoes cost m cents each and papayas cost p cents each. If I buy 5 mangoes and 2 papayas, what does 5m + 2p stand for?

9. Mary's basic wage is $500 per month. She is also paid another $2 for each hour of overtime that she works. If h stands for the number of hours of overtime that she works and if w stands for her total wage in $,
   i) write down an equation connecting w and h,
   ii) what would Mary's total wage be if she worked 10 hours of overtime?

10 A mother promised to pay her son 20 cents for every Maths problem that he got right and fined him 5 cents for each wrong answer. After 10 problems, the mother had to pay the boy $1.50. How many problems did the boy answer correctly?

11 A woman has altogether 50 cows and chickens. These animals have a total of 172 legs. How many cows are there?

Figure 4. Open-Ended Algebra Test (Cheung, et al., 1991, p.80)

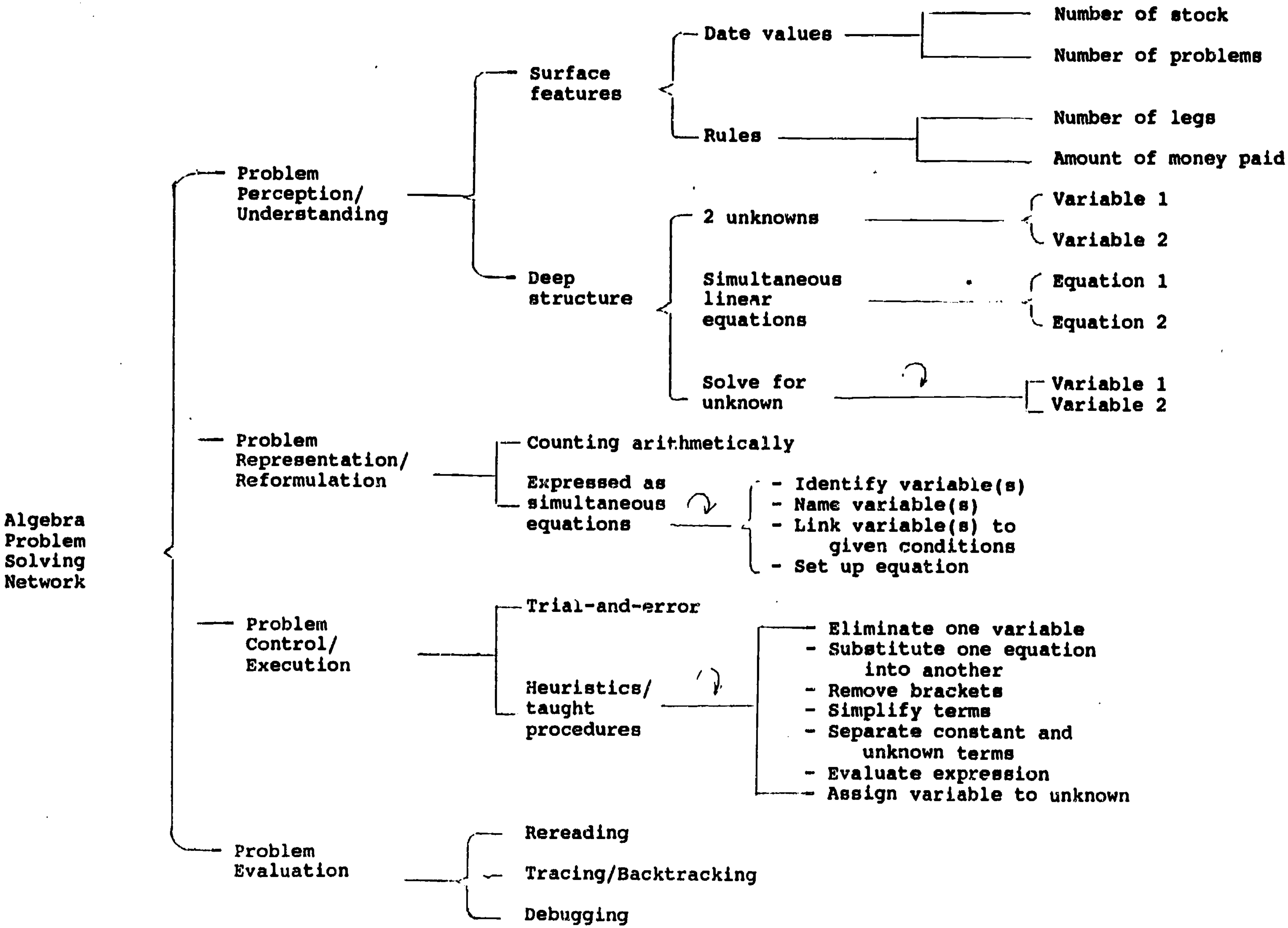


Figure 5. Common Designed Deep-Structure of the Core Items (Loh & Cheung, 1991, p.36)

