TYPES AND LEVELS OF DATA ARRANGEMENT AND REPRESENTATION IN STATISTICS AS MODELED BY GRADE 4 – 7 LEARNERS

TABLE OF CONTENTS

CHAPTER 3: REPRESENTATION IN MATHEMATICAL AND STATISTICAL MODELING AND PROBLEM SOLVING

CHAPTER 4: METHOD OF RESEARCH

CHAPTER 5: ANALYSIS OF RESULTS

CHAPTER 6: FINDINGS AND RECOMMENDATIONS

LIST OF FIGURES

LIST OF TABLES

CHAPTER 1
ORIENTATION, PROBLEM STATEMENT AND RESEARCH DESIGN

1.1 INTRODUCTION AND ORIENTATION

Statistics traditionally has been a subject and practice for university-level courses and professional statisticians. Statistics, however, being the science of collecting, organising and interpretation of data and the study of probability, plays a major role in modern society: everyday we are deluged with statistical data generated by computers. Data sheets, tables, pictograms, circle, line and column graphs, histograms, stem and leaf plots, box plots and scattergrams are some examples of data representations used in the field of medicine, business, education, public administration, social work, policy studies, management, urban and regional planning, labour relations and many more. One just has to open a newspaper to realise the truth of the statement "…to create graphs and analyze data have become essential skills in our technological society" (Parker & Widmer 1992:48). The prophetical words of H.G. Wells that "… statistical thinking will one day be as necessary for efficient citizenship as the ability to read and write" (Burrill 1990:118) may have become true in the twenty first century. Data handling has become a key part of education for responsible citizenship (Shaughnessy, Garfield & Greer 1996:206). The authors stress the fact that living in an information age, it is essential that learners develop conceptual and practical tools to make sense of that information. The aim is not to turn all learners into competent statisticians, but to "…have them understand enough statistics to be able to respond intelligently to claims based on statistics" (Schaeffer, Watkins & Landwehr 1998:31). In response to the crucial role of statistics in our society, worldwide calls for reform in statistics education at all levels have been heard in recent years. The NCTM has included "explorations of statistics in real-world situations" (1989:105, 167) in their Principles and Standards for School Mathematics and there have been ongoing calls for reform in statistics education, with emphasis on the primary grades (NCTM 2000).

The inclusion of statistics in the mathematics curriculum is a controversial subject. Some statisticians argue that statistics should not be taught as one of the strands of mathematics, but as subject on its own, because where the focus in mathematics is on abstraction, statistics can never be without context (Cobb & Moore 1997:801). The emergence of constructivism and exploratory data analysis in which data are explored with graphing techniques, supports this advocated difference between mathematics and statistics well. Statistics however is included in the school mathematics curriculum of most countries and is not regarded as a subject on its own. In the South African mathematics curriculum Statistics, or Data handling and Probability, is included as the fifth learning outcome, the other four being

- Number, Operations and Relationships;
- Patterns, Functions and Algebra;
- Shape and Space; and
- Measurement.

At the Sixth International Conference on Teaching Statistics (ICOTS 6) in 2002 in Cape Town the emphasis of the organisers of ICOTS, the International Association for Statistical Education (IASE), was on *Developing a Statistically Literate Society.*

> In selecting the theme *Developing a Statistically Literate Society*, the IASE wanted to emphasise that statistical concepts, abilities, reasoning and understanding are important for citizens at large. This has been emphasised in many recent curricula for schools and universities around the world. Statistics literacy can help everyone in their understanding of the world, in taking informed decisions, in successfully carrying out a variety of tasks that require dealing with data, and in being critical consumers. Consequently, the statistics education community has been challenged to adapt the teaching of statistics to

> the different cognitive capacities, interests, contextual and cultural factors of children and adults around the world, so that each of them can enjoy and profit from the learning and understanding of statistics (Carmen Batanero IASE president, Preface ICOTS Proceedings CD 2002)

Lajoie and Romberg (Lajoie 1998:xv) state that part of the difficulty with the pedagogy of statistics is that the content is as new for many teachers as it is for learners. Teacher knowledge of basic statistics and the statistical thinking of learners in the primary school, is insufficient and in some cases non-existent. Different authors have discussed this lack of appropriate statistical background of teachers at all levels (Burrill & Romberg 1998:57; Bright & Friel 1998:63; Friel & Bright 1998:89; Gal 1998:275; Watson 1998:271). Furthermore, very little research in South Africa has been done up to now about the statistical thinking of learners in the intermediate phase. Not enough is known about how statistical reasoning can be continuingly actualised through the appropriate developmental timing of introducing instruction involving statistical situations (Schaeffer, Watkins & Landwehr 1998:4). Shaughnessy, Garfield and Greer (1996) comment about the relatively little research that has been done on the efficacy of instructional programmes in data exploration worldwide. According to Fennema, Carpenter, Franke, Levi, Jacobs, and Empson (1996) most of the research was not based on sound cognitive structures. While statistics education received attention in the USA, Britain, the Netherlands, Italy, Australia and New Zealand, South Africa has fallen behind. Although statistics and probability is included as one of the basic strands in mathematics education, teacher knowledge of subject matter and of the development of statistical thinking, especially in the forming years in primary school, is not yet up to standard.

Statistics education in South Africa is still in the process of building an identity and learning programmes as well support materials are being designed for the Revised National Curriculum Statement. It is an open field that presents many challenges in the unique diversity of the country. University lecturers who are training pre-service and in-service mathematics teachers (Prof Dirk Wessels, Unisa; Prof Delia North, UKZN; and Ms Erna Lampen, Wits) agree that most mathematics teachers, especially in the Intermediate and Senior Phase, do not have sufficient statistical content knowledge or pedagogical content knowledge to teach statistics with confidence (personal conversations). Information about the level of professional development of pre-service and in-service teachers in statistics and statistics education in South Africa is not available. Only some of the universities in South Africa include statistics education as module in their teacher training courses, and most of these courses have been introduced very recently, therefore most practicing teachers have not had any statistics training, and may not be cognisant about learners' statistical thinking. Considering the role of statistics in the lives of ordinary citizens and in the Revised National Curriculum Statement, it is imperative that the statistical thinking of learners be researched to inform teachers and coordinators of teacher training programmes.

1.2 MOTIVATION FOR THE STUDY

The aim of mathematics education (and therefore statistics education) must be to "equip students to use the mathematical skills and insights they have gained, ranging from rediscovered to self-invented, in solving a whole range of problems from both daily life and the world of mathematics" (Van den Heuvel-Panhuizen 2001). Through problem solving learners deal with real-world situations and make connections between their lives, other school subjects and mathematics. In using their mathematical and statistical skills in the problem solving process, they revert to known strategies or invent new ones appropriate to the problem.

Traditional problem solving usually involves only one modeling cycle to go from the given to the answer, which is a narrow view that does not allow for the mathematisation of real world contexts (Lesh & Doerr 2000:379). The shift to think of learning from a modeling perspective shed new light on learners' mathematisation of real-world problems. Learners' understanding of these problems becomes apparent from their representations in the modeling process. Teachers and researchers realised that a learner's responses in modeling tasks and other problem solving activities are not always on the same level.

Frameworks labeling learners as functioning on a specific cognitive level do not provide answers for this phenomenon. In the last decade and a half there is a growing trend to investigate learner understanding through analysing learner responses. This means that learners are not labeled according to a certain level of cognitive functioning but that there is an acknowledgement that learners respond differently to tasks on different occasions and in different contexts and circumstances. One of the prominent tools that are used world wide to simplify analysis of learner responses is the SOLO Taxonomy. This neo-Piagetian taxonomy has been used in a variety of fields, e.g. statistics, science, technology, and in different strands in mathematics such as geometry, measurement and fractions. The SOLO Taxonomy also forms the basis of a number of the statistical thinking frameworks that could be found in the literature. Learners' understanding is evident in their representations (see Chapter 3. 1) therefore the SOLO Taxonomy can be adapted to form a SOLO Taxonomy framework with which learner representations in a specific research project can be categorised in order to gain insight in their thinking.

During a series of data handling activities that formed part of a modeling task, the researcher has observed Grade 5 learners model the different components in the statistical investigation process, namely question posing, data collection, analysis of data and interpreting results. The researcher became interested in the modeling process. How does mathematical modeling link to the process of problem solving? What role do mental models play in modeling? What does data modeling consist of? Another aspect crystallised from the learners' modeling of the problem. The only way in which the researcher could make sense of the learners' modeling of the problem was through studying their representations albeit spoken words, written symbols, pictures or diagrams. This fact pointed to the crucial role of representation and led to the question of where representation fits into the framework of modeling and problem solving.

The way learners arranged and represented the data was also of particular interest to the researcher. The same task elicited a rich variety of different representations in which learners used different ways to organise or arrange the data. Questions emerged about this process of transforming and representing the data. What types of arrangement do learners use when transforming the data? What types of representations do learners spontaneously use when modeling data? What is the relevance and significance of data arrangement in the process of the representation of data? Friel, Curcio and Bright (2001:150) echo this question when they highlight the need to investigate learners' invented, reinvented and nonstandard representations:

> Exploring the ways children (in particular) when not limited to standard representations, choose to represent data may be worthwhile. Invented or reinvented representations may better convey explicit understandings about data and the relationship between analyzing data and answering questions that have been posed.

The authors also express the need for further study focused on learners' inventing or reinventing representations as tools for use and understanding data when not limited to standard representations.

The product and not only the process of the representation process intrigued the researcher. The learners' representations or products of the modeling process clearly exhibited a difference in statistical level. More questions surfaced. What statistical elements are fundamental to this specific task? How can the statistical level of a representation be evaluated? How would learners in other grades spontaneously represent data and on what levels would these representations be?

These questions in time condensed into one primary question that became the main focus of this research project: What are the types and levels of data arrangement and representation as modeled by Grade 4 – 7 learners?

1.3 RESEARCH FOCUS

The questions discussed above lead to the primary research focus, which is to acquire an understanding of the crucial role of representation in statistical modeling and problem solving. The study focused on the way learners organised and represented statistical data while engaged in open-ended data handling tasks. The focus was on spontaneous representations, therefore there was no mention of the word *graph* in the tasks, nor was any representations specifically taught or shown to the learners as examples of possible representations.

The areas of research evident from the reasoned exposition are

- modeling and problem solving in mathematics and statistics
- the nature and roles of representation
- data arrangement types
- data representation types
- levels of representation of Grade 4 – 7 learners
- the categorising of the SOLO level of the representation responses

The areas of the research outlined above, lead to research objectives that will now be discussed.

1.4 AIM AND OBJECTIVES OF THE RESEARCH

The aim of this study is to determine the types and levels of data arrangement and representation as modeled by Grade 4 to 7 learners in open-ended data tasks.

The following objectives are addressed:

- A review of the nature of problem solving and modeling
- The investigation of representation in mathematics and statistics
- A discussion of the types of data arrangement in learner responses
- A scrutiny of the types of spontaneous representations in learner responses
- A perusal of the SOLO Taxonomy
- A discussion of the research design of the empirical study to investigate the types and levels of data arrangement and representation as evident in learner responses.
- An analysis of the levels of statistical thinking in data representation by using the SOLO Taxonomy to categorise learner responses.
- The processing of data and the synthesis of the empirical investigation and research question.

The implications of the study for classroom practice and teacher training will be detailed.

1.5 RESEARCH DESIGN

1.5.1 Research methodology

The research methodology of this study consists of a literature study and an empirical study. The first leg of the research comprised a literature review. The literature study reply to the first five objectives stated in 1.4. A qualitative research design, more specifically, descriptive research, underpins the empirical study, which is strengthened by a limited quantitative analysis. The empirical study answers to the next three objectives in the list.

1.5.2 Qualitative research

The focus of the qualitative research in this study is to describe and interpret conditions and events of the present (Charles & Mertler 2002:265). Several distinct features emerge in qualitative research. The study described here focuses on the data arrangement and representation of the learners at a specific point in time, thus describing and interpreting events of the present. Written data and not numerical data were collected. Data sources are documents in the form of learners' written responses to two tasks. The tools for data generation are open-ended tasks and the treatment and analysis of data will be done by data conversion and limited statistical treatment. The study therefore satisfies the first condition for qualitative research, namely the reliance on written or spoken data in stead of numbers to document variables and inductive analysis of the collected information. Hittleman and Simon (2002:38) state that the basic qualitative research purposes are to "…describe, interpret, verify and evaluate" and further elaborate by saying that "…in interpretive analysis, the researcher explains or creates generalizations". The documented variables and analysis in the current study rely on all written responses of learners.

Secondly, context forms a central focus of qualitative research. The conceptual theoretical context refers to theories, assumptions, biases and beliefs that support the investigator's work. Contexts also include individual classrooms and school wide situations, but can be considered as the total life situations of teachers and learners involved. Learners' actions must be studied in the context of their natural setting, because a basic premise of this type of research is that people do not act in isolation. The classroom and school will form the setting for the research described in this document.

Thirdly, qualitative research is concerned not only with the outcomes of an activity, but also the process. The researcher is interested in the learners' thinking as evident in their data arrangement and the representation of statistical data, which shows the process through which they organise the data. The whole process of organising and representing the given data in the tasks will be considered and not only the final response.

A fourth issue is that data are rationally rather than statistically analysed.

> "The outcomes of much qualitative research are the generation of research questions and conjectures, not the verification of predicted mathematical relationships or outcomes. This is an additional key feature of qualitative research. Because of the descriptive nature of qualitative research, many of its data collection procedures are similar to those found in quantitative descriptive research. A distinguishing feature between the two is the use in qualitative research of the search for logical patterns within and among aspects of the research setting" (Hittleman & Simon, 2002:39).

The contribution of qualitative research then is the identification and interpretation of patterns of human responses resulting from knowledge, experiences and theoretical orientations to education. In this study the main contribution is the identification of arrangement and representational types as well as the adaptation of the SOLO Taxonomy with which learner responses can be categorised. The adapted SOLO Taxonomy will be referred to as the SOLO Taxonomy framework.

In this study no predictions of mathematical relationships or outcomes were made beforehand. Although the data generation method bears resemblance to quantitative data generation methods, the treatment of the generated data differs from quantitative methods. Learners' responses were analysed to identify their spontaneous arrangement and representation of statistical data. These responses are the result of social knowledge and different experiences in the classroom over years. The context and range of questions however are broader than the experience and knowledge gained from activities embedded in the curriculum.

In qualitative research the data are mostly written or verbal, analysis is a logico-inductive process and the purpose is to discover patterns (Charles & Mertler, 2002:178). The authors warn against the errors of subjectivity and imprecision in both the generation and analysis of qualitative data. Care should be taken that perceptions and interpretations are not influenced improperly by prejudices and preconceived notions. The researcher strived at objectivity and impartiality as well as the realistic depiction of context in the research. Qualitative methodology as described by Filstead (1990:6) will be followed, getting 'close' to the data and "…developing the analytical, conceptual, and categorical components of explanation from the data itself".

1.5.3 Research principles

The following operating rules of research emphasized by Charles and Mertler (2002:12-21), were adhered to. Measures to ensure the adherence of this research to the principles mentioned below are discussed under each topic:

1.5.3.1 Legal principles

When human participants are being used, the following apply

- **Protection**
  Individuals are by law protected against physical, mental or emotional harm.

- **Confidentiality**
  The anonymity of human participants must be maintained except when express permission to the contrary has been given.

Consent of the Gauteng Department of Education was sought and consent of parents of participants was obtained before conducting the experiment. Confidentiality was guaranteed to participants and parents not only in this document but also in all research reports or articles that might ensue from the research.

1.5.3.2 Ethical principles

These principles concern the moral aspect of research and include

- **Beneficience**
  This educational research was not conducted to do harm, denigrate, cast blame, find fault, deny opportunity or stifle progress, but to gain knowledge and shed light on the statistical thinking of learners.

- **Honesty**
  The data were not manipulated. Data were reported exactly as obtained, with no alterations, suppressions or procedural exceptions in collection methods, which would render the research misleading and meaningless.

- **Accurate disclosure**

  Accurate information about the topic and procedures of the research was given to participants in the research. Accurate disclosure does not necessarily mean full disclosure, for full disclosure would in some instances render the research invalid because knowing all detail of the research would introduce the possibility of error. In the described study learners may have changed their strategies of arrangement and representation if they knew the exact purpose of the study.

1.5.3.3 Philosophical principles

The following philosophical principles (Charles & Mertler 2002:17-21) pertain to the anticipated value of this particular investigation.

- **Importance**

  Research topics that are trivial, superficial or that have potentially inconsequential findings are not permissible. This research intends to contribute to human knowledge and aims to be useful to researchers, policy makers, teacher educators and teachers.

- **Generalisability**

  The purpose of the study is not to be generalisable, but to gain insight in the spontaneous arrangement and representation strategies of the learners and the levels on which their responses were created.

- **Replicability**

  This principle is one of the prime means of establishing credibility. This research should be replicable or repeatable, because if another researcher follows the “same recipe”, the results should be comparable. The kind of descriptive research that are reported in this study can be repeated but factors such as the data handling activities the learners were exposed to, the social knowledge of learners, the time allowed for completing the tasks, and so forth will influence comparison of results.

- **Probability**

  "Research deals in probabilities, or the best answers among a variety of possibilities. It almost never provides certainty" (Charles & Mertler 2002:18). For reliable research findings, there must be a very strong probability that the findings would almost always be approximately the same if the research is repeated numerous times. This research is reliable in the sense that the findings would probably be approximately the same if repeated in another school under the same conditions.

1.5.3.4 Procedural principles

These principles relate to the steps followed in the study to obtain, analyse and interpret data.

- **Researchability**

  The research design of this study makes provision for the answering of the following questions pertaining to researchability:

  * Can a scientific method be used to investigate the chosen topic? The topic could be examined by using qualitative measures, specifically a descriptive research method.
  * Is the scope of the topic comprehensive enough for a doctoral thesis, but at the same time limited to such an extent to make it researchable? The scope of the research here described covers an extensive review and analysis of representation of statistical data as presented in model-eliciting problems, with special attention on types and levels of arrangement and representation as categorised with the use of

the adapted SOLO Taxonomy framework (see 5.4). The research focuses on just one of the components of the process of statistical inquiry (see 2.5.4), namely analysis of data by the learners, thereby limiting the research field to make the study attainable.

* Can the investigation be done within existing practical constraints such as time, facilities, money, distance, and so forth? The study was planned to happen within time constraints imposed by the Department of Education and the school; the distance to the school and facilities available made it viable for the researcher to do the research project at the specific school.

- **Parsimony**

  The best research procedures obtain data and provide analysis through measures that are clear, simple, efficient, and to the point as far as conditions allow. This principle holds that "…the simpler a theory is, the better it is, provided it adequately explains the phenomena involved" (Charles & Mertler 2002:20). The researcher strived at using simpler rather than complicated measures to conduct the research, for example, used the existing SOLO Taxonomy in adapted form, the SOLO Taxonomy framework, to categorise learner responses and did not formulate a new theory to classify responses; likewise existing arrangement categories were used in adapted form.

- **Credibility**

  Established procedures of research, such as significance, reliability and validity, were adhered to in order to ensure that the research is credible. This principle was established by

  * selecting a significant and researchable topic
  * adhering to the mentioned operating principles (see 1.5.3)
  * obtaining on-target (valid) and consistent (reliable) data

* analysing the data according to appropriate methods, e.g. using the SOLO Taxonomy (see 3.7)
* reporting findings supported by the generated data
* clearly and accurately reporting conclusions which are related to the research questions and which are logically persuasive.

- **Rival explanations**

  The researcher took into account other interpretations and criticism from researchers with different viewpoints. Aspects such as accounting for confounding variables, following procedures properly, analysing data appropriately, and foreseeing and ruling out possible alternative interpretations were taken into account to anticipate such problems.

1.5.4 Criteria for data generation

Generated data were scrutinised according to the criteria for data generation to ensure authenticity, believability, validity and reliability.

- **Authenticity and believability of data**

  External criticism and *internal criticism* are according to Charles and Mertler (2002:40) the two informal, unstructured means of assessing data for authenticity and believability. *External criticism* must be used to verify whether data were obtained from legitimate sources, while *internal criticism* concerns data accuracy and bias. The data generation method was discussed with various researchers in the same field to corroborate the legitimacy of the method and sources. The data came from real people, learners in Grades 4 to 7, and are accurate in as far as they are the real responses from the learners. The data are however not measurable to a standard or criterion as they are the spontaneous responses of learners to open-ended questions and therefore varied from learner to learner.

- **Validity of data**

  Data are considered valid if the topic under consideration is directly dealt with, i.e. did the research actually measure what it was intended to measure? There are four types of validity to consider, namely content validity, predictive validity, concurrent validity and construct validity.

  * Content validity is determined by expert judgment and is present

  > "…when the content of an instrument such as an achievement test appear to very similar to the information contained in a course or training programme" (Charles & Mertler 2002:157).

  The two open-ended tasks have been adapted from the interview protocol-tasks of Mooney, Langrall, Hofbauer and Johnson (2001). Teachers and researchers have been consulted to establish content validity, which in this case was confirmation that the content of the open-ended tasks fall in the scope of the curriculum of the age group involved in the research.

  * Predictive validity deals with the prediction value of one set of data or measurement for future scores of the same participants on the same test. This type of validity is not applicable in this study as the same participants will not be tested on the same tasks again.

  * Concurrent validity is present when a particular measurement instrument yields results that relate closely with other tests of high acclaim. The researcher strived to achieve concurrent validity through adapting the SOLO Taxonomy for the purpose of categorising learner responses and through discussion of the adapted framework with other researchers who have experience in using used an adapted SOLO Taxonomy framework.

  * Construct validity is according to Charles and Mertler (2002:157) present when an instrument appears to measure a particular mental construct. A construct is an underlying, non-measurable characteristic or trait. In this study, it is not a particular construct that was measured, but spontaneous responses of learners to the data tasks were analysed and categorised.

- **Reliability of data**

  Data is considered reliable to the extent they are consistent. Scores from an instrument should be consistent and free from sources of measurement error. Charles and Mertler (2002:159) emphasise the importance of the relationship between validity and reliability: "A valid test is always reliable, but a reliable test is not always valid". Reliability is a necessity for validity, but not always sufficient. The tools for generating data on the types and levels of data arrangement and data representation are criterion-referenced tests. Each test item relates directly to an instructional objective. In the case of the data tasks each task is related to a type of data and could be arranged and represented in different ways because the tasks were open-ended. The tasks were also set in different contexts to ensure that responses of participants not relating to a specific context would not impede results. Triangulation of the adapted SOLO Taxonomy framework and the use thereof was ensured by conducting discussions with researchers who are experienced in using the SOLO Taxonomy as well as other researchers who were asked to check the coding of learner responses for reliability (see 4.7.4 and 5.4).

1.5.5 Phases of the research

The research was conducted in three phases, a literature study, an empirical investigation and the analysis and interpretation of results.

1.5.5.1 Phase One: Literature Study

A Dialog literature search was performed with the following descriptors: problem solving; modeling; mental models; data handling; data modeling; statistical thinking/development; statistics education; representation; data arrangement;

data representation, representation. The purpose of this literature search was to identify all relevant research projects and publications related to the field of study in this research project.

Primary and secondary sources were perused. Relevant literature and research projects about the following topics were studied:

- Modeling and problem solving
- Data handling and data modeling
- Data arrangement types
- Representational types
- Categorising learner responses by using the SOLO Taxonomy framework
- Different views on the nature and levels of statistical development and thinking
- Levels of statistical thinking in the representation of statistical data

1.5.5.2 Phase Two: Empirical study

The investigation

The research was conducted in a government school in Pretoria with the consent of the Gauteng Department of Education (see Appendix A). Learners of mixed ability in Grade 4 to 7 formed the population for the study.

Data generation

Two open-ended data tasks in different contexts were administered with the aim of eliciting the spontaneous arrangements and representations from the learners (see Appendix 2). One task included categorical data and the other numerical data. The data representations of the learners were collected and categorised according to the types of data arrangement and representational types found. The purpose was to determine the types of spontaneous arrangements and representations of statistical data and learners' level of statistical thinking according to the SOLO Taxonomy framework (see 4.7.4 and 5.4).

1.5.5.3 Phase Three: The analysis and interpretation of findings

All the data that were gathered during the empirical study were interpreted qualitatively. A limited explorative quantitative analysis was conducted to complement the qualitative findings. The qualitative analysis was done in three different stages:

- Data tasks were analysed firstly according to the different types of arrangement found.
- Tasks were then categorised according to types of representation.
- The statistics elements fundamental to each task were determined and the level of representations was categorised by using the SOLO Taxonomy

The limited quantitative analysis comprised the following:

- A limited Rasch analysis to investigate validity and reliability
- The compilation of tables comparing results over different categories and grades.

1.6 VALUE OF THE RESEARCH

Statistics education is not just of importance as preparation for a career in mathematics or statistics. The ordinary citizen is bombarded with statistics and graphs in the media and in almost every conceivable area of life. It is therefore of crucial importance to equip all learners with skills in and understanding of data handling and probability.

Teachers' lack of content knowledge in basic statistics and of statistical thinking of learners in the primary grades necessitates research in this field to inform pre-service and in-service training of mathematics teachers in the intermediate phase.

The planning and development of mathematics curriculum and instruction is dependent on the knowledge of learners' thinking (Cobb, Wood, Yackel, Nicholls, Wheatley, Trigatti & Perlwitz 1991; Resnick 1983). The exploration of levels of statistical thinking in the experiment shed light on this relatively new field in the mathematics curriculum. Knowledge of interdisciplinary links and influences between statistics and mathematics education is also necessary for the planning of curricula, instruction, teacher training and the development of learning support materials.

This research study not only provides an insight into the types of data arrangement and representation of learners in Grade 4 to 7, but also into the levels of data representation. It informs the reader of the nature and level of statistical thinking of young learners and therefore is helpful in exploring the development of statistical thinking that underpins the more complex statistical reasoning needed in the higher grades.

1.7 AN OVERVIEW OF THE RESEARCH

The study was recorded in six chapters. In this chapter the research problem was discussed and a broad outline of the research design and investigation were given.

Chapter 2 discusses the links between modeling, problem solving and mental representations. The role of mental models in learners' understanding receives attention and the nature of data modeling is investigated. Different statistical thinking frameworks are perused.

Chapter 3 focuses on representation in mathematics and statistics, and the role of representations in understanding is investigated. Specific consideration is given to arrangement and representation of data. Statistical thinking levels are explored and the use of the SOLO Taxonomy for the categorisation of responses scrutinised.

In Chapter 4 the research design is revisited and the design and progress of the empirical investigation described. The sample, instrument and analysis are explicated. Measures taken to ensure that the research conforms to the requirements are spelled out.

The analysis of the arrangement types, representational types and the SOLO levels of responses are described in Chapter 5, illustrated with examples from learner responses and tables summarising the use of the different types of arrangement and representation. The use of the SOLO Taxonomy framework for categorising learner responses is explained and elucidated with examples from the learners' modeling attempts.

The study is concluded in Chapter 6, which takes the findings under scrutiny, draws conclusions and spells out the implications of the research findings for classroom practice, the design of learner projects, learning materials and teacher training. Recommendations and suggestions for further research are given.

CHAPTER 2
MODELING AND PROBLEM SOLVING IN MATHEMATICS AND STATISTICS EDUCATION

2.1 INTRODUCTION

Modeling[1] has become a very popular term in the industry, publications and conversations of the day. This is also true in the field of mathematics education where it became widely used over the last decade. Modeling in mathematics education earned the status as one of the fruits of the development of constructivism and with it came the continuing emphasis on 'developing understanding' in the teaching and learning of mathematics that dominated mathematics education since the early 1980's. Constructivism brought the shift in emphasis from teaching to learning and the accompanying sharper focus on how learners learn and understand mathematics (Wessels 2000:142-145).

Over the past three decades consensus has been reached that teaching is a complex activity that cannot be regarded only in the light of the teacher's contribution to the didactic situation. Teacher knowledge came under scrutiny in this process and can be described along three dimensions: knowledge of subject matter, knowledge of pedagogy and pedagogical content knowledge. This view of the complex character of what a teacher should know was first spelt out by Schulman (1986). Pedagogical content knowledge includes knowledge of the conceptual and procedural knowledge that learners bring to the situation, misconceptions they may have developed, and the different stages of understanding they may pass through in the process of learning (Schulman, 1986; Carpenter, Fennema & Franke 1996). This understanding of learners'

[1] The word 'modeling' is generally spelt with a single 'l' by people from North America and with a double 'l' in the United Kingdom. The spelling with one 'l' will be used in this document, except when a citation is used where it was spelt with a double 'l'.

mathematical thinking forms the core of the construct of *teaching for understanding.* Teaching for understanding is one of the hallmarks of the reform movement in Mathematics Education in the United States (Eisenhart, Borko, Underhill, Brown, Jones & Agard 1993:8) and the focus of the teacher development program Cognitively Guided Instruction (CGI) (Carpenter et al 1996). Teaching mathematics for understanding is a process in which a teacher needs considerable mathematical and pedagogical knowledge and skills (Eisenhart et al 1993:9). Modeling is crucial in this process of teaching for understanding. Lesh and Doerr (2000:376) state “to develop models is to learn”. To teach for understanding, therefore requires knowledge of modeling and model-eliciting activities. When learners model a problem, they do not only produce answers to questions, but also create powerful conceptual tools that can be communicated to others and reused in other situations. A teacher needs to take cognisance of what modeling is and how to use model-eliciting activities to help learners gain understanding, and also to help him/her to become conscious of how learners think, and to help them to modify, refine and extend their ways of thinking.

The processes of problem solving and modeling are closely linked. In contrast with a narrow view of problem solving where a question requires a one-cycle process to arrive at the answer, problem solving should rather concern non-routine problems that require more than one cycle to go from the givens to the goals. This view of problem solving brings the process in line with modeling, where multiple cycles are needed in the process of making sense of a problem (see 2.2.2). In the early eighties the focus was mainly on problem solving, but with the growing emphasis on the development of pedagogical content knowledge and *teaching for understanding*, attention shifted to modeling in the nineties. Modeling became an important way to foster understanding and to understand learners’ thinking. This understanding of children’s learning can only happen when a teacher observes and studies representations of learners’ modeling. When they model a problem, they document what they are learning, in other words, they represent their understanding internally and externally in some way. From these external representations, the teacher is able to make sense of learners' models.

This chapter is organised in four main parts. Modeling as an important way of making sense of problem situations in the real world is first discussed. The role of problem solving in the teaching and learning of mathematics is then perused, highlighting the solving of problems as goal of mathematics education but also as a major means of learning mathematics, which links it to modeling. Thirdly the role of mental models as internal representations of concepts, knowledge and models then is investigated. The last part comprises data modeling as evident in the statistical process and different frameworks of statistical thinking are perused. Representation and the interrelationship between problem solving, modeling and representation will be investigated in Chapter 3.

2.2 MODELING IN mathematics education

In this section the concept of modeling will be defined in the light of different views of the nature of mathematics, after which the modeling process as well as the process of emergent modeling as used in Realistic Mathematics Education will be scrutinised. The role of modeling in different curricula will be considered, and the link between modeling and representation will receive attention.

2.2.1 DEFINITIONS OF MODELING

When considering modeling, we need to discuss the nature of mathematics. When asked about the nature of mathematics answers will vary. Some answers will be about on the applications of mathematics, in other words on how mathematics can be used in everyday situations. Other interpretations will connect mathematics with numbers or geometric shapes. Still others will think of mathematics as problem solving, reasoning or representations. More current views about mathematics focus on the importance of a connected and balanced view of mathematics. The various mathematical ideas are not isolated and unrelated, but integrated in a significant way. Mathematical ideas can be connected in many different ways that leads to a much clearer view of their coherence and of the coherence of their applications. Connections can be found between different representations of an idea or problem, between mathematical generalisations and between mathematics and the real world (O'Daffer, Charles, Cooney, Dossey & Schielack 2002:48). When mathematics is viewed according to this interpretation, it is considered a logical, coherent subject where learning about one aspect builds upon other knowledge and creates the foundation for other ideas. A balanced view of mathematics looks upon mathematics as an activity concerning skills, concepts, relationships and higher-level processes (O'Daffer et al 2002:48-51). Higher order processes such as reasoning; problem solving and pattern finding are essential to doing mathematics and stress the importance of communicating mathematical ideas. This development of higher order reasoning processes are realised through structured frameworks.

Having highlighted some aspects of mathematics, it is now necessary to shift the attention to structured frameworks such as mathematical models and modeling. Professional mathematicians and scientists consider modeling fundamental to their everyday work. Romberg, Carpenter and Kwako (2005:13) contend that

> … meaningful inquiry, involving cycles of model construction, model evaluation, and model revision, is central both to understanding in a domain and the professional practice of both mathematicians and scientists.

The fact that models are diverse and widely used in these disciplines, indicate that modeling can help learners to develop understanding about a wide range of important mathematical and scientific ideas, therefore modeling practices can and should be fostered at every age and grade.

The concept of modeling is interwoven with a balanced and connected view of mathematics. Although specialists in a field rarely use dictionary explanations of words, it is necessary to look at the basic explanation of the word *model*. The Reader's Digest Illustrated Oxford Dictionary (1998:523) distinguishes between six meanings of the noun *model*: a three dimensional representation (on smaller scale) of an existing person, thing or structure; a figure in wax, clay, etc. to be reproduced in another material; a particular design or style; an exemplary person or thing; a person posing for an artist or photographer; a person employed to display clothes by wearing them; and then the definition that is most often used in connection with mathematics: "a simplified (often mathematical) description of a system, etc., to assist calculations and predictions". This explanation rightly refers to the simplification of the matter under scrutiny and the use of the model or description for assisting calculations and predictions, but lacks reference to real-world connections and connections to the development of understanding found in more elaborate definitions. Romberg et al (2005:15) maintain that

> … models are conceptual systems that represent phenomena in the world by means of system of theoretically specified objects, relations, operations, and rules governing interactions.

They distinguish between two different types of models (2005:13). The two conceptions are described as "model as <u>a natural process</u> used to construct an explanation of natural phenomena" and "model as <u>a representational tool</u> for communicating about the conceptual referent". Both of these models start with phenomenological context such as an event, question and problem situation, identifying key attributes or features of the phenomena and how these features are related. Both conceptions also use representations as tools to support disciplinary practices such as communication, mobility, combination, selection and predictability. There are, however, differences in the use of the term ‘model’, the features emphasised and the validation of the models.

A variety of technical and everyday meanings can be assigned to the term *model,* depending on the perspective of the person defining it. A definition common to the fields of mathematics, physics, chemistry and other physical sciences contends that:

> A model is a system consisting of (a) *elements*, (b) *relationships* among elements, (c) *operations* that describe how the elements interact, and (d) *patterns or rules*, such as symmetry, commutativity, or transitivity, that apply to the preceding relationships and operations (Lesh & Doerr 2000:362).

Doerr and English (2003:112) argue that with these models, being systems of elements, operations, relationships and rules, the behaviour of a familiar system can be predicted. Models may be represented by physical or iconic images, mathematical symbols and so on, but models are ideas and not simply physical images. They are representations representing ideas, and the ideas are at the heart of modeling (Romberg et al 2005:15).

When a modeling approach is adopted in the teaching and learning of mathematics, the focal point is the mathematisation of realistic situations that are meaningful to the learner. This approach implies three important changes in the teaching and learning of mathematics (Doerr & English PME 2001 CD:1-2):
(a) the quantities and operations needed to mathematise situations must be useful; (b) meaningful contexts must be used to create a need for the development of a model to describe, explain and predict the behaviour of an experienced system; (c) generalisations, in stead of just an answer to a given problem, must be developed so that learners can use and reuse it to find solutions. In a modeling approach to learning mathematics, generalising and refining models are the key activities:

> Thus, a modeling perspective leads to the design of an instructional sequence of activities that begins by engaging students with nonroutine problem situations that elicit

> the development of significant mathematical constructs and then extending, exploring, and applying those constructs in other problem situations leading to a system or model that is reusable in a range of contexts (Doerr & English 2003:113).

2.2.2 THE PROCESS OF MODELING

Dossey, McCrone, Giordano and Weir (2002:114) describe modeling as a process of presenting real-world situations through mathematics, in the process developing worthwhile mathematics with which events in the real world can be understood, predicted and controlled. They define a mathematical model as "…a mathematical construct designed to study a particular real-world system or phenomenon" including graphical, symbolic, simulation and experimental constructs. A mathematics model is designed to study representations and form new ones. When a situation is not difficult to make sense of, an existing model can usually be applied with minor adaptations; in the case of a complicated situation existing models usually have to be refined to be usable. Hence modeling is a process developing higher order thinking skills with which events from the real world can be modeled in order to describe them, make sense of them, use them to solve problems and to predict how other systems or models can be understood.

In solving typical school "word problems", learners usually engage in one- or two-cycle problem solving steps to solve the problem, in the process mapping problem information onto arithmetic quantities and operations. The teacher has usually carefully planned the problem so that it will be easily computable and the learner only has to "unmask" the mathematics by mapping the information on to the strategy learnt. In a modeling task, however, a real-world or realistic setting is the starting point and the learners' goal is to make sense of the problem so that he can mathematise it in ways that make sense to him. This process is cyclic, and relevant quantities will be selected, meaningful representations created and operations defined, which in turn may lead to new quantities (Doerr & English

2001:362). The mental challenge of this cyclic process leads to the development of higher order thinking skills. The model constructed in the process will be refined and adapted to be reusable in other contexts. The learner will also be able to use the model to describe, make sense of, or make predictions about another model (Lesh & Doerr 2000:362).

Mathematical models are representations of reality and facilitate understanding of the environment, helping individuals deal with problems. A real-world situation may be a complex phenomenon that sometimes must be oversimplified by the learner in order to create a model. Modeling is central to understanding the real world and is a closed process starting with a real-world phenomenon from which data are gathered to formulate a model. The model is then analysed and conclusions reached. Interpretations of the model lead to predictions or explanations after which the conclusions about the real-world system are tested against new data or observations. The model might need refinement or might not be suitable at all, which then requires the formulation of a new model. The six steps in the process of constructing a model are described by Dossey et al (2002:116):

Step 1: Identify the problem
The question of what particular aspect of the situation it is that needs to be studied must be sorted out and the problem must be formulated in such a way that it can be translated into mathematical language or statements.

Step 2: Make assumptions
The number of factors to be considered needs to be reduced to make the problem or situation to be modeled manageable and the relationship between the remaining variables must be determined. If the situation is complex, it may not be possible to see the relationship between the

variables and submodels must then be studied. These submodels will later be connected.

Step 3: Solve or interpret the model

All the submodels are put together and to make sense of the situation. In some cases the mathematical equations or inequalities have to be solved in order to find the information needed to make sense of the model.

Step 4: Verify the model

Three tests can be applied to test the model:

- Does the model answer the question asked in the first step?
- Is the model usable in a practical way, can you gather the data necessary to operate the model?
- Does the model make sense?

The reasonableness of the model must be corroborated by the data collected.

Step 5: Implement the model

It must be possible to explain the model to potential users and it must be user-friendly.

Step 6: Maintain the model

Determine whether the model must be refined or simplified, or if adjustments have to be made.

These modeling steps are not a once-off process. Model construction is an iterative process. Multiple modeling cycles are used to construct, modify, refine or extend the model (Dossey et al 2002:118; Lesh & Doerr 2000:380).

Romberg (1999:2-6) describes the modeling process with reference to the different stages as the mathematisation of a problem. The first stage of identification of a “problem” receptive to mathematical treatment may be long and requires many skills that are not related to mathematics, e.g. when learners are asked to describe the growth of a plant. In the identification process, the essential or significant features of the problem situation must be sorted out. Romberg refers to the identification or idealisation as a crucial part of the process because the general problem may be very complex and entailing many processes. Simplification is acquired when the insignificant or irrelevant features in the situation are ignored and the original complex problem is reduced to one that is mathematically pursuable. Once the significant features have been identified, they have to be translated into a conceptual model in the second stage. Thirdly, after the initial conceptual model has been established, each variable must be mathematised to create a mathematical model. Romberg contends that the mathematisation of the variables is often the most difficult stage of the modeling process because of the nature of the possible mathematisation of the variables. One class of variables is called deterministic because they can, at least in theory, be precisely measured. Another class of variables cannot be known precisely, and are therefore called stochastic variables, referring to their uncertainty. A model containing stochastic variables will need statistical and probabilistic techniques while a model with deterministic variables will require the use of algebra or often calculus. The significant or critical features of some situations are initially not recognised to reveal how they can be mathematised; while in other situations both deterministic and stochastic variables can be found. Once a model has been constructed, it needs to be validated, which represents the fourth stage. Some validation is usually carried out throughout the formulation, because the formulas or other mathematical relations set up in the model are continually checked with the initial situation. All mathematics used in the model must yield to the usual rules of mathematical logic and must be self-consistent. The ultimate test for the validity of a model rests in its ability to represent the initial situation. Judgment on validity however, is subjective because a model only represents some features of reality and not reality itself. It is not about whether it is a “true” model that accurately represents the working of the system at all stages, but if it is an adequate model in which the results obtained were sufficiently representative of the situation for the purpose of the problem at hand. Romberg et al note that “…simple, adequate but incomplete models are less costly and sometimes more useful than elaborate models requiring elegant analytic procedures for solutions” (2005:17). Interpretation of the results of the model is the last or fifth step in the process described by Romberg where results must be re-interpreted in terms of the problem situation. This usually entails the restating of the problem in terms of the situation rather than in the language of mathematics.

To be a model, a system should be usable to describe, think, make sense, explain, or make predictions about some other system. The model becomes mathematically significant when its focus is on the underlying characteristics of the system being described. Learners need to be exposed to multiple experiences in which their models can be applied in new settings and extended or modified. In each stage of the development of the model, multiple cycles of interpretations, descriptions, conjectures, explanations and justifications are constructed and refined by the learner. Social constructivistic practices play an important part in model building. Although learners can develop models on their own, interaction and communication are valuable in this process. The sharing of models does not, however, mean that each individual in the class or group who participated in the model building activity, share the same model or understanding of the system (Lesh & Doerr 2000:366).