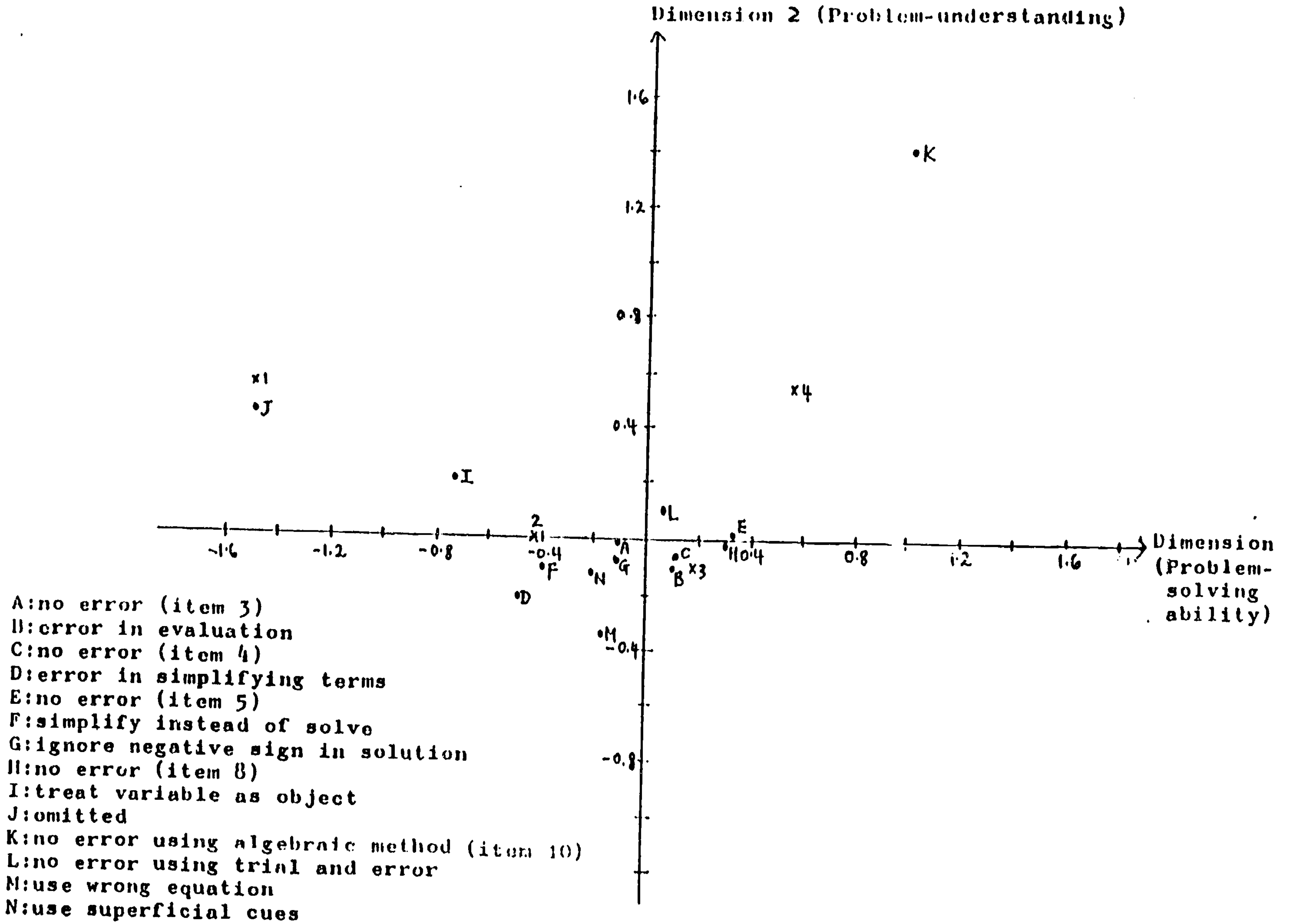


Figure 6. Hierarchical Structure of Key Types of Errors in Relation to Levels of Problem-Solving Proficiency (Cheung & Loh, 1991, p.63).