2.2.3 EMERGENT MODELS

Realistic Mathematics Education (RME) is rooted in the interpretation of the renowned Dutch mathematician and mathematics educationist, Hans Freudenthal. He based his philosophy of mathematics education on the premise that mathematics is a human activity (Freudenthal 1973; 1983) and that learners should not be presented with ready-made structures and strategies to deal with problems. He stated that learners should learn mathematics by *doing* mathematics, reinventing mathematical insights, knowledge and procedures from real situations. These real situations can include mathematically authentic contexts and contextual problems, allowing learners to participate in the mathematisation of reality. Lesh and Doerr (2000:366) state that mathematising (e.g. quantifying, visualising or coordinatising) is a form of modeling.

Mathematisation has two faces: *horizontal mathematisation* in which learners use informal strategies to describe and solve contextual problems and *vertical mathematisation* in which these informal strategies or models guide them to solve problems by using mathematical language, progressing to more formal models. In this process of horizontal and vertical mathematisation informal models thus evolve into formal models. Gravemeijer (1998:32; 1999:240-243) calls this a process of emergent modeling. He gives a description of emergent models for progressive mathematisation from the perspective of RME in which the term *model* is perceived in a holistic, dynamic way (Gravemeijer, Cobb, Bowers & Whitenack 2000:240). The contextual situation of the problem is the starting point from which the problem is modeled. The learner later on uses this model as the foundation of the more formal mathematics, which is the ultimate goal of mathematising. Formal mathematics is in this way being developed from learners' informal mathematical activities. This shift of *model of* to *model for* correspond with a shift in learners' thinking from the modeled problem situation to mathematical relations (Gravemeijer 1998:34).

Gravemeijer identifies four levels of activity related to the distinction between *model of* and *model for* and describes how a model *of* a certain situation can become a model *for* more formal reasoning. This distinction involves a developmental progression but is not a strictly ordered hierarchy. The first level is activity in instructional settings or task settings that involves situation-specific imagery. In the second or referential level the model is rooted in learners' understandings of paradigmatic settings in experientially based activities. Learners use their models in explanations or descriptions of how problems were interpreted and solved from the starting point. Thus far the modeling is still rooted in concrete experiences. The third level or general activity emerges as learners are no longer dependent on situation-specific imagery. The fourth level of activity results from this process and becomes evident when the model is used for mathematical reasoning, therefore progressing to a formal level:

> This transition can be seen as a process of reification wherein the students begin to collectively reflect on their referential activity. In the process, the model becomes an entity in its own right and serves more as a means of mathematical reasoning than as a way of symbolizing mathematical activity grounded in particular settings (Gravemeijer 1998:35).

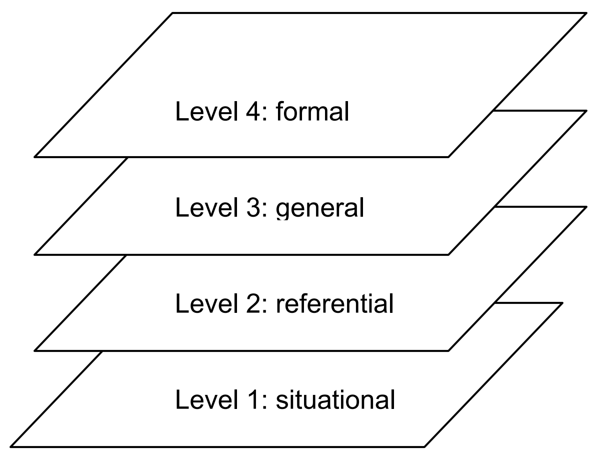


Figure 2.1 Levels of emergent modeling from situational to formal reasoning

The four levels can be summarised as task setting or situational, referential, general and formal (Fig 2.1). The overarching concept is evident in what can be described as a chain of significance.

From a mathematics education perspective, this chain of significance shows how formal mathematical signs are grounded in concrete activities of learners.

Gravemeijer (1998:40) discusses the dynamic character of the chain of significance and points to the dual meaning of the term 'emergent' where on the one side it refers to the process by which a model emerges in RME but on the other side also to the process by which the emergence of formal mathematics is supported by these models.

The use of multiple problem solving strategies and representations and the making of connections between different representations, is not just a one cycle problem solving process, but rather points to the multiple and iterative cycles in the construction of a model. Learners model problems or problem situations by using existing knowledge and strategies to make sense of the problem and to find a solution. When this initial model is used again and adapted or refined, a system of relationships that is reusable and generalisable is created. In a modeling approach to mathematics the reusing and generalising of models are central activities (Doerr & English 2003:113). Thus modeling proceeds beyond a one-cycle problem going from the given to the goal, it is an iterative process in which the model is simplified as required. In the process learners construct new knowledge by progressing from the concrete (real-world) to the abstract (representation).

2.2.4 MODELING IN THE SCHOOL CURRICULUM

In the Mathematics Learning Outcomes of the RNCS of South Africa (DoE 2002), there is no explicit mention of modeling or the use of models in the Intermediate or Senior Phase (Grade 4-9). The four steps of Polya's problem solving model (discussed in 2.3) and also the first four steps in the construction process of a model (Dossey et al 2002:115-117), however, correspond with the four steps in the problem solving process described by the RNCS. Each of the four steps detailed in the RNCS is given with the corresponding steps in the construction of a model in brackets: making sense of the problem (identifying the problem); analysing and synthesising (make assumptions – classify the variables and determine interrelationships between the variables); determining and executing solution strategies (solve and interpret the model); validating and interpreting solutions appropriate to the context (verify the model). The National Curriculum Statement (NCS) states that Grade 10 to 12 learners should use models and represent these mathematical models in different ways in the Learning Outcome *Functions and Algebra*. The RNCS for the Foundation, Intermediate and Senior Phases (Grade 0-9) as well as the NCS for Further Education and Training (FET: Grade 10-12) is set in Outcomes-Based Education (OBE) which has constructivism as premise. The constructivist approach to teaching and learning asserts that conceptual knowledge cannot be transferred from one person to another, but must be constructed by each learner solely on the basis of own experiences (English & Halford 1995:11). The construction of appropriate and internally consistent understandings and knowledge happens through modeling, which means that the use of modeling and models are implied in the OBE curriculum. As mentioned before, models are not only constructed by individuals, but are often developed by groups, involving cognitive as well as social functions. Thus models are related to social constructivism, linking modeling with one of the critical outcomes of the RNCS, specifically to work effectively with others as members of a group or team.

The NCTM focuses on modeling as mathematical representation of the elements and relationships of a simplified version of reality, and the use of mathematical models for problem solving. The word *model* in mathematics education is used in different ways (NCTM 2000:70). The NCTM reviews different uses of the word in mathematics education, ranging from manipulative models used in class; exemplification or simulation when a teacher ‘models’ a problem for the learners; to being a synonym for representation. The term ‘mathematical model’ in the context of school mathematics is seen as a mathematical representation of the relationships and elements within an idealised interpretation of a complex phenomenon and can be used to solve problems and better understand the phenomenon. This implies the use of representations not only as reproduction of the situation, but also as tool to interpret the model and the phenomenon modeled.

The link between modeling and representation is strongly emphasised in this statement of the NCTM. Mathematising is depicted as a form of modeling, involving the use of specialised languages, symbols, graphs, pictures, concrete materials, and other notation systems to develop mathematical descriptions that make obviously make heavy demands on learners' representational capabilities (Lesh & Doerr 2002:366, 367). The meaning of a model or conceptual system is closely interwoven with representation systems that interact with each other. The representation system may be written symbols, spoken words, pictures or diagrams, concrete manipulatives or experience-based metaphors. Lesh and Doerr (2000:363) characterise the difference between models and representations in terms of systems and objects functioning within these systems:

> …although the term *models* tends to emphasize the dynamic and interacting characteristics of the systems being modeled, the term *representations* tends to draw attention to the objects within these systems. Models tend to refer to functioning whole systems, whereas representations tend to be treated as inert collections of objects to which manipulations and relationships must be added in order to function.

The relationship between models and representations will be further discussed in Chapter 3.

2.2.5 MODELS AS TOOLS TO UNDERSTAND LEARNING

Models are not only used to study and represent real-world situations through mathematics, but can be used as conceptual and representational tools to understand the way in which mathematical concepts are acquired, developed, and applied by learners, teachers, researchers and educators (Lesh, Carmona & Post 2002:89).

Van der Walle (2004:13) provides a number of verbs to indicate the nature of the activities in an elementary mathematics classroom where learners are doing mathematics. These thinking tools are: explore, investigate, conjecture, solve, justify, represent, formulate, discover, construct, verify, explain, predict, develop, describe and use. These activities foster mathematical thinking, indicating the process of *making sense* and *figuring out,* which is in essence modeling. Modeling in mathematics cannot take place without the happening of one or more of these typical mathematical processes, which points to the close relationship between problem solving and modeling. The learner directly relates modeling as well as problem solving to the process of making sense of a mathematical situation. Without well-chosen model-eliciting problems modeling cannot successfully take place. Learner responses to model-eliciting problems or activities are not just solutions, but powerful conceptual tools. To model means to learn. Learners' descriptions, explanations and justifications reveal their thinking and they tend to learn and document what they are learning. This documentation leaves a trail showing the development of learners' thinking, so that the teacher guides learners to go "…beyond thinking *with* these conceptual systems to also thinking *about* them" (Lesh & Doerr 2000:376). This description corresponds with Gravemeijer's description of emergent models where a model *of* a certain situation becomes a model *for* more formal thinking. The more formal model or mode of thinking is the product of reflective practice where learners think and talk about their informal models, acting upon them and in the process progressing to a more abstract level of thinking. The modeling perspective advocated by Lesh and Doerr (2000:376) is based on the tenet that the most important goals of mathematics instruction are to help learners make sense of the kind of the complex systems that are pervasive in our technology-driven society and learners need to develop powerful mathematics models to do this. In summary it could be said that models provide a focus to inquiry and interaction that helps scientists and learners to understand and communicate about the phenomena being modeled.

When perusing the literature on modeling and problem solving, there is a significant correspondence in the terms used to describe modeling and problem solving, such as: making sense and figuring out; real-world problems; application in different contexts; central to understanding; from the concrete to the abstract; constructing knowledge - in keeping with constructivism; building from what is already known; etc. In the next section the influence of a modeling approach on how problem solving is viewed, will be discussed.

2.3 MODELING AND problem solving

Modeling and problem solving are closely connected. The application of a modeling perspective to problem solving is a relatively new occurrence in Mathematics Education. The use of the same terminology for describing modeling and problem solving was mentioned in 2.2. In order to understand the relationship between the two processes, one needs to investigate problem solving as process.

2.3.1 The process of problem solving

Problem solving is a process by which questions are answered or situations are dealt with, and has been defined as "…what to do when you don't know what to do" (Johnson & Herr 2001:5). O'Daffer et al (2002:39) offer a more sober definition of problem solving as

> … a process by which an individual uses previously learned concepts, facts, and relationships, along with various reasoning skills and strategies, to answer a question or questions about a situation.

Problem solving became the focus in the teaching and learning of mathematics since the 1980s (Cockcroft 1982) and brought a shift in thinking about mathematics instruction. Traditionally, instruction started where the teacher was instead of where the children were. The teacher taught the mathematics, the learners practised for a while and then had to solve problems using the new practised skills or ideas. This methodology is strongly engrained in our culture but rarely works well (Van de Walle 2004:37). It takes as starting point the teacher's knowledge and strategies, assuming the learners at that time possess the knowledge and ideas to make sense of the teacher's explanation and that there is only the teacher's way or no way. Dewey, as early as 1926, said that school instruction is plagued by a push for quick answers and criticized the evasion of a feeling of uncertainty by teachers and learners, which may lead to the search for alternative methods of solution. Understanding and the quality of methods are then replaced by a single, mechanically executed procedure. Dewey viewed the quality of mental process, not just correct answers, as something that could cause a revolution in teaching:

> Probably the chief cause of devotion to rigidity of method is, however, that it seems to promise speedy, accurately measurable, correct results … Were all instructors to realize that the quality of mental process, not the production of correct answers, is the measure of educative growth, something hardly less than a revolution in teaching would be worked (1926:206-207).

Dewey's opinion is aligned with the aims expressed in the mathematics curriculum statements of many countries, such as in the United States and South Africa. Since the publication of the 1989 Curriculum Standards of the NCTM in the United States, evidence accumulated that problem solving is a powerful and effective vehicle for learning and the development of mathematical thinking. Over the last twenty five years, the influence of constructivism in the structuring of mathematics curricula and teacher training programmes brought a growing awareness of the nature of the processes involved in the individual and social construction of mathematical knowledge. The role of problem solving in building new mathematical knowledge came under the spotlight and soon became one of the five pillars of the teaching and learning approach of the NCTM, as expressed in the Problem Solving Standard:

Instructional programs from Prekindergarten through grade 12 should enable students to:

- build new mathematical knowledge through problem solving
- solve problems that arise in mathematics and in other contexts
- apply and adapt a variety of appropriate strategies to solve problems
- monitor and reflect on the process of mathematical problem solving (NCTM 2000:52)

Modeling and problem solving were at first regarded as two separate processes, existing along each other. Later educators and researchers however realised that the two processes were related. In traditional problem solving only one modeling cycle is needed, while applied problem solving is thought of as "a special case of generalized, content-independent, problem solving processes" (Lesh & Doerr 2000:379). The way learning is realised through problem solving depends on the teaching approach of the facilitator. Three teaching approaches to problem solving will now be described with reference to how these approaches influence learning.

2.3.2 Approaches to problem solving

Three approaches to problem solving can be distinguished (Schroeder & Lester 1989:32,33), being teaching *about* problem solving; teaching *for* problem solving and teaching *via* or *through* problem solving. These approaches will now be discussed, with reference to their influence on learners' understanding.

- Teaching *about* problem solving

This approach uses Polya's four-step model as starting point. The four phases are: understanding the problem; making a plan; carrying out the plan and reflecting on the results. These phases are directly taught, as well as a number of strategies to from which learners can choose to solve the problem. The limitation of this approach is that problem solving becomes yet another topic in the curriculum and may be taught in isolation from other content and relationships. This approach does not foster original thinking, because learners are given a variety of solutions to choose from and problem solving becomes an exercise in choosing one of the supplied solutions.

- Teaching *for* problem solving

  The teaching *for* problem solving approach is about applying acquired knowledge to solve routine and nonroutine problems, so teachers want to prepare learners to transfer gained knowledge to other contexts by exposing them to many instances of the mathematical concept and structures they are studying. Problem solving according to this approach is an activity learners become involved in only after they have studied a new concept or algorithm, and then just to apply recently learned knowledge and skills. This teaching approach limits learners' thinking, as they are not encouraged to find their own solutions, but are supplied with one algorithm that they have to practise using.

- Teaching *via* or *through* problem solving

  The approach of teaching via problem solving uses problems both as purpose for learning mathematics and as primary means of learning mathematics. A problem situation is the point of departure and strategies and techniques for solving the problem are developed by the learners themselves, in the process moving from

> … the concrete (a real-world problem that serves as an instance of the mathematical concept or technique) to the abstract (a symbolic representation of

> a class of problems and techniques for operating with these symbols) (Schroeder & Lester 1989:33).

> This approach requires an inquiry-orientated classroom atmosphere where higher order thinking processes are fostered through problem solving experiences and where learners are encouraged to reflect on their own solutions and that of others.

The approach of teaching through problem solving is consistent with the ideas of reflective inquiry (Hiebert, Carpenter, Fennema, Fuson, Human, Murray, Olivier & Wearne 1997:156-159) in which knowledge is constructed through active participation. The approach of teaching mathematics through problem solving is also congruous with the view expressed in the NCTM's Principles and Standards (2000:52), highlighting the dual role of problem solving as an integral part of mathematics learning: "Solving problems is not only a goal of learning mathematics but also a major means of doing so". Van de Walle (2004:36,38) supports the ideas reflected by Hiebert et al (1997:25) "… that mathematics teaching and learning should happen *through* problem solving, and not *for* problem solving". This means that problem-based tasks or activities are the vehicle by which mathematical concepts and understanding is developed. This view coincides fully with the NCTM's interpretations of problem solving. Van de Walle continues by remarking that problem solving is the way in which "… most, if not all, important mathematics concepts and procedures can best be taught" (Van de Walle 2004: 36). This does not mean that every moment of every lesson will be spend solving problems, but rather that problems and problem solving situations rising naturally from studying key concepts, should be used.

2.3.3 Problem solving in the Revised National Curriculum Statement

The RNCS (DoE 2002:5) expounds that one of the seven critical outcomes for Outcomes-based Education in South Africa "…envisages learners who are able to identify and solve problems and make decisions using critical and creative thinking". The Statement about Mathematics as learning area declares that one of the unique features of learning and teaching Mathematics is problem solving. The learning of mathematics through problem solving has a strong real-life connection and is evident in the curricula of different countries. In the RNCS in South Africa (DoE 2002:43,71) the real-life connection is stressed in the Assessment Standards for Mathematics, stating that problems should be solved in context, including contexts that may be used to build awareness of other learning areas, as well as human rights, social, economic and environmental issues. The Realistic Mathematics Education (RME) approach in the Netherlands, rooted in Freudenthal's interpretation of what mathematics is, has reality as starting point. RME is based on the idea that mathematics is an activity that involves "…solving problems, looking for problems, and organizing a subject matter resulting from prior mathematizations or from reality" (Gravemeijer et al 2000:236). Freudenthal (as quoted by Treffers 1993:94), wrote the following about mathematising:

> The globally structuring force, as we called it, should be lived through reality. Only this way can we teach mathematics fraught with relations, can we be sure that the student integrates the mathematics he has learned, and can we guarantee the applicability of mathematics.

The use of problems contextualised in real-life is not only the focus of the RME approach from the Netherlands, but is also upheld in the approach in the 2000 version of the *Principles and Standards for School Mathematics* of the NCTM as reported above. Holmes (1995:2) also draw attention to the link between problem solving and real life by explaining that the context for problem solving in the elementary school is found in stories, text material, school assignments, and real-life situations. She then elaborates on the value of problem solving in learning, as both a means and an end in elementary mathematics instruction. Problem solving enables learners to see the relevance of mathematics to other subjects and the real world. Children should have the ability to solve real-world problems that have mathematical dimensions, as well as textbook or teacher-posed problems. Problems should help learners make connections between mathematics, their lives, and other school subjects.

Dossey et al (2002:71) maintain that problem solving "is at the very heart of *doing mathematics*" and gives an example of the different levels on which the same problem can be solved. To find the least common multiple of 9 and 24, in sixth grade a learner might write down the first 10 multiples of both numbers and then find the smallest match. An eleventh grader might find the factors of the two numbers, which might be a routine computation on this level. A college student might divide the product of 9 and 24 by the greatest common divisor of the two numbers, 3, to get the quotient of 72. The theorem the college student applied to get to the solution, is written as

$$\gcd(a,b)\cdot \mathrm{lcm}(a,b)=ab$$

To compute the least common multiple of 9 and 24 using this theorem, we get

$$\mathrm{lcm}(a,b)= \frac{ab}{\gcd(a,b)}$$

As illustrated above, problems can be solved on different levels and by using a variety of different strategies. Problem solving then is the process by which we deal with the problem situation or answer the question. The strategy used to solve a problem, will vary from learner to learner. What is a routine and quick computation problem for one person may be a tiresome and difficult problem for another.

Polya (1945), in his book "How to solve it", postulates a broad, flexible problem solving model without detailing specific strategies. Polya's problem solving model shows similarities to the steps in model construction offered by Dossey et al (2002:115-117), which brings us to the connection between modeling and problem solving.

2.3.4 The contextualisation of problem solving in the modeling process

Problem solving took on a deeper dimension when it was contextualised in the modeling process. Modeling and the use of models was mainly used in fields of Applied Mathematics; it is only in the past decade or so that models and modeling became the lens through which the construction of knowledge was studied in Mathematics Education (Lesh et al 2002; Dossey et al 2002; Lesh & Doerr 2000; Gravemeijer et al 2000).

The word *model* in connection with problem solving is used by Carpenter, Fennema, Franke, Levi and Empson (1999:55) who describe problem solving as modeling. Counting and direct modeling strategies are cited as specific examples of the fundamental principle of modeling and should be seen "as attempts to model problems rather than as a collection of distinct strategies". Carpenter et al (1999:55) illustrate problem solving by modeling by the following example:

> 19 children are taking a minibus to the zoo. They will have to sit either 2 or 3 to a seat. The bus has 7 seats. How many children will have to sit 3 to a seat and how many can sit 2 to a seat?

The typical way for young children to solve the problem is to take 19 counters and attempt to place them in 7 groups of either 2 or 3 in a group until all the counters are used up. Some children use up all the counters by systematically sorting them into 7 groups while others achieve the outcome by trial and error. Even though children will not place the same number of counters in a group and the answer is not the number of counters in a group, the strategies used are similar to those used to solve a partitive division problem (the idea of equal sharing). Central to understanding children's solutions to problems is the fact that they model the problem situations directly and when children have difficulty solving a problem, it often is because they can not figure out how to model the problem. The focus on problem solving as modeling not only sheds light on children's strategies for solving specific problems of addition, subtraction, multiplication and division which are discussed as examples, but also provides a synthesised framework for problem solving in the primary school:

> This conception of problem solving provides a foundation for integrating instruction in problem solving with instruction in fundamental mathematics concepts and skills. Not only can symbols and procedures be presented as ways of representing problem situations, but the construction of procedures for calculating answers can be presented as a problem solving task (Carpenter et al1999:55).

Young children exhibit intuitive modeling skills to analyse and solve problems, but the principle applies to more complex problem situations as well. Carpenter et al (1999:56) suggests that regarding problem solving as modeling and thus as a "meaning-making" activity, will influence children's conceptions of problem solving and of themselves as problem solvers. If children use problem solving as means of making sense of problems and problem situations from an early age, they may come to regard doing mathematics as a meaningful activity.

Teaching through problem solving shows similarities to a modeling approach. The description of the movement in problem solving from "… the concrete (a real-world problem that serves as an instance of the mathematical concept or technique) to the abstract (a symbolic representation of a class of problems and techniques for operating with these symbols)" (Schroeder & Lester 1989:33) is in accord with Gravemeijer's notion of emergent modeling from a situational level with concrete experiences to more formal (abstract) levels of reasoning.

Lesh and Doerr (2000:379-380) however, contend that modeling embodies much more than the traditional narrow understanding of problem solving. Problem solving research in the past three decades largely centered on Polya's four step process proceeding from givens to goals, with metacognitive processes of expert problem solvers added in more recent research. Applied problem solving then is envisaged as a special case of generalised, content-independent, problem solving processes. The kind of task or problem plays a crucial role in modeling activities. The problems presented to learners should be problems that they find interesting and meaningful but which they cannot easily solve using routinised procedures or drilled responses (Hiebert et al 1997:115). Useful responses to modeling tasks involve different kinds of heuristics and strategies than those that have been emphasised in traditional problems with only a single interpretation cycle. In traditional problem solving, being a special case of modeling, only one modeling cycle is necessary to get from the givens to the goals. In model-eliciting tasks, there are multiple modeling cycles with multiple ways of thinking about givens, goals and solutions (Lesh & Doerr 2000:380). These authors describe this process in the following way:

> It is crucial to recognize that these ways of thinking evolve over the course of the activity in ways that are increasingly stable. Modeling activities are much more than the mapping of problem information (givens) onto an invariant model in order to reach a solution (goals). In modeling, it is the interpretation and the model itself that are constructed, modified, refined, or extended.

This multi-cycling modeling process should not be considered a continuous linear sequence, but rather as a "back-and-forth" sequence. When adopting a modeling perspective on learning and problem solving in Mathematics Education, the learner is not viewed as a traditional problem solver, but rather as a model builder.

Problem solving and modeling are both linked to internal and external representations (Dossey et al 2002; Lesh et al 2002; Lesh & Doerr 2000; Cifarelli 1998). A learner's problem solving strategies and models become accessible to others to the degree that internal and external representations of these models and strategies become accessible. Learners form internal representations, which they then express in external form. These internal and external systems may be expressed as spoken language, written symbols, pictures, diagrams, and concrete models. Lesh and Doerr (2000:363,364) propose that modeling involves the interactions of three kinds of systems: (a) internal conceptual systems, (b) representational systems that function both as externalisations of internal conceptual systems and as internalisations of external systems and (c) external systems that are experienced in nature or that are artifacts that were constructed by humans. The third kind of external system that they describe, are systems or artifacts such as economic systems, communication systems and mechanistic systems. The three systems, conceptual (internal) systems, notation (external) systems and systems or artifacts (external) are distinct but also partly overlapping, interdependent and interacting. The authors describe the boundaries of these systems as "…fluid, shifting and at times ambiguous" as depicted in Figure 2.2.

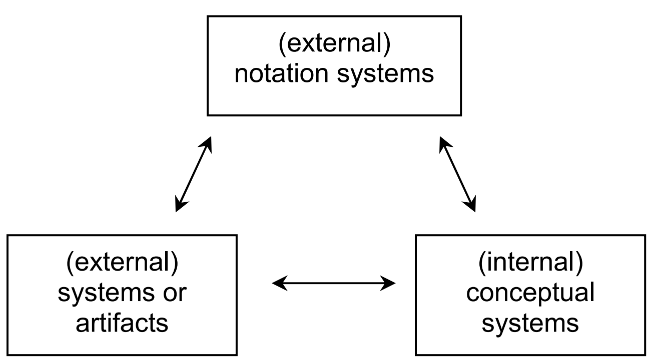


Fig. 2.2: Modeling interactions among three types of systems

The question of whether models refer to systems inside or outside the mind is often asked, because mathematical models and conceptual systems are often described as though they have no connection to representations, tools or external artifacts. In practice, however, it is clear that models and conceptual systems almost always function through the use of tools or representational systems, shedding light on different aspects of the specific system (Lesh & Doerr 2000:362). Lesh et al (2002:89) maintain that models are conceptual (largely internal) systems that are expressed by using external notation systems, which in turn are used to construct, describe or explain the behaviours of other systems and reside in the minds of the learners. These models manifest in equations, diagrams, computer programs and other representational media and are comparable to cognitive structures described in the cognitive psychology:

> Because models are conceptual systems, they are partly internal and are similar to the conceptual systems that cognitive scientists refer to as cognitive structures (Lesh et al 2002:89).

As cognitive systems seem to be largely internal and external representation systems seem to consist outside a learner's mind in a form that are accessible to others, it is also true that part of the meaning of these representation systems can not be shared with others (Lesh & Doerr 2000:364). Conceptual systems may only later become accessible or partly accessible to others when they become stable systems:

> …the constructs and conceptual systems that are in human minds today may be used to create systems that function as objects in the world tomorrow, and systems that are created for their own sake today may be used to make sense of other systems tomorrow (Lesh & Doerr 2000:364).

In the next section internal conceptual representations or mental models and the cognitive processes by which such models are constructed, will be discussed.

2.4 MENTAL MODELS

The relationship between models and the representation thereof surfaced in the discussion of models and modeling. The notion of models being internal conceptual systems that are expressed in external notation systems and thereby becoming accessible to others, were discussed. A question that ensues is about the nature of mental representations and the role they play in mathematical understanding. David Bartholomew as quoted in Wild and Pfannkuch (1999:223) connect mental models with our interpretation of everyday experiences:

> We all depend on models to interpret our everyday experiences. We interpret what we see in terms of mental models constructed on past experience and education. They are constructs that we use to understand the pattern of our experiences.

One of the most important goals of mathematics education is that learners should understand the mathematics they encounter. Extensive research in the field of cognition and cognitive science has been done in which the important role of mental representations in cognitive processes and understanding has been described.

Mental models or mental representations are internal mental structures that are related to a section of the real world (Halford 1993; English & Halford 1995; Dossey et al 2002). Hiebert and Carpenter (1992:78) contend that mental representations should be considered in terms of networks of interrelated ideas where the depth of understanding is determined by the number and strength of connections. When talking or thinking about any mathematical object or process, we all relate to something we have in mind, a mental representation of the object or process under consideration. Dreyfuss (1991:31) describes the representation of a concept as a generation of an instance, specimen, example or image of it. This mental representation refers to internal schemata or frames of reference that a person uses to interact with the world, in other words, to model reality.

Cognitive processes involve operations on mental representations and mental representations in turn depend on structural correspondence and are not just pictures in the mind but have different codes. Halford (1995:21) discusses different ways of representing knowledge to give an insight into the way children understand and operate with mathematics. One of the ways of representing knowledge is by means of mental models. The process of the understanding of mathematics points to the existence of a mental model or internal cognitive representation reflecting the structure of a concept (Halford 1995:48). He defines mental models as "representations that are active while solving a particular problem and that provide the workspace for inference and mental operations" (1995:23) and suggests that "to understand a concept entails having an internal cognitive representation or mental model that reflects the structure of that model". Mental models are content specific and usually reflect an individual's experience. Mental models may, according to the requirements of the task, be retrieved from memory or constructed to meet these requirements. If the task is analogous to a problem previously modeled, the model can be retrieved from memory and adapted if necessary. If the task involves an unfamiliar problem situation, a new mental model has to be constructed.

Mathematical understanding gives rise to the development of appropriate strategies for a task (Halford 1995:48). Mental models thus serve as guides in the development of learning strategies and the acquisition of cognitive skills. The important characteristics of a performance are to be found in the way the strategy was planned or developed from the person's concept of the task, rather than in the surface form of a strategy. Strategies and mental models can be adapted when circumstances change; therefore skills based on understanding can be transferred from one domain to another and be used for the generation of new appropriate strategies. Some learning occurs without understanding, for example the driver of a car does not necessarily understand how the car works, but can still drive the car. When learning is based on understanding however, it can be extended and adapted. Learning processes can be metacognitive, that is based on mental models, or associative, that is based on practice. Halford (1995:55) summarises the power of learning with understanding, which is learning based on mental models, as follows:

> Associative learning provides relatively effortless accumulation of large amounts of information, but it is subject to associative interference, and transfers only to similar contexts. Learning based on mental models, or concepts of the task, is more effortful, but potentially more powerful, and can mediate cross-domain transfer.

Mental models vary in adequacy, and the adequacy of the mental model is consistent with the quality of understanding. The role of mental models in the understanding of mathematics and in the problem solving or modeling process is crucial. Halford (1995:48) argues that understanding means to have a mental model and that "to understand a concept entails having an internal cognitive representation or mental model that reflects the structure of that concept. The representation defines the workspace for problem solving and decision making with respect to the concept". The view that understanding has a direct link to mental models is also maintained by Johnson-Laird (1983:2). Mental models can be considered as a type of representation (Martinez 1999:28) and real understanding will surface and be discernable to others when a learner can produce multiple representations and translate between these representations, where representations are equated to forms of knowing (Gardner 1991:18).

As discussed in 2.3.4 internal representations are not under normal circumstances directly observable by other people. It is only when expressed in observable form such as words, graphs, pictures, equations, etc., that they are accessible to anyone with suitable knowledge. The teacher only comes to understand a learner's modeling of a mathematical concept when observing the external representations the learner produces. The focus on understanding is essentially the question of which model the learner used and how he modeled the problem. Likewise, when the learner gain new insights, the teacher would like to know which new model substitutes the previous model in the thought framework of the child.

These representations are the internal (a mental image) images at first, which through manipulation can become external representations. These internal and external representations are interacting and interdependent, causing the meaning of the representations to be changing.

> Thus, the meanings and functions of students' representations are not static; they are continually evolving. The same is true for the underlying mathematical constructs that the representations embody, as well as for the external systems that they describe (Lesh & Doerr 2000:368).

Representations will be discussed at length in Chapter 3.

In this section, the relationship between modeling, mental models and understanding has been discussed. This study however is not about modeling, mental models and representation in general, but specifically deals with the organisation and representation of statistical data. Against this background, the question now emerges: What is data modeling and what role does data modeling play in the statistical thinking and understanding of young children? These questions will be discussed in the next section.

2.5 DATA MODELING IN statistics

Different views of the nature of mathematics were explored when modeling as a way to make sense of the real world was discussed (see 2.2.1). When considering data modeling, it is necessary to first explore the nature of statistics.

2.5.1 The nature of statistics

Statistics is described by Moore (Wild & Pfannkuch 1999:250) as the science of data and further qualifies it by saying that it is a science of variability and a way to deal with uncertainty that surrounds us in our daily life, in the workplace and in science (Moore 1997:123). Bakker (2004:1) state that statistics '… is used to describe and predict phenomena that require collections of measurements". Schaeffer (2000:173) also refers to statistics in terms of quantitative information; he describes statistics as a way of thinking about quantitative information, a process of "…thinking through a problem from inception, to clarification, to data, to analysis, to conclusion" and points out that the process is more important than the parts. In this process different areas of mathematics such as number concepts, geometry, algebra and functions are integrated and ideas are communicated, from understanding the initial practical problem to the statement of conclusions in such a way that others can understand it.

2.5.2 Statistics in school curricula

Statistics are in most school curricula considered as a branch of mathematics dealing with collection, analysis, interpretation and representation of data. Statistics and probability are sometimes referred to as *stochastics* (Reading & Pegg 1995:140; Truran 1997:538; Truran, Greer & Truran 2001:258; Shaughnessy & Watson 2003:192). As mentioned in 2.2, the term stochastic refer to uncertainty, as opposed to the term deterministic, which refers to certainty. The term *stochastics* is not used in school curricula and even the word statistics does not appear in most national curricula, but rather the terms data handling and probability. Shaughnessy, Garfield and Greer (1996:208) hold that various national curriculum documents agree substantially that learners should

- collect, organise and describe data
- construct, read, and interpret displays of data
- explore chance and random phenomena
- formulate and solve problems that involve collecting and analysing data
- describe and interpret data
- create visual and graphical representations of data and
- develop a critical attitude towards data

The RNCS (DoE 2002:56) for the Intermediate Phase (Grade 4-6) states that the learner in the fifth Learning Outcome, which is Data Handling, should be able to collect, summarise, display and critically analyse data in order to draw conclusions and make predictions, and to interpret and determine chance variation. This should enable learners to participate meaningfully and responsibly in political, social and economical activities:

> In this Learning Outcome, the learner will develop a sense of how Mathematics can be used to manipulate data, to represent or misrepresent trends and patterns. The learner will develop a sense of how Mathematics can provide solutions that sustain or destroy the environment, and promote or harm the health of others. The learner is thereby able to use Mathematics effectively and critically, showing responsibility towards the environment and health of others (DoE 2002:38).

The process of data modeling or statistical investigation comprises collecting, summarising, displaying and critically analysing data. “Selecting the appropriate type of analysis and designing a study to support this analysis are major components of statistical problem solving” (Lajoie 1998: xix).

2.5.3 Data modeling

Model development as learning was discussed in 2.1. This connection between model development and learning is also pointed out by Doerr and English (2003:111), but with specific reference to data modeling contexts. Lehrer and Schauble (2000:52) define data modeling as

> a multicomponential process of posing questions; developing attributes of phenomena; measuring and structuring these attributes; and then composing, refining, and displaying models of their relations.

The process described in this definition corresponds with Lehrer and Romberg’s description of data modeling (1996:70) as “the construction and use of data”.

Lehrer and Schauble's definition of data modeling discussed above touches upon the basic components of the statistical process, which will next be scrutinised.

2.5.4 The process of statistical investigation

The process of statistical investigation lies at the heart of statistics and modeling. There are four basic processes that are linked to all the elements in statistics content. A concept map of the process of statistical investigation is displayed in Figure 2.3. In the concept map, the four different components of the process and the elements of statistics content related to them are included and can be explained as follows (Friel & Bright 1998:96):

- **A question is posed** because a problem must be solved. The problem involves exploring one or more of the following: describing a data set, summarising what is known about a data set, comparing and/or contrasting two or more data sets, or generalising from a set of data in order to make predictions about the next case or the population as a whole.
- **Data collection** involves the identification of the population to be studied and the methods for data collection. In the case of sampling, different types of sampling can be considered, e.g. random sampling, convenience sampling, or a census. Randomness, representativeness and bias must be taken into account.

- **Analysis of the data** may include the following: describing and/or summarising a single data set, comparing and/or contrasting two or more data sets, or making predictions and/or assessing implications from one or more sets of data. This is done by organising, sorting, classifying and displaying data using tables, diagrams, and graphs. It also involves the determining of measures of central tendency (mean, median and mode), measures of variation (range and standard deviation) and measures of association (line of best fit and correlation coefficients).

- **Interpreting results** takes us back to the purpose of the investigation and to the question posed at the beginning of the process. How do the collected and analysed data help us find answers to the original question?

The concepts displayed show what teachers and learners need to know about statistics and the process defines the skills needed. In the context of the school, questioning and exploration in hands-on activities with open-ended questions are needed for learners to construct the necessary knowledge and develop the appropriate skills.

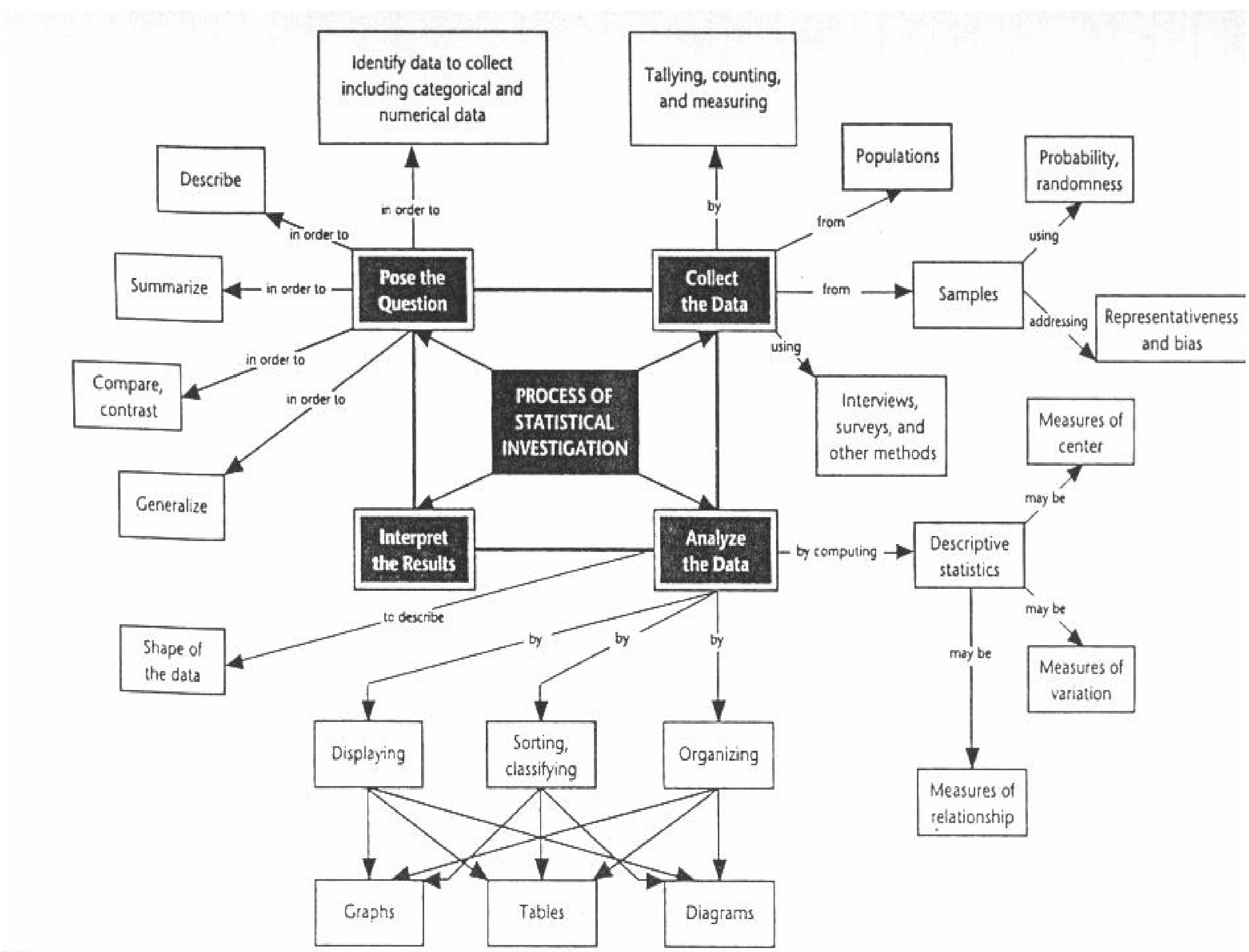


Figure 2.3 Concept map of the process of statistical investigation
(Friel & Bright 1998:95)

The four components in the process of statistical investigation show in broad sense similarities to Polya's problem solving process. The first component, the posing of a question to solve a problem, points to the analysis of the problem hence relating to Polya's first phase where the problem in question is analysed. The second component is the collection of data with which the problem can be answered. This component corresponds with Polya's second phase, the making of a plan to solve the problem. The third component on the concept statistical map comprises of the analysis of data, which is in line with the carrying out of the plan (Polya). Interpreting the results takes one back to the purpose of the investigation, namely the problem with which the investigation started. Reflection on whether the collected data and the analysis thereof helped find answers to the initial question shows a direct correspondence to the last phase in Polya's problem solving model, namely reflecting on the results. The statistical process can thus be described as statistical problem solving through which statistical knowledge and skills are developed and fostered.

The statistical process as postulated by Friel and Bright (1998:95) in the concept map is at the heart of statistics and consistent with the definitions of statistics as discussed in the first paragraph of this section about data modeling.
Wild and Pfannkuch (1999:223) contend, "… the thinking and problem solving performance of most people can be improved by suitable structured frameworks". They developed a framework for thinking patterns involved in problem solving, strategies for problem solving, and the integration of statistical elements within problem solving. They argue that all thinking uses models and that the main contribution of the discipline of Statistics to thinking has been its own distinctive set of models or frameworks for thinking about the aspects in a statistical investigation (Wild & Pfannkuch 1999:227). The basis of teaching in any area is the development of a theoretical structure with which to make sense of

experience, to learn from it and to transfer insights to others. In a data modeling approach to statistical thinking is the cornerstone. Before data modeling and the mapping of information in data and context knowledge can take place throughout the whole statistical process, an enormous amount of statistical thinking must be done (Wild & Pfannkuch 1999:224).

2.5.5 Statistical thinking

There is a scarcity of literature on statistical thinking. Existing descriptions and definitions are also focusing on different levels or serve different purposes, e.g. for professional statisticians, entry-level university students or for school learners. The American Society for Quality defines statistical thinking against the background of three fundamentals: All work occurs in a system of interconnected processes, variation exists in all processes, and understanding and reducing variation are keys to success (Wild & Pfannkuch 1999:257). Snee (Wild & Pfannkuch 1999:256) identifies process, variation and data as key elements of statistical thinking and defines statistical thinking as

> … thought processes which recognise that variation is all around us and present in everything we do, all work is a series of interconnected processes, and identifying, characterising, quantifying, controlling, and reducing variation provide opportunities for improvement (Snee 1990:118).

The ultimate goal of statistical investigation is learning in the context sphere. Learning in Statistics is not just the collecting of information, but involves synthesising the new ideas and information with existing ideas and information into an improved understanding. Wild and Pfannkuch (1999:225) formulated a four-dimensional framework for statistical thinking in empirical enquiry to organise some of the elements of statistical thinking. Dimension 1 comprises of the investigative cycle; Dimension 2 of the types of thinking; Dimension 3 of the interrogative cycle and Dimension 4 of dispositions (Fig. 2.4).

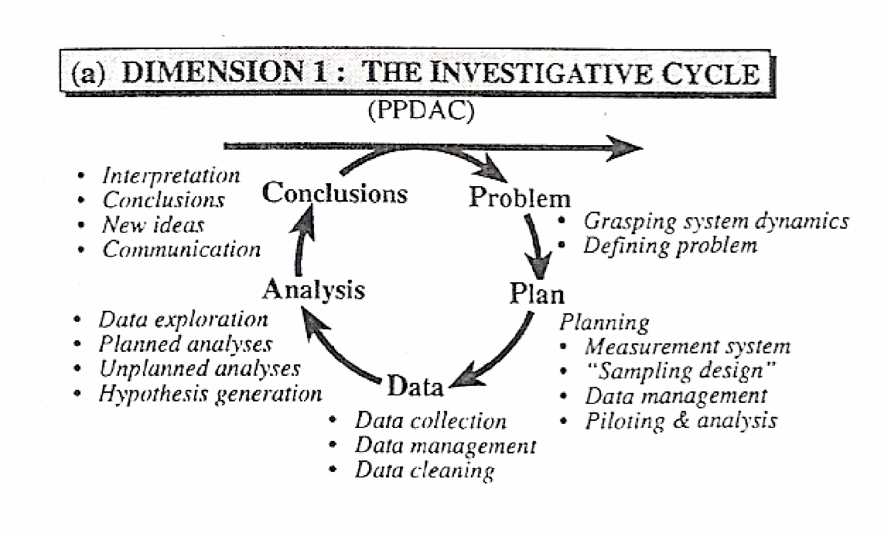


Figure 2.4 First dimension of the statistical thinking framework
(Wild & Pfannkuch 1999:226)

The **first dimension** relates to the way one acts and what one thinks during the course of a statistical investigation. The PPDAC model in Statistics adapted as core of this dimension incorporates the following aspects:

- **P**roblem (P)
- **P**lan (P)
- **D**ata (D)
- **A**nalysis (A)
- **C**onclusions (C)

The PPDAC model shows similarities to Polya's problem solving model discussed in 2.3.3. This problem solving model fosters the quality of mathematical thought just as the PPDAC model fosters quality of statistical thought. The PPDAC model is about abstracting and solving a statistical problem grounded in a 'larger' real problem.

Knowledge gained and needs identified within the cycles in this model may initiate further investigative cycles. This aspect of the statistical model reminds us of mathematical models situated in real-world problems and the iterative nature of model construction as described in 2.2.2.

The **second dimension** specifies general types of thinking common to all problem solving, e.g.

- strategic thinking
- seeking explanations
- modeling
- applying techniques

(b) DIMENSION 2 : TYPES OF THINKING

GENERAL TYPES

- **Strategic**
  - — planning, anticipating problems
  - — awareness of practical constraints
- **Seeking Explanations**
- **Modelling**
  - — construction followed by use
- **Applying Techniques**
  - — following precedents
  - — recognition and use of archetypes
  - — use of problem solving tools

TYPES FUNDAMENTAL TO *STATISTICAL* THINKING (Foundations)

- **Recognition of need for data**
- **Transnumeration**
  (Changing representations to engender understanding)
  - — capturing "measures" from real system
  - — changing data representations
  - — communicating messages in data
- **Consideration of variation**
  - — noticing and acknowledging
  - — measuring and modelling for the purposes of prediction, explanation, or control
  - — explaining and dealing with
  - — investigative strategies
- **Reasoning with statistical models**
- **Integrating the statistical and contextual**
  - — information, knowledge, conceptions

Figure 2.5 Second dimension of the statistical thinking framework

(Wild & Pfannkuch 1999:226)

There are five types of thinking that is the essence of statistical thinking, namely

- recognition of the need for data: the inadequacy of personal experience and verbal evidence leads to a desire to base decisions on deliberately collected and analysed data.
- transnumeration, which is a dynamic process of forming and changing data representations to arrive at a better understanding. Wild and Pfannkuch (1999:227) describes transnumeration as the most fundamental idea in a statistical approach to learning.
- considering of variation: the importance of variation in statistical thinking is stressed by various authors (Snee 1990; Snee in Wild & Pfannkuch 1999; Moore in Wild & Pfannkuch 1999; Breslow in Wild & Pfannkuch 1999; Biehler in Wild & Pfannkuch 1999). Statistical thinking in a modern sense is concerned with learning and making decisions under uncertainty, which mainly originates from omnipresent variation and which all aspects of life and everything we observe.
- reasoning with statistical models: Statistics uses a distinctive set of tools to think about and model problem situations. These frameworks or models are needed to reason about data and arrive at conclusions.
- integrating the statistical and the contextual: the raw material on which statistical thinking operates is statistical knowledge, content knowledge and information in data. These elements are synthesised in the thinking process to produce implications, insights and conjectures.

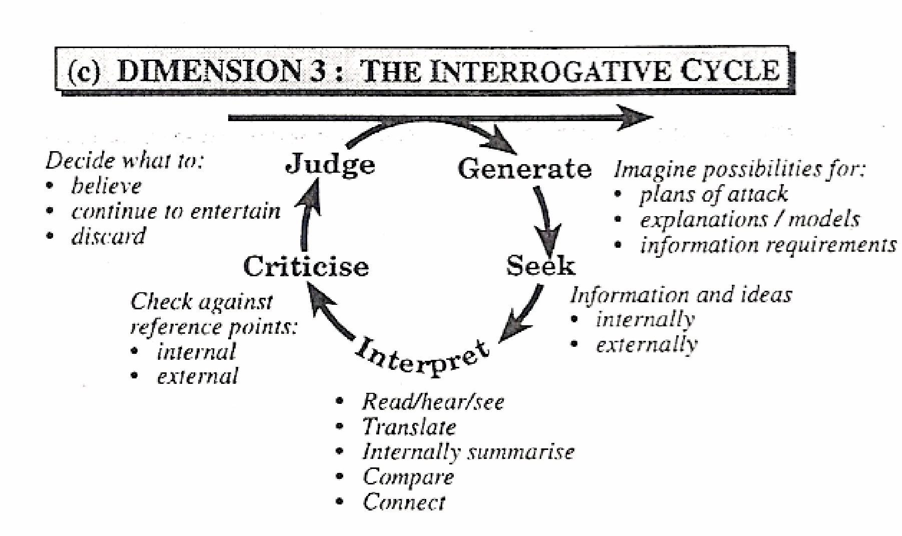
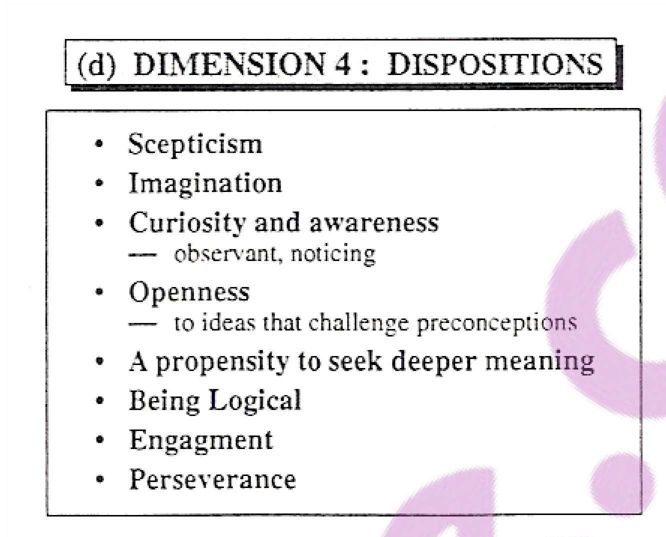


Figure 2.6 Third dimension of the statistical thinking framework
(Wild & Pfannkuch 1999:226)

The **third dimension** concerns a generic thinking process that is constantly used in statistical problem solving. The components of this process are

- the generation of possibilities for plans of attack, explanations or models with which to understand the data and information requirements
- the seeking of information (internally and externally)
- interpretation through a process of read/see/hear → translate → internally summarise → compare → connect, applying to all components of the statistical process
- checking and criticising incoming information against reference points and for internal consistency

- judgment, the endpoint of the criticising process, leading to decisions of what to keep and what to discard; applied to reliability of information, usefulness of ideas, practicality of plans, conformance with both context-matter and statistical understanding, etc.


(d) DIMENSION 4 : DISPOSITIONS

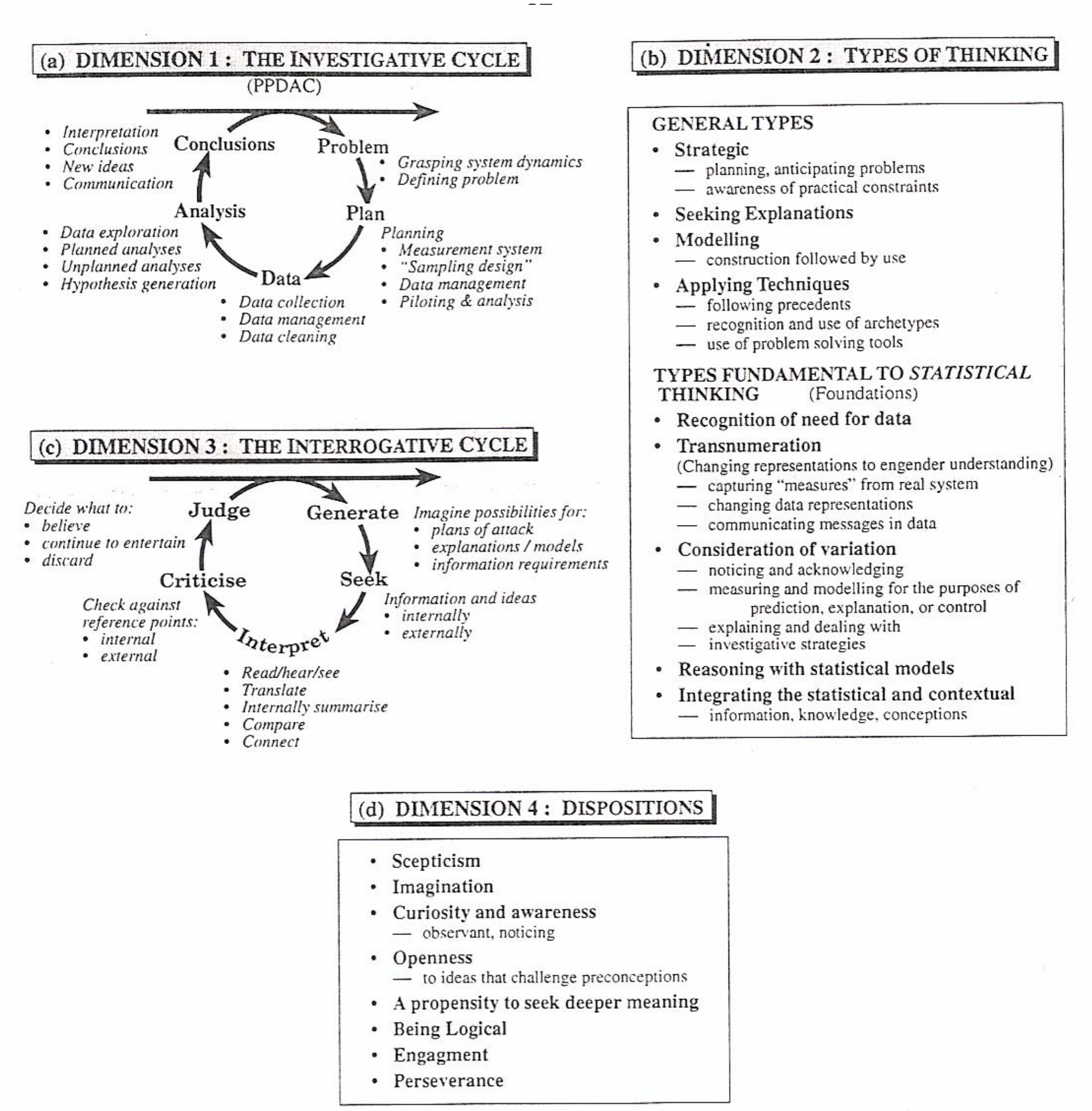
- Scepticism
- Imagination
- Curiosity and awareness
  — observant, noticing
- Openness
  — to ideas that challenge preconceptions
- A propensity to seek deeper meaning
- Being Logical
- Engagment
- Perseverance

Figure 2.7 Fourth dimension of the statistical thinking framework
(Wild & Pfannkuch 1999:226)

The **fourth dimension** involves personal qualities, which plays a role in statistical thinking. Dispositions influences or even initiate a person's entry into a thinking mode. The 'dispositions' of a person is problem dependent, because they can change to the extent that the individual is engaged by the problem. The relevant dispositions are:

- scepticism
- imagination
- curiosity and awareness
- openness
- a propensity to seek deeper meaning
- engagement and perseverance.

Figure 2.8 Four-dimensional framework for statistical thinking in empirical enquiry
(Wild & Pfannkuch 1999:226)



Statistical thinking takes place in all four dimensions at once, as explained by the authors: A thinker could be categorised as currently operating in the planning stage of the Investigative Cycle (Dimension 1), dealing with some aspect of variation in Dimension 2 (Types of thinking) by criticising a tentative plan in Dimension 3 (Interrogative Cycle) driven by skepticism in Dimension 4 (Dispositions) (Fig.2.8).

Snee proposes a simplified model for statistical thinking (Wild & Pfannkuch 1999:256), comprising of the three key elements, namely process, variation and data. He explains that all activity is a process (work or other). A process can be defined as any activity that converts inputs to outputs. The problems of empirical enquiry are connected with one or more processes; hence the process or processes concerned provide the ‘context’ for statistical work. He explains further that all processes vary which accounts for the fact that process improvement and problem solving get complicated quickly. The need to deal with this variation, guides us to make measurements as a way of characterising the process being studied and thus creating a (numerical) basis for comparison. The result of the measurement process is data. Statistical tools are used to analyse process data, which also shows variation because of the process and measurement system. Therefore, the elements of statistical methods are variation, data and statistical tools. The relationship between statistical thinking and statistical methods can be expressed in a diagram (Figure 2.5):

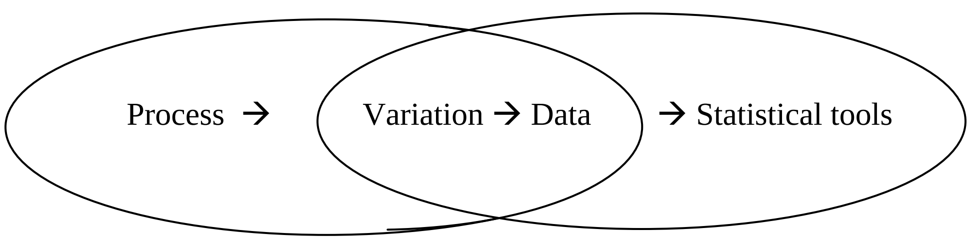


Figure 2.9: Relationship between statistical thinking and statistical methods

When considering statistical thinking, Moore (Wild & Pfannkuch 1999:250) also suggests a less complex structure of statistical thinking for beginners, because they lack intellectual maturity and content knowledge needed for full statistical problem solving. The age group involved in the research described in this thesis is still at a basic level in Statistics or Data Handling and can be expected to deal with only elementary levels of the aspects of statistical thinking. Mooney, Langrall, Hofbauer and Johnson (2001:438) have postulated a more appropriate framework for learners in Middle School, which is also appropriate for the target group of this research, the Intermediate Phase (Grade 4-6). The four statistical components included in the framework are:

- describing data
- organising and reducing data
- representing data
- analysing and interpreting data

This framework does not include the phases of question posing and data collection that are contained in the description of the statistical process of Friel and Bright (1998:95). Mooney et al (2001:438) characterise statistical thinking for middle school learners as

> …the cognitive actions that students engage in during the data-handling processes of describing, organizing and reducing, representing, and analyzing and interpreting data.

The processes involve the following:

- Describing data involves the explicit reading of data in the form it is presented, such as tables, charts and graphical representations.
- Organising and reducing data is about arranging, categorising or consolidating data to summarise.
- Representing data entails displaying data in graphical form.
- Analysing and interpreting data involves the identification of trends and making predictions or inferences from a data set.

The SOLO Taxonomy is used to describe statistical thinking in various projects in Australia and the United States. The System of the Observed Learning Outcome (SOLO) is a neo-Piagetian taxonomy developed by Biggs and Collis (1982, 1991). The taxonomy postulates that all learning occurs in one of five modes of functioning or thinking. These modes correspond with Piaget's stages of development, but the SOLO classification does not suggest that learners' responses in different situations are an indication of the level of their cognitive development or necessarily related to their age. The five modes of thinking are: sensimotor; ikonic; concrete symbolic; formal and postformal. There are five levels of response applicable to each mode of functioning. These levels measure the increasing sophistication with which learners deal with tasks. The five levels in order of sophistication are:

- **Prestructural:** The learner focuses on irrelevant aspects of the situation and does not engage in the task, resulting in a response that is below the target mode.
- **Unistructural:** The learner engages in the task and concentrates on the problem, but considers only one piece of data.
- **Multistructural:** The learner focuses on two or more pieces of data, but does not recognise any relationships between them, so that no integration takes place.
- **Relational:** The learner uses all available data perceiving all pieces as interrelated in a coherent structure.
- **Extended abstract:** the learner can reason beyond the data, generalising from new and abstract features.

The SOLO Taxonomy resulted from the analysis of learner responses in various fields and subjects, such as number and operations, history, geography, poetry amongst others. It was since adapted for use in many areas, including geometry and spatial development, statistical thinking, fractions, technology and more. Mooney (2002:29), pursuing the SOLO model, hypothesises that the statistical thinking of learners could exhibit the following five levels: *idiosyncratic* (associated with the prestructural level and representing thinking in the ikonic mode), *transitional, quantitative* and *analytical* (associated respectively with unistructural, multistructural and relational levels; representing thinking in the concrete symbolic mode) and *extended analytical* (associated with the extended abstract level; representing thinking in the formal mode). The ikonic and concrete-symbolic modes are most appropriate for learners in the Intermediate Phase. The SOLO model as tool for categorising different aspects of statistical thinking in learner responses will be discussed in more detail in Chapter 3 (3.7).

Statistical thinking is further explicated by Wild and Pfannkuch (1999: 230). The constructing and use of models to understand and predict the behaviour of aspects of the world that involve us are part of a general way of thinking. Modeling is included in the second dimension of the Wild and Pfannkuch framework for statistical thinking. We all need models to interpret and understand our everyday experiences, thus understanding is built up in mental models of the context reality and these mental models are informed by information from the context reality: “In an ideal world, we would be continually checking the adequacy of the mapping between model and reality by ‘interrogating’ the context reality” (Wild & Pfannkuch 1999:230). Statistical data is one kind of information we seek and get from context reality. The statistical models that we build help us to gain insights from and interpret this information that is then fed back into the mental model. The term ‘statistical models’ is used in a general sense, referring to all our statistical conceptions of the problem that influences the way in which we collect data and analyse data about the system. Statistical knowledge and experience plays a major role in the statistical conceptions that we form in order to collect and analyse data. Statistical elements can also be part of the way we perceive the world and can therefore become an integral part of our mental models of the context reality (Fig. 2.10). This however, depends on the problem, the education and the experience of the thinker.

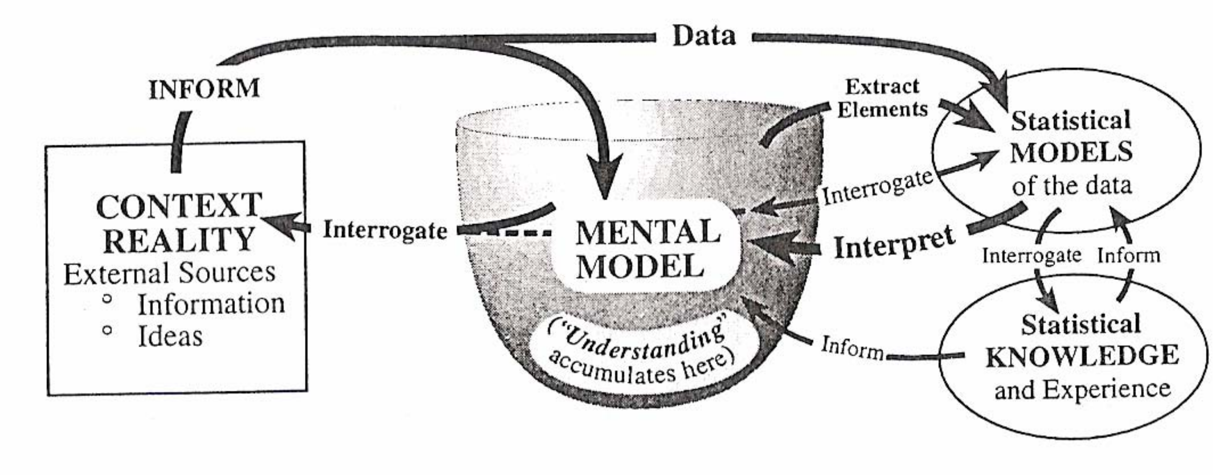


Figure 2.10: Learning via statistics

As mentioned in the beginning of this discussion, statistical thinking plays a crucial role in data modeling and problem solving in Statistics. Statistics is all about the modeling of data. Data modeling can be characterised as

> …a multicomponential process of posing questions; developing attributes of phenomena; measuring and structuring these phenomena; and then composing, refining and displaying models of their relations (Lehrer & Schauble 2000:52).

In this process, the different types of general and statistical thinking are employed to model real-world situations.

The construction and use of data are closely connected to mathematical models as the very idea of data embodies a separation between the world and a representation of that world. The constructing and use of data are referred to as data modeling (Lehrer & Romberg 1996:70; Horvath & Lehrer 1998:147).

A modeling approach to the teaching and learning of mathematics and statistics focuses on the mathematisation of realistic situations that are meaningful to the learner. In this approach meaningful contexts are explicitly used to elicit the creation of useful systems or models:

> …a modeling perspective leads to the design of an instructional sequence of activities that begins by engaging students with nonroutine problem situation that elicit the development of significant mathematical constructs and then extending, exploring and refining those constructs in other problem situations leading to a generalisable system (or model) that can be used in a range of contexts (Doerr & English 2001:362).

The sequence of data analysis problems provides a background against which the development of learners' interpretations of the problem situation, their reasoning about relevant elements of the system, their selection of quantities, operations, and representations, and their multiple cycles of interpretation can be investigated. The way learners represent a problem will determine how well they solve it. One of the most important factors in problem representation is the ability to understand the problem statement (Lajoie 1998:viii). Problem representation and specifically the representation of data tasks will be discussed in detail in Chapter 3.

2.6 CONCLUSION

Although some controversy surrounds the question of whether or not statistics is an independent science or a part of mathematics, data handling is for the purposes of this study regarded as an integral part of the school mathematics curriculum. In this chapter the construct of modeling in mathematics was discussed in general and with special reference to data modeling.

The nature of modeling and its role in presenting real-world situations through mathematics, developing worthwhile mathematics with which real-world events can be understood, controlled and predicted, were investigated. The six steps in the construction of a model and the iterative nature of modeling were explored, leading to a discussion of levels of emergent models starting in concrete experiences and progressing to models that can be used in formal mathematical reasoning. The role of social constructivism in model building was touched upon and with that, the difference in individual's models and understanding constructed in a group setting. A modeling perspective on learning and problem solving necessitates new ways of thinking about effective teaching. The learner as model builder rather than traditional problem solver implies that the teacher needs to develop a classroom atmosphere and collaborative classroom settings in which learners can engage in model eliciting activities.

The review of models or internal conceptual systems indicates representation, both internal and external. Conceptual systems or models almost always function with the support of powerful tools or representational systems. The three interacting systems involved in modeling, namely internal conceptual systems, external notation systems and external systems or artifacts, and their role in sense making of the world were perused. Representations and representation systems are of primary importance for teachers and researchers, because it is the only way that learners' understanding can be studied. The different forms of external representations were described. Representation is also the main focus of Chapter 3 and the study as a whole.

The relationship between models and problem solving were subsequently investigated. Different approaches to problem solving in mathematics education was scrutinised and its connectedness with real-life problems spelt out. Problem solving as a cognitive activity associated with the teaching, learning and understanding of mathematics was also discussed.

The analysis of both modeling and problem solving was set against the background of constructivism. The shift in emphasis from traditional teaching to teaching for understanding' was highlighted. The real life connection of problem solving and modeling in this teaching for understanding runs like a golden thread through all literature perused. This is also true for data modeling.

The nature of statistics and the statistical process of investigation received attention as background to the investigation into data modeling. Although Data

Handling or Statistics forms part of most school curricula, teacher content knowledge as well as pedagogical knowledge on the subject is not up to standard. In South Africa there is a lack of knowledge and research specifically about the statistical development and thinking of the learner in the Intermediate Phase. From the discussion in this chapter, the importance of representation in modeling, problem solving and therefore data modeling, crystallised. The focal point of the study described in this document is data arrangement and representation. As study focus and integral part of the process of statistical investigation, representation is investigated in Chapter 3.

CHAPTER 3
REPRESENTATION IN MATHEMATICAL AND STATISTICAL MODELING AND PROBLEM SOLVING

INTRODUCTION

Representation plays a crucial role in mathematics, constituting an essential component in teachers' understanding of how learners thinking and develop mathematically:

> The study of representation in mathematical learning allows us - at least potentially - to describe in some detail students' mathematical development in interaction with school environments and to create teaching methods capable of developing mathematical power (English, 2002:198).

In Chapter 2, problem solving and modeling were considered to shed light on the way in which learners model problems. In 2.3.4 and 2.4 the close relationship between modeling, problem solving and representation was considered. Modeling, being an iterative process, goes through multiple cycles in which representations play an all-important role. When modeling a problem, learners form internal representations to make sense of the problem, which are then expressed in external form and changed in the modeling process (Dossey et al 2002:114). The end product of each modeling cycle is a representation in which a real-world problem is expressed as an idealised version of a complex phenomenon (NCTM 2000:70, 71). A learner's problem solving strategies and models become accessible to others to the degree that internal representations of these models and strategies become accessible as external representations. Teachers can only gain insights into how learners think about and interpret mathematical ideas when studying the external representations of their problem solving strategies or models (NCTM 2000:68). These internal and external representations may be expressed as spoken language, written symbols, pictures and diagrams, concrete manipulatives or experience-based metaphors.

As this study endeavours to gain understanding of the spontaneous data representations of learners, representation stands central to the whole study and will therefore be investigated in this chapter. Different fundamental concepts in the representation theory will be considered as related in a broad sense to mathematics. Current views of representation as well as internal and external representations will be scrutinized. The role of representation systems and the relationship between modeling, problem solving and representation will be considered. As the tasks in the research instrument are set in the context of data handling, data representation will also be investigated, with special attention to the nature and types of data arrangement and representation. When studying learners' spontaneous representations, one of the questions that emerge concerns the level of the representations. As described in 1.1, the SOLO Taxonomy is one of the prominent tools used worldwide to simplify the analysis of learner responses and this neo-Piagetian taxonomy also forms the basis of a number of the statistical thinking frameworks that could be found in the literature. The SOLO Taxonomy will be adapted to form a framework for categorising the statistical thinking level of learner responses in the empirical study (see 4.7.3) and will therefore be reviewed in this chapter.

Different PERSPECTIVES INFLUENCING THE CONCEPT OF representation

The perspective on mathematical learning and problem solving from which it is studied influences the concept of representation. Goldin (1998:137-140) explores the influence of different theoretical perspectives and ideas on the study of mathematical learning and problem solving and therefore on the construct of representation and representational systems. One of the strands of thought that had a profound effect on views about mathematical learning ensued from empiricism and behaviourism. Behaviourists regarded environments and empirically observable behaviours as the essential components in their theory in which stimulus-response, operant conditioning and rule-governing learning also played an important role (Skinner 1953, 1974). In mathematics classrooms behaviourism resulted in an emphasis on procedural skills and observable performance, thus focusing on the external while deemphasising the internal. Neo-behaviourist notions in time included the permissibility of internal responses by learners (Skinner 1974; Gagné 1970). The behaviourist set of ideas had an impact on the consideration of strategies as patterns in behaviour during problem solving and resulted in the analysis of formal mathematical problem structures, the creation of sophisticated strategy scoring systems and the study of interactions of problem structure with strategy use as evidenced in the behaviour patterns of problem solvers (Branca & Kilpatrick 1972; Dienes & Jeeves 1965, 1970; Goldin & Gramick 1980; Lester & Garofalo 1982).

Another theme described by Goldin (1998:138) is the characterisation of problem solving heuristics and the effort to typify the structure and development of mathematical problem solving ability as one of the forces that shaped views on mathematical thinking and learning. Learner beliefs were held to be either powerful facilitators or obstacles to problem solving success (Goldin 1983; Krutetskii 1976; McKlintock 1984; Nesher & Kilpatrick 1991; Polya 1965; Schoenfeld 1985, 1987).

The developing field of study of cognitive science influenced the psychology of mathematics education through findings from the artificial intelligence research, heuristic programming and computer tools in which learning and problem solving are computer simulated. After an initial behaviourist inclination, the term cognitive science now refers to a broader set of ideas from developmental psychology, mathematics and linguistics. Self-regulated learning, including self-monitoring, reflection, and an awareness of one's own knowledge and beliefs about this knowledge became part of the cognitive science discussion (Davis 1984; Goldin 1984; Johnson-Laird 1983; Pylyshyn 1973; Silver 1985).

The genetic epistemology emphasised epistemological and structural analyses of children's mathematical behaviour. Stages in the mathematical development of children were identified and analysed in detail (Piaget 1965, 1967, 1969, 1970; Piaget & Inhelder 1971). The characterisation of cognitive structures and schemata became the focus of research and children's initial mathematical development was investigated (Carpenter & Moser 1984; Fuson 1986; Herscovics 1989; Herscovics & Bergeron 1983, 1984, 1988).

Representation and symbol systems in mathematics education became the focus with progress in the fields of psychology, formal linguistics, semantics and semiotics. Learners' interaction with computer environments and the study of mathematical structures also necessitated the study of representation and symbol systems. Critical factors in the understanding of mathematical concept forming were recognised as being visualisation, spatial and kinesthetic representation, image schemata and imagistic representation in general (di Sessa 1983; Goldin 1983, 1988; Goldin & Kaput 1996; Janvier 1987; Kaput 1987, 1991; Kosslyn 1980; Lesh 1981; Lesh, Landau, & Hamilton 1983; Presmeg 1986, 1992).

The influence of affect, attitudes, belief systems and emotional states in an adequate theoretical model of mathematical learning and problem solving have been substantiated by research which proposes that important information during problem solving is in fact encoded affectively (DeBellis & Goldin 1993, 1997; Goldin 1987; McLeod & Adams 1989).

Constructivist perspectives emerged as maybe the most prominent perspective opposing behaviourist ideas (Goldin 1996). Moderate constructivists regarded meaningful learning as the result of an internal process of constructing knowledge (Confrey 2000; Ernest 1991; von Glasersfeld 1990,1996). Radical constructivists assigned tremendous emphasis to constructive processes, adopting a subjective, relativistic approach. Social constructivists viewed mathematical truth as negotiated knowledge in a social environment and focused on the cultural and sociological processes through which knowledge is constructed. In the mathematics classroom constructivism led to an emphasis in non-routine problem solving and group activities, while mathematical exploration and discovery, open-ended questions, alternate solution methods, contextualised understandings and the use of technology received attention. Constructivists emphasised the internal, in contrast to the behaviourists' emphasis on the external.

Along with the above discussed broad constructs and themes influencing views of mathematical teaching and learning, research in specific content domains of mathematics, such as additive and multiplicative structures, "story problem" tasks, rational number learning, algebraic reasoning and so forth, resulted in descriptions of specific knowledge structures and their development. Cognitive barriers and misconceptions occurring among learners concerning specific mathematical concepts have also been identified.

On the South African scene the different broad views on mathematics learning and problem solving discussed above played an important role in the classroom. Behavioural ideas led to what is known as traditional teaching methods with an emphasis on external products to the detriment of the process in which this product was created – answers to problems were all important and not the process by which it was derived. The Outcomes-based curriculum has constructivism as one of its pillars and since the implementation of the new and the revised curriculum teachers have become sensitive to the processes of knowledge construction and the value of group work. Problem solving is an inherent component of an outcomes-based curriculum and teaching in general with and through problem solving (a problem-centered approach) is gaining increased recognition in South Africa (Wessels & Kwari 2003). Findings from cognitive science research are influencing education in South Africa as is the case all over the world. In this regard, self-monitoring, self-regulated learning, reflection and the effective use of technology closely connect with the critical and developmental outcomes of Outcomes-based education in South Africa (DoE 2002:1, 2). The influence of affect, attitudes, belief systems and emotional states on learning and problem solving are acknowledged in the curriculum and the important role of representation are spelled out in discussions of the five learning outcomes in all phases of the curriculum (DoE 2002:1-5, 7-13, 33-39, 61-67; DoE 2003:2-5, 12-14).

With the discussion of the different broad perspectives and research directions, the need for a unifying model of mathematical teaching, learning and problem solving has become apparent. Cocking (1999:xii) points to the fact that representation is becoming an important unifying concept in the behavioural and neural sciences. Goldin (2002:139) suggests that the notion of representational systems and their construction can unify all the constructs above to form a

> …theoretical foundation for mathematics education, one that can accommodate the most helpful and applicable constructs from a variety of approaches, including those discussed above… For this I think that a framework based on the study of representations and representational systems is of great assistance.

These notions of a unified model for representation have become evident through a new generation of research studies throughout the diverse approaches to representation.

CONCEPTS IN A THEORY OF REPRESENTATION

In this section some of the fundamental notions related to representation in mathematics education will be investigated.

Representational systems

The term *representation* has a variety of everyday and technical meanings. Several authors discuss the term representation with regard to mathematics learning and understanding (Byrnes 1999; Goldin 1998, 2002; Goldin & Kaput 1996; Goldin & Shteingold 2001; Janvier 1987; Lesh & Doerr 2000; Martinez 1999; Sigel 1999; Vergnaud 1987; Von Glasersfeld 1987b). Different concepts in the representation theory will now be explored with reference to a number of these different definitions.

A representation can in a general sense be described as a configuration that can represent something else in some manner (Goldin 2002:208). For example, a real-life object can be represented by a word and the same numeral can in one instance represent the cardinality of a set or in another instance a position on a number line. The relation between two configurations representing the same entity must eventually be made explicit and should rather be seen as bidirectional in stead of distinct in a fixed or final way. That implies that when one configuration represents another, the latter can often be considered as useful in representing the former. For example, in mathematics the solution set of an algebraic equation can be represented as a Cartesian graph or the graph can be represented by an equation satisfying the coordinates of the points on the graph. Individual conventions, which over time became shared conventions, became normative amongst people involved in mathematics, resulting in coherent interaction between participants. Individual representations can seldom be understood in isolation, but as configurations belonging to a wider system, they can be used to communicate understanding.

Studies previously regarded representation from a single systems view, but have currently shifted to the idea of representation as multiple systems that have overlapping functions that at the same time possess unique individual functions and properties (Cocking 1999:viii). Cocking contends that the convergence of information from many sources led to the change in the belief that

> …representation is not unitary, that it has structure, it has domain specificity, is different from metacognition, that individuals actively participate in the formation of representations, that there is a role played by privileged classes of information, that culture plays an important role, and that there are systems of representation (Cocking 1999:xii).

In the technology-based society that we live in, one of the requirements for success is the ability to use a multiplicity of systems in representation (Lesh & Doerr 2000:382). The multiple systems view of representation will form the basis of discussions on representation in the rest of this chapter.

Attention will now be given to primitive components, configurations and structures in representation systems.

3.3.1.1 Characters, configurations and structures

Goldin (1998:143; 2002:208,209) contends that a representational system consists of primitive characters or signs that can be discrete identities from a well-defined set, for example Roman numerals or bases in a DNA molecule. Conversely, these characters can also be less well-defined or partly defined entities, such as physical objects and their attributes or words in the English language.

Representational systems possess more complex structures, such as networks, configurations of configurations, partial or total orderings on the class of configurations, mathematical operations, and so forth. Additionally, representational systems involve rules for moving from one configuration to another, or for combining the signs into permitted configurations, for example, sentences are permitted by rules of grammar and syntax, and single-digit numerals may be combined to form multi-digit numerals according to rules of place value. Therefore, the meaning of the signs of a representational system can exist only within the structures of the system.

3.3.1.2 Symbolic relationships

Characters, configurations or structures in one system can represent, evoke, stand for, encode, produce or symbolise those in another, subject to certain rules, which is the main reason for these systems to be called *representational* systems (Goldin 1998:144). A representational system can therefore be said to have both intrinsic relations (between the sign or configurations of the systems itself) and extrinsic symbolic relations (that is, with other systems of representation). For example, signs in numeration systems such as numerals and arithmetic symbols, have syntactic links with each other, but can also stand for something else such as action sequences related to the counting of a set of objects.

External representation systems for mathematics are structured by shared conventions and assumptions, and when using them, one has to conform to conventional norms. Since rules for the order of operations have been established specifying that multiplication has to be performed before subtraction, an expression such as 15 – 2 x 3 is evaluated by these rules. Once these norms or rules are established, the patterns in it are no longer optional. Goldin stresses the difference between that which is conventional and that which is not:

> Having assumed the conventional properties of natural numbers, our base ten notational system, the conventional definitions of addition and multiplication, and the conventional definition of a prime number, it is *true* that 23 is a prime number and 35 is not. We invoke here no metaphysical or Platonic notions of absolute truth. Rather we highlight the important and elementary mathematical distinction between that which is conventional and that which is (objectively) no longer so, once the context of mathematical assumptions is established (2002:210).

Since representational systems are conventional constructs, intrinsic and extrinsic symbolic relations in them sometimes are ambiguous. This ambiguity is the next point of discussion.

3.3.1.3 Ambiguity in representation

As representational systems are conventional constructs, it may not always be easy to determine the boundaries between different representational systems. Convenience and simplicity of description play an important role in the distinction between where one representational system begins and where another leaves off, or whether to view additional structures as intrinsic to a given representational system or as arising from the symbolic relationship between two systems. This typically leads to ambiguities in the description of representational systems. Exceptions to almost all syntactic and semantic rules exist, complicating the structure of a descriptive model of learning or problem solving. Goldin (1998:145) regards ambiguity as a necessary feature in the concept of a representational system. He points out that the initial family of signs may be well-defined as in the case of Roman or Arabic numerals; close to well-defined as in the case of words in a specific language or highly ambiguous as in the case of real-life objects. In addition, symbolic relationships between two representational systems may also be very precise or highly ambiguous. Examples are the precision in representation of abstract groups by matrices acting on vector spaces in contrast with ambiguity in possible imagery linked to words or the metaphorical interpretation of mathematical ideas.

Resolving ambiguities in representation involves the context in which the ambiguous sign, configuration or symbolic relationship occurs. For example, homonyms such as “pale” or “pail” in spoken language requires semantic interpretation of the words to settle the ambiguity. “Context” in this case points to that which is not part of the representational system of words in the example. The ambiguity within the system of words requires the considering of the words outside this system in another unambiguous system. Goldin explains this resolving of ambiguity as follows:

> To resolve ambiguity in the symbolic relationship between two representational systems, contextual information can sometimes be incorporated by going to a *third* system, to which each of the first two bears a symbolic relationship, or where the symbolic relationship between the first two is itself represented (1998:145).

Ambiguity thus may be regarded as a necessary feature in the characterisation of representational systems or their relations to other systems.

Internal and external representation

When scrutinising concepts in a theory of representation, another significant distinction that has to be considered is that between internal and external processes and products in representation. Internal and external representations and the interactions between them will therefore be discussed next.

Scholnick (1999:113) describes representation as “the way the mind encodes the world”, alluding to the fact that different representations are generated by theories of the mind while conversely, diverse views of cognition and its development are shaped by different views of representation. In an investigation of the term representation, von Glasersfeld (1987b:216-219) distinguished between four German words for representation: darstellen, vorstellen, vertreten and bedeuten. The words Darstellung and Vorstellung pertain to our discussion. The word Darstellung closely corresponds with the term external representation while Vorstellung corresponds with internal representation.

Before reflecting on definitions of internal and external representation in a mathematical context, it is important to consider the distinction between representation as a process and representation as a product. Representation as process refers to the act of capturing a mathematical concept or relationship in some form and representation as product refers to the form itself, a physical object or external representation, or a strictly cognitive entity or internal representation (Denis 1994:1; NCTM 2000:67). The term *representation* therefore alludes to externally observable processes and products as well as to the internal products and processes in the minds of people doing mathematics. Dossey et al (2002:83) contend that the process of representing is just as important as the product or object and define a representation as an object that describes or models a situation where the process of representation is “the act of capturing a mathematical concept or relationship in some form that conveys an idea, a picture or a mathematical connection to the viewer”.

Internal representation and representational systems

Internal representation will first be scrutinised. When investigating internal representation, it is necessary to study different authors’ interpretation of the term. Because of Piaget’s crucial role in the study of child development, his views on representation need to be considered. Piaget (1951:67) claims that representation is a result of individual activity that takes on qualitatively different forms throughout ontogenesis and that it manifests in two forms:

> In its broad sense representation is identical with thought, i.e. with all intelligence which is based on a system of concepts or mental schemas and not merely on perceptions and actions. In its narrow sense, representation is restricted to the mental or memory image, i.e. the symbolic evocation of absent realities.

Piaget's opinion countered the associationist views of the image as a copy of perception with the notion that the mind's representations are not facsimiles of experience but that they rather reflect what we know than what we see (or hear or feel) (Martinez 1999:21, 22).

Another interpretation is of representation as a mental model and focuses on the relationship between mental models and understanding. Mental models consist according to Johnson-Laird (1983:156) of knowledge that "plays a direct representational role since it is analogous to the structure of the corresponding state of affairs in the world – as we perceive or conceive it". He views mental models not as a single image, but as a combination of images that forms a dynamic image system that can be used to make predictions and asserts that a direct connection between mental models and understanding exist: "The psychological core of understanding, I shall assume, consists in having a 'working model' of the phenomenon in your mind" (Johnson-Laird 1983:2). The relationship between representation, understanding and knowledge play an important role in the definitions in the field of cognition. Martinez (1999:18) states that "mental models are a subset of all possible representations" and contends that the human mind does not process and store countless sensory bits like a video camera, but rather constructs an inner and outer world according to the organising principle of meaning:

> The fact that knowledge can be represented in different ways implies that knowledge is not a sensory transcription of the external world into the inner world of the mind (1999:21).

The fundamental nature of cognitive or internal representations in our understanding of how the mind works is also of interest. Martinez (1999:13) describes representations as "…the means by which we think and behave intelligently" and highlights different features of internal representations:

- They are knowledge structures that symbolise some state of affairs.
- Each is likely to be one of a set of alternative structures.
- They may differ in their ability to facilitate solving particular problems.

The first feature points to the relationship between knowledge and representation while the second feature alludes to translations between and transformation within representations, which will later be discussed in the section on representational fluency (3.3.5). The role of representations in problem solving will be touched upon in 3.4.

To elucidate the concept of internal representation, the contributions of two cognoscenti in the field, namely Byrnes and Goldin in conjunction with a number of other researchers, will now be examined. Byrnes (1999:274), in the vein of representation and knowledge define representation with regard to cortical activity as

> … a pattern of recurrent cortical activity that can be evoked or elicited by another pattern of cortical or subcortical activity. These patterns of activation, in turn, correspond to entities in the "real" world or are themselves components of an imagined world.

Byrnes (1999:274-290) examines the nature and development of representation by discussing eight important distinctions that resulted from different views through the last few decades:

- Knowledge versus Thinking

  The terms knowledge and representation are fundamentally coextensive. When a person knows a fact or skill it implies that the person has created a representation of the fact or skill. Verification that a person has specific knowledge lies in the fact that the individual can evoke the relevant representation when prompted in some way (e.g. asked a question). However, when a person has knowledge, it does not necessarily mean that the knowledge is *used* to recognise something, make inferences, or solve problems. “Knowledge is the grist for the thinking mill, but it is not the same as thinking” (Byrnes 1999:275). Very elegant thinking sometimes is based on flawed knowledge. Much of Piagetian research was aimed at discovering how children think and how they often generate incorrect answers because of insufficiently developed knowledge (Inhelder & Piaget 1964). Piagetians and post-Piagetians differ on knowledge and thinking abilities of children. The former hold that the thinking abilities of preschoolers are not well developed while post-Piagetians argue that preschoolers can think well about many or most things. Byrnes suggest that both these research traditions only are half right (1999:276) and that preschoolers can think very well when given a handful of topics that they have mastered. More difficult ideas, such as those presented in the Foundation Phase at school may cause them to perform more poorly. ”The problem is not with their thinking per se, as much as inadequacies in their knowledge” (Byrnes 1999:276). When only one half of the argument is taken into consideration, knowledge growth that does occur is often overlooked.

- Declarative versus Procedural

  Declarative knowledge is the collection of all the facts one knows. Procedural knowledge on the other hand is the collection of all of the strategies, skills and algorithms one knows. Byrnes adds in his argument a third kind of knowledge, namely conceptual knowledge which entails understanding of the meaning of facts and the outcomes of procedures (Byrnes 1992:236, 1999:227). Conceptual knowledge is relational and helps the individual to link facts or knowledge and procedures. Byrnes states three reasons why the distinction between declarative, procedural and conceptual knowledge should be maintained: firstly, careful lines of philosophical argumentation show that declarative knowledge can not be reduced to procedural knowledge; secondly, findings of cognitive neuropsychologists have confirmed that separate declarative and procedural representation systems exist and thirdly, that educational studies in which both conceptual and procedural knowledge have been assessed, have revealed that procedural knowledge are underpinned by well developed conceptual knowledge. Procedures are learnt better by children with better conceptual understanding than children with worse conceptual understanding. Furthermore, research on brain-damaged individuals points to the fact that separate declarative and procedural representational systems exist.

- Implicit versus Explicit

  It is possible to have both implicit and explicit representations of conceptual and procedural knowledge. An example of implicit conceptual knowledge can be found in the context of categorisation. Although a learner may be able to implicitly abstract the physical properties needed to distinguish between one object and another (e.g. a dog versus a horse), he or she may be unaware of the properties used to categorise the two objects. This is an illustration that an individual may not always be aware of his or her knowledge or be able to articulate what he or she knows. Implicit knowledge associated with categorisation and other conceptual initiatives can over time become the object of reflection and therefore can become explicit knowledge. Byrnes maintains that a corresponding pattern of implicit use with ensuing explicit awareness has been found in diverse domains such as mathematics and languages (1999:279). Vergnaud (1998:175) is of the opinion that the status of knowledge is very different when it is made explicit, rather than being totally disguised by behaviour and states that explicit knowledge can be communicated, but implicit knowledge cannot.

- Concrete versus Abstract

  Piaget, Vygotsky and schema theorists agree that knowledge can exist at different levels of abstraction and that the brain over time only retains some of the information gained through experience:

  > The brain somehow abstracts what is common to related experiences and retains only this abstracted information; details and specific sensory-based representations are not retained (Byrnes 1999:280).

  Byrnes also points to the fact that Piaget, Vygotsky and schema theorists support the same definitions of concrete and abstract knowledge, where concrete knowledge is regarded as representations linked to immediate action or perception and abstract knowledge as representations that exhibit commonalities across concrete representations. They may however differ about whether this knowledge is a single type of abstract knowledge or multiple levels of abstract knowledge. Byrnes (1999:280) is of opinion that knowledge does exist at different levels of abstraction. He gives an example of three levels of abstraction in the Piagetian sense, indicating the difference between a mental image of a set of objects such as three apples, mathematical symbols corresponding to numbers (for example “3”), and mathematical symbols corresponding to variables (for example “X” or “Y”). Because of the fact that knowledge can exist at different levels of abstraction, internal and external representations of knowledge also occur at different levels of abstraction.

- Domain-specific versus Domain-general

  Domain-specific knowledge pertains to concepts, processes, or procedures that can be applied to content and facts from a specific domain whereas with domain-general knowledge it can be applied to content and facts from multiple domains. The concept of domain-general is applicable to concepts, processes, or procedures that can be applied to content, including facts from multiple domains. Byrnes (1999:282) holds that

> Any given concept, process or procedure may be (a) completely domain-general (i.e., it can be applied to any content), (b) only partially domain-general (i.e., it can be applied to several but not all types of content), or (c) domain-specific (e.g., parsing for language).

In summary, representations of declarative knowledge can only be domain-specific because declarative knowledge is inherently domain-specific with facts generally contributing to the core of a single domain. Representations of cognitive processes, conceptual knowledge and procedural knowledge on the other hand can both be domain-specific and domain-general.

- Ad Hoc (short term) versus Permanent

  Ad hoc knowledge results when a new fact is created or derived at on line as response to a question. The facts derived in this way are working memory representations that have no equal in long-term or permanent memory. Permanent knowledge on the contrary is permanent records existing in long-term memory and manifests in a pattern of synaptic connections, having been constructed over time. When children have to categorise objects, they are able to note similarities in objects, which is not equivalent to the actual construction or organisation of knowledge in permanent memory. Noting similarities involves having knowledge in memory, but pursuing links between items is not identical to having these links represented in permanent memory (Byrnes 1999:285).

- Can versus Do

  Byrnes (1999:286,287) describes the distinction between *can* versus *do* in relation to the Piagetian notion of graduated levels of performance and the Vygotskian notions of proximal development and scaffolding. He points out that a child usually is able to do more when prompted than without help or prodding. The fact that learners *can* demonstrate a competence at a specific time does not mean that they *do* demonstrate this competence regularly at home. The implication is amongst others that if a learner is able to represent some knowledge in a certain way at a specific time, that he or she will do it in a similar way at another point in time. The opposite is also true: when a learner does not create a representation at some point in time does not necessarily mean that he or she is not able to do it.

- Innate versus Constructed versus Learned
  To say that knowledge that is innate as advocated by the nativists, means according to Byrnes's definition of knowledge discussed earlier, that children are born into the world with a representation consisting of pattern of synaptic connections already configured, which instantly enables them to recognise an object or class of objects. Byrnes (1999:288) is of the opinion that this argument does not hold true and explains that a baby would only form a pattern of synaptic connections that corresponds with an object after several exposures to the object. The notion of constructed knowledge is in keeping with the constructivist view and implies that a child may introduce a conceptual link that does not seem to have emerged because of direct instruction or maturation. Learning is the forming of a pattern of synaptic connections that is related to specific environmental or internal stimulation. This forming of an un-elaborated, sensation-based representation for an object is consistent with the empiricist view. More research about the links between neurology and knowledge is needed before it will be possible to say in which of these three ways knowledge was conceived of in a specific instance.

The discussion of the eight distinctions clearly shows the complexity of representation in the human mind. These distinctions are however not mutually exclusive and any given representation may be a particular combination of the elements of the eight distinctions, for example, a single representation may exemplify implicit, domain-general, procedural knowledge (Byrnes 1999:290).

Byrnes (1999:290, 291) concludes that representational change lies at the heart of cognitive development and that from the perspective of the eight distinctions seemingly contradictive views and research findings may make perfect sense and cohere together nicely.

Goldin and Kaput interpret an internal representation as “a construct arrived at by an observer from the observation of behaviour (including, of course, verbal and mathematical behaviour)” (1996:400). Goldin describes the nature of representation by claiming that internal psychological representational systems involve an individual’s natural language, personal symbolisation constructs, visual and spatial imagery, problem solving heuristics, and affect (Goldin 2002:210). The definition of Goldin and Shteingold (2001:2) is similar to this explication, but adds a new dimension by specifically referring to mathematics:

> Internal systems … include students’ personal symbolization constructs and assignments of meaning to mathematical notations, as well as their natural language, their visual imagery and spatial representation, their problem solving strategies and heuristics, and (very important) their affect in relation to mathematics.

Five psychological fundamental types of internal representation, which are typical of mature cognitive internal systems can be distinguished (Goldin & Kaput 1996:417-420; Goldin & Shteingold 2001:5):

- A verbal syntactic system: describes an individual’s natural language processing capabilities, including mathematical and non-mathematical vocabulary, grammar and syntax. It is a dynamic representational system and is culturally provided and yet universal in occurrence as internal system.
- Several different imagistic cognitive representational systems: Imagistic abilities are essential for meaningful interpretation of verbal communication and for describing mathematics learning and problem solving the most important are visual/spatial, auditory/rhythmic and tactile/kinesthetic. Tactile/kinesthetic encoding is associated with actual or imagined hand gestures and body movements, while auditory /rhythmic

are vital since learners learn counting sequences and letters, clap in rhythm, and so forth.

- Formal mathematical notations, which are usually constructed from culturally provided, conventional systems and may be static or dynamic. These representation systems may be static or dynamic and may be imagistic in nature, for example the internal construct of a Cartesian graph. Formal notational representation occurs internally when learners mentally manipulate numerals, visualise geometrical figures or symbolic steps in solving an algebraic equation.
- The planning, monitoring and control of mathematical problem solving processes organised into heuristic processes: the strategic and heuristic processes are represented as the mental development and organisation of methods such as “trial and error” or “working backward” takes place.
- A system of affective representation, which is neither formal nor imagistic. This system is necessary for effective learning and problem solving refers to the rapidly changing feelings a learner experience during problem solving, their beliefs and values about mathematics, or about themselves in relation to mathematics.

These five systems do not function mutually exclusive. From the discussion above, it is clear that connections do not only exist between all five of these internal representation systems, but also between these systems and mathematical learning and problem solving.

The next important feature of internal representations that needs consideration is the fact that they are by nature not directly observable. One of the main differences between internal representations (“mental configurations”) and external representations (“physically embodied configurations”) lies in their accessibility (Goldin 1998:145, 2002:210; Goldin & Kaput 1996:399).

> Internal configurations are those characteristics of the reasoning individual that are encoded in the human brain and nervous system and are to inferred from observation (Goldin & Kaput 1996:402).

The implication is that teachers and researchers have to infer learners' internal representations, mathematical conceptions and misconceptions from their external behaviour, which can include their actions or words, and their interaction with or production of external representations. These inferences are often made more tacitly than explicitly (Goldin & Kaput 1996:399). External configurations in contrast are accessible to direct observation and involve among others written words, speech, concrete manipulatives, formulas and computer microworlds as they appear on a computer screen.

External representation and representational systems

After the discussion of a number of different definitions of internal representation it becomes necessary to distinguish external representations from internal psychological representations of an individual. In this discussion the views of three researchers in particular will be investigated, namely Goldin, Lesh and Martinez. The former two authors have both researched and published in this field individually as well as in association with other researchers.

Goldin and Shteingold (2001:3), referring mainly to externally observable representations, contend that a representation typically is a sign or a configuration of signs, characters or objects, and point to the fact that it can stand for (symbolise, depict, encode, or represent) something other than itself. The represented 'thing' can vary according to the context or use of the representation, for example a Cartesian graph can represent a function, the solution set of an algebraic equation or a data set. They contrast their definition of internal representation systems referred to in the previous paragraph with external representation systems by stating:

> External representations range from the conventional symbol systems of mathematics (such as Cartesian coordinate representation, the real number line and formal algebraic notation) to structured learning environments (for example, those involving concrete manipulative materials or computer-based microworlds) (Goldin & Shteingold 2001:2).

From the examples given, it is clear that Goldin (2002:208) is referring to external representation when defining a representation as "… a configuration that can *represent* something else in some manner" and elucidates his definition with the following examples: a real life object can be represented by a word; a numeral can represent the cardinality of a set or the same number can represent a position on a number line. A number of authors regard external representations as externalisations of internal systems of thought (Goldin 2002:211, 1998:147; Lesh 1999:331; Lesh, Post & Behr 1987:33). External representations cannot however just be regarded as externalisations of internal representations and vice versa. Internal representations can also act as mediators in the process where learners translate from one representational system to another or construct entirely new representations in external task environments (Goldin 1998:146). The position of Lesh (1999:331) is that external representations serve as much as externalisations of internal systems of thought than as simplifications of external systems when learners mathematise problem solving situations.

The interaction between internal and external systems of representation (Fig.3.1) plays an important role in effective teaching and learning (Goldin & Kaput 1996:399; Goldin & Shteingold 2001:2; Lesh & Doerr 2000:364).

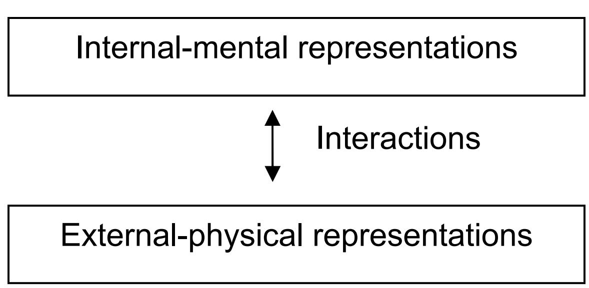


Figure 3.1: Internal versus external representations
(Goldin & Kaput 1996:399)

Representation plays an all-important role in mathematics. Goldin and Janvier (1998:1,2) were involved in the Working Group on Representations of the International Group for the Psychology of Mathematics Education (PME) from1989 until 1998 and summarise the various interpretations of the terms *representation* and *representation systems* in connection with mathematics learning, teaching, and development through those years as comprising the following:

- An external, structured physical situation, or structured set of situations in the physical environment, that can be described mathematically or seen as embodying mathematical ideas;
- A system of language or linguistic manifestation, in problem posing or mathematical discourse, with emphasis on syntactic and semantic structural characteristics.
- The representation of situations through symbols or through a system of symbols, of a formal mathematics construct or system of constructs. The construct or system of constructs usually adheres to precise definitions or axioms, including mathematical constructs that may represent aspects of other mathematical constructs.
- An internal, individual cognitive configuration, or a complex system of such configurations, inferred from behaviour or introspection, describing some aspects of the processes of mathematical thinking and problem solving.

The essential role of internal representation in the thinking of an individual has been discussed earlier. External representation plays just as important role in reasoning as is explained by Cox and Brna (1995:82) with reference to the properties of a chosen external representation and its effectiveness in problem solving:

> External representation emerges as a crucial phase of reasoning – selecting an appropriate ER is often very difficult because the requirements of tasks vary considerably between and within problems. The expressive properties of the chosen ER must be capable of representing the semantics of the problem (Cox & Brna 1995:282) (The authors use the abbreviation ER for external representation - HMW).

Earlier in this section, a similar view of Martinez (1999:13) was investigated, in which he pointed out that not all representations are suitable or conducive to the solving of a problem. Learners who are successful in problem solving seem to be able to construct appropriate problem representations and use these representations as support for understanding the information and relationships of the situation (Cifarelli 1998:239).

Lesh, Post and Behr (1987:33) describe different roles of representations and translations between representations in mathematical learning and problem solving. They interpret the term representations as “external embodiments of students internal conceptualizations”, admitting that this view of representations is artificial, naive and restricted. The authors (1987:33) identify 5 distinct types of external representation systems occurring in mathematics learning and problem solving, namely

- Experience-based scripts in which “real world” events act as general contexts for interpreting and solving other types of problem situations.
- Manipulative models, for example Dienes blocks, fraction bars and number lines in which the elements of the system do not have much meaning per se, but the “built in” relationships and operations fit many everyday settings.
- Pictures or diagrams that are static figural models, which can be internalised as “images”.
- Spoken languages, including specialised sublanguages related to domains like logic and so forth.
- Written symbols, which like the spoken languages mentioned, can entail specialised sentences and phrases.

These representation systems are important in mathematics, but translations among them and transformations within them are also significant. External representations such as words, symbols and graphs that learners use are partly descriptions or simplifications of external systems. These representations however are also externalisations of internal systems because they focus on hypothesised relationships, patterns and regularities that are assigned to external systems rather than being derived from them.

The role of context and content in representation

Representations can not be understood in isolation (Goldin 2002:208, 1998:143; Goldin & Kaput 1996:398; Goldin & Shteingold 2001:1; Von Glasersfeld 1987b:216; Cocking 1999:viii,ix). A mathematical representation, as its general counterpart, is content and context sensitive, constructed in the present to meet present demands. It depends on content and context for meaning; for example, a graph is meaningless unless understood in the system to which it belongs. External representations do not stand alone, but depend on content and context for meaning, an individual numeral or a graph for example is almost insignificant or meaningless apart from the system to which it belongs (Goldin & Shteingold 2001:3).

Scholnick (1999:113) regards any theory of representation as connected to definite parameters:

> A theory of representation must account for the modality in which the event is encoded, its format, content, and connections with other representations, the mechanisms producing the representations, the consequences of representations for the cognitive system, and the psychological/physical nature of the representation.

Von Glasersfeld (1987b:216) describes the involved nature of a representation when saying that when trying to define a representation, it must be kept in mind that “ a representation does not represent by itself – it needs interpreting and, to be interpreted, it needs an interpreter”.

Contextualised understanding of mathematics from a representational view refers to the internal encoding of familiar contexts as representational configurations in common words, images, formal notations, strategies and operations and comfortable affect. Such internal structures that are familiar and have a common-sense nature are likely to be widely shared, based on everyday experiences that can easily be referred to, coded in redundant ways, developmentally prior to the mathematics being learned in the given context and culturally supported (Goldin 2002:214).

Decontextualised representation is according to Goldin often found in traditional teaching practices. He defines decontextualised representation as "… formal mathematical notations and rules of procedure introduced as syntax without semantics, or rules and methods without context" (2002:215). Decontextualised representation intends to avoid contextual restraints, to exemplify that which is abstract in mathematics, but is likely to result in the construction of a formal internal system without semantic connections. Goldin warns against regarding decontextualised representation as abstraction, pointing out that decontextualised representation may be limiting, but that insisting that all mathematics be contextualised may also be limiting as some contexts may pose natural obstacles to later abstraction (Goldin 2002:216). As structure is built, progressive detachment of representations from their original contexts should take place to prevent initial contexts to become cognitive obstacles. Goldin advocates the process of contextualisation in which new semiotic acts allow the same familiar representational configurations to acquire new meanings in new semantic domains and identifies abstraction and contextualisation as complimentary representational processes that develops depth of understanding in mathematics (2002:216).

Representational fluency

Representation is not a static, but a dynamic process. Vergnaud (1998:167, 175) contends that knowledge is "action and adaptation". In this process, active representations are “under construction” as with addition of new knowledge. The underlying mathematical constructs that the representations embody, as well as the external systems that they describe, are likewise not static, but continually evolving (Lesh & Doerr 2000:368). This dynamic evolving nature of representations and systems of representation is a critical component of representational fluency. Representational fluency involves the ability to represent a problem in more than one way and to be able to translate fluently between different representations. In our technology-based society, representational fluency is becoming increasingly important. Learners need to develop representational fluency to mathematise systems that entail more than simple counts and measures. They need to be able to deal with mathematical entities such as signed quantities (for example, positive or negative); directed quantities (for example, simple or intuitive uses of vectors); ratios of quantities; rates (for example, per quantities or intensive quantities); coordinates; accumulating quantities; continuously changing quantities; derived measures (for example, based on a formula); learners invented constructs; measures connected to frequencies of events (probabilities); measures linked to sets of data (statistics); patterns (trends, sequences, series); and so forth Lesh (1999:347). Representational fluency needs to be extended to include computer and calculator generated representations, such as animations, graphs, tables and notation systems.

Genuine understanding will most likely emerge and be accessible to others when learners are able to represent knowledge of a concept or skill in a number of different ways and can translate back and forth among these different representations:

> An important symptom of an emerging understanding is the capacity to represent a problem in a number of different ways and to approach its solution from varied vantage points, a single rigid representation is unlikely to suffice (Gardner 1991:18).

Martinez (1999:28) supports this view when pointing out that understanding is much more likely to occur when the same phenomenon can be represented by way of multiple integrated representations and continues to say that when a concept or idea is represented with a single representation, understanding is unlikely to be communicated.

Lesh (1999:331) accentuates the importance of representational fluency in the analyses of problems and planning of solutions that entail multiple steps, resources and constraints; the justification and explanation of proposed actions and the prediction of their consequences; the monitoring and assessment of progress and the integration and communication of results in useful ways. Multiple representations are not only useful for an individual in communication with others, but also in communication with himself:

> …the purpose of representations is not simply for students to communicate with one another, it is also for students to communicate with themselves and to externalize their own ways of thinking so that they can be examined and improved (Lesh 1999:331).

Not only does the use of multiple representations engender better understanding in an individual, it also results in better communication of his thinking and understanding to others. From the above discussion the need for mathematics teachers to foster the development and use of multiple representations is evident.

THE relationship between MODELING, problem solving AND representation IN mathematicS

Problem solving and modeling are both linked to internal and external representations (Dossey et al 2002; Lesh et al 2002; Lesh & Doerr 2000; Cifarelli 1998). Lesh and Doerr (2000:363) describe the meaning of a model with reference to representation systems:

> The meaning of a model, or conceptual system, tends to be distributed across a variety of interacting representation systems that may involve written symbols, spoken languages, pictures or diagrams, concrete manipulatives, or experience-based metaphors.

The term *models* points to the dynamic and interacting characteristics of the systems that are modeled, whereas the term *representations* refers to the objects within these systems. Models pertain to functioning whole systems, while representations pertain to inert collections of objects that must be manipulated and related in order to function. Modeling and problem solving entails interactions between three types of systems:

- (internal) conceptual systems existing in learners' minds;
- (external) models or representational systems that function both as externalisations of internal conceptual systems and as internalisations of external systems and
- (external) systems given in nature or that were designed by humans (Lesh 1999:335,336; Lesh & Doerr 2000:363, 364).

Conceptual systems are in existence only in learners' minds. Systems that function as externalisations of internal systems and vice versa, seem to be rooted in spoken language, diagrams, written symbols, pictures and concrete models. Some examples of the external functioning systems created by humans described above are economic systems, mechanistic systems and communications systems. The boundaries of the three systems are not distinctly defined, but tend to be fluid, shifting and may be ambiguous. These systems, though in some ways distinct, partly overlap and are interdependent and interacting (Fig. 3.2). Lesh and Doerr (2000:382) emphasise that the most useful representation systems are those that are functionally and dynamically related.

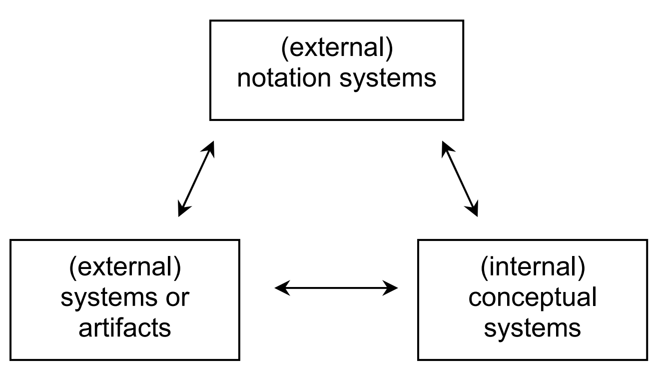


Figure 3.2: Modeling interactions among three types of systems

Although learners' cognitive representation systems can be shared with others by using external representations, part of the meaning of these internal systems are not accessible to others. Teachers have to create cognitive conflict by introducing model-eliciting tasks in order to create a need for learners to share their conceptual tools and representations. Cognitive conflict, which is the need to develop increased conceptual stability, is a fundamental determinant in the process of conceptual adaptation. In this process representation systems promote the progressive differentiation and integration of relevant conceptual systems (Lesh & Doerr 2000:379).

In model-eliciting tasks learners have to think beyond just one representation or model to also consider alternative representations with their strengths and weaknesses. Model-eliciting activities challenge learners to introduce, modify and adapt useful representations, thus reflecting on their own thinking and communicating their thinking to others. The fostering of discussions, allowing and nurturing of alternative approaches and building of powerful connection with non-mathematical experience out of the school environment are crucial in the model building process and the meaningful use of appropriate representations. Additionally, final products or representations of the modeling process show little of the process of development. Teachers need to study learners' representations used in the development of their models to be able to assess their knowledge and understanding.

Dreyfuss (1991:34) describes modeling in terms of a mathematical representation for a non-mathematical object or process:

> Typically, the term modeling refers to finding a mathematical representation for a non-mathematical object or process. In this case, it means constructing a mathematical structure or theory which incorporates essential features of the object, system or process to be described. This structure or theory, the model, can then be used to study the behavior of the object or process being modeled.

Vergnaud (1987:227) connects with this view when stating the importance of representation as vital component of mathematics teaching and learning and points to the relationship between representation and modeling as the process of conceptualising the real world:

> Representation is a crucial element for a theory of mathematical teaching and learning, not only because the use of symbolic systems is so important in mathematics, the syntax and semantic of which are rich, varied, and universal, but also for two strong epistemological reasons:
>
> (1) Mathematics plays an essential part in conceptualizing the real world;
>
> (2) Mathematics makes a wide use of homomorphisms in which the reduction of structures to one another is essential.

Dossey et al (2002:114) refers to mathematical modeling as the process of presenting real-world situations through mathematics and that mathematical models facilitate understanding of the environment and help individuals deal with problems (see 2.2.2). The first reason for the importance of representation Vergnaud's states is in this sense connected to mathematical modeling. Representations as signposts throughout the modeling process and also as a tool to express the final product of the process thus play an important part in the conceptualisation of real-world situations. The second epistemological reason for representation as crucial element in a theory of mathematical teaching and learning is the reduction of structures to one another. When one structure is reduced to another by using homomorphisms, representations usually are changed to give rise to better understanding of the structures themselves and of the situation represented. This idea connects with the dynamic process of transnumeration in which (data) representations are changed to engender understanding (Chick 2003:207; Wild & Pfannkuch 1999:227). The construct of transnumeration will be discussed in 3.6.2.

representation IN THE mathematicS CURRICULUM

The crucial role of representation in mathematics is explicated in the NCTM's Principles and Standards (2000:67). The creation and use of representations to organise, record and communicate ideas stand central in mathematics. Learners should be able to solve problems by selecting, applying and translating among mathematical representations and use representations to model and interpret physical, social and mathematical phenomena. Representations not only help learners to understand mathematical concepts and relationships, but also enable them to communicate mathematically, sharing their approaches, arguments and understanding with others. In this way learners are able to apply mathematics to real-world problems and recognise connections among related concepts. Teaching for representations as ends in themselves, puts limitations on the power and utility of representations as tools for learning and teaching and is counterproductive (NCTM 2000:69). Learners should learn to express themselves through the representations, even if the representations are unconventional. However, to be able to learn mathematics and facilitate communication with others, learners should also learn conventional forms of representation. Technology should be integrated into mathematics teaching and learning to open up new possibilities of expression and representation to learners.

Representation is explicitly part of Mathematics in the Revised National Curriculum Statement in South Africa. Representation of numbers and their relationships; of patterns and their relationships; characteristics of 2-D shapes and 3-D objects and their relationships and the representation of statistical data are stated in Learning Outcomes 1, 2, 3 and 5 (DoE 2002:6).

There are however contradictory tensions in the current curriculum. For example, the curriculum espouses constructivist views, yet these are not compatible with an inherently behaviouristic view of producing "measurable outcomes". An outcomes-based curriculum focusing on testable behaviour is not at all compatible with constructivist view that learning is conceptual and cognitive, which cannot be measured directly.

DATA REPRESENTATION

3.6.1 Introductory notes

Key skills in statistics education are grouped into two distinct but interrelated clusters, namely generative skills and interpretive skills (Gal 1998:276). Gal gives a description of each category of skills: generative skills include the generation or gathering and organising of data, the execution of needed computations, the construction of graphs and charts and the carrying out of possible statistical significance testing. Interpretive skills entail the evaluation and communication of the meaning and implications of data. The context of the data will determine the specific nature of interpretive skills needed. In this study the focus is on generative skills, specifically the organisation and representation of data.

Data representation is a crucial component of statistical thinking (Friel & Bright[2] 1998:94, 95; Wild & Pfannkuch 1999:227; Mooney 2002:27). Bright and Friel (1998:64) contend that "…statistical understanding is not useful unless that understanding can be communicated to others". The implication is that statistical understanding must be represented in some way to be communicated. Bakker (2004:31), in the same vein, emphasises the need for conceptual structures in the process of sense-making in statistics, and states that communication about concepts is impossible without representations. Data representation refers to the way in which

> …data are summarized, presented, and interpreted and whether or not the type of table, charts and/or graphs that a student constructed to represent data were appropriate (Lajoie, Lavign, Munsie & Wilkie 1998:222).

[2] Note that the authors Friel and Bright have published different articles in the same publication (1998), each author acting as first author in one of the articles.

The RNCS (DoE 2002:38,66) emphasises the critical role of representation in data handling, referring to the way in which different representations can either highlight or hide features of a situation. In the Curriculum Statement it is recommended that when the data are collected, special attention must be given to the representation thereof. Learners should be guided to understand how to organise the data in a manner that allows them to conduct the proper data analysis to answer the question posed in the beginning.

Choosing the best way in which to represent data has in a number of studies shown to be difficult (Chick 2000; Cox & Brna 1995; Friel, Curcio & Bright 2001; Gerber, Boulton-Lewis & Bruce 1995; Li & Shen 1992). Effective representations are only possible if the data are transformed appropriately (Chick 2003:213). Chick and Watson (2001:106) suggest that learners find it easier to interpret data than to represent it in an appropriate way and that learners are able to interpret data on a higher level than their representational skills. Chick (2003:212) recounts that learners need to progress through four phases in order to obtain an appropriate and effective representation. These four phases in which data are transformed, embody the process of transnumeration and will be scrutinised in the next section.

3.6.2 Transnumeration

The process of translating between data representations is captured by the term *transnumeration.* The term was first used by Wild and Pfannkuch (1999:227) and is described as: "… a dynamic process of changing representations to engender understanding". The authors argue that the key idea in a statistical approach to learning is to form and change data representations of aspects of a system to arrive at a better understanding of that system.

Chick (2003:207) points out that the choice of representation in data tasks has proven problematic and that the transnumerative process of representing data may be more difficult than the process of interpreting the data. She suggests that learners must go through four linked processes of transnumeration in order to represent data:

- first a decision must be made what message to convey from the data
- the second step is to determine what kind of representation is needed
- thirdly a choice of computation method to transform the data must be made (this phase includes data arrangement)
- finally the data as transformed in the third step is used in the representation.

The first two processes may happen in reverse order or even simultaneously. The last three processes appear to be particularly intertwined. If learners do not have a clear sense of the message that the data is conveying, they will have difficulty to decide what kind of transnumerative processes or representation to use. The process of data arrangement can take place in either the second or third steps, differing from learner to learner. Genuine understanding will most probably emerge and become evident to others when transnumeration takes place, that is, when a learner is able to represent knowledge in a number of different ways and is capable of translating between these different representations. One of the reasons for problematic data representation stems from a failure to understand how to represent different data types appropriately (Chick 2003:207), therefore different data types will next be investigated.

3.6.3 Data types

There are two main data types, namely quantitative and qualitative or categorical data (Steyn, Smit, Du Toit & Strasheim 2000:6,7). Quantitative data consists of continuous data and discrete data. There are also two qualitative or categorical data types, namely ordinal and nominal data. The different data types can be schematically represented as follows:

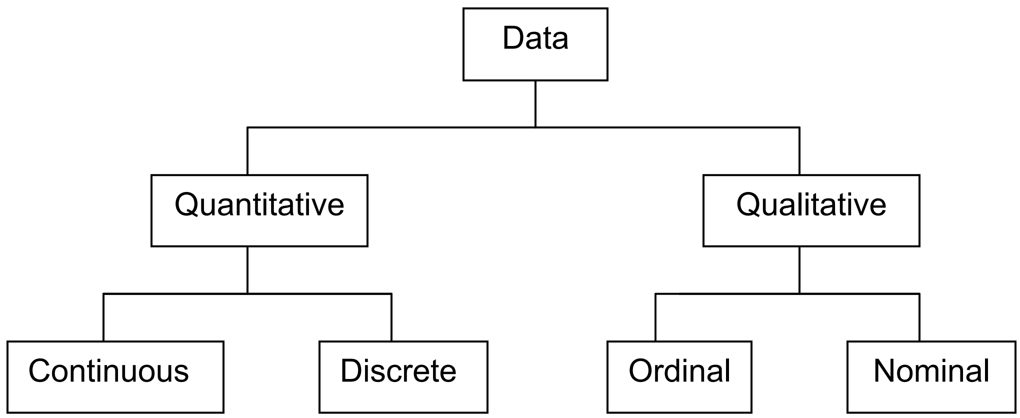


Figure 3.3: Schematic representation of data types (Steyn et al 2000:6)

3.6.3.1 Quantitative data

Information about any characteristic that is measurable on a numerical scale is called quantitative data (Steyn et al 2000:7). Examples of quantitative data are the number of siblings in a family; and information about the age and salaries of workers. Two types of quantitative data are distinguished by Steyn et al. (2000:7, 8):

- Discrete data
  When observations of a characteristic can take on only fixed, isolated values, this information is called discrete data. Examples are the number of learners in a class; the number of classrooms in a school and the possible outcomes when a die is cast repeatedly.
- Continuous data

When observations of a characteristic can take on all possible values within a logic interval, the information is called continuous data. Examples of this data type are the fuel consumption of cars; the surface area of a floor; and the weight or length of a group of people. Most quantitative data are continuous.

Quantitative data can in some cases be grouped in categories, for example small, medium sized, and big.

3.6.3.2 Qualitative data

Many characteristics, such as favourite colour, eye colour, favourite sport cars, and so forth, can not be described as a number. This data type is called qualitative or categorical data.

- Ordinal data

  Categorical data can often be quantified, for example when a teacher classifies his learners' disposition in class as negative, indifferent, good or excellent and then quantifies it by allocating a number to it such as 1, 2, 3 or 4. Data that display a definite order or position in the categories are called ordinal data. The numbers do not have any physical meaning but are only used to order disposition.
- Nominal data

  Characteristics such as hair colour and data about preferences like favourite pets are classified as nominal data.

Different data types influence the kind of representation used to communicate data. Multiple representations and the translation between these different representations enhance clear communication and understanding of data. Data types are one of the critical factors influencing the way in which data are organised in order to represent it. Organisation or arrangement of data constitutes part of the processes of transnumeration and will be the focus of the next section.

3.6.4 Data arrangement

Data arrangement refers to the way in which learners organise, categorise or consolidate data into summary form during the process of data reduction (Mooney 2002:26). Data arrangement or organisation plays an important role in the transformation of data during the phases of transnumeration as described by Chick (2003:212). This topic has however not received much attention in the past. Virtually no research has been documented on data arrangement types of primary and middle school learners (Mooney 2002:26; Johnson & Hofbauer 2002:1282).

Johnson & Hofbauer (2002:1284-1288) describe 5 categories of data arrangement found in a study of Grade 6-8 learners making sense of raw unorganised data. In this research study learners were asked to arrange and represent data in a suitable format for the school newspaper. The arrangement categories used by these authors are: no arrangement; clustered arrangement; sequential arrangement, summative arrangement and regrouped summative arrangement. When learners did not attempt to arrange the data or arranged data inappropriately, their responses were categorised as ‘no responses’. Clustered arrangement involved the sorting of data into groups with no totals, while in sequential arrangement learners listed data in alphabetical or numerical order. Summative arrangement comprised the sorting of data into groups with totals provided. When learners regrouped the data and provided totals, their responses were classified as regrouped summative arrangement. These data arrangement types identified by Johnson & Hofbauer (2002:1284-1288) reveal an increasing level of sophistication and present guidelines of what can be expected from learners.

3.6.5 Representational types

Different representations are used as tools for thinking about and solving problems, as well as instruments for communication. The development and use of a variety of representations to model problem situations, investigate mathematical relationships and to justify or disprove conjectures, is emphasised in the Representation Standard of the NCTM (2000: 68):

> Instructional programs from Kindergarten through grade 12 should enable all students to-
>
> - ♦ Create and use representations to organize, record, and communicate mathematical ideas;
> - ♦ Select, apply and translate among mathematical representations to solve problems;
> - ♦ Use representations to model and interpret physical, social and mathematical phenomena.

Representational types that can be expected from the age group in this study and that fall within curriculum guidelines (Department of Education 2002:56, 57, 88) are the following:

- Idiosyncratic, invented or nonstandard representations
  These representations include learners' own idiosyncratic attempts to represent the data to make sense thereof, and may be in the form of pictures, pairing off of data points, and so forth.

- Lists

A list exists of a number of connected items, names, etc. written or printed together usually consecutively to form a record or aid to memory (Readers' Digest Illustrated Oxford Dictionary 1998:472).

- Stem-and-leaf plots

  A stem-and-leaf display is a statistical representation resembling the shape of a leaf which is used to organise and display a set of numerical data to make it easier to order the numbers (Department of Education 2002:108).
- Tables

  A table is an arrangement of numerals, letters or signs, usually in rows and columns, to show facts or relationships between them in a compact form (Bendick & Levin 1973:190). Relatively few data points are involved compared to a complex graph (Gal 1998:278).
- Pictogram or pictograph

  A graph that makes use of pictures (for example, people, cars) to represent data (Department of Education 2002:107).
- Frequency graphs

  An arrangement of data according to the number of times an event occurs (Bendick & Levin 1973:90). The frequency can be depicted with crosses or similar symbols.
- Bar graphs

  A bar graph is a graph that uses vertical or horizontal bars on a set axes to represent information (Department of Education 2002:103).
- Histograms

  A histogram is a bar graph which shows the frequencies of grouped data as rectangles or bars (Department of Education 2002:105; Watkins, Schaeffer & Cobb 2004:41).

- Pie charts

A pie graph or pie chart is a graph which uses the sectors of a circle to display the ratio between different categories in the data (Department of Education 2002:107).

- Line and broken-line graphs

  A line graph is a graph representing continuous information, for example an event occurring over time (Cassim, Geju, Nel, Wessels & Wessels (2006:258). A broken-line graph is a display where plotted points are joined by line segments (Department of Education 2002:103).

The kind of representation used depends on a number of factors: content and context of the question; data type; what ‘story’ or message of the data needs to be conveyed; the repertoire of statistical tools of the respondent or his/her creativity in inventing a non-standard display, and so forth. One of the most important factors may be the decision about how to effectively use a graph to tell the story of the data, in other words the message of the data (Chick 2003:207). Depending on the kind, variety and richness of activities a learner has been exposed to, the expectation is that sophistication of representations that may be used should increase with grade level. Processing capacity usually increases with age because representations become differentiated into more dimensions, enabling more complex relations to be represented. The different levels of complexity shown by the different representational types can however not be regarded as hierarchical because the distinguishing factor is not the level of complexity a learner can display in a representation, but how effectively the message in the data can be conveyed. The appropriate way in which a statistician who knows all the standard ways of displaying data may represent a specific set of data may be a less sophisticated display such as a pie graph or a table and not a complex representational type, because the message he wants to communicate may be better conveyed by a less complex representation.

Elementary teachers should however not only focus on graphing activities, but on the characteristics of and trends in a set of data. Generative as well as interpretive skills should be developed:

> By explicitly directing attention to the nature of data, alternative representations, and prediction, the focus of a graphing activity changes from the activity of drawing and tabulating data to underlying elements …. If these elements are developed at the primary level, it will provide the necessary base on which secondary teachers can build (Pereira-Mendoza 1995:6).

When analysing learner responses, teachers and researchers need to make qualitative judgements about the level of each of the responses. One of the notable tools that are used world wide to simplify analysis of learner responses is the SOLO Taxonomy and will now be examined.

the solo taxonomy

Piaget postulated three stages of representation, namely a topological phase (2-7 years); projective or Euclidean phase (7-12 years) and an explicit, formal phase in which representational systems is mastered (12-18 years) (Pegg & Davey 1998:120). These stages build on one another and Jerome Bruner, building in part on Piaget's theory, contends that learners move through three stages of representation as they learn: the enactive, iconic[3] and symbolic stages. Each of these developmental stages builds on the previous stage. A child directly manipulates objects in the enactive stage, during the second stage mental imagery with visualisation of operations or concrete manipulation takes place and in the third stage the manipulating of symbols in stead of objects or mental images of objects occurs. Therefore concept forming results through

1. manipulating objects
2. pictorially representing them

[3] Note that the word iconic in this context is spelled with a 'c' rather than a 'k' as in the SOLO context.

3. symbolically representing them.

These three modes of representation constitute a learning cycle and connect with what is described by the term multimodal functioning (Pegg & Davey 1998:120). In the theories postulated by Piaget and Bruner learner responses are regarded as an indication of the cognitive level on which they operate. Their theories however failed to explain different levels of response for different tasks of the same learner and thus the different levels of understanding evident in the responses. In Chapter 1 the suitability of the SOLO Taxonomy as evaluative tool to capture the differences in responses was discussed. This neo-Piagetian taxonomy has been used in a variety of fields, e.g. statistics, science, technology, and in different strands in mathematics such as geometry, measurement and fractions. The SOLO Taxonomy also forms the basis of a number of the statistical thinking frameworks that could be found in the literature and was developed by Biggs and Collis (1982) to offer a better appreciation of learners' understanding. This categorisation system is referred to as the Structure of the Observed Learning Outcome or SOLO (Biggs & Collis 1982) and focuses on learner responses rather than on their thinking level or stage of development. The SOLO Taxonomy categorises the level of a learner's response in a specific task and situation, thus "the learner's current state of understanding of some particular content or process" (Killen 2004:80) and is not an indication of a cognitive level on which the learner operates. The taxonomy provides a general framework for the systematic assessment of the quality of learning (Collis & Biggs 1986:1). SOLO levels describe a particular performance at a particular time and "…are not meant as labels to tag students" (Biggs & Collis 1982:23). Pegg and Davey (1998:116) state that in contrast to the views of Piaget and Bruner, the SOLO classification

> … does not imply that the way students perform in different situations is typical of their stage of cognitive development, nor that this is necessarily related to their age. In particular, student growth in understanding is not seen in terms of stages related to some overall logical structures that exist within the mind.

The SOLO Taxonomy view understanding as a much more individual feature that is both content and context specific. This taxonomy is based strongly on the significance of working memory capacity and information-processing theories. This taxonomy is concerned with observable behaviours and the determining of the response category depends on learners' familiarity with content and context (Pegg & Davey 1998:110).

The SOLO Taxonomy, first postulated by Biggs & Collis (1982), evolved through later modifications (Biggs & Collis 1991; Pegg 1992). The SOLO Taxonomy theorise that all learning takes place in one of five modes of functioning, a characteristic that corresponds with Piaget's stages of development (Fig. 3.4).

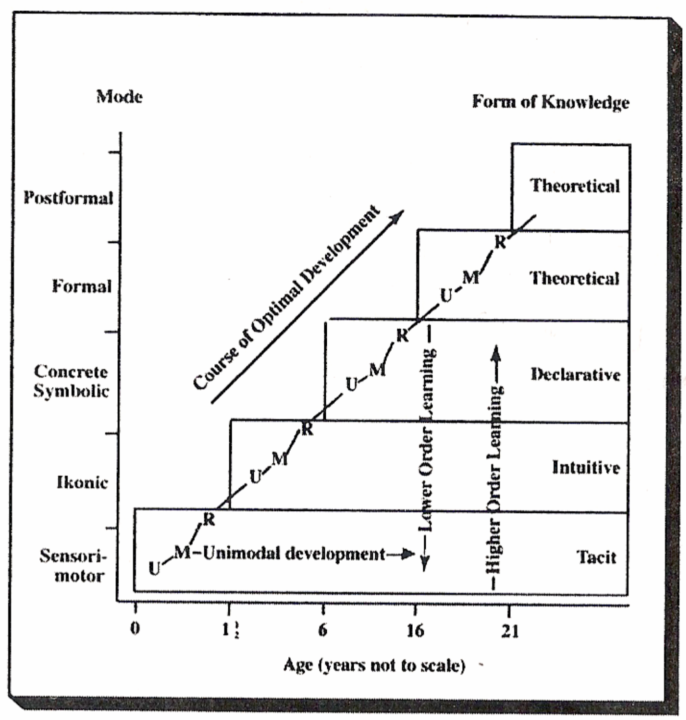


Figure 3.4: The SOLO Model: modes, learning cycles and forms of knowledge (Pegg & Davey 1998:119)

3.7.1 Modes of functioning

The five modes are now described with mention of the age at which they generally begin to appear:

- Sensorimotor (soon after birth)
  The child reacts to his physical environment and acquires motor skills in this mode. Knowledge acquired in this mode is referred to as tacit knowledge.
- Ikonic (from 2 years)
  Actions are internalised in the form of images. The individual develops words and images that can stand for events and objects. The form of knowledge acquired in this mode is referred to as intuitive knowledge.
- Concrete symbolic (from 6 to 7 years)
  The child learns to use symbol systems such as the number system and written language. This mode is most commonly addressed in the upper primary and secondary school. Knowledge acquired in this mode is called declarative knowledge.
- Formal (from 15 or 16 years)
  More abstract concepts can be dealt with and the individual can work in terms of “principles” and “theories” (Pegg & Davey 1998:117). The kind of knowledge in this mode is described as theoretical knowledge.
- Postformal (from about 22 years)
  The fundamental structure of theories or disciplines are questioned or challenged. The kind of knowledge in this mode is as in the formal mode referred to as theoretical knowledge.

These five modes correspond with Piaget’s stages of development. An important adjustment of Piaget’s model lies in the placing of the *early formal stage* of the 13-15 year olds into the earlier group of stages called *concrete operations.* Learners in this age range are satisfied by a few specific instances that a rule is reliable, they are not yet “formal thinkers” but “concrete generalizers” (Pegg &

Davey 1998:116), they are at this age thus still connected to their concrete experiences.

3.7.2 Multimodal functioning

Another contrasting aspect of the SOLO model and Piaget's stages of development is the concept of multimodal functioning. Piaget's stages represent a single-path development where one stage is replaced by another. In the perspective of multimodal functioning one mode is not subsumed or replaced by another. Instead, the development of a mode is supported by the continued development of earlier modes and growth in later modes is often connected to thinking and associations of earlier modes. Pegg and Davey (1998:120) give two examples of multimodal functioning, one where functioning in an earlier mode is supported by higher modes and the other where the target mode is supported by learning in earlier acquired modes. The first example is of an athlete, striving to improve his performance (sensorimotor level of response) by practicing the skills (sensorimotor mode). He can also gain better insight into his own performance by (a) watching elite performances in action through which he can build mental images (ikonic mode); (b) reading about techniques to improve related aspects of the skill to help him in his performance (concrete symbolic mode) and (c) analysing problems in his own performance and generalising principles about performance or competition (formal mode). If quick reactions are needed, responses should become automatic and it will take time to build skills and incorporate new techniques.

A second example is where the target mode is the concrete symbolic mode and earlier modes support functioning in this mode such as the development of generalised rules concerning the four operations. The use of Dienes blocks to support learning in this regard represents the ikonic mode as support for the concrete symbolic mode. Another example of using the sensorimotor and ikonic modes as support for the concrete symbolic mode is the use of concrete materials in the development of pre-algebra. The main focus, however, should be the target mode. Too heavy an emphasis on supporting modes may lead to the development of two independent structures, which may give rise to confusion and defeat the purpose of learning.

Most learners in the primary and secondary school are capable of responding in the concrete symbolic mode (Pegg & Davey 1998:117). It is however not implied that a learner who responds in one mode of functioning in a specific task, will respond in the same mode for another task. Although the concrete symbolic mode is target mode in the primary school and teaching techniques are adapted to this mode, learners may still respond to stimuli in the ikonic mode and also respond in the formal mode in some tasks.

3.7.3 Levels of response

A second important characteristic of the taxonomy is the five levels of response that measure increasing sophistication in handling certain tasks within a particular mode (Biggs & Collis 1982, 1991). The levels are:

- Prestructural responses (P) (lower than the target mode)
  The individual is not engaging in the task at hand and often focuses on irrelevant aspects of the situation. There is no use of the elements required to identify the mode in question.
- Unistructural responses (U)
  The learner is focusing on the problem but uses only one piece of relevant data.
- Multistructural responses (M)
  No integration occurs on this level. Although the learner uses two or more pieces of data, no relationships between them are observed. The

processing of several disjoint aspects of the data are usually done in sequence.

- Relational responses (R)
  The learner focuses on several aspects of the data and perceives relationships between different aspects in the data.
- Extended abstract responses (EA)
  The individual can go beyond the data, generalising from new and abstract features. Integration is accomplished to such an extent as to enter the unistructural level of a higher mode.

Prestructural responses are an indication of functioning in the previous mode while extended abstract responses indicates functioning in the next mode, therefore the unistructural, multistructural and relational levels are considered as the basic levels in the concrete symbolic mode.

In summary the common features of the three basic levels are when learners focus on (a) the context rather than the data (ikonic or prestructural); (b) the data as single values (unistructural); (c) the data as a series of values (multistructural) and (d) the data as belonging to an entire data set (relational) (Watson & Moritz 2001:52).

3.7.4 Intramodal development

In findings of recent research studies an intramodal development pattern has been identified (Campbell, Watson & Collis 1992; Levins & Pegg 1993; Pegg 1992; Watson, Collis & Campbell 1995; Watson, Collis, Callingham & Moritz 1995). This intramodal development pattern concerns at least two U-M-R cycles in the concrete symbolic mode. The first cycle of growth may be linked to the *development* of a particular concept ($U_1M_1R_1$) while the following second cycle ($U_2M_2R_2$) is associated with the *consolidation* and *application* of the concept. Figure 3.5 represents the intramodal cycles of growth in the concrete symbolic mode. The first cycle represents the intuitive responses which are building blocks for the second cycle, and explains the variability within the U_2 responses.

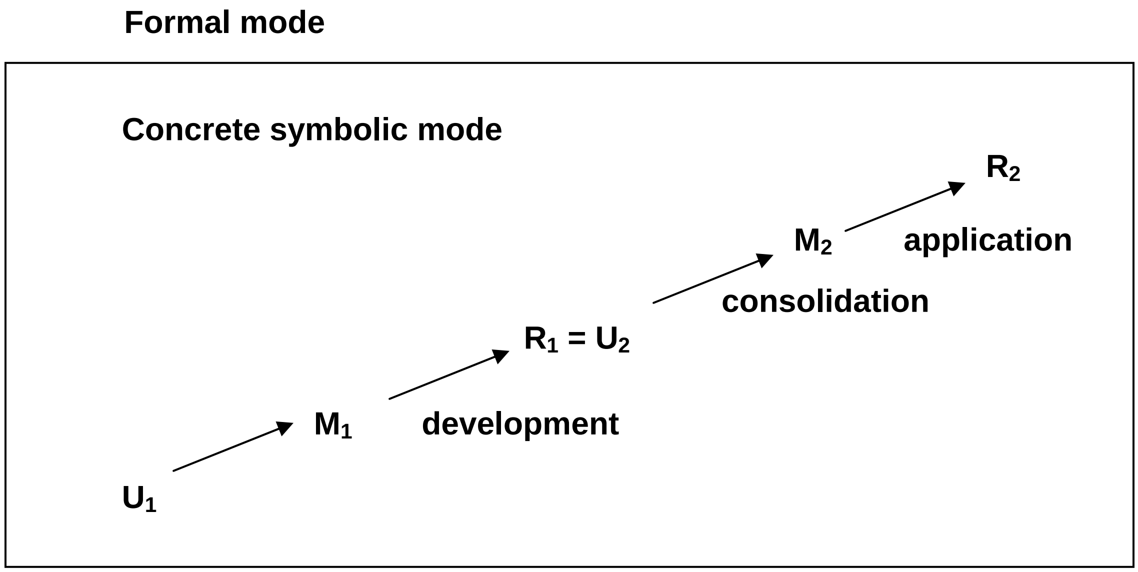


Figure 3.5: U-M-R cycles within a mode
(Watson, Collis, Callingham & Moritz 1995:251)

3.7.5 The problem solving path

In addition to the intramodal development pattern, another way of thinking about ikonic and concrete symbolic functioning lies in the problem solving path suggested by Collis and Romberg (1991). Watson et al (1995:252) have adapted this problem solving path and indicate that this path describes the relationship between ikonic and concrete symbolic functioning during problem solving (Fig. 3.6). In this problem solving path a learner at the onset chooses an ikonic or concrete symbolic route, with possible interaction taking place at stages B and C of the problem solving process. Although concrete symbolic functioning is associated with successful problem solving, some ikonic functioning may also potentially result in successful problem solving.

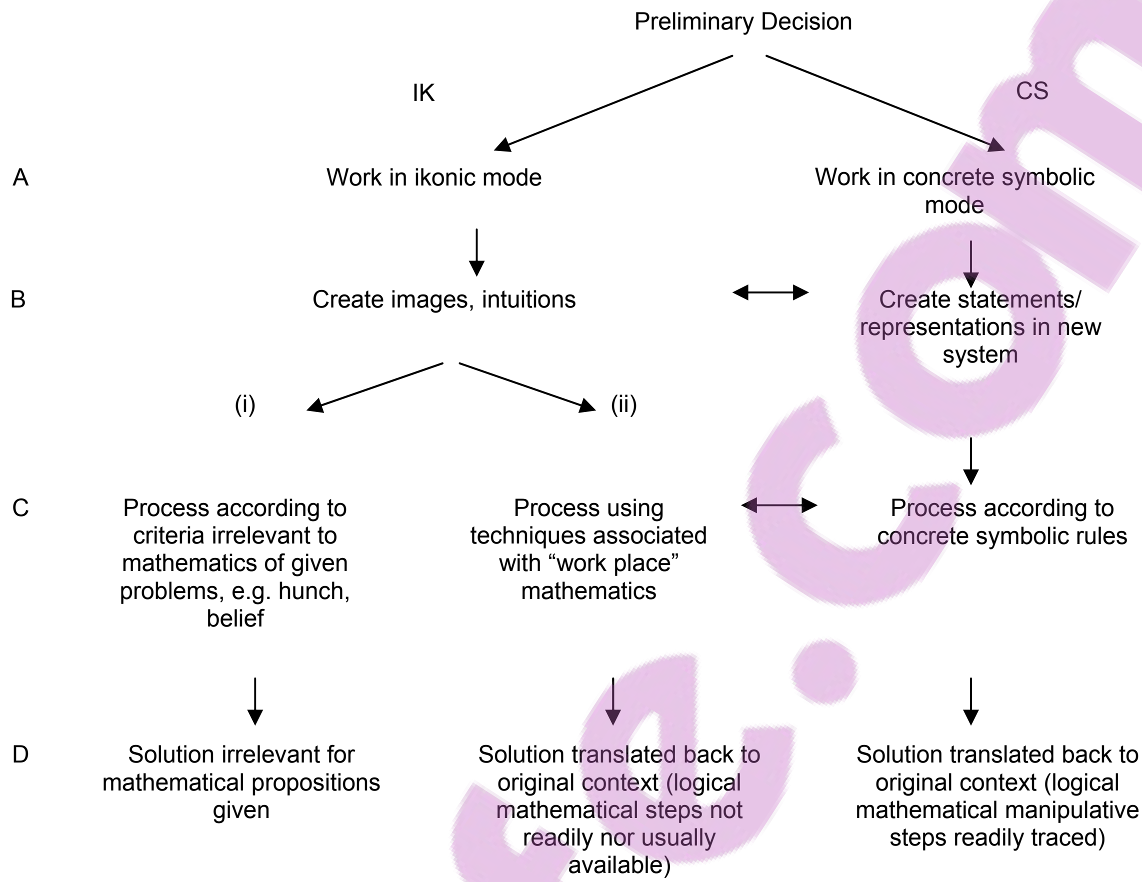


Figure 3.6: The problem sol Co

(Watson et al 1995:252)

In summary it can be said that a framework such as the SOLO Taxonomy can provide a developmental perspective describing the interaction between “the specific, concrete imagery of the ikonic mode and the logical structures of the concrete symbolic mode” (Campbell, Collis & Watson 1995:180). From the discussion the enormous potential of the SOLO Taxonomy as an evaluation tool and as model to explore and explain learner growth is clear and that the value of the SOLO Taxonomy is to be found in the depth of analysis it provides for interpreting learner responses.

CONCLUSION

Representation can not be regarded in isolation. Different perspectives on mathematical learning and problem solving that influenced mathematics education during the past years, lead to diverse views on representation. In this chapter the concept of representation has been examined from the context of a unified model of representational systems.

From a modeling perspective, representations are the tools with which an individual communicates his or her understanding of a complex real-world situation to himself or herself and to other people, including the teacher. Internal representations can only become accessible to others when it is communicated in external form. Internal and external representations are continually interacting with each other, leading to better understanding of the represented concept or entity. Representational fluency is of critical importance in representation and usually is an indication of integrated understanding. Our technology-based society requires the use of a multiplicity of systems, which are functionally and dynamically linked.

Data representation as generative skill is an integral element of statistical thinking and is the end result of the four phases in the process of transnumeration. It is crucial that learners have a good sense of the message of the data in order to appropriately represent it and it is equally critical to realise that some representations are better than others for communicating the data or “telling the story in the data”.

The SOLO Taxonomy has proved to be a valuable evaluative tool for qualitatively judging learner responses. The taxonomy categorise learner responses to determine the individual’s current state of understanding of particular content or processes. Through using the SOLO model the intuitive understanding of a concept can be categorised as well as the consolidation and application thereof, facilitating a depth of analysis for interpreting learner responses by teachers and researchers not easily achieved by other means.

In Chapter 4 the empirical investigation will be described, focusing on detail such as population, samples, instruments, analysis categories and coding of solutions. The research method and analysing techniques will also be stated.

CHAPTER 4
METHOD OF RESEARCH

4.1 INTRODUCTION

This chapter describes the planning and execution of the empirical investigation. The focus of the research is revisited, the population and samples are described and the selection of test items is discussed. An explanation of the coding of solutions and the chosen categories of analysis is given while the method of research and the techniques used to analyse the data are stated.

4.2 REVISITING THE RESEARCH QUESTION

The main research question of the study as posed in 1.2 concerns the types and levels of data arrangement and representation as modeled by Grade 4 to 7 learners, and can be condensed to five main sub questions to guide the investigation:

- What is the role of modeling and problem solving in mathematics and data handling?
- What is the role of representation in mathematics and statistics teaching and learning?
- How will learners spontaneously arrange statistical data?
- What kind of representations will learners spontaneously use when arranging statistical data?
- What is the observed SOLO level of data representation of each learner?

The first two sub questions were dealt with in the literature study in chapters two and three. An empirical investigation was necessary to obtain answers for the last three questions.

The quest of the researcher was to determine the statistics elements fundamental to each task, as well as to understand and describe the arrangement strategies and representation of learners. The explanatory nature of a study of this kind necessitates a qualitative study, more specifically descriptive research (see 1.5.2). The researcher wanted to gain insight in the learners' understanding and modeling of the tasks as evident in their representations. A limited quantitative analysis was done to support the quantitative investigation (see 1.5.5.3). The planning and execution of the empirical study will now be reported.

4.3 STUDY POPULATION AND SAMPLE

4.3.1 Study population

The experiment was conducted with permission of the Gauteng Department of Education in a suburban government school in Pretoria where hundred and forty four learners completed the data tasks of the study. This particular school was selected because mathematics teaching and learning are conducted according to the problem centered approach in this school. The participants were grade 4 to 7 learners of mixed ability. The language of instruction, Afrikaans, was used in all the tasks in the experiment. The gender distribution for the tasks respectively was 80 boys and 64 girls.

4.3.2 Sample

In this study, non-probability sampling was used, in particular convenience sampling. The specific school chosen to conduct the experiment in follows a problem-centered approach in mathematics. Furthermore, permission from the principal and logistical arrangements made it convenient for the researcher to conduct the research at this particular school. Therefore, availability, convenience and a problem-centered background in mathematics were the key factors in determining the sample of the study. In another school that permission was granted to do the research, a large group of children could not complete the tasks because of transport problems or extra-curricular activities. The data collected at this school could therefore not be used in the study.

Grade 4 to 7 learners were chosen for this study because in the South African mathematics curriculum learners have received a limited amount of instruction in statistics at that age. As one of the goals of the study is to elicit learners' *spontaneous* efforts to arrange and represent statistical data, this age group presented a better choice than older learners who may have received more instruction in statistics and whose responses might not have been spontaneous, but influenced by instruction. The author also believed that Grade 4 to 7 learners would be able to model the task better than could learners in the Foundation Phase because of their exposure to a problem-centered environment in mathematics. A summary of the sample is presented in Table 4.1.

GRADE	GIRLS	BOYS	TOTAL
4	15	21	36
5	19	22	41
6	18	26	44
7	12	11	23
TOTAL	64	80	144

Table 4.1: Summary of sample

4.4 INSTRUMENT

The instrument consisted of two data representation tasks. The tasks were open-ended, designed to include categorical as well as numerical data and provide different contexts as learners may not relate well with a certain context. The tasks have been adapted from interview protocol tasks used in the research of Mooney, Langrall, Hofbauer & Johnson (2001), Jones, Mooney, Langrall & Thornton (2002) and Chick & Watson (1998, 2001). Because of the attention span of the age group and time limitations set by the school, only two tasks could be used.

Task 1 (categorical data)

23 Grade 4 learners have been asked about their favourite colour. The data was collected on cards and a list of the data has afterwards been drawn up. You have received the list. Your job is to arrange the data to be presented on a poster that will be exhibited in the class. The poster must give a good idea of what the favourite colours of learners are, even when one is not looking at the poster from close up. You may not copy the list, but must think up some other presentation of the data (information).

<u>List</u>

Susan	Brown	Talitha	Blue
Morgan	Yellow	Sam	Blue
Johann	Red	Lindiwe	Orange
Lee-Anne	Brown	Charl	Yellow
Dalene	Yellow	Hans	Red
Nomvula	Green	Peter	Blue
Sally	Pink	Chané	Green
Mpho	Green	Naledi	Green
Darren	Red	Sanette	Blue
Nomsa	Red	Rudolf	Green
Shirley	Blue	Themba	Blue
Lida	**Pink**		

Figure 4.1: Task 1

Task 2 (numerical data)
A company that manufactures beach sandals needs to know more about the size of Grade 4 to 6 learners' feet. A list was made of the feet size to the nearest cm (centimeter) of 8 boys and 8 girls. Present the data on a poster that will be exhibited together with 30 other posters at a meeting of beach sandal manufacturers (you may not copy the list).

List

	Boys			Girls	
John	Gr 6	17 cm	Patricia	Gr 5	15 cm
Jannes	Gr 5	16 cm	Kendra	Gr 4	14 cm
Rudi	Gr 4	16 cm	Sandra	Gr 6	18 cm
Cassim	Gr 5	17 cm	Joyce	Gr 6	16 cm
Tony	Gr 5	16 cm	Thandi	Gr 6	18 cm
Sipho	Gr 4	13 cm	Tina	Gr 4	15 cm
Hassan	Gr 6	18 cm	Odette	Gr 5	16 cm
Pieter	Gr 6	19 cm	Nomsa	Gr 4	15 cm

Figure 4.2: Task 2

Both assignments clearly state that the given list might not be copied onto the poster. The first task consists of categorical data, while the second task comprises numerical data (see 3.6.3).

Interviews with the Grade 4 -7 teachers will be conducted if necessary to shed light on findings.

4.5 VALIDITY AND RELIABILITY OF THE INSTRUMENT

Legal, ethical, philosophical and procedural principles as stated in 1.4.3 have all been adhered to in the research project to ensure the validity and reliability of the research.

Face validity of the tasks was established by judgment of two independent researchers in the field. The test items were judged to represent items from the content of Learning Outcome 5 (Data Handling) of the Mathematics curriculum of the target age group as well as of the area of research.

In the quantitative analysis, validity was established by using the Rasch model. The Rasch analysis was done by Dr John Barnard, an academic of Melbourne, Australia. The Rasch model allows one to create person-item maps and provides indices of separation (person separation reliability, item separation reliability). By using both statistics and visual plots, such maps and indices allow one to evaluate the construct validity of an instrument (Boone & Rogan 2005:35). A person-item map was created to evaluate construct validity by examining the distribution of items along the latent trait.

> If the items in a test or questionnaire are sufficiently well separated do define several statistically distinct levels, and hence a direction, we are ready to examine their ordering to see whether it makes sense. The pattern of item calibrations provides a description of the reach and hierarchy of the variable. This pattern can be compared with the intentions of the item writers to see if it confirms their expectations concerning the variable they wanted to construct. To the extent that it does, it affirms the *construct validity* of the variable (Wright & Masters 1982:93).

Rasch item infit and outfit statistics were calculated for each item on the test. Learner ability and item difficulty show a good fit, and none of the items were identified being misfits. When there is a lack of misfitting items on a test, strong content as well as construct validity for the test is suggested (Boone & Rogan 2005:35). Through validity, reliability is also established, because validity always implies reliability even though the opposite is not true (Charles & Mertler 2002:159).

4.6 DATA SOURCES AND DATA GENERATION

4.6.1 Data sources

The data sources consisted of

- arrangements and representations of the statistical data in the two data tasks and
- tables, descriptions and summaries generated during analysis.

4.6.2 Data generation

Permission of the Education Department, the principal of the school and the parents were obtained (see Appendix A, B and C) before the data collection took place. Learners did the tasks in classes of between 15 and 25 learners each, but responded individually and were not permitted to discuss the tasks with their class mates. Teachers who invigilated in the different classes underwent a short training/information session of 20 minutes during which the researcher explained the purpose and logistics of the experiment as well as their role in the class. Some of the teachers nevertheless did not give clear instructions to learners. This was evident from the fact that some learners did not use the paper provided (one sheet of paper for each task), but responded on the typed task sheets itself or gave both responses on one sheet instead of one response per sheet. Space on the typed task sheets was limited, with the result that some learners, who responded on the task sheets in stead of on the answering sheets, ran out of space. Learners took between 40 and 60 minutes to complete the data tasks.

4.7 DATA ANALYSIS

The data will be analysed qualitatively. The qualitative analysis comprises three parts:

- analysis and description of the data arrangement types
- analysis and description of the representational types
- analysis and description of the levels of data representation according to the SOLO Taxonomy framework.

A limited quantitative analysis will be done to establish validity and reliability using the Rasch model and to support the qualitative findings by tabling and comparing results for different categories and grades. The results of the quantitative measures used can show trends and problems that may otherwise not be evident. The quantitative analysis adds detail to complete the picture.

4.7.1 Data arrangement types

Coding of data arrangement types

Data arrangement types described in the literature (see 3.6.4) adhere to five main categories:

- **No arrangement:** no response or an incomplete response given
- **Clustered arrangement:** data sorted in groups with no totals
- **Sequential arrangement:** data sorted in alphabetical or numerical order
- **Summative arrangement:** groups or categories provided with totals
- **Regrouped summative arrangement:** when data were regrouped with totals provided

In the analysis of the data combinations of these categories were however found and new categories had to be created. The categories used for the analysis are:

0 = No arrangement: no attempt to arrange the data, leaving it as raw data, or copied data as given

1 = Inappropriate arrangement: statistically inappropriate arrangements of data

2 = Clustered arrangement: data sorted in groups with no totals

3 = Sequential clustered arrangement: data sorted in groups with no totals and groups sorted in either alphabetical or numerical order

4 = Summative arrangement: groups or categories provided with totals

5 = Sequential summative arrangement: groups or categories provided with totals and groups listed in alphabetical or numerical order

6 = Regrouped summative arrangement: data sorted into new groups or categories and totals provided

7 = Regrouped sequential summative arrangement: data sorted into new groups or categories and totals provided; groups sequentially sorted

4.7.2 Representational types

Coding of representational types

The representational types of the data tasks appropriate for Grade 4 – 7 in the South African mathematics curriculum as described in Chapter 3.6.5 learners are

1 = idiosyncratic, invented or nonstandard representations

2 = lists

3 = tables

4 = pictograms and frequency graphs

5 = bar graphs

6 = pie charts

7 = line graphs and broken-line graphs

As was the case with arrangement categories, learner responses displayed a wider variety than expected and categories had to be added:

0 = no representation

1 = pictures/shapes/names/numbers

2 = lists

3 = tables

4 = pictograms and frequency tables

5 = bar graphs

6 = pie charts

7 = line graphs

8 = anomalous representation

4.7.3 SOLO levels of representation

The concrete symbolic mode is the target mode under consideration in this study as learners in the Intersen Phase generally respond in the ikonic or concrete symbolic mode (Pegg & Davey 1998:118). The focus is more specifically on the unistructural and multistructural levels in this mode, although responses on the relational level may also be present.

A framework according to the SOLO Taxonomy described in Chapter 3 will be developed and refined to assess the level of learner responses. The framework describes the general characteristics of cycles of levels within the concrete symbolic mode. The main purpose is to use the framework to investigate a possible hierarchy of responses. The basic framework that will serve as starting point for the analysis includes multimodal functioning, that is two U-M-R cycles (see 3.7.2), and consist of the following categories:

- Prestructural level (P)
- Unistructural level, first cycle (U_1)
- Multistructural level, first cycle (M_1)
- Relational level, first cycle (R_1)
- Unistructural level, second cycle (U_2)
- Multistructural level, second cycle (M_2)
- Relational level, second cycle (R_2)

4.7.4 Coding reliability check

Triangulation, being qualitative cross-validation, was used to corroborate coding and interpretation of results. Triangulation was achieved by double coding as well as supplementary coding reliability checks:

- Double coding of all responses by a second researcher was done after working through the coding rationale and categories. The researcher involved has very wide experience in all aspects of mathematics and statistics education research and has a well developed sense for coding of learner responses. Coders had 86% agreement on coding before discussions. A single numerical coding for each level was assigned as described in 4.8.1.3. Differences were discussed until consensus was reached. Coders managed to reach consensus on all disputed codings.
- An independent researcher analysed and coded the arrangement and representation of a 10% sample of the data tasks for a reliability check. The researcher concerned did post graduate studies on the development of symbolism in Algebra and his wide literature study and experience in the field of symbols and symbolism gave him a good background for the coding reliability check of the arrangement and representation types. The coding system was explained to the second coder and three of each task was coded together. The second coder then coded three tasks independently. The results were compared and differences discussed. The second coder then coded the 10% sample of the data tasks independently. These results were also discussed and where interpretations differed, discussions took place until consensus was reached. Slightly different interpretations of definitions were the reason for most of the differences. In over 80% of the cases, agreement was reached without any discussion being necessary.
- Discussions with cognoscenti in the field of Statistics Education at the 25th, 27th and 29th Annual Conferences of the International Group for the Psychology of Mathematics (PME) (2001, 2003, 2005) yielded a sound orientation for the analysis of the results. Discussions with various researchers at the Universities of Wisconsin-Madison (USA) and Georgia during a four month visit as doctoral fellow in the US provided the opportunity to discuss alternative interpretations and tease out valid explanations. During a visit to the University of New England in Armidale,

Australia (July 2005), the SOLO coding was discussed with three researchers who are world renowned for their knowledge and use of the SOLO Taxonomy, namely Professor John Pegg, Doctor Chris Reading and Doctor Rosemary Callingham. These researchers have all been using the SOLO Taxonomy to categorise learner responses for more than ten years and have all published internationally in this field. Discourse with these researchers provided the opportunity to refine the SOLO Taxonomy framework and verify coding and interpretations. Differences in coding were as in the other coding reliability checks discussed until consensus was reached and coding was changed if necessary.

4.8 SUMMARY

This chapter described the course of the empirical experiment. The research question was revisited to sharpen the focus on the purpose of the study and the sample was described. The basic categories in the three sections of the analysis, namely arrangement types, representational types and SOLO levels of responses were given. The measures that were taken to ensure reliability and validity were stated, for example the coding reliability check. The results of the data analysis are presented in Chapter 5.

CHAPTER 5
ANALYSIS OF RESULTS

5.1 INTRODUCTION

In the preceding chapter the design of the empirical study was presented. This chapter gives the results of the study. The results and discussion of the analysis of data arrangement, representation and SOLO levels of representation are given separately. Some preliminary conclusions are given.

Two tasks were used as instruments to determine the types and levels of data arrangement and representation. Learners had to represent the data on a poster so that an audience could have a sense of “the story the data are telling”, even when they are not close up. The aim was to elicit spontaneous representations from learners. To achieve that, the wording of the task specifically made no mention of graphs and invigilators were told not to use the word ‘graph’ when learners asked questions about the tasks. Some learners were able to choose an appropriate representational form to “tell the story of the data” while others had difficulty interpreting and transforming the data in the tasks. Some learners “… developed sketchy, self-invented ways of symbolizing that did not resemble commonly accepted mathematical language” as did learners in other research projects described by Gravemeijer, Cobb, Bowers, and Whitenack (2000:238). A third task was given, but almost half of the learners did not respond to Task 3, therefore Task 3 was not included in the final analysis and description of results. While it is not quite clear why this is the case, some explanation can be found in the fact that some learners spent so much time on detailed decorated responses to the first two tasks that they did not have time to do the third (see 5.2.1).

Arrangements at the school where the sample was taken, did not allow for extra time to complete all tasks. The context of the task as well as the format in which it was done, may also have contributed to the difficulty learners experienced in answering the question: “In statistics, the context motivates procedures and is the source of meaning and basis for interpretation of results of such activities” (Gal & Garfield 1997:6).

Although arrangement and representation are two linked phases in the process of transnumeration (see 3.6.2), the two aspects were analysed separately to shed more light on the different ways in which learners organise data and which representational types they spontaneously use to display data.

5.2 DATA ARRANGEMENT

The first area of investigation was the data arrangement of learners. As discussed in Chapter 3, a successful representation can only be obtained if the data are transformed in an appropriate way. Arrangement or transformation of data comprises one of the four steps in the process of representing data (see 3.6.2).

As Task 1 comprised categorical data, a summative arrangement is considered to be the most appropriate way of arranging the data. As discussed in 4.7.1, a summative arrangement is a grouped arrangement in which totals for each group are given. The data in Task 2 is numerical and requires different treatment. Regrouping of the data according to feet length is essential to make sense of the data in a meaningful way. As the task states that the data must be displayed for manufacturers of beach sandals, an appropriate way would be to regroup data according to feet length and display the data in a bar graph. Feet length should be shown sequentially so that an observer could get a good idea of the spread of feet lengths.

Johnson and Hofbauer (2002: 1284-1286) describe five broad categories of arrangement namely

- No arrangement
- Sequential arrangement
- Clustered arrangement
- Summative arrangement
- Regrouped summative arrangement (see 3.6.4)

The responses to the two data tasks of this study however revealed combinations of these broad categories and a refined framework for arrangement types had to be designed, including the following combinations of the main arrangement categories: sequential clustered arrangement, sequential summative arrangement, regrouped summative arrangement and sequential regrouped summative arrangement. Cases where no responses were given or where a response was incomplete in such a manner that the researcher could not tell the learners' intent, a 0 on the nominal scale was assigned. Inappropriate responses, however sophisticated, were coded a 1 on the nominal scale. Appropriate responses according to the task requirements were assigned a number 2 to 7. Although sequential arrangement (where data are sorted in alphabetical or numerical order) is one of the possible main categories for analysis, no responses were found where sequential arrangement was the only way of arranging the data, so this category was thus dropped from the list of categories. Sequential arrangement however was found in combination with other arrangement strategies, namely sequential clustered and sequential summative, as well as sequential regrouped summative and were included as such. The eight categories of arrangement types were coded nominally in the following way:

0 = No arrangement: no attempt to arrange the data, leaving it as raw data, or copied data as given

1 = Inappropriate arrangement: statistically inappropriate arrangements of data

2 = Clustered arrangement: data sorted in groups with no totals

3 = Sequential clustered arrangement: data sorted in groups with no totals and either groups sorted in alphabetical or numerical order

4 = Summative arrangement: groups or categories provided with totals

5 = Sequential summative arrangement: groups or categories provided with totals and groups listed in alphabetical or numerical order

6 = Regrouped summative arrangement: data sorted into new groups or categories and totals provided

7 = Regrouped sequential summative arrangement: data sorted into new groups or categories and totals provided; groups sequentially sorted

As the data in Task 1 is categorical, data could not be sequentially arranged, sequentially clustered or regrouped and therefore none of the Task 1 responses could be categorised in arrangement categories 3, 6, and 7. Although the data in Task 2 could be arranged sequentially, no learner chose to only arrange data this way. The category of sequential arrangement was only retained in combination with other arrangement strategies. The analysis of responses in Task 1 and 2 will be considered next.

5.2.1 No response

In the first category, coded by a 0, no response was given. All learners responded to Task 1. For Task 2, 12 learners did not respond. All Grade 7 learners responded to this task, but 19,4% of the Grade 4's; 9,8% of Grade 5's and 2,3% of Grade 6's did not respond (Table 5.1). Learners' ability to interpret the task and/or their working speed seems to improve with age.

Grade	Number of no responses	% of no responses	% of all possible responses in grade
4	7	58	19
5	4	33	10
6	1	8	2
7	0	0	0
Total	12	≈100	8

Table 5.1: Analysis of no response across grades (T2)

There may be different reasons why learners did not give a response to Task 2. Some learners may have had difficulty interpreting the task and therefore may not have responded. Some learners in the lower grades engaged in very detailed pictures and a lot of decorating in Task 1 and may not have had time to give a response for Task 2, e.g. L27[1] (Gr 4) who produced two responses to Task 1 (Figure 5.1 and Figure 5.2), indicating that the first response (Figure 5.1) is wrong.

Figure 5.1: First attempt T1 (L27, Gr 4)

[1] Throughout this chapter references to a specific learner will be written with a capital L followed by the number assigned to the learner, e.g. Learner #27 will be referred to as L27. References in brackets to Grade will be abbreviated as Gr and references in brackets referring to Task 1 and Task 2 will be abbreviated as T1 and T2.

Figure 5.2: Second attempt T1 (L27, Gr 4)

This learner's responses are a good example of the modeling process. The learner's first response to, or modeling of the problem situation was to draw a little figure for each child, grouping those who like the same colour and indicating the total number of children liking each colour. Some of the children's names were added to their pictures. This constructed model was then modified to a refined arrangement where the children's names were omitted in the summative arrangement in which only the colours and totals for each category were given. This detailed and neat modeling response however probably caused the learner to run out of time for Task 2. According to the teacher this learner is of average or even of below average ability, but nevertheless invented a final representation that, although ikonic, is no longer on the intuitive level.

5.2.2 Inappropriate arrangement

An arrangement is judged to be inappropriate if it is not representative of the data, not suitable for the type of data or when summary statistics are computed that either do not represent the data or do not solve the task. The different inappropriate arrangements identified in the analysis and coded by a 1, are:

- Copying of the given list, although the task specified that the list may not be copied
- Pictures with personal comments
- Pictures with own data supplied
- Inappropriate clustering, e.g. by grade (T2)
- Inappropriate graphs, e.g. line graph with sequential arrangement of data, bar graphs with one bar for each child, bar graphs with clustering by grade
- Inappropriate statistical treatment: calculations of the sum of all feet lengths, range per grade, upper limit for each grade, mean.

For Task 1 the most inappropriate arrangements (Table 5.2) were produced by Grade 5 learners (67,4% of all T1 inappropriate responses). In Grade 4 there were 7 inappropriate responses (16,3% of all T1 inappropriate responses), in Grade 6 there were 6 (14,3% of all T1 inappropriate responses) and in Grade 7 only 1 (0,02% of all T1 inappropriate responses). It is not clear why almost 71% Grade 5 learners responded inappropriately to this task. The difficulty level and context of the task was judged appropriate for all the grades in the sample (see 4.5). Interviews with the Grade 4-7 teachers yielded that one of the reasons might be that the data collection was done in August and that they have not done any data handling activities yet during that year, as the teacher planned to do this in the fourth quarter. They may also not have been exposed to rich learning activities in Data Handling during their grade 4 year. The Grade 5 teacher at the school where the research was conducted each year involves learners in well-planned model-eliciting data handling activities. The Grade 6 and 7 learners therefore have been exposed to rich learning activities in this Learning Outcome in their Grade 5 year.

Grade	Number of inappropriate responses	% of inappropriate responses	% of all possible responses in grade
4	7	16	19
5	29	67	71
6	6	14	14
7	1	2	4
Total	43	≈100	30

Table 5.2: Analysis of inappropriate arrangement across grades (T1)

In 69,4% of all responses in Task 2, learners arranged data inappropriately (Table 5.3). A significantly high number of Grade 7 learners created inappropriate arrangements. Teachers are not well trained in Data Handling (Statistics) and usually choose data handling tasks comprising categorical data for classroom activities. Learners in this sample may not have dealt with numerical data often, if at all, and this fact could have contributed to the high number of learners who responded inappropriately to this task.

Grade	Number of inappropriate responses	% of inappropriate responses	% of all possible responses in grade
4	21	21	58
5	30	30	73
6	29	29	66
7	20	20	87
Total	100	100	69

Table 5.3: Analysis of inappropriate arrangement across grades (T2)

The category of inappropriate arrangement, coded by a 1, includes the following arrangements observed in the analysis:

- Copying of the given list, although the task specified that the list may not be copied.

- Pictures with personal comments (If you are in Gr 6 and wear 19 cm beach thongs like Peter, make a tick next to your name) - (T2, L73, Gr 5, Fig. 5.3):

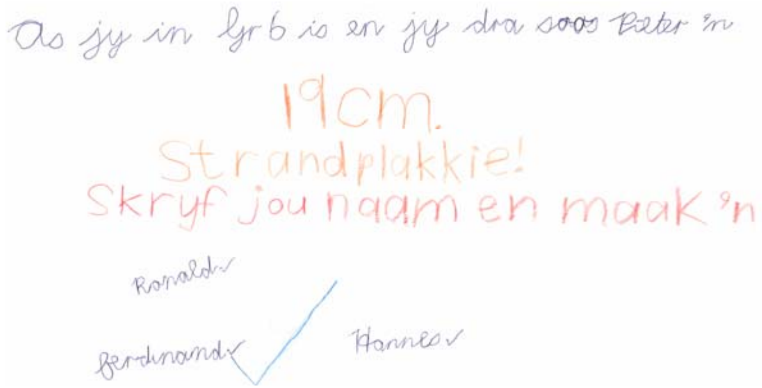


Figure 5.3: Inappropriate arrangement
(T2, L73, Gr 5)

- Pictures with own data supplied (T2, L59, Gr 5, Fig. 5.4)

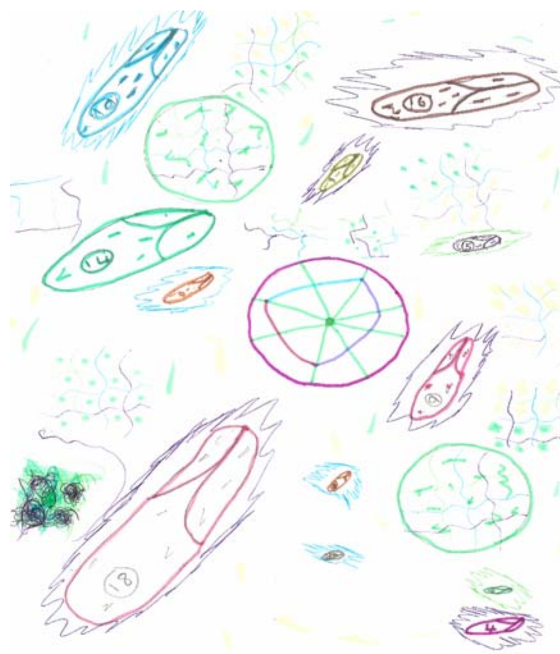


Figure 5.4: Inappropriate arrangement (T2, L59, Gr 5)

- Inappropriate clustering, including by grade (lists, tables, pictures)
- Inappropriate graphs (e.g. line graph with sequential arrangement of data: Task 2, L168, Grade 7, Fig. 5.5; bar graphs with one bar for each child; bar graphs with clustering by grade)

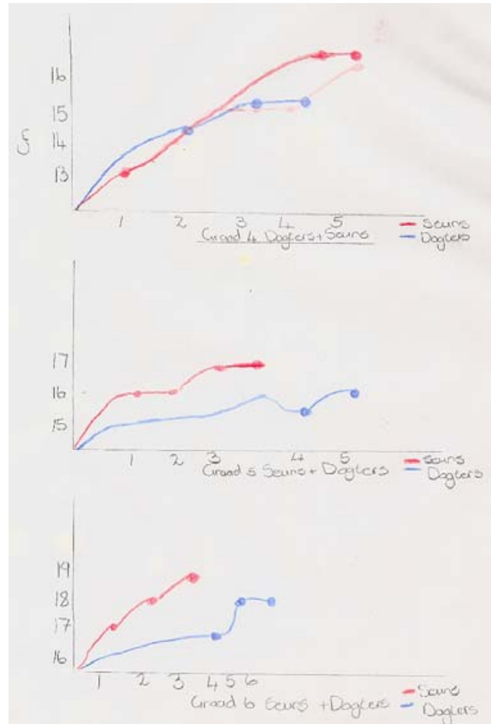
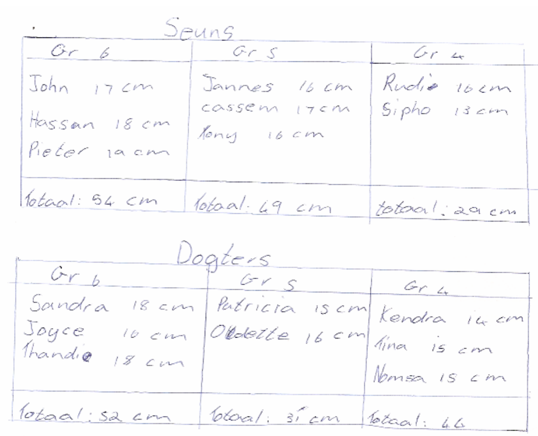


Figure 5.5: Inappropriate arrangement
(T2, L168, Gr 7)

- Inappropriate statistical treatment: calculations of the sum of all feet lengths (T2, L81, Gr 5, Fig. 5.6); range per grade (T2, L150, Gr 7, Fig. 5.7); upper limit for each grade; mean (T2, L152, Gr 7, Fig. 5.8); mode(T2, L161, Gr 7, Fig. 5.9).



Seuns

Gr 6	Gr 5	Gr 4
John 17 cm Hassan 18 cm Pieter 19 cm	Jannes 16 cm cassem 17 cm Tony 16 cm	Rudie 16 cm Sipho 13 cm
Totaal: 54 cm	Totaal: 49 cm	totaal: 29 cm

Dogters

Gr 6	Gr 5	Gr 4
Sandra 18 cm Joyce 16 cm Thandie 18 cm	Patricia 15 cm Odelette 16 cm	Kendra 14 cm Tina 15 cm Nomsa 15 cm
Totaal: 52 cm	Totaal: 31 cm	Totaal: 44

Figure 5.6: Inappropriate arrangement
(T 2, L81, Gr 5)

Figure 5.7: Inappropriate arrangement
(T 2, L150, Gr 7)

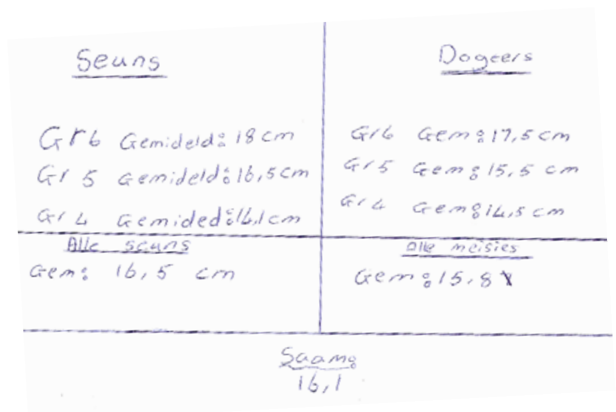


Figure 5.8: Inappropriate arrangement
(T2,L152, Gr 7)

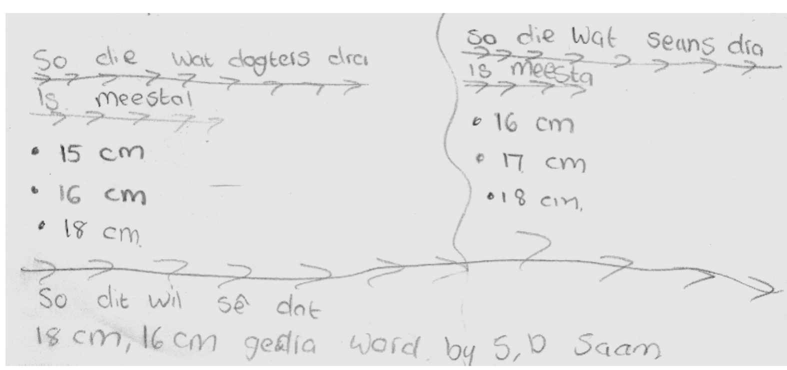


Figure 5.9: Inappropriate arrangement
(T2, L161, Gr 7)

5.2.3 Sequential arrangement

There was only one sequential response for L57, T1, Gr 5 (Figure 5.10). The learner indicated that the children's names were arranged in alphabetical order by writing "Alfabetiese volgorde" at the left hand top of the list but then omitted the names and only wrote down the colours with a little drawing next to each of the colours. The arrangement was judged to be inappropriate and was coded with a 1 and not as a sequential arrangement per se. As mentioned in the introduction of Data Arrangement (5.2), as no response showed sequential arrangement per se, it was not included in the list of categories.

Figure 5.10: Sequential arrangement
(T 1, L57, Gr 5)

For Task 2, a number of learners arranged the feet lengths sequentially, but because they have also clustered the data without giving totals, their arrangements were coded in the category for sequential clustered responses.

As a result, no responses in either task were categorised as sequential responses.

5.2.4 Clustered arrangement

In this category, coded by a 2, responses were clustered, but no totals given. In Task 1, clustering was done according to favourite colours. In Task 2, clustering was done according to different variables: grade, gender or feet length, or by combinations of variables such as grade and gender or feet length and gender. Clustering according to grade or gender was categorised as inappropriate, while clustering according to feet length was classified as appropriate. An important difference between clustered arrangement and summative arrangement is that summative arrangements, besides being clustered, also have totals given for each group. Clustered arrangements were evident in representations such as pictures (T2, L68, Gr 5, Fig. 5.11), lists (T1, L157, Gr 7, Fig. 5.12); frequency tables (T2, L114, Gr 6, Fig. 5.13); pictograms T1, L77, Gr 5, Fig. 5.14) and pie graphs (T1, L150, Gr 7, Fig. 5.15).

Figure 5.11: Clustered arrangement (T2, L68, Gr 5)

Figure 5.12: Clustered arrangement (T1, L157, Gr 7)

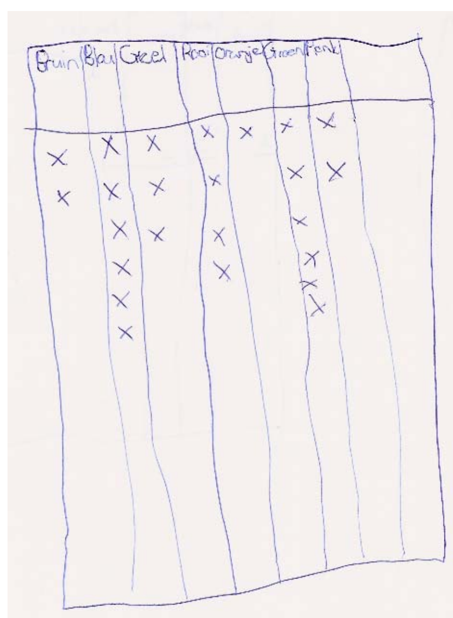


Figure 5.13: Clustered arrangement (T2, L114, Gr 6)

Figure 5.14: Clustered arrangement
(T1, L77, Gr 5)

Figure 5.15:
Clustered arrangement
(T1, L150, Gr 7)

Clustered arrangement was the second most popular arrangement type for appropriate arrangements in Task 1. Almost 31% of Gr 4's (Table 5.4) chose to cluster the data in this task, while clustered arrangement was not a preferred approach for the other grades (12,2% for Gr 5's, 15,9% for Gr 6's and 21,7% for Gr 7's). For Task 2 no clustered arrangements were found for Grade 4 and 7. The Grade 5's produced two thirds of the clustered arrangements, the other third by the Grade 6's.

Grade	Number of clustered arrangements	% of clustered arrangements	% of all possible responses in grade
4	11	39	31
5	5	18	12
6	7	25	16
7	5	18	22
Total	28	100	19

Table 5.4: Analysis of clustered arrangement across grades (T1)

Grade	Number of clustered arrangements	% of clustered arrangements	% of all possible responses in grade
4	0	0	0
5	2	67	5
6	1	33	2
7	0	0	0
Total	3	100	2

Table 5.5: Analysis of clustered arrangement across grades (T2)

5.2.5 Sequential clustered arrangement

Responses in this category were coded by a 3. There were no responses in Task 1 in this category, as categorical data can not be sequentially clustered. In Task

2 no Grade 4's used sequential clustering while an equal number of learners in grade 5, 6, and 7 chose to arrange the data in this way (Table 5.6).

Grade	Number of sequential clustered arrangements	% of sequential clustered arrangements	% of all possible responses in grade
4	0	0	0
5	2	33	5
6	2	33	5
7	2	33	9
Total	6	≈100	4

Table 5.6: Analysis of sequential clustered arrangement across grades (T2)

The six learners who chose to arrange the data of Task 2 clustered sequentially without giving totals responded with lists, a frequency table and pictograms. The pictograms differ as ikonic support is evident in some of the arrangements (T2, L148, Gr 7, Fig. 5.16), while others are more abstract arrangements, like the one in Figure 5.17, bordering on a bar graph (T2, L89, Gr 5).

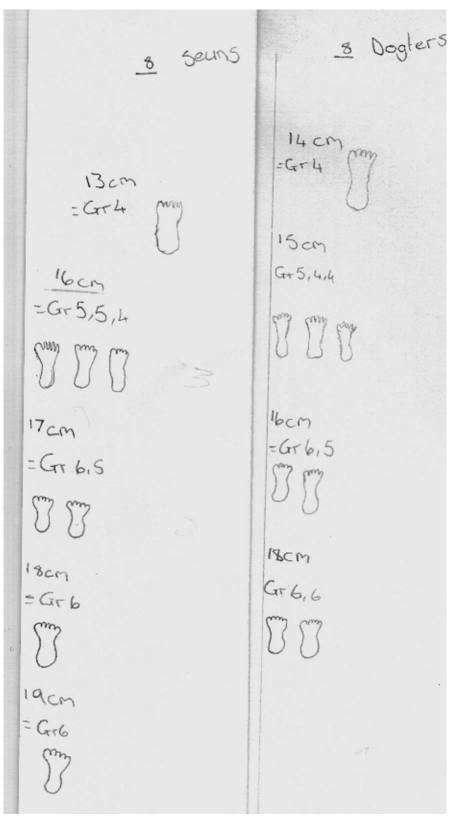


Figure 5.16: Sequential clustered
(T2, L148, Gr 7)

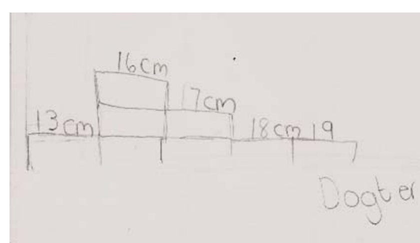


Figure 5.17: Sequential clustered
(T2, L89, Gr 5)

5.2.6 Summative arrangement

Summative arrangement was coded by a 4. There were no responses for Task 2 that were categorised as appropriate summative arrangements. There was only one inappropriate summative arrangement for Task 2 (L142, Gr 6) where the learner gave the total number of learners in each grade, adding a bar graph with one bar for each child. Of 144 responses for Task 1 46,5% were categorised as summative arrangements where the data were grouped and totals for each group given (Table 5.7). These responses included pictures (T1, L13, Gr 4, Fig. 5.18), lists (T1, L106, Gr 6, Fig. 5.19), tables (T1, L7, Gr 4, Fig. 5.20), pictograms (T1, L118, Gr 6, Fig. 5.21), frequency tables (T1, L133, Gr 6, Fig. 5.22), and bar graphs (T1, L125, Gr 6, Fig. 5.23).

Figure 5.18: Summative arrangement (T1, L13, Gr 4)

Figure 5.19: Summative arrangement
(T1, L106, Gr 6)

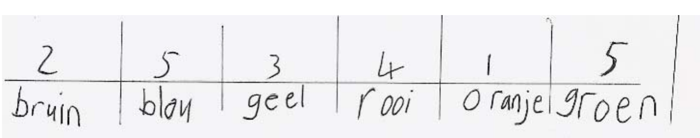


Figure 5.20: Summative arrangement (T1, L 7, Gr 4)

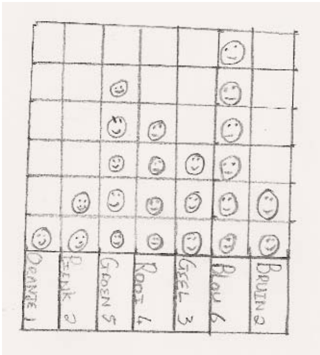


Figure 5.21: Summative arrangement
(T1, L118, Gr 6)

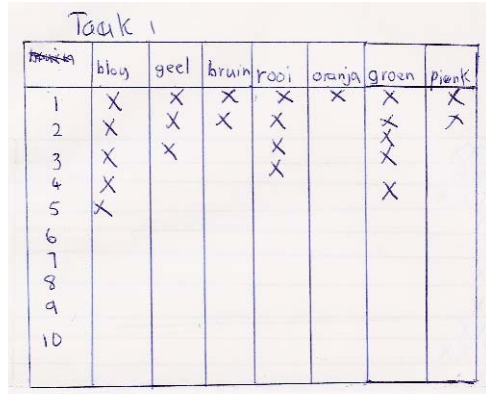


Figure 5.22: Summative arrangement
(T1, L133, Gr 6)

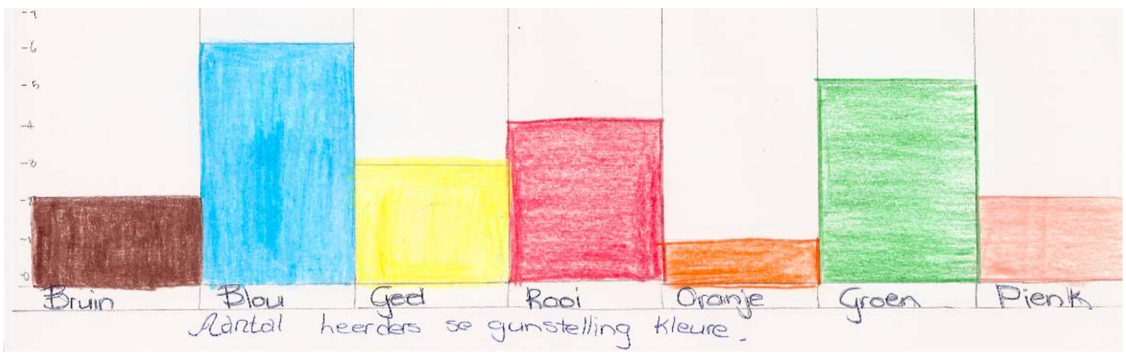


Figure 5.23: Summative arrangement (T1, L125, Gr 6)

As mentioned before, a summative arrangement of the data in Task 1 is considered the appropriate response in the specific context. 43,8% of all learners responded to Task 1 with a summative arrangement (Table 5.7). Summative arrangement was a more popular way of dealing with the data amongst Grade 6's than in other grades. A significant number of Grade 7's chose to arrange the data in this way, which may point to a better understanding of how to appropriately deal with the data than the earlier grades.

Grade	Number of summative arrangements	% of summative arrangements	% of all possible responses in grade
4	16	25	44
5	7	11	17
6	24	38	55
7	16	25	70
Total	63	≈100	44

Table 5.7: Analysis of summative arrangement across grades (T1)

5.2.7 Sequential summative arrangement

In this category, coded by a 5, data were sorted in groups with totals, and groups were listed in numerical order of totals. Such arrangements were only found for Task 1 and resulted in lists (T1, L26, Gr 4, Fig. 5.24), tables (T1, L127, Gr 6, Fig. 5.25), pie graphs (T1, L170, Gr 7, Fig. 5.26) and bar graphs (T1, L128, Gr 6, Fig. 5.27). The sequential summative arrangements for Task 2 were also regrouped and thus not coded in this category.

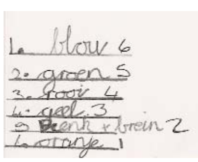


Figure 5.24: Sequential summative
(T1, L26, Gr 4)

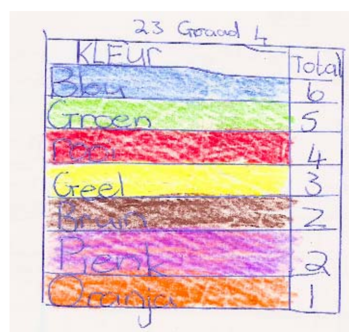


Figure 5.25: Sequential summative
(T 1, L127, Gr 6)

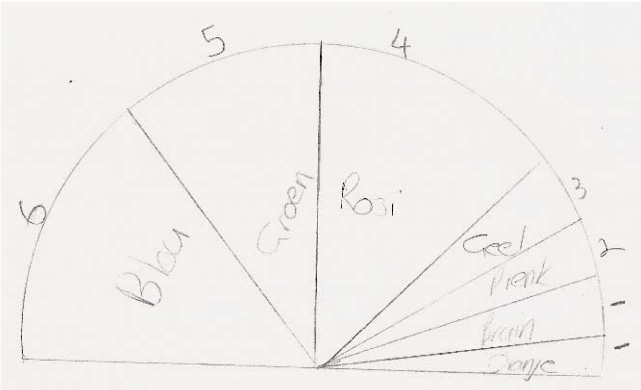


Figure 5.26: Sequential summative (T1, L170, Gr 7)

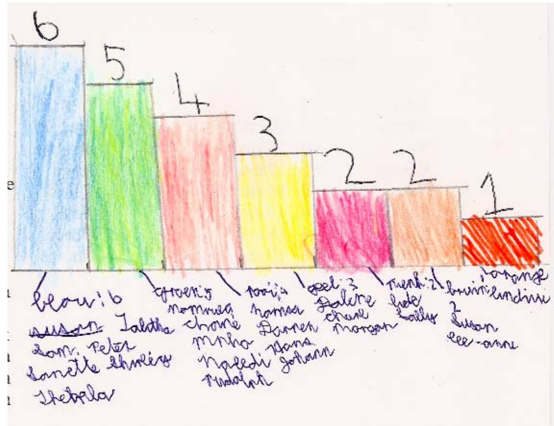


Figure 5.27: Sequential summative (T1, L128, Gr 6)

Sequential summative arrangement was not a popular way of arranging data. No Grade 5’s used this way of arrangement, while only 6,9% of all learners arranged the data of Task 1 in this way (Table 5.8). While 63 learners used summative arrangement for Task 1, only 10 learners arranged the data in a sequential summative way according to the totals.

Grade	**Number of sequential summative arrangements**	**% of sequential summative arrangements**	**% of all possible responses in grade**
4	2	20	6
5	0	0	0
6	7	70	16
7	1	10	4
Total	10	100	7

Table 5.8: Analysis of sequential summative arrangement across grades (T1)

5.2.8 Regrouped summative arrangement

Regrouped summative arrangement was coded by a 6. The categorical data in Task 1 could not be regrouped when arranged; therefore all 15 responses in this category were from Task 2. 11,4% of all responses in Task 2 were classified in this category (Table 5.9). A surprising number of Grade 4's, the same number as Grade 6's, showed insight in the arrangement of the data in this task, regrouping it summatively.

Grade	Number of regrouped summative arrangements	% of regrouped summative arrangements	% of all possible responses in grade
4	6	40	17
5	2	13	5
6	6	40	14
7	1	7	4
Total	15	100	10

Table 5.9: Analysis of sequential summative arrangement across grades (T1)

The data in Task 2 were given grouped according to gender and were regrouped according to feet length by learners in the form of pictures (T2, L110, Gr 6, Fig. 5.28), lists (T2, L17, Gr 4, Fig. 5.29), tables (T2, L8, Gr 4, Fig. 5.30) and bar graphs (T2, L127, Gr 6, Fig. 5.31). Responses grouped according to grade or gender were regarded as inappropriate and coded not with a 6 but with a 1.

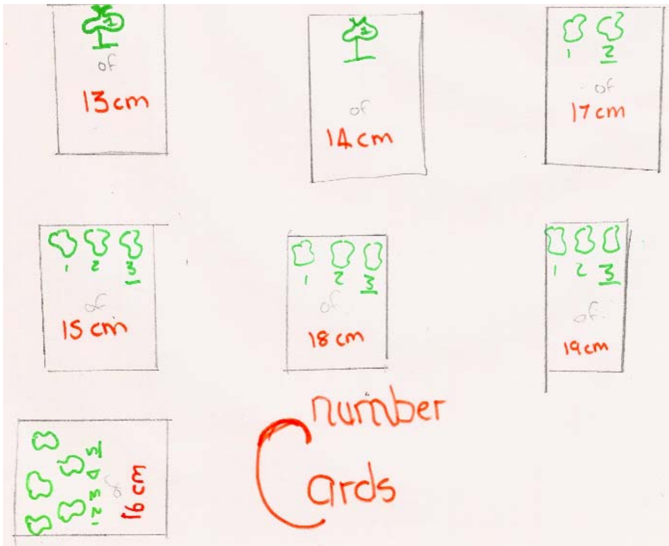


Figure 5.28: Regrouped summative arrangement (T2, L110, Gr 6)

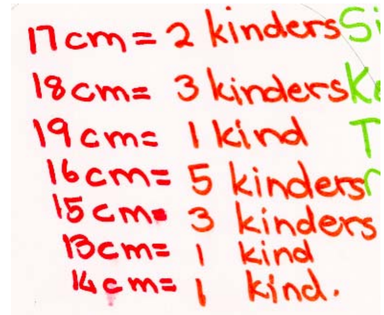


Figure 5.29: Regrouped summative arrangement (T2, L17, Gr 4)

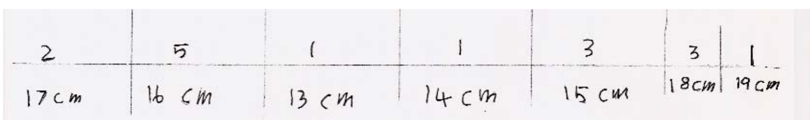


Figure 5.30: Regrouped summative arrangement (T2, L8, Gr 4)

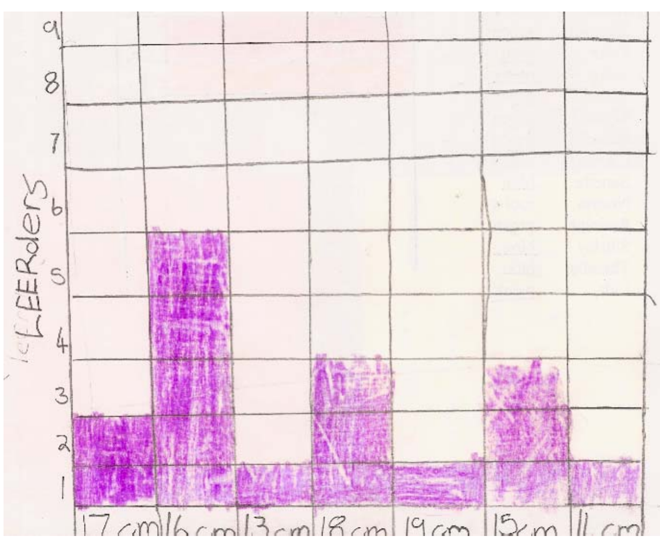


Figure 5.31: Regrouped summative arrangement (T2, L127, Gr 6)

5.2.9 Regrouped sequential summative arrangement

The data in Task 1 could not be regrouped as it was categorical data. Eight learners which is 5,6% of all possible responses for Task 2, not only regrouped the data summatively, but also gave it in sequential order according to the totals of each group (Table 5.10).

Grade	**Number of sequential regrouped summative arrangements**	**% of sequential regrouped summative arrangements**	**% of all possible responses in grade**
4	2	25	6
5	1	13	2
6	5	63	11
7	0	0	0
Total	8	≈100	6

Table 5.10: Analysis of sequential summative arrangement across grades (T2)

These responses were coded with a 7 and realised in lists (T2, L16, Gr 4, Fig. 5.32), tables (T2, L70, Gr 5, Fig. 5.33) and bar graphs (T2, L122, Gr 6, Fig. 5.34).

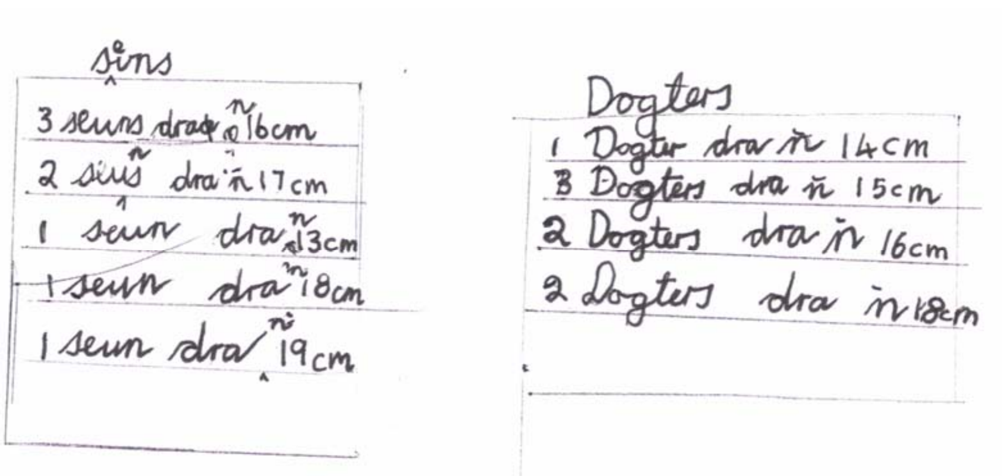
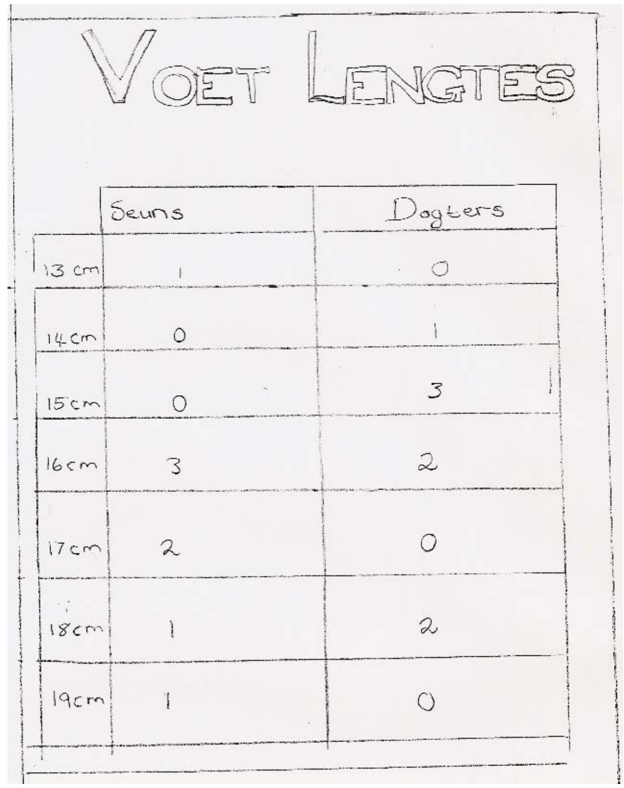


Figure 5.32: Regrouped sequential summative arrangement (T2, L16, Gr 4)



Voet Lengtes

	Seuns	Dogters
13 cm	1	0
14cm	0	1
15cm	0	3
16cm	3	2
17cm	2	0
18cm	1	2
19cm	1	0

Figure 5.33:
Regrouped sequential summative arrangement (T2, L70, Gr 5)

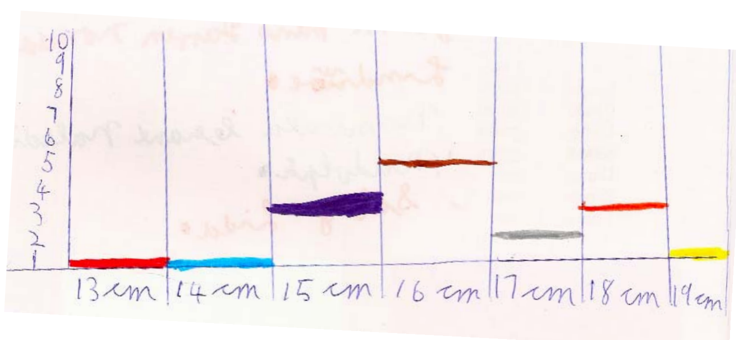


Figure 5.34: Regrouped sequential summative arrangement (T2, L122, Gr 6)

5.2.10 Summary of Data Arrangement Types

Arrangement is a key transnumerative process prior to the summarising and representation of data (see 3.6.2). Learners had to make sense of the given data and had to transform the data appropriately in order to create an effective representation.

Almost half of all responses in the two tasks were inappropriate, with 29,9% of Task 1 responses inappropriate and 69,4% of Task 2 responses inappropriate (Table 5.11). Learners clearly had difficulty to interpret the numerical task. Their unfamiliarity with such tasks may be one of the contributing factors. Of the three possible appropriate arrangement types for Task 1, namely clustered, summative and sequential summative, 43,8% of learners chose to arrange the data summatively; 19,4% clustered the data and 6,9% arranged the data sequentially according to the totals of the groups (sequential summative). For Task 2 there are seven appropriate ways in which data could be arranged. Of these seven, the category of regrouped summative arrangement was used by most learners who produced appropriate responses (10,4%). Sequential regrouped summative arrangements were the second most popular way of arranging the data (5,6% of learners). Clustered arrangement (2,1%) and sequential clustered arrangement (4,2%) were the other two arrangement types used.

The unexpectedly large bulge of inappropriate responses may be due to the fact that learners were not shown any examples of representations and were not given any hints (also see 5.3.10) because the aim was to elicit spontaneous representations.

The analysis of arrangement strategies showed that learners produced all kinds of different representations for each kind of arrangement type.

Task	Grade	No response	Inappropriate	Clustered	Sequential clustered	Summative	Sequential summative	Regrouped summative	Sequential regrouped summative	Total
		0	1	2	3	4	5	6	7	
1	4	0	7	11	-	16	2	-	-	**36**
	5	0	29	5	-	7	0	-	-	**41**
	6	0	6	7	-	24	7	-	-	**44**
	7	0	1	5	-	16	1	-	-	**23**
	Total	**0**	**43**	**28**	**-**	**63**	**10**	**-**	**-**	**144**
2	4	7	21	0	0	0	0	6	2	**36**
	5	4	30	2	2	0	0	2	1	**41**
	6	1	29	1	2	0	0	6	5	**44**
	7	0	20	0	2	0	0	1	0	**23**
	Total	**12**	**100**	**3**	**6**	**0**	**0**	**15**	**8**	**144**

Table 5.11: Summary of arrangement types

The refined framework for categorising data arrangement types that was designed to accommodate all the different combinations of the broad arrangement categories, proved an adequate tool to shed light on learners' strategies to deal with data in the representation process. A relationship between the arrangement strategy and the SOLO level of a response became evident as will be discussed in 5.5. As mentioned in paragraph 3.6.4, learners have to have a sense of the message the data are conveying to be able to appropriately arrange and display the data. Three of the four linked processes of transnumeration are especially entwined, namely identifying the message in the data, choice of representation and the process of transforming the data. The ultimate success of the representation is dependent on the first of the abovementioned processes. The second phase in the linked process of transnumeration, namely representation, will next be considered.

5.3 REPRESENTATION OF DATA

The second aspect that was investigated is how learners spontaneously represent data. In the process of representation, the arrangement and representation end product are inseparable parts of the process.

Although the arrangement type and representational type were analysed in a seemingly disconnected way, this approach was chosen because looking at both at the same time, compounds the analysis and makes it difficult to appreciate all the different aspects of each type. Interesting observations could be made by analysing the two aspects separately. One would for example expect that certain kinds of arrangement would result in a specific representation, but contrary to expectations, the different arrangement types each yielded a number of different types of representation. Clustered arrangement for example resulted in all the identified types in the list of representations in the analysis, namely different kinds of pictures, lists, tables, pictograms, frequency tables, bar graphs and pie graphs. Some learners produced unsophisticated or idiosyncratic representations for a lack of exposure to more advanced representational types and the lack of statistical tools to display the data in a more useful way. More sophisticated types of arrangement therefore did not necessarily result in sophisticated representational types. The nature of the tasks also limited the kind of possible representations.

As Task 1 comprises categorical data, a summative arrangement is considered to be the most appropriate way of arranging the data. As discussed in 4.7.1, a summative arrangement is a grouped arrangement in which totals for each group are given. The data in Task 2 is numerical and requires different treatment. Regrouping the data according to feet length is essential to make sense of the data. As the task states that the data must be displayed for manufacturers of beach sandals, an appropriate way would be to regroup data according to feet length and display the data in a bar graph. Feet lengths should be shown sequentially so that an observer could get a good idea of the spread.

Whereas arrangement types were categorised according to their statistical appropriateness, representational types were classified according to the type of representation, even if inappropriate. The number of appropriate and inappropriate representations in each category is however given and examples of each type discussed. It is the arrangement tenet that determines whether a representation makes sense or not, but as the data are communicated via the representation, the choice of representation is crucial.

Some learners produced more than one representation as they modeled the task, in most cases ending with a more sophisticated response than their first attempt (see 5.2.1 and 5.4.6). This is in line with the idea of modeling in which vertical mathematisation of the problem takes place (see 2.2). The more sophisticated representation of the set was coded, which in most cases was the refined or adjusted product of the modeling process.

Each representational type has characteristics unique to that type, but it is important to realise that some representations are transitional in nature. These representations exhibit some – but not all – characteristics of the next more sophisticated representational type, and can therefore not yet be classified as the following type. Where applicable, examples of transitional responses are given.

The six expected representational types (see 4.7.1.2) were

1 = lists
2 = tables

3 = pictograms and frequency tables

4 = bar graphs

5 = pie charts

6 = line graphs

Three more types had however been added to provide for the whole spectrum of learner responses. When a learner did not give a response or left the response incomplete in such a way that the intent is not clear, the response was coded with a 0 on the nominal scale. Many responses displayed a predominantly ikonic character and were categorised as pictures. This added category was coded with a 1. Responses that did not display characteristics to fit into one of the discussed eight categories were categorised as anomalous responses and coded with an 8.

The comprehensive coding structure for representational types includes 9 different categories, coded from 0 to 8 on a nominal scale:

0 = no representation

1 = pictures/shapes/names/numbers

2 = lists

3 = tables

4 = pictograms and frequency tables

5 = bar graphs

6 = pie charts

7 = line graphs

8 = anomalous representation

5.3.1 No representation

As in the case of data arrangement, all learners responded to Task 1 while 11 learners did not respond to Task 2. Four of these were incomplete responses in which one can not be sure what the learner intended to do. The percentage of learners not responding to Task 2 is given/expounded in the following table:

Grade	Number of no responses	% of no responses	% of all possible responses in grade
4	7	64	19
5	3	27	7
6	1	9	2
7	0	0	0
Total	11	100	8

Table 5.12: Analysis of no response across grades (T2)

From the table it is clear that most learners who did not respond, were in Grade 4, and that the number of no responses decreased with age. This is an indication that older learners work faster and/or have more insight into the task and were therefore able to produce a representation.

5.3.2 Pictures (shapes/names/numbers represented pictorially)

Representations in which the ikonic element dominates are regarded as pictures. The first impression of such a representation is of a picture, with data given in shapes, words or numbers. These representations differ from others where ikonic support is visible, but the overall impression is not of a picture, but of some other representational type. L119 (Gr 6, Figure 5.35) for example, represented the data of T1 in a bar graph, filling the bars with different drawings and decorating the space above the bars. Likewise, L144 (Gr 6, Figure 3.36) used ikonic support in representing the data of T2 in a bar graph.

Figure 5.35: Bar graph with ikonic support (T1, L119, Gr 6)

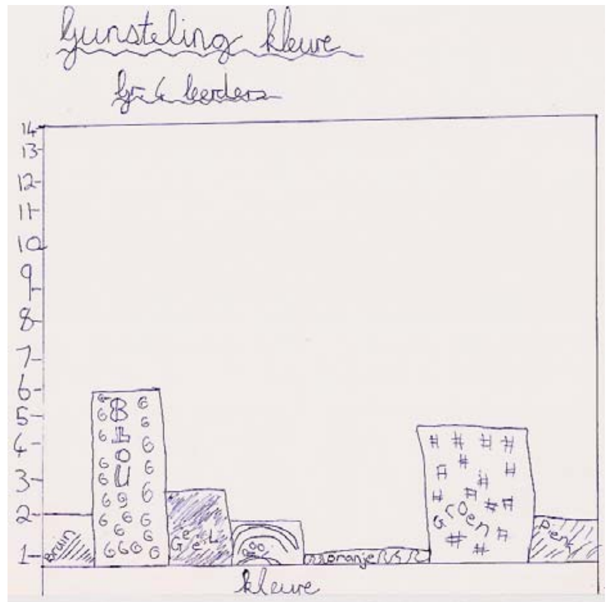
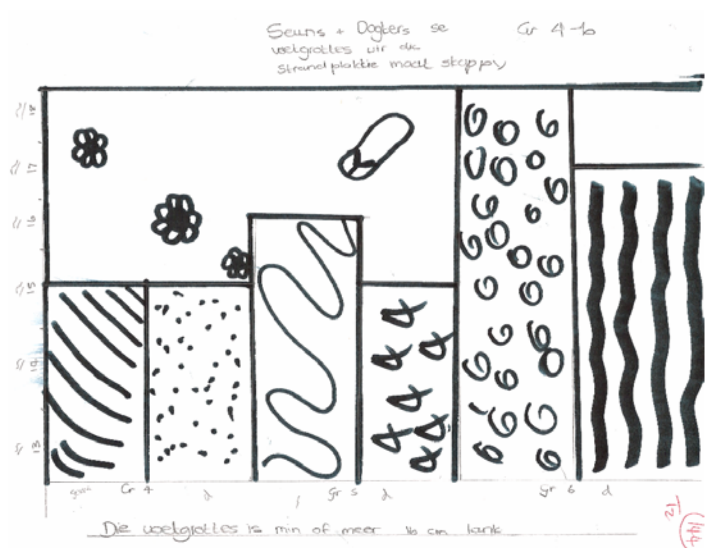


Figure 5.36: Bar graph with ikonic support (T2, L144, Gr 6)



The pictorial responses were found to belong to five different groups namely

- random pictures
- words, names or numbers randomly represented
- names and numbers or words paired off , which includes the 1-on-1 pairing off or grouping of data
- lists presented in pictures

A distinct difference could be found between random pictures and names, numbers and/or words represented randomly. The last group of responses does not have pictures; but names, words and numbers are represented randomly across the page, giving an impression of a picture rather than a list or other representation. In the next group, names, words and numbers are not represented randomly, but ordered either with

- the pairing off of one list with another, or
- a one-on-one pairing off of data or
- a one-more pairing off of or grouping of data.

Examples of data represented in random pictures (Fig. 5.37, T1, L68, Gr 5; Fig. 5.38, T1, L54, Gr 5; and Fig. 5.39, T2, L72, Gr 5) show the ikonic nature of the responses.

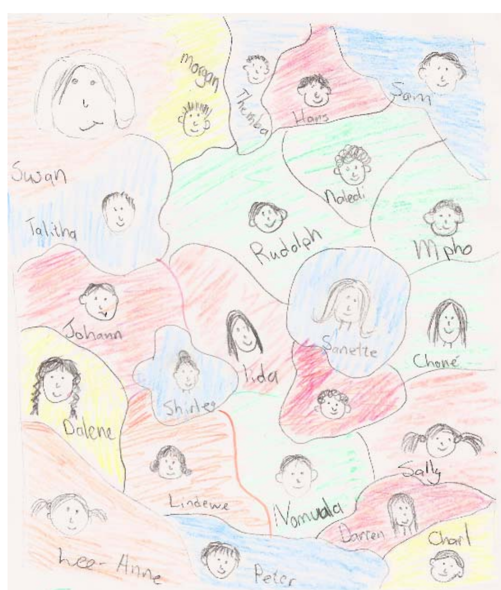


Figure 5. 37: Random pictures
(T1, L68, Gr 5)

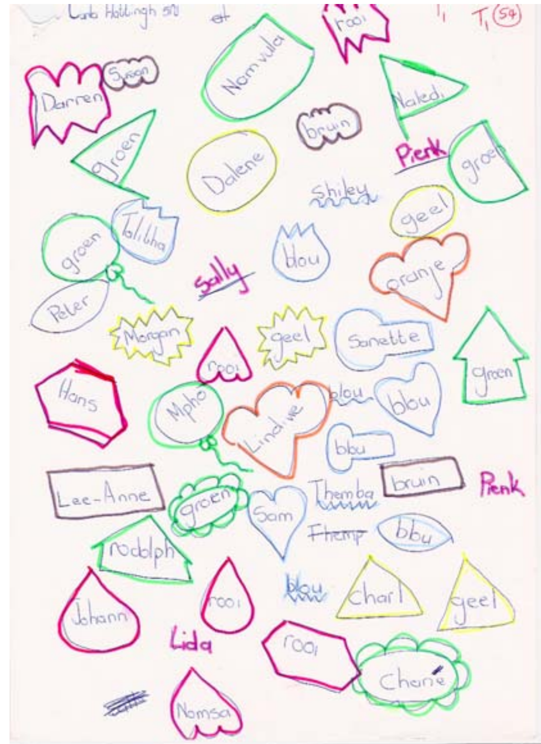


Figure 5.38: Random pictures
(T1, L54, Gr 5)

Figure 5.39: Random pictures
(T1, L72, Gr 5)

Figure 5.40: Random words and numbers (T1, L113, Gr 6)

Figure 5.40 shows the random representation of words and numbers. Lists, as described in 5.3.3 is an ordered writing down of data points, usually underneath each other. The representation in Fig. 5.40 can therefore not be regarded as a list, as the data are not written down ordered underneath each other. The representation resembles a picture, hence it is categorised as random representation in the category 'pictures'.

Responses in the group that comprise the pairing off of data, display ikonic characteristics. In some cases data points in two lists are paired off (Fig. 5.41, T1, L110, Gr 6), in other cases data points are randomly drawn and paired off on a 1-on-1 basis (Fig. 5.42, T1, L60, Gr 5). The first impression of these representations is of pictures.

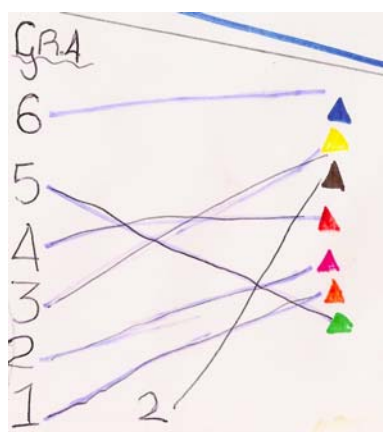


Figure 5.41: Numbers and pictures paired off (T1, L110, Gr 6)

Figure 5.42: One-on-one pairing off (T1, L60, Gr 5)

In still other cases, data are sorted into a table-like representation, but the ikonic features dominate and the representation does not satisfy the conditions of a table therefore the response is categorised as a picture with data grouped in a one-on-one way (Fig. 5.43, T1, L18, Gr 4).

Figure 5.43: One-on-one paring off (T1, L18, Gr 4)

The last group in the category of ikonic representations is the ‘picture lists’. As discussed in 3.6.5 a list is a number of connected names, numbers or items, written or printed together, usually consecutively. The representations in this group contain lists, but the ikonic aspect is so strong that the first impression of the representation is that of a picture and not primarily of a list (Fig. 5.44, T1, L19, Gr 4; Fig. 5.45, T2, L126, Gr 6).

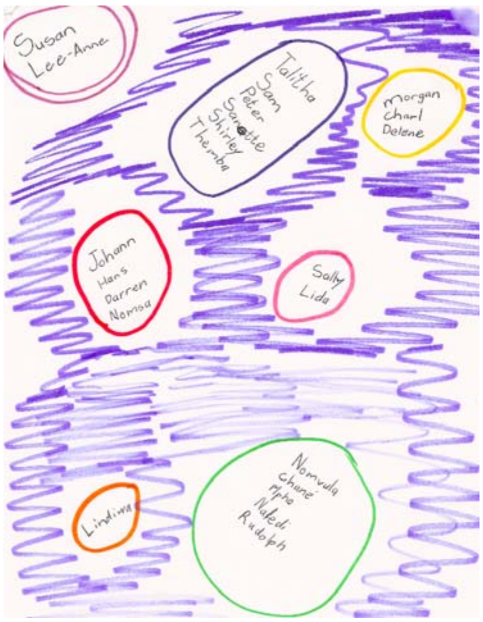


Figure 5.44: Picture list (T1, L19, Gr 4)

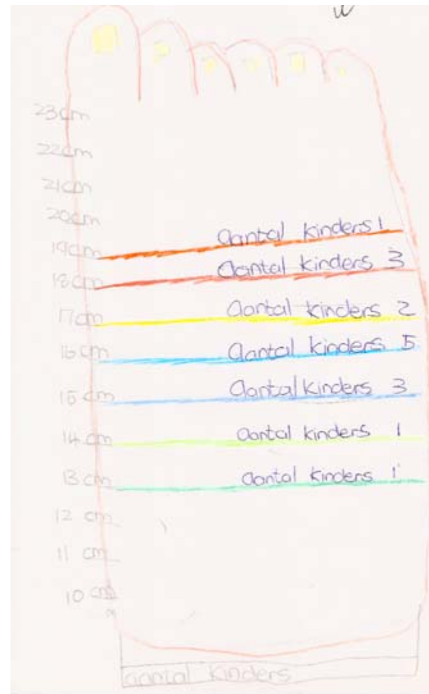


Figure 5.45: Picture list (T2, L126,Gr 6)

Figure 5.46: Picture list (T1, L31, Gr 4)

When looking at the table summarising the responses in representational type, it is clear that ikonic displays were favoured more by younger learners. The categorical data in Task 1 and the context of the task may have contributed to the fact that 36% of Task 1 responses in comparison with 14% of Task 2 responses were pictorial representations. 20% of all possible responses in both tasks were pictorial responses.

Grade	Number of pictures	% of pictures	% of all possible responses in grade
4	21	40	58
5	22	42	54
6	5	10	11
7	4	8	17
Total	52	100	36

Table 5.13: Analysis of representation (pictures) across grades (T1)

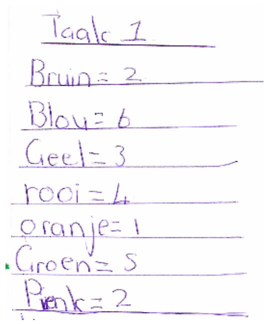
Grade	Number of pictures	% of pictures	% of all possible responses in grade
4	11	55	31
5	6	30	15
6	2	10	5
7	1	5	4
Total	20	100	14

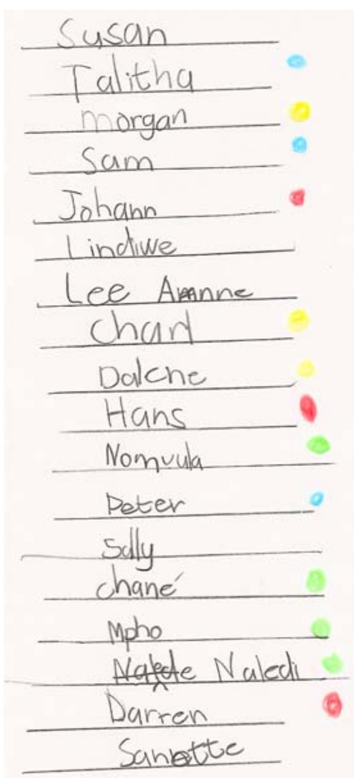
Table 5.14: Analysis of representation (pictures) across grades (T2)

5.3.3 Lists

The lists in this category are distinguishable from 'Picture lists' in the previous one because of the more abstract way in which they are displayed. 'Picture lists' with their strong ikonic component resembles pictures, while the lists in this category do not look like pictures, even though ikonic support and colour was used in some of the displays. The data in the category “Lists” are organised underneath each other. L25 (Fig. 5.47, T1, Gr 4) produced a summative list in a more symbolic abstract form than those in the previous category. Other examples of lists are L122 (Fig. 5.48, T1); L70 (Fig. 5.49, T1) and L75 (Fig. 5.50, T2).

Fig. 5.47: List (T1, L25, Gr 4)





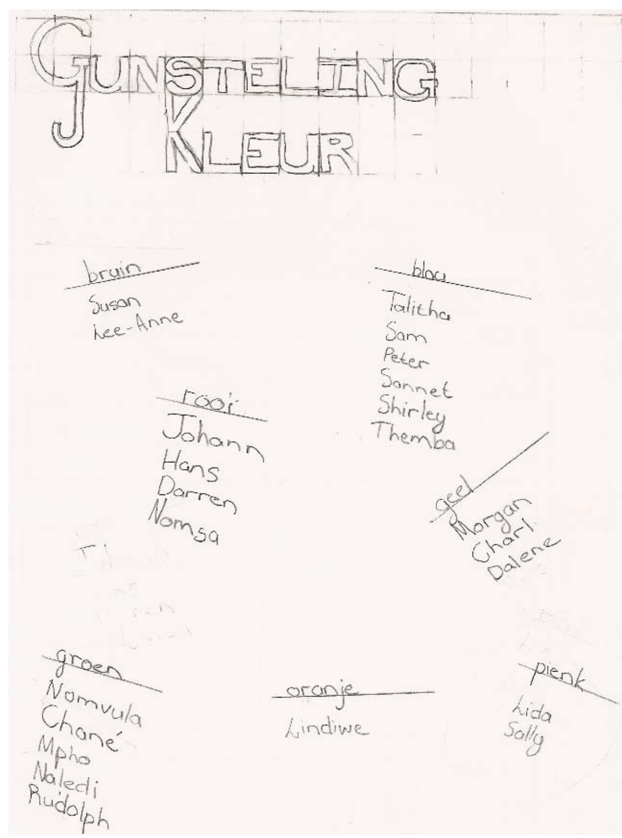


Fig. 5.48: List (T1, L20, Gr 4)

Fig. 5.49: List (T1, L70, Gr 5)

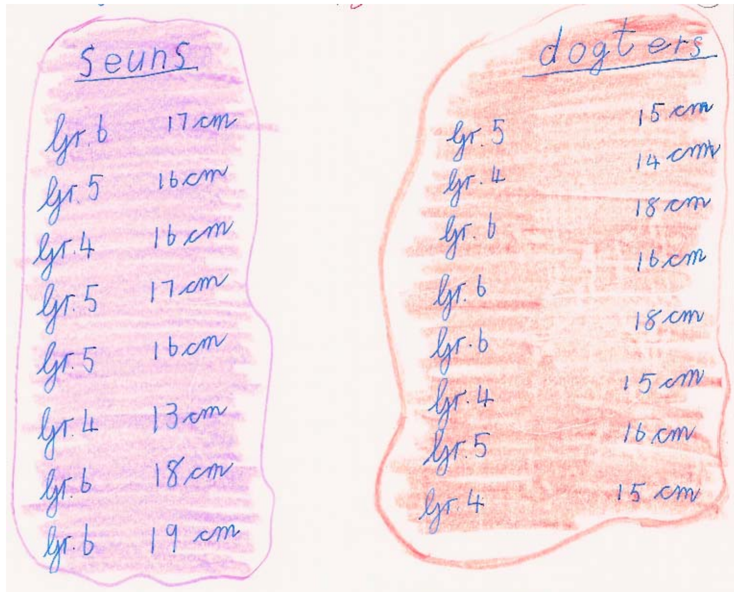


Fig. 5.50: List (T2, L75, Gr 5)

The display in Fig. 5.51 (T1, L20, Gr 4) is judged to be a list because the data are sorted in consecutive order, although not underneath each other.

Fig. 5.51: List (T1, L122, Gr 6)

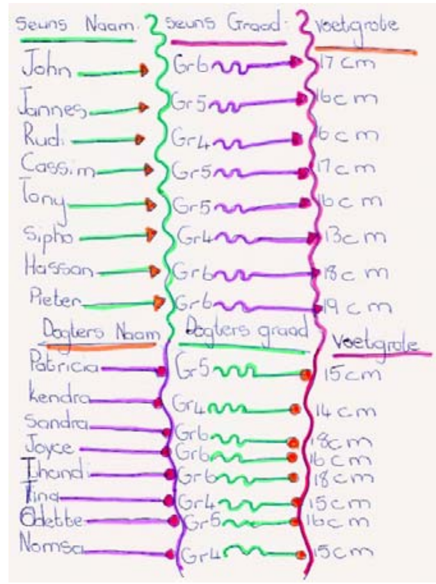


Figure 5.52: List (T2, L54, Gr 5)

From Tables 5.15 and 5.16 it is clear that more learners in Grade 5 than in other grades in both tasks chose to represent the data in lists (38% for T1 and 32% for T2). More than 20% of Grade 4's and 5's used lists for Task 1. The number of responses in which lists were used, are significantly high for Task 2 in all grades. In Grade 7 more than half of all responses for Task 2 were lists. Many learners copied the given list even though the task explicitly stated that it may not be copied. As no interviews were conducted, there can only be speculated about the reason for the popularity of lists as representational type. Some learners may not have been able to decide what message from the data they want to convey or may not have known how to convey the message and therefore just copied the given list (Fig. 5.52, T2, L54, Gr 5). L54 (Figure 5.52, T2) copied the list in an ikonic way.

Grade	Number of lists	% of lists	% of all possible responses in grade
4	8	28	22
5	11	38	27
6	7	24	16
7	3	10	13
Total	29	100	20

Table 5.15: Analysis of representation (lists) across grades (T1)

Grade	Number of lists	% of lists	% of all possible responses in grade
4	13	23	36
5	18	32	44
6	14	25	32
7	12	21	52
Total	57	≈100	40

Table 5.16: Analysis of representation (lists) across grades (T2)

5.3.4 Tables

A table is defined as “an arrangement of numerals, letters or signs, usually in rows and columns, to show facts or relationships between them in compact form” (Bedick & Levin 1973:190). Only few representations satisfied these requirements for a table. Some lists were transitional in nature, displaying columns with names and/or numbers ordered underneath each other, but not yet a horizontal and vertical association as required for a table (Fig. 5.53, T2, L167, Gr 5). Colour played an important role in 4 of the 7 tables for T1, contrary to the 3 out of 16 tables for T2. The context of the first task, being about the favourite colours of learners, may have contributed to this phenomenon.

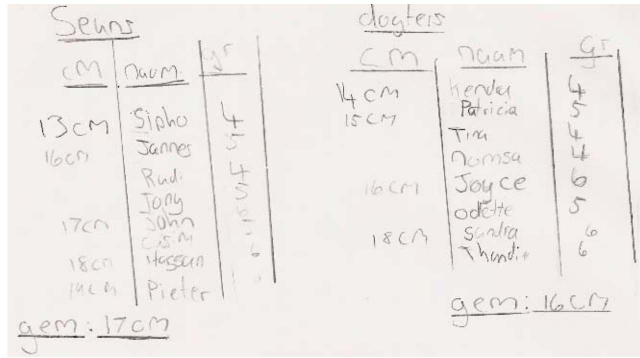


Figure 5.53: Transitional list (T2, L167, Gr 7)

L7 (Figure 5.54, T1, Gr 4) and L131 (Figure 5.55, T2, Gr 6) used displays that satisfy the conditions for a table, having rows and columns with a direct relationship.

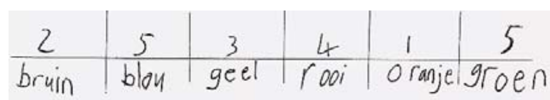


Figure 5.54: Table (T1, L7, Gr 4)

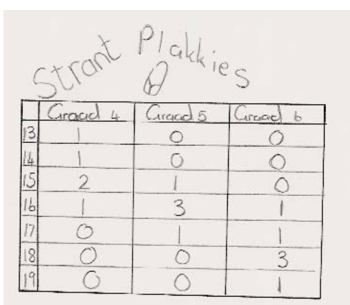


Figure 5.55: Table (T2, L131, Gr 6)

The summarising tables (5.17 and 5.18) show that more or less half of the Grade 6's used tables to display the data for Task 1 and 2, which is much more than for other grades. Tables also made out a sizable percentage of all Grade 6 responses for Task 2. As no interviews were conducted, the reason for the popularity of tables as representational type in Grade 6 is not clear. In general, tables were not a favoured way of displaying the data.

Grade	Number of tables	% of tables	% of all possible responses in grade
4	2	29	6
5	1	14	2
6	3	43	7
7	1	14	4
Total	7	100	5

Table 5.17: Analysis of representation (tables) across grades (T1)

Grade	Number of tables	% of tables	% of all possible responses in grade
4	2	13	6
5	3	19	7
6	8	50	18
7	3	19	13
Total	16	≈100	11

Table 5.18: Analysis of representation (tables) across grades (T2)

5.3.5 Pictograms and frequency graphs

The pictograms show varying degrees of abstractness. In the pictogram of L9 (Figure 5.56, T1, Gr 4), different pictures were used to depict the number of learners liking each colour, each row of pictures followed by an equal-to sign and a coloured block to show the preferred colour. The display is filled with colourful waving lines and ikonic symbols. L77 (Fig. 5.57, T1, Gr 5) used drawings of girls and boys in a display that is more abstract than the previous one. Next to each row of figures the favourite colour is written in that specific colour. L110 (Fig. 5.58: T2, Gr 6) produced a pictogram with a legend, black for girls and green for boys. The figures used are less ikonic than in the previous example and a Y-axis give the feet lengths. No X-axis is given and a heading explains that the numbers are to the nearest centimeter and that the figures give the number of children. A higher degree of abstractness is thus evident. L118 (Fig. 5.59, T1, Gr 6) displayed the data in an even more abstract pictogram. The different favourite colours are written on an X-axis, giving the number of learners who like the colour. The faces depicting the number of children liking each colour are drawn in a table. The last two representations are transitional in nature, moving towards a bar graph, but not yet classifiable as bar graphs as they only have one axis. Frequency graphs (Fig. 5.60, T1, L114, Gr 6) were coded in the same category as pictograms, as they have only one axis, and the data points are depicted with some kind of symbol, usually a cross.

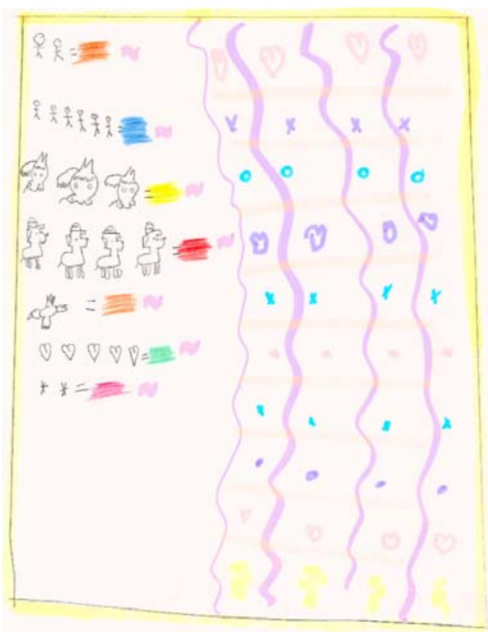


Figure 5.56: Pictogram
(T1, L9, Gr 5)

Figure 5.57: Pictogram
(T1, L77, Gr 5)

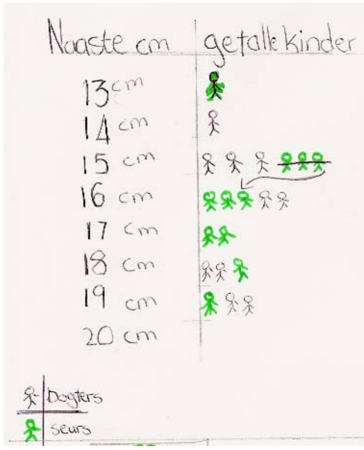


Figure 5.58: Pictogram
(T2, L110, Gr 6)

Figure 5.59: Pictogram
(T1, L118, Gr 6)

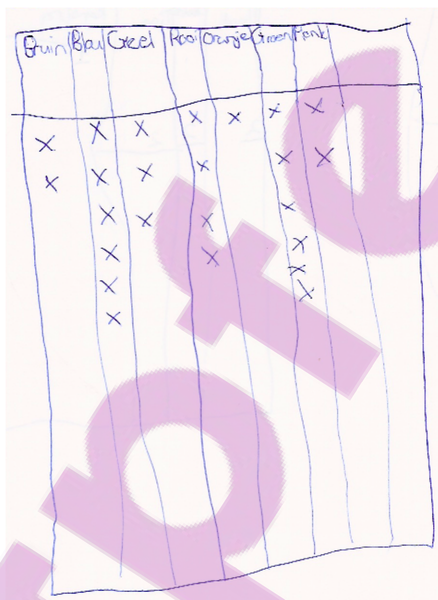


Figure 5.60: Frequency graph (T1, L114, Gr 6)

In general pictograms were not a popular representational type. None of the Grade 7 learners used pictograms for T1 or T2. This might be because they have had more exposure to other kinds of representations and because they are older, and are less inclined to ikonic displays than younger learners. Most of the pictograms for T1 however, were created by Grade 6 learners, who are just one year younger than the Grade 7's, so the explanation may not lie in their age, but rather in teaching activities they have been exposed to. 11% of the pictograms for T1 were produced by Grade 4's, but none of them used a pictogram to display the data of T2. For T2, only Grade 5's (60%) and 6's (40%) used pictograms.

Grade	**Number of pictograms and frequency tables**	**% of pictograms and frequency tables**	**% of all possible responses in grade**
4	1	11	3
5	1	11	2
6	7	78	16
7	0	0	0
Total	9	100	6

Table 5.19: Analysis of representation (pictograms and frequency tables) across grades (T1)

Grade	**Number of pictograms and frequency tables**	**% of pictograms and frequency tables**	**% of all possible responses in grade**
4	0	0	0
5	3	60	7
6	2	40	5
7	0	0	0
Total	5	100	4

Table 5.20: Analysis of representation (pictograms and frequency tables) across grades (T2)

5.3.6 Bar graphs

The bar graphs also showed different levels of correctness and abstractness. The first two examples of transitional bar graphs for T1 and T2 were produced by L141 in Gr 6 (Fig. 5.61, T1 and Fig. 5.62, T2) and show representations with two axes, though all the axes are not explicitly drawn. The T1 representation gives the favourite colours on the x-axis, but no line marks this axis. The numbers of learners liking each colour are given on the y-axis, but stand detached from the line where the first bar commences. A line in each bar is drawn to indicate the top of the bar. The bars are filled with crosses as one would expect in a frequency table, but the crosses are not evenly spread and are just to fill the space, not giving the number depicted by the individual bars, e.g. blue is the preferred colour of 6 learners, but 7 crosses fill the bar for a blue preference and 6 crosses for yellow in stead of the correct number of 3. The T2 representation shows two implicit axes though there is no line demarcating the two axes. It seems as though the learner first drew the crosses and then drew a line around the contours to form a bar graph.

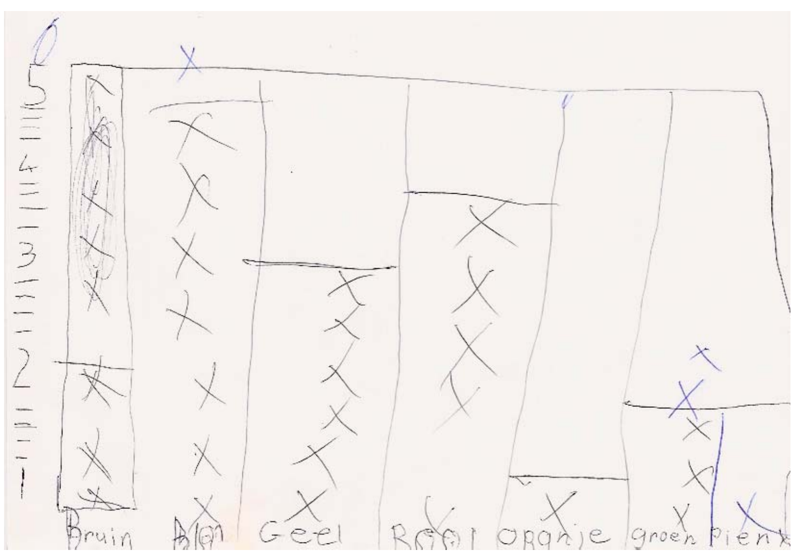


Figure 5.61: Transitional bar graph (T1, L141, Gr 6)

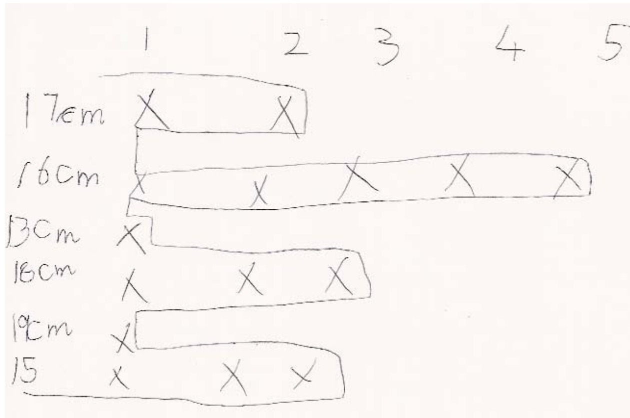


Figure 5.62: Transitional bar graph (T2, L141, Gr 6)

Some learners displayed the data inappropriately in a bar graph, displaying one bar for each child (Fig. 5.63, T2, L82, Gr 5); another drew separate bars for boys and girls in each grade, using incorrect averages for the Grade 4's (Fig. 5.64, T2, L136, Gr 6). L119 (Fig. 5.65, T2, Gr 6) also drew separate bars for boys and girls, but incorrectly indicated the different feet length of more than one child per bar.

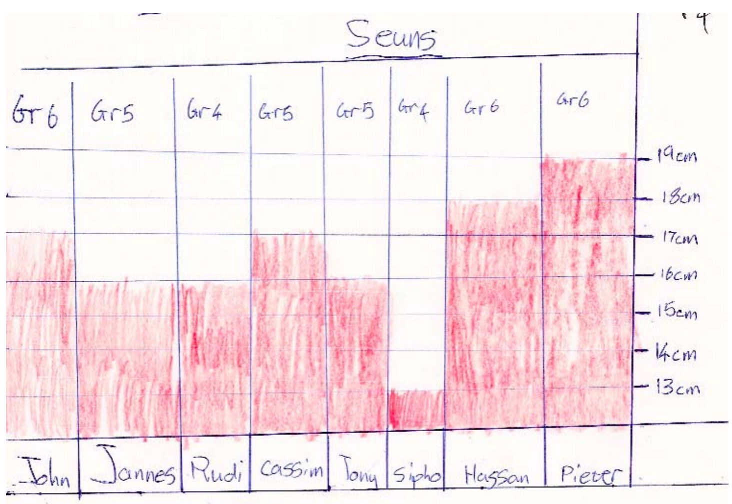
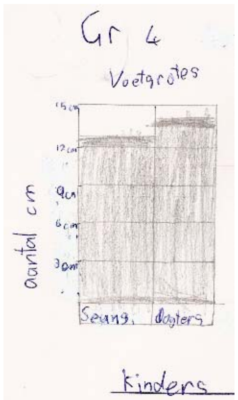


Figure 5.63: Bar graph (T2, L82, Gr 5)



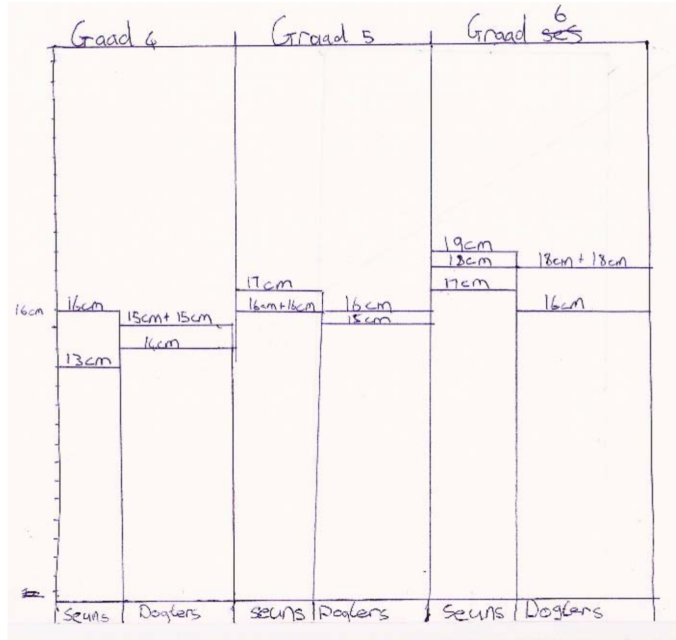


Figure 5.64: Bar graph (T2, L136, Gr 6)

Figure 5.65: Bar graph (T2, L119, Gr 6)

L90 (Fig. 5.66, T1, Gr 5) drew a correct bar graph, but retained the names of learners on the x-axis, indicating their favourite colour using the corresponding colour in each bar.

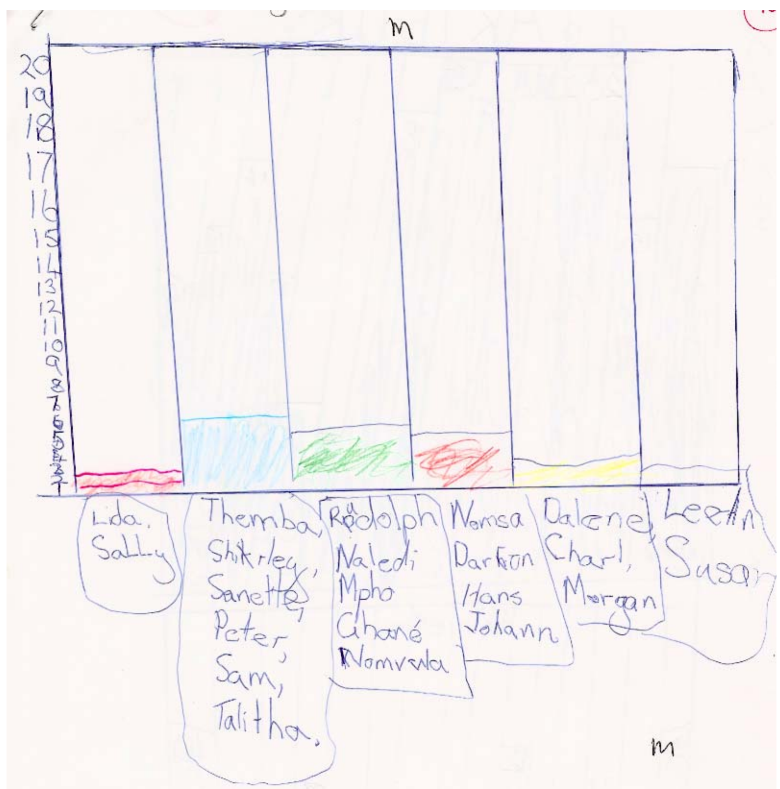


Figure 5.66: Bar graph (T1, L90, Gr 5)

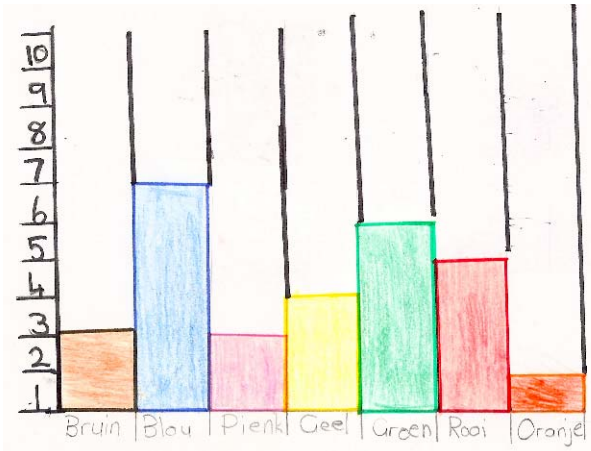


Figure 5.67: Bar graph
(T1, L128, Gr 6)

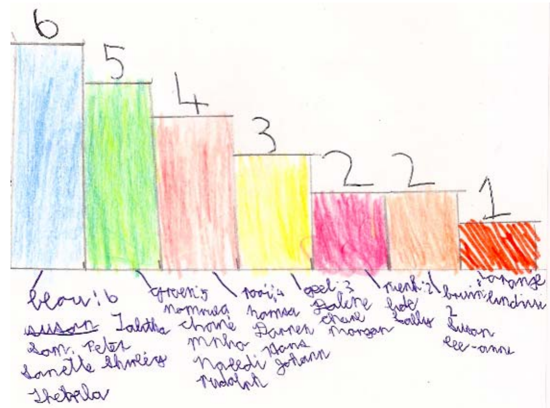


Figure 5.68: Bar graph
(T1, L5, Gr 4)

L128 (Fig. 5.67, Gr 6) drew a correct and very neat bar graph to display the data of T1. L5 (Fig. 5.68, T1, Gr 4) sequentially arranged the bars in the graph, but retained the names of individual cases on the x-axis, and as L90, indicating their favourite colour using the corresponding colour in each bar.

Bar graphs were an extremely popular representational type. 51% of T1 and 61% of T2 bar graphs were produced by Grade 6's. 50% of all Grade 6 responses for T1 and 39% of T2 responses were bar graphs. A significant percentage of Grade 7 responses (48%) were also bar graphs. Bar graphs are a very familiar representation used in data handling activities, which is one of the factors that explain the increasing number of bar graphs with age. Another contributing factor is exposure to bar graphs in the media.

Grade	Number of bar graphs	% of bar graphs	% of all possible responses in grade
4	4	9	11
5	6	14	15
6	22	51	50
7	11	26	48
Total	43	100	30

Table 5.21: Analysis of representation (bar graphs) across grades (T1)

Grade	Number of bar graphs	% of bar graphs	% of all possible responses in grade
4	3	11	8
5	4	14	10
6	17	61	39
7	4	14	17
Total	28	100	19

Table 5.22: Analysis of representation (bar graphs) across grades (T2)

5.3.7 Pie charts

Pie charts are only introduced in the curriculum in Grade 7, but learners often see pie charts in the media. Learners usually have more success in interpreting pie charts than drawing them. L167 (Fig. 5.69, T1, Gr 7) drew a rough but reasonably accurate pie chart.

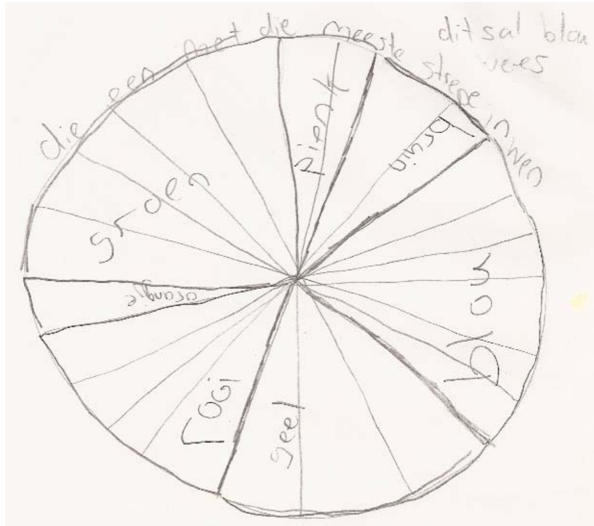


Figure 5.69: Pie chart (T1, L167, Gr 7)

All 4 pie charts for T1 were drawn by Grade 7 learners, which is an indication of the difficulty level of this representational type. No learners drew pie charts to represent the data in T2.

Grade	Number of pie charts	% of pie charts	% of all possible responses in grade
4	0	0	0
5	0	0	0
6	0	0	0
7	4	100	17
Total	4	100	3

Table 5.23: Analysis of representation (pie charts) across grades (T1)

5.3.8 Line graphs

There was only one response in which a line graph was used to represent data. A grade 7 learner chose to represent the data in Task 2 with an inappropriate line graph. This one line graph response represents 4% of the Grade 7 responses for Task 2. The graph is equivalent to a horizontal bar graph with dots showing the number of learners who have a certain foot length, the dots then joined together to form a broken line. The learner separated gender by colour, drawing two line graphs for each grade. This one line graph represents 4% of Grade 7 responses.

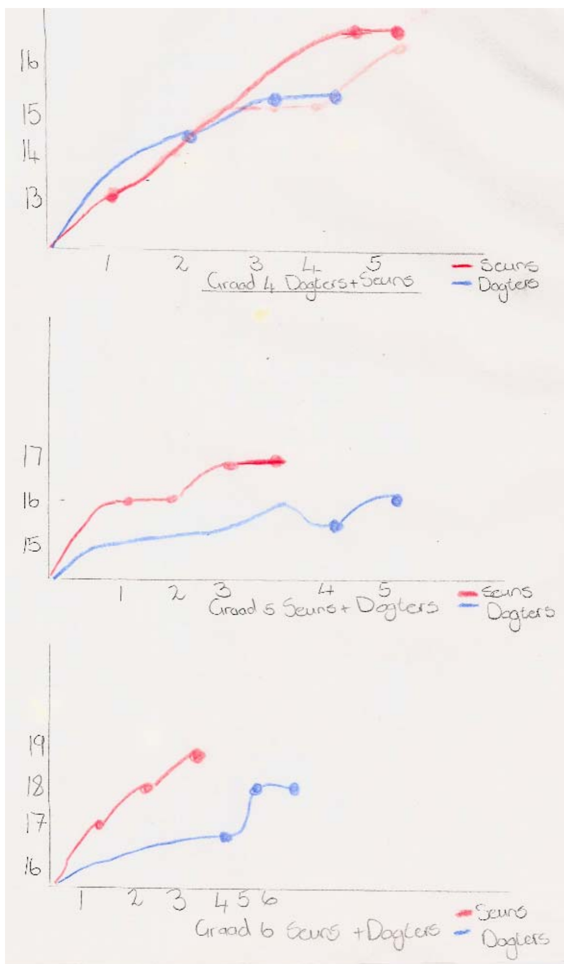


Figure 5.70: Line graph (T2, L168, Gr 7)

5.3.9 Anomalous representations

There were no anomalous representations for Task 1. Except for one anomalous representation in Task 2 which was ikonic in nature (Fig. 5.71, L73, Gr 5), all the anomalous representations were attempted quantitative summaries of the data, where learners just added up all the feet lengths or tried to calculate the mean of all feet lengths (Fig. 5.72, T2, L74, Gr 5 and Fig. 5.73, T2, L78, Gr 5).

.

Figure 5.71: Anomalous representation (T2, L73, Gr 5)

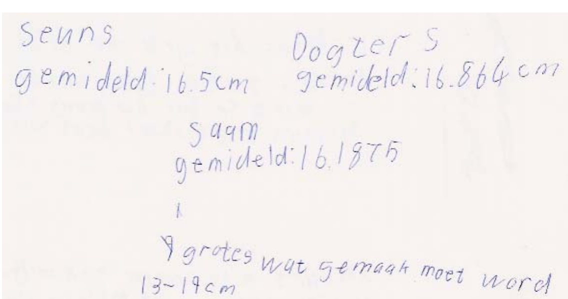


Figure 5.72: Anomalous representation, (T2, L74, Gr 5)

Figure 5.73: Anomalous representation (T2, L78, Gr 5)

Anomalous responses represent only 4% of all learner responses. There were no anomalous responses in Grade 4 or 6. 67% of Grade 5's and 33% of Grade 7's produced anomalous responses, which is approximately 10% of responses in each of the two grades.

Grade	Number anomalous representations	% of anomalous representations	% of all possible responses in grade
4	0	0	0
5	4	67	11
6	0	0	0
7	2	33	9
Total	6	100	4

Table 5.24: Analysis of representation (anomalous representations) across grades (T2)

5.3.10 Summary of analysis of representational types

Task	Grade	Percentages (%)									Number of possible reponses
		No response/incomplete (can not tell)	Pictures (shapes/names/numbers)	Lists	Tables	Pictograms and frequency tables	Bar graphs	Pie charts	Line graphs	Anomalous representations	
		0	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	
1	**4**	0	58	22	6	3	11	0	0	0	**36**
	5	0	54	27	2	2	15	0	0	0	**41**
	6	0	11	16	7	16	50	0	0	0	**44**
	7	0	17	13	4	0	48	17	0	0	**23**
	Total	**0**	**36**	**20**	**5**	**6**	**30**	**3**	**0**	**0**	**144**
2	**4**	19	31	36	6	0	8	0	0	0	**36**
	5	7	15	44	7	7	10	0	0	10	**41**
	6	2	5	32	18	5	39	0	0	0	**44**
	7	0	4	52	13	0	17	0	4	9	**23**
	Total	**8**	**14**	**40**	**11**	**4**	**19**	**0**	**1**	**4**	**144**
1 & 2	**Total**	**4**	**25**	**30**	**8**	**5**	**24,5**	**1,5**	**0,5**	**2**	**288**

Table 5.25: Summary of representational types across grades (percentages do not add up to 100 in all cases because the values have been rounded, except for the total percentage for T1 & T2).

Having studied all the different representational types that were found in the analysis, it becomes necessary to give a few summarising remarks. An unexpected range of different representational types were found in learner responses. The fact that no reference to graphs were made in the wording of the tasks, allowed learners to spontaneously arrange and represent the data, which was one of the aims of the research. In another research project about representations (Chick & Watson 2001:100), very few learners came up with the idea to use graphs to represent the data in the first lesson. The learners were then shown various graphs and tables and this then prompted most learners to use graphical approaches to the data for their posters. In yet another project, learners were introduced to more formal ways of representing data, which included summarising tables, two-by-two tables, Venn diagrams, and scattergrams (Watson & Callingham 1997). The question could be asked if the learners considered the graphical representations shown to them as valuable tools for representing data, or if they used them because they thought they ought to because the researchers value such representations. In a personal conversation with the renowned Dr Rosemary Callingham at the University of New England, Armidale, Australia (July 2005), she expressed her amazement at the wide variety of different representations that learners produced. She believed the reason that they did not get such a variety was the fact that they did not allow learners to produce spontaneous representations, but introduced them to different representations before they made their posters. Another result of teaching specific kinds of representation before allowing the learners to model the task is that their options for representation become limited, but on the other hand the representations may be more appropriate or effective if more formal representational types are used. Dr Callingham has published many articles and is widely quoted in the field of representation and the SOLO categorising of responses.

As mentioned in 5.2.1, the process of modeling becomes evident when a learner produced more than one representation. The second representation mostly was on a higher level than the first and is an indication that the learner had rethought and refined the first response. This however does not imply that no modeling took place when only one representation was created. Some learners modeled the task by doing the planning and going through different stages of solving the problem without showing a paper trail. The only way in which it would be possible to identify the different steps in the modeling process the learner went through, would be to do an interview task. Because of time and logistic constraints at the school where the data collection took place this option was not available.

A large percentage of all responses were pictures, which shows that although the target mode for the Grade 4 to 6 learners are the concrete symbolic mode, many learners tend to respond ikonically to tasks, or use ikonic support in the concrete symbolic mode. The fact that 40% of all learners used lists for T2, many of which were just copies of the given list, indicates that they had trouble interpreting the task or that they were not able to transform the data in an appropriate way to effectively represent it (Chick 2003:212). Representations using names and pictures, suggest that it can not be assumed that all learners are aware of the need to display data in a way that allows visual counting (Watson & Moritz 2001:73).

Three noticeable trends emerged in the analysis of representational types. Table 5.26 shows that pictures (36%), bar graphs (30%) and lists (20%) were the representational types mostly used for T1 while for T2 lists (40%), bar graphs (19%) and pictures (14%) were used most. A sizable percentage of tables were also used in T2. The representational types that proved the most popular for both tasks together were lists (30%), pictures (25%) and bar graphs (24,5%). As already mentioned (5.3.1; 5.3.9) no responses and anomalous representations were only produced for T2, while pie charts are only possible for categorical data (5.3.7) and were thus only used for T1. The process of drawing a pie graph is not an easy task for learners of this age and is also time consuming which may be the reasons for the unpopularity of this representational type. Only one line graph (incorrect) were used (T2) and only few tables, pictograms and frequency tables were created. Pictograms, although usually easy to understand and draw, are also time consuming and may therefore not been chosen as a way to display the data. In the school curriculum, line graphs are not included until Grade 8, therefore learners are usually unfamiliar with this kind of representation.

5.4 SOLO LEVELS OF DATA REPRESENTATION

As discussed in 3.7.1 the SOLO Taxonomy theorises that all learning occurs in one of five modes of functioning of which the ikonic and concrete symbolic modes are applicable to this study. The target mode in the primary school is the concrete symbolic mode (Pegg & Davey 1998:117,118), but many learners still respond to tasks in the ikonic mode. A major shift in abstraction takes place in the move from the ikonic to the concrete symbolic mode because symbolic systems have a logic and an internal order as well as order in relation to the particular context. Many learners respond in the concrete symbolic mode but still use ikonic support, thus moving to a higher level of abstraction while using images as support. The SOLO Taxonomy was adapted to create a framework to assist with the in depth analysis of the data.

In the SOLO analysis it is imperative to consider not only the arrangement and representation, but to also carefully consider the statistical requirements of each task and the statistical level of each response. As the first task comprises categorical (qualitative) data, the levels of responses are limited. Apart from counting the number of kids who like each colour, no other calculations are meaningful. One of the best ways to display the data for an observer, who may not be close up, is to summatively arrange data, showing learners' favourite colours in a bar graph. In the second task, numerical (quantitative) data are given, but the data have to be dealt with in a categorical way. The statistical requirement for the task calls for the sequential regrouping of the data according to feet length in a bar graph, with discussion of patterns in the data.

Some learners produced more than one representation as they modeled the tasks. As in the analysis of responses for arrangement and representational types, the highest level of functioning which was observed was recorded. The recorded level is not regarded as the highest level of a learners' understanding, but only an indication of the level of response for the specific task.

The first level of response was prestructural or lower than the target mode. In the target mode, the concrete symbolic mode[2], an intramodal development pattern with two distinct U-M-R-cycles[3] was identified in the analysis of all learner responses. These two cycles are not evident in each response of each learner because they develop over time. When analysing all responses of all learners, the two cycles do however become evident as levels on which learners respond.

The first U-M-R-cycle ($U_1M_1R_1$) is characterised by intuitive responses in which concepts are still being formed. This cycle shows intuitive statistical representations with pairing-off or connecting of variables, grouping, inappropriate bar graphs and pictograms, as well as incorrect or inappropriate statistical treatment of the data. In the second U-M-R-cycle ($U_2M_2R_2$) cycle concepts that have already been formed, are consolidated and used in a more formalised or "statisticalised" approach. This cycle goes beyond the intuitive considering of the data to exhibit a more quantitative handling of the data, with summative lists/tables/groups, appropriate bar graphs and summary statistics. The first intuitive cycle is precursory to the more quantitative handling of the data in the second cycle and in the first cycle learners focus on the data in an individual sense rather than in a more aggrevated sense as in the second cycle. Responses on each of these levels will now be discussed in detail.

---

[2] The ikonic and concrete symbolic mode will be abbreviated as IK and CS in the detailed discussion of responses in these modes.

[3] The cycle of unistructural, multistructural and relational levels will be abbreviated in the text by U-M-R when referring to the general levels and $U_1M_1R_1$ or $U_2M_2R_2$ respectively for the first and second cycles.

Responses were classified according to the six different levels according to the two hypothesised U-M-R cycles (see 4.7.1.3) with an added category to provide for learners who offered no responses. At the R_2 level, learners should as on the M_2 level have re-organised data (feet lengths) into intervals that will make sense to manufacturers of beach sandals, displaying it in a bar graph, but interpretation and discussion of the data and graph should have been added. No learners however responded on this level and this hypothesised category was therefore dropped from the list. The responses were coded from 0 to 6 on a nominal scale in the following way (in each case all the different possibilities found are listed):

0 = * No response
* Incomplete – cannot tell

1 = Prestructural/Ikonic:

* Uses no relevant information
* Supplied own data
* Copied given list (told not to do it)
* Incorrect, inappropriate answer
* Pictures with personal comments

The first U-M-R cycle (U_1 M_1 R_1):

2 = U_1 (Unistructural first cycle):

* Uses only colour
* Represent list in pictures without names
* copied given list – omitted names
* copied given list – omitted grade (incomplete)
* names or colours separately with no connection
* pairing off names and favourite colour
* connecting names and colour (incomplete)

3 = M_1 (Multistructural first cycle):

* pictogram
* frequency table
* pie graph
* inappropriate bar graph (one bar for each child)
* inappropriate bar graph (one bar for each child) – gender separate
* inappropriate clustering (according to grade)
* appropriate clustering by one characteristic, not summative

4 = R_1 (Relational first cycle):

* inappropriate bar graph + attempt to summarise (mean, range, biggest shoe, etc.)
* attempt at summative grouping, no clustering, retain names
* pictogram with extra variable (gender)
* sum of feet length

The second U-M-R cycle (U_2 M_2 R_2):

5 = U_2 (Unistructural second cycle):

* Summative list
* summative table
* summative grouping

6 = M_2 (Multistructural second cycle):

* appropriate bar graph/line graph (equivalent to bar graph)
* bar graph (percentages incorrect)

Each category will now be discussed in detail, giving examples of learner responses and a discussion of trends in categories and grades.

5.4.1 No response

All learners responded to and completed Task 1. Grade level played a role in the completion of Task 2, most learners who did not respond to the task were in Grade 4, less in Grade 5 and 6, and all Grade 7's responded to Task 2.

Grade	Number of no responses	% of no responses	% of all possible responses in grade(s)
4	7	58	19
5	4	33	10
6	1	8	2
7	0	0	0
Total	12	≈100	8

Table 5.26: Analysis of no response across grades (T2)

5.4.2 Prestructural level (P)

Prestructural responses are on a level lower than the target mode. The individual is not engaging in the task at hand and often focuses on irrelevant aspects of the situation. Own data may be supplied, or pictures with personal comments given. Some learners copied the given list though it was explicitly stated that it may not be done, while others gave incorrect, inappropriate answers. A prestructural response indicates functioning in a previous mode. Characteristics of responses on this level include:

* no relevant information used
* own data supplied
* given list copied (told not to do it)

* pictures with personal comments
* incorrect, inappropriate response

Grade	Number of prestructural responses	% of prestructural responses	% of all possible responses in grade(s)
4	0	0	0
5	4	80	10
6	1	20	2
7	0	0	0
Total	5	100	4

Table 5.27: Analysis of prestructural response across grades (T1)

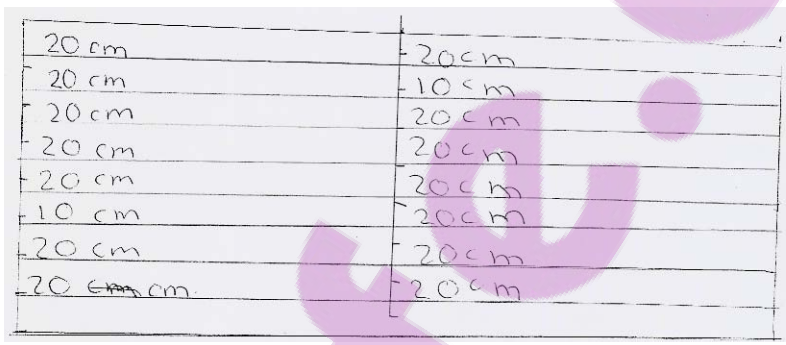
Four of the five prestructural responses in Task 1 came from Grade 5 learners (Table 5.28) and included a list (L71), two pictures with personal comments (L72, L73). One incomplete response was from a Grade 6 learner (L56) in which it is not clear if the learner intended to add more data to the picture.

Grade	Number of prestructural responses	% of prestructural responses	% of all possible responses in grade(s)
4	5	23	14
5	11	50	27
6	4	18	9
7	2	9	9
Total	22	100	15

Table 5.28: Analysis of prestructural response across grades (T2)

The most prestructural responses in all grades (50%) for Task 2 also came from Grade 5 learners (Table 5.29). It is not clear why Grade 5's responded on a lower level than even the Grade 4's. One of the reasons for the low levels of response may be that the Grade 5's have not done any data handling activities up to August of the year in which the research was done, but that is also true of Grade 6 and 7 learners. Grade 4 learners were exposed to data handling activities earlier that year. Examples of learner responses on the prestructural level:

- L25 (Fig.5.74, Gr 4) rounded off all the feet lengths and represented these numbers in a table, which is an example of a response where information used was not relevant.



20 cm	20cm
20 cm	10cm
20 cm	20 cm
20 cm	20 cm
20 cm	20 cm
10 cm	20cm
20 cm	20cm
20 ~~cm~~ cm	20 cm

Figure 5.74: Prestructural response (T2, L25, Gr 4)

- L10 (Fig 5.75, T2, Gr 4) copied the given list (although instructions were explicit that it may not be copied) and paired off names, grades and feet length incorrectly. L109 (Fig. 5.76, T1, Gr 6) supplied her own data and paired off the names and colour.

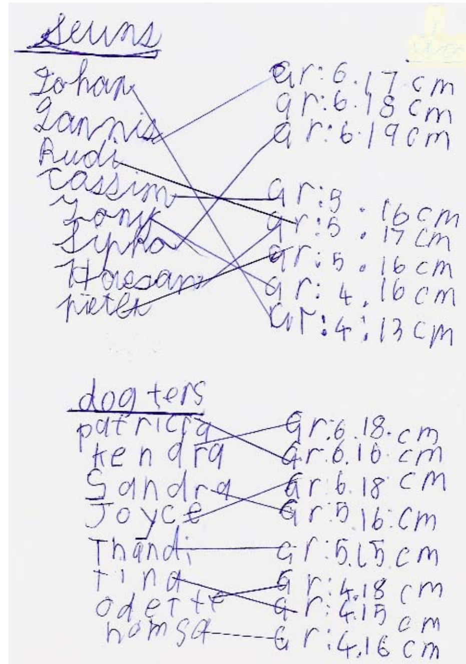


Figure 5.75: Prestructural response

(T2, L10, Gr 4)

Figure 5.76: Prestructural response

(T1, L109, Gr 6)

- The response in Figure 5.77 shows an ikonic representation with personal comments ("Blue rules!") and is considered a prestructural response.

Figure 5.77: Prestructural response (T1, L72, Gr 5)

- An example of an incorrect, inappropriate and may be incomplete response can be seen in the attempt of L56 (Fig. 5.78, T1, Gr 5), writing "Susan likes a mud colour". It is not clear whether the learner intended to add to the picture or considered it complete.

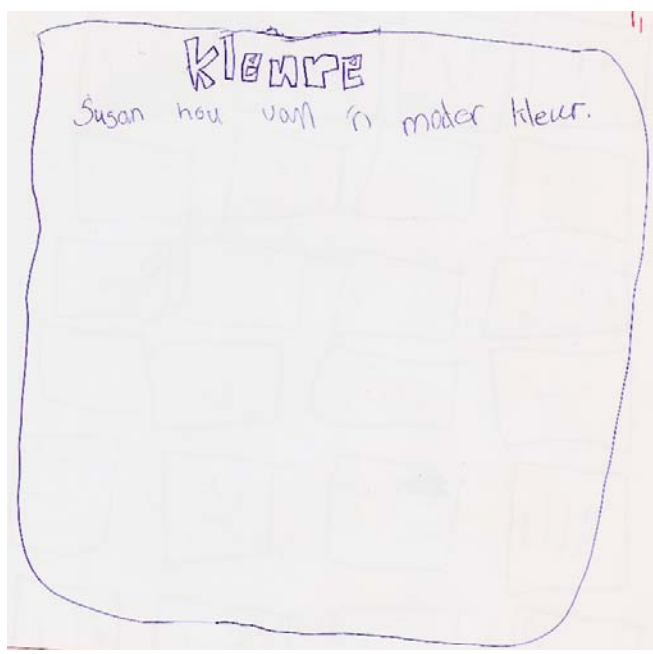


Figure 5.78: Prestructural response (T1, L56, Gr 5)

5.4.3 The first U-M-R cycle

The three levels in this cycle represent the intuitive statistical thinking of learners. Responses on the U_1 level show a focus on individual data values. On the M_1 level learners start clustering data, using two or more data values in appropriate and inappropriate ways. On the R_1 level learners feel the need for quantitative treatment of the data, but do not succeed in appropriately computing the data.

5.4.3.1 Unistructural level, first cycle (U_1)

On this level, the learner engages in the task but uses only one piece of relevant data, focusing on the data in an individual sense. Individual data values are used as principal element in responses, e.g. placing the name of the colour or a coloured dot next to the name of a child. Responses on this level include the following:

* uses colour only
* represent list in pictures without names
* names or colours separately with no connection
* copied given list – omitted names or grade (Fig. 5.81, T1, L69, Gr 5)
* copied given list – omitted grade (incomplete) (Fig. 5.82, T2, L71, Gr 5)
* correct pairing off of names and favourite colour
* connecting names and colour (incomplete) (Fig. 5.85, T1, L28, Gr 4)

Examples of responses:

- L81 (Fig. 5.79, T1, Gr 5) omitted names and only focused on the colour, drawing the colour corresponding to each names in successive blocks, while L51 (Fig. 5.80, T1, Gr 5) and L51 (Fig. 5.81, T1, Gr 5) also omitted the names but chose to draw pictures with the names of the favourite colours.

Figure 5.79: U_1 response (T1, L81, Gr 5)

Fig. 5.80: U_1 response (T1, L51, Gr 5)

- Some learners created an ikonic display with names and favourite colour represented in a disjoint way (Fig. 5.81, T1, L54, Gr 5).

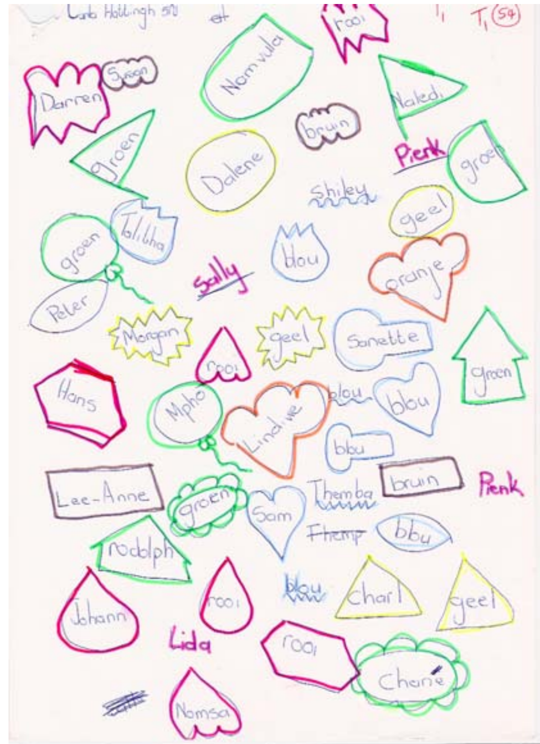


Figure 5.81: U_1 response (T1, L54, Gr 5)

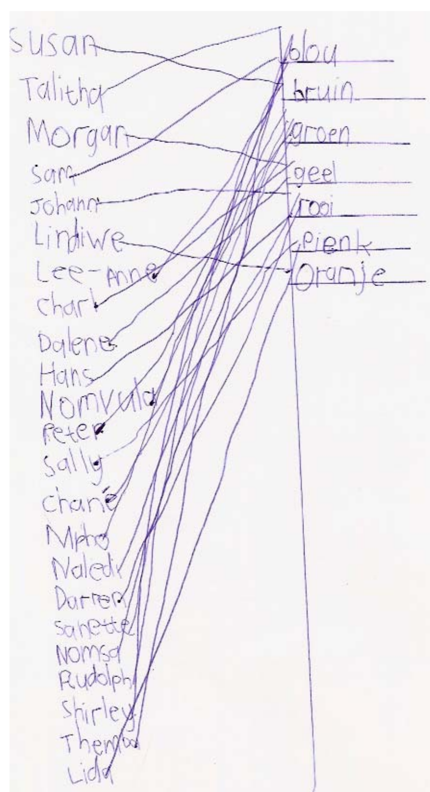


Figure 5.82: U_1 response (T1, L12, Gr 4)

- In contrast to the incorrect pairing off in the prestructural mode, pairing off in the unistructural mode is done correctly (Fig. 5.82, T1, L12, Gr 4).
- Another kind of unistructural response is an ikonic representation in which names and colours are linked (Fig. 5.83, T1, L28, Gr 4).

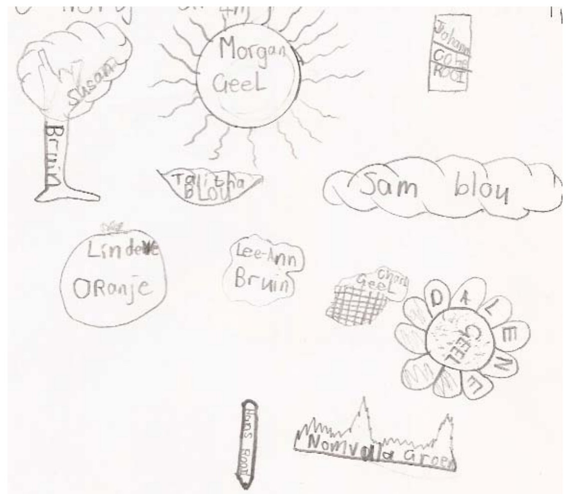


Figure 5.83: U_1 response (T1, L28, Gr 4)

Table 5.30 and 5.31 show that more Grade 5's responded on the unistructural level than other Grades. In Task 1 more than half of all Grade 5's responded on a unistructural level. Only a few Grade 7 learners responded on this level in Task 1 (3% of U_1 responses and 4% of all Grade 7 responses). In Task 2 no Grade 6 or 7 learners responded on the unistructural level.

Grade	**Number of U_1 responses**	**% of U_1 responses**	**% of all possible responses in grade(s)**
4	6	16	17
5	24	65	59
6	6	16	14
7	1	3	4
Total	37	100	26

Table 5.29: Analysis of U_1 response across grades (T1)

Grade	Number of U_1 responses	% of U_1 responses	% of all possible responses in grade
4	2	33	6
5	4	67	10
6	0	0	0
7	0	0	0
Total	6	100	4

Table 5.30: Analysis of U_1 response across grades (T2)

5.4.3.2 Multistructural level, first cycle (M_1)

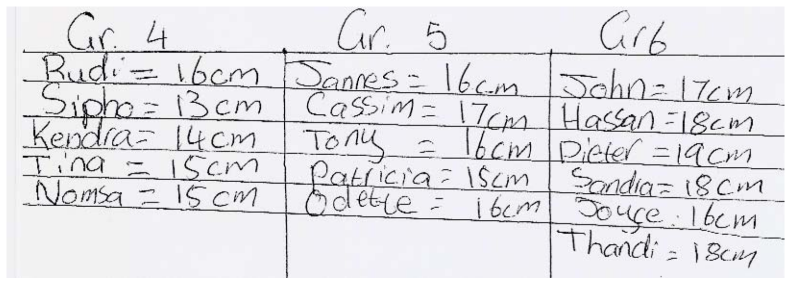
No integration occurs on this level and although the learner uses two or more aspects of the data, there are no relationships between them. The processing of several disjoint aspects of the data is usually done in sequence. Names and colour/ feet length are connected in inappropriate bar graphs, clustering/grouping of names and favourite colour or feet length, pictograms and frequency graphs. The fact that the data are clustered is an indication that the learner is not just focusing on individual data values, but is looking at more than one aspect of the data.

Examples of responses observed:

* inappropriate clustering (according to grade or gender)
* appropriate clustering by one characteristic, not summative
* pictograms
* frequency tables
* pie graphs
* inappropriate bar graphs (one bar for each child)
* inappropriate bar graph (one bar for each child) – gender separate

* inappropriate line graph

- L123 (Fig. 5.84, T2, Gr 6) clustered the names of learners by grade with feet length next to the names. The data were thus regrouped and clustered but the result is a statistically inappropriate response because the data should have been regrouped according to feet length to make sense to sandal manufacturers.



Gr. 4	Gr. 5	Gr6
Rudi = 16cm	Jannes = 16cm	John = 17cm
Sipho = 13cm	Cassim = 17cm	Hassan = 18cm
Kendra = 14cm	Tony = 16cm	Pieter = 19cm
Tina = 15cm	Patricia = 15cm	Sandia = 18cm
Nomsa = 15cm	Odette = 16cm	Joyce : 16cm
		Thandi = 18cm

Figure 5.84: M_1 response (T2, L123, Gr 6)

- L84 represented the list in pictures without names (Fig. 5.85, T2, Gr 5) but indicated feet length and gender in a quite complicated way: the data are inappropriately clustered by gender, shown by drawings of pants for the boys and blouses for the girls while feet length is indicated by colour as explained in a legend at the bottom of the response (Fig. 5.86, T2, Gr 5). Although the learner in fact distinguishes in a complicated way between different variables, the ikonic nature of the response tend to hide the message of the data rather than make it explicit.

Figure 5.85: M_1 response
(T2, L84, Gr 5)

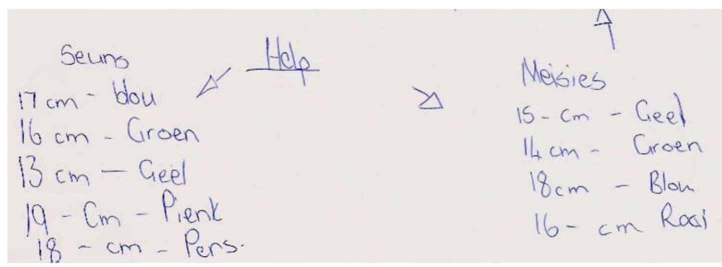


Figure 5.86: legend for Fig. 5.85
(T2, L84, Gr 5)

- An example of appropriately clustered data is found in the display of L15 (Fig. 5.87, T1, Gr 4). The names of learners are clustered according to their favourite colour.

Figure 5.87: M_1 response
(T1, L15, Gr 4)

- Appropriate clustering is also found in pictograms (Fig. 5.88, T1, L136, Gr 6); frequency tables (Fig. 5.89, T2, L133, Gr 6); pie graphs (Fig. 5.90, T1, L153, Gr 7) and bar graphs (Fig. 5.91, T2, L136, Gr 6 and Fig. 5. 92, T2, L3, Gr 4)

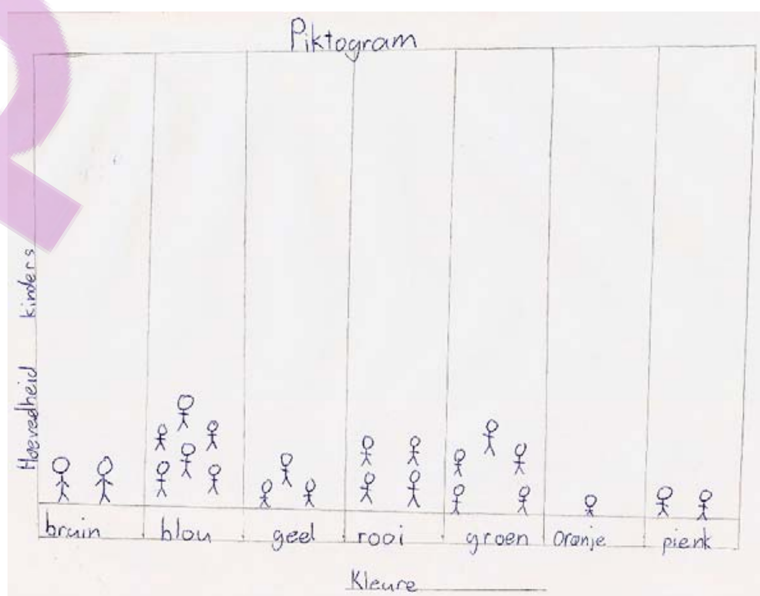


Figure 5.88: M_1 response
(T1, L136, Gr 6)

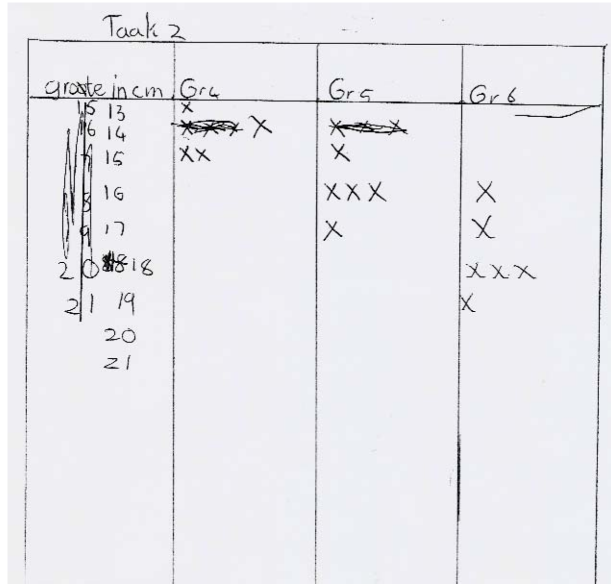


Figure 5.89: M_1 response
(T2, L133, Gr 6)

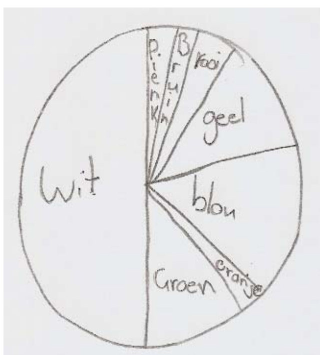


Figure 5.90: M_1 response
(T1, L153, Gr 7)

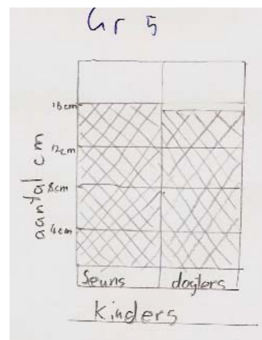


Figure 5.91: M_1 response
(T2, L136, Gr 6)

Fig. 5.92: M_1 response
(T2, L3, Gr 4)

- As clarified in 5.3.8 the inappropriate line graph is equivalent to a horizontal bar graph with dots showing the number of learners who have a certain foot length, the dots then joined together to form a broken line (Fig. 5.93, T2, L168, Gr 7). The learner separated gender by colour, drawing two line graphs for each grade. The graph shows inappropriate clustering by grade and gender and is categorised as a M_1 response.

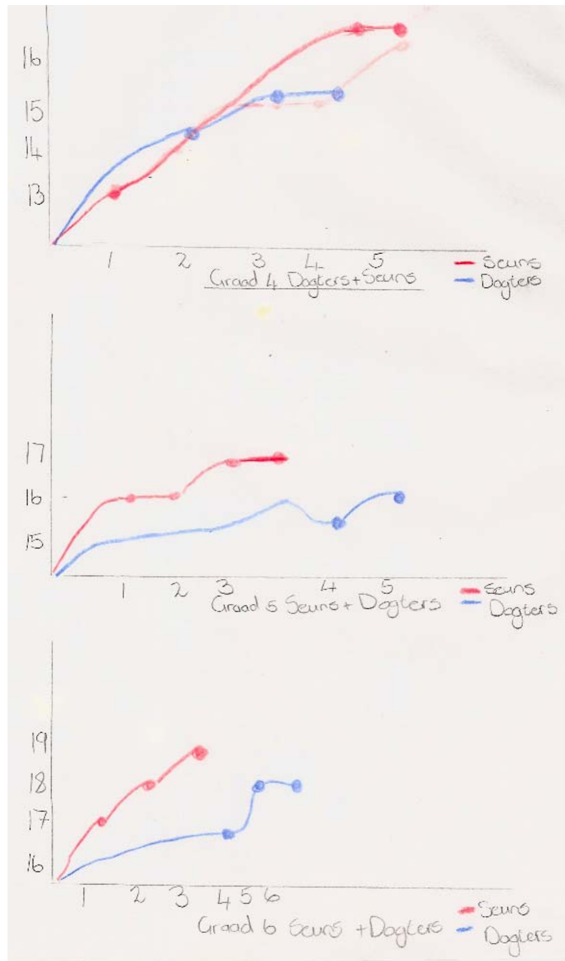


Figure 5.93: M_1 response (T2, L168, Gr 7)

18% of all possible responses for T1 and 33% of responses for T2 were on the M_1 level, which gives a total of 25% of all possible responses for both tasks on this level. In T1 significantly more Grade 4 learners responded on the M_1 level than other grades. For T2 the percentage of responses does not indicate a big difference between grades on this level.

Grade	Number of M_1 responses	% of M_1 responses	% of all possible responses in grade(s)
4	11	42	31
5	5	19	12
6	6	23	14
7	4	15	17
Total	26	≈100	18

Table 5.31: Analysis of M_1 response across grades (T1)

Grade	Number of M_1 responses	% of M_1 responses	% of all possible responses in grade(s)
4	11	23	31
5	12	26	29
6	15	32	34
7	9	19	39
Total	47	100	33

Table 5.32: Analysis of M_1 response across grades (T2)

5.4.3.3 Relational level, first cycle (R_1)

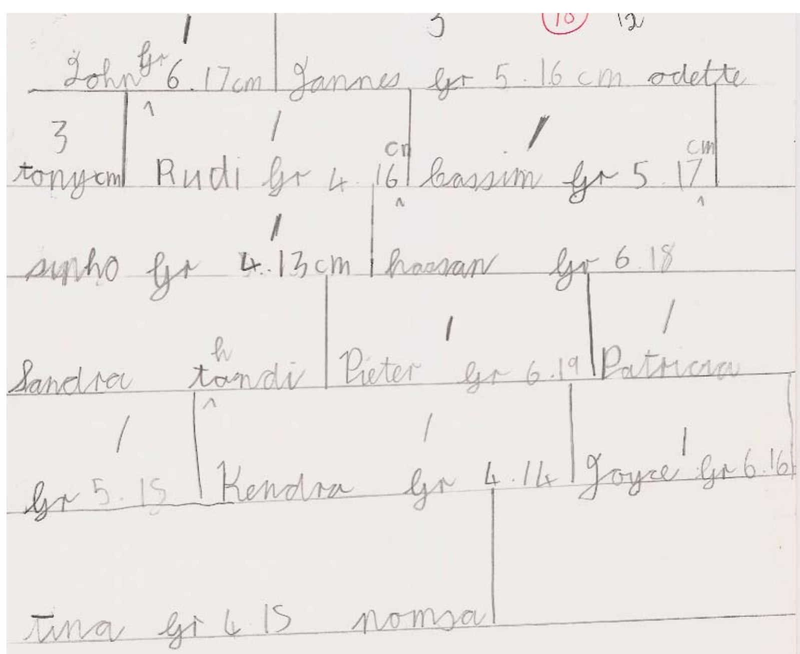
On this level the learner focuses on several aspects of the data and perceives relationships between different aspects in the data. An example is an appropriate pictogram according to feet length distinguishing between boys and girls with colour, thus adding an extra variable in the display. One learner drew an inappropriate bar graph clustering by grade, but then gave the range, the biggest and smallest feet length, trying to summarise the data statistically. Some learners added up the feet length, feeling the need to compute some statistics, but then didn’t know how to proceed, or computed the average, which is inappropriate in the context of the task.

Examples of responses on this level:

* attempt at summative grouping, no clustering, retain names
* attempt at summative grouping of colours with incorrect percentages (Fig. 5.98, L105, T1, Gr 6)
* inappropriate bar graph + attempt to summarise (mean, range, biggest shoe) (Fig. 5.95, L139, T2, Gr 6)
* pictogram with extra variable (gender) (Fig. 5.96, L110, T2, Gr 6)
* sum of feet length (Fig. 5.97, L81, T2, Gr 5)
* attempt to summarise (mean, range, biggest shoe, etc.)

- L18 computed the total number of learners who have the same feet length but could not let go of individual data values such as names (Fig. 5.94, L18, T2, Gr 4). Instead of groups with totals, he gave the names of individuals with feet length and the number of learners with the same feet length in blocks, obscuring the data and not elucidating it. The intent of the learner is to treat the data quantitatively but it is done intuitively and inappropriately.

Figure 5.94: R_1 response (L18, T2, Gr 4)



- L105 attempted to treat the data quantitatively and computed incorrect percentages of learners liking each colour (Fig. 5.95, T1, Gr 6).

Figure 5.95: R_1 response (L105, T1, Gr 6)

- Another way of responding on the R_1 level was to give the mean for each grade and draw an inappropriate bar graph (Fig. 5.96, L139, T2, Gr 6)

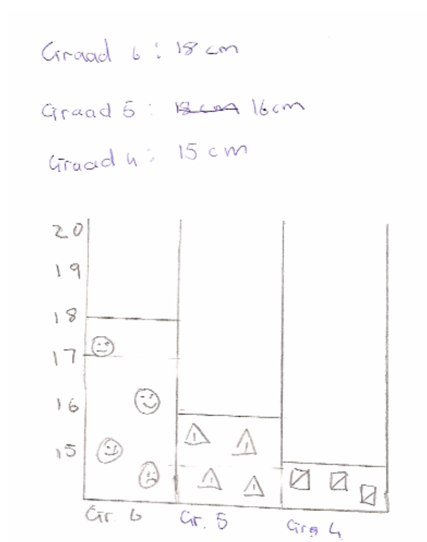


Figure 5.96: R_1 response (L139, T2, Gr 6)

- L110 rearranged the data correctly according to feet length, and drew a pictogram (Fig. 5.97, T2, Gr 6). She distinguished between boys and girls in the graph by colour, thus adding an extra variable in the graph. A pictogram is a clustered representation and therefore on the M_1 level, but because of the extra variable added, this representation is on the R_1 level.

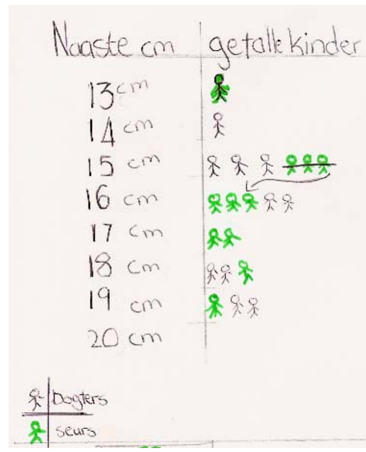


Figure 5.97: R_1 response
(L110, T2, Gr 6)

- Some learners just added up the feet length of learners, feeling the need to compute the data, but not realising what would be appropriate for this specific task (Fig. 5.98, L81, T2, Gr 5), others calculated the mean (Fig. 5.99, L151, T2, Gr 7), gave the upper limit of feet lengths in a grade (Fig. 5.100, L85, T2, Gr 5) or gave the range of feet lengths in each grade or of all learners' feet (Fig. 5.101, L145, T2, Gr 6). These learners felt a need to treat the data quantitatively, but gave inappropriate or incorrect statistics for the tasks.

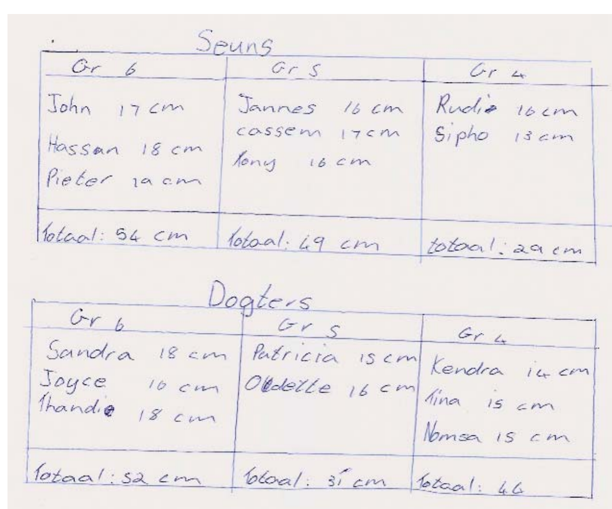


Figure 5.98: R_1 response
(L81, T2, Gr 5)

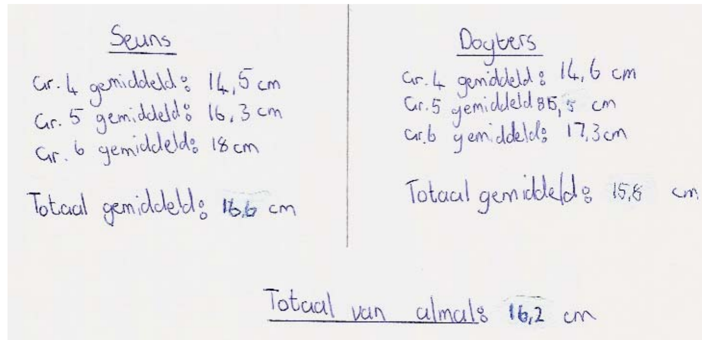


Figure 5.99: R_1 response
L151, T2, Gr 7

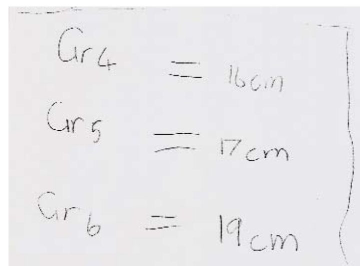


Figure 5.100: R_1 response
(L85, T2, Gr 5)

Figure 5.101: R_1 response
(L145, T2, Gr 6)

One learner in each grade created a R_1 response for T1 while Grade 6 (38%) and 7 learners (32%) produced more R_1 responses than younger learners for T2 The fact that more responses created by older learners are on a higher level shows that older learners might have had more exposure to data representation in the media and in class activities as one would expect.

Grade	Number of R_1 responses	% of R_1 responses	% of all possible responses in grade
4	1	25	3
5	1	25	2
6	1	25	2
7	1	25	4
Total	4	100	3

Table 5.33: Analysis of R_1 response across grades (T1)

Grade	Number of R_1 responses	% of R_1 responses	% of all possible responses in grade
4	3	9	8
5	7	21	17
6	13	38	30
7	11	32	48
Total	34	100	24

Table 5.34: Analysis of R_1 response across grades (T2)

5.4.4 The second U-M-R cycle ($U_2M_2R_2$)

This cycle shows appropriate quantitative treatment of data. On the U_2 level learners not only cluster the data as in M_1, but they treat the groups summatively. On the M_2 level the summative groups are represented in bar graphs. As discussed in 5.4, no responses were on the R_2 level therefore this category is not included in the description of responses in the second U-M-R cycle.

5.4.4.1 Unistructural level, second cycle (U_2)

As mentioned in 5.4.4, the distinguishing characteristic for U_2 level representations is that the data are grouped with totals, evident in the quantitative treatment of data values with summative clustering such as groups or pictures, lists and tables.

- Examples of summative pictures can be found in the display of L27 and L158. L27 writes at the top of her display: “List of colours and numbers” (Fig. 5.102, T1, Gr 4) but creates an ikonic representation, drawing a picture with a section for each colour and filling each section with numbers indicating the number of learners preferring that colour