Table 1. Competence Schemes with Suggested Levels of Understanding the Part-Whole Concept of Fractions (Koh & Cheung, 1990, p. 13)

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 of 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 theset 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.

Table 2. Problem-Solving Proficiency Continuum of Algebraic Word Problems (Loh & Cheung, 1991, p.44)

Scale Performance Level	Raw Score	Progressive Forms Of Knowing	Ability (Logits)	Items 10	9i	11	8	2	5	7ii	7i	9ii	4	6	3	1
		Know relationships	3.0	4	3	4	2	2	3	2	2	2	2	1	2	1
		between variables	2.8	4	3	4	2	2	3	2	2	2	2	1	2	1
4		from word problem.	2.6	4	3	4	2	2	3	2	2	2	2	1	2	1
		Write equation to	2.4	4	3	4	2	2	3	2	2	2	2	1	2	1
		represent problem.	2.2	3	3	4	2	2	3	2	2	2	2	1	2	1
		Solve simultaneous	2.0	3	3	4	2	2	3	2	2	2	2	1	2	1
		equations.	1.8	2	3	4	2	2	3	2	2	2	2	1	2	1
			1.6	2	3	4	2	2	3	2	2	2	2	1	2	1
	25	Remove bracket &	1.4	2	3	3	2	2	3	2	2	2	2	1	2	1
		simplify terms.	1.2	2	3	3	2	2	3	2	2	2	2	1	2	1
		Solve equation	1.0	2	3	2	2	2	3	2	2	2	2	1	2	1
3		with brackets &	.8	1	2	2	2	2	3	2	2	2	2	1	2	1
		negative solution.	.6	0	1	2	2	2	3	2	2	2	2	1	2	1
		Interpret algebraic	.4	0	0	2	2	1	3	2	2	2	2	1	2	1
		expression.														
	16	Separate constant	.2	0	0	2	0	1	2	2	2	2	2	1	2	1
		from unknowns to	.0	0	0	0	0	1	2	2	2	2	2	1	2	1
2		solve equation.	-.2	0	0	0	0	1	2	1	2	2	2	1	2	1
		Understand	-.4	0	0	0	0	1	1	0	2	2	2	1	2	1
		algebraic relation	-.6	0	0	0	0	0	1	0	0	1	1	1	2	1
		ship between known	-.8	0	0	0	0	0	0	0	0	0	1	1	2	1
		& unknown terms.														
	6	Simplify terms.	-1.0	0	0	0	0	0	0	0	0	0	0	1	2	1
		Evaluate expres-	-1.2	0	0	0	0	0	0	0	0	0	0	1	2	1
		sion.	-1.4	0	0	0	0	0	0	0	0	0	0	0	1	1
1		Translate simple	-1.6	0	0	0	0	0	0	0	0	0	0	0	1	1
		sentence to	-1.8	0	0	0	0	0	0	0	0	0	0	0	1	1
		algebraic expres-	-2.0	0	0	0	0	0	0	0	0	0	0	0	1	1
		sion.	-2.2	0	0	0	0	0	0	0	0	0	0	0	1	0
			-2.4	0	0	0	0	0	0	0	0	0	0	0	1	0
			-2.6	0	0	0	0	0	0	0	0	0	0	0	0	0
			-2.5	0	0	0	0	0	0	0	0	0	0	0	0	0

Table 3. Crosstabulation of Key Types of Errors against scale Performance Levels (Cheung & Loh, 1991, p.54)

Scale Performance Level	Item 3		Item 4		Item 5			Item 6			Item 10			
	A	B	C	D	E	F	G	H	I	J	K	L	M	N
4 (n=11)	10	1	11	0	10	0	1	10	1	0	5	4	0	1
3 (n=70)	63	7	65	4	48	7	6	50	6	1	3	19	8	23
2 (n=44)	40	3	28	7	11	4	5	13	10	12	0	8	5	11
1 (n=5)	3	0	0	0	0	1	0	0	2	3	0	1	0	2
Total	116	11	104	11	69	12	12	73	19	16	8	32	13	37

A : No error
B : Error in evaluation
C : No error
D : Error in simplifying terms
E : No error
F : Simplify instead of solve
G : Ignore negative sign in solution
H : No error
I : Treat variable as object
J : Omitted
K : No error using algebraic method
L : No error using trial and error
M : Use wrong equation
N : Use superficial cues

Table 4. Intertwining Anxiety Schemes and Progressive Forms of Anxiety (Mooi & Cheung, 1990-91, p.339)

Subscale	*1*	*2*	*3*	*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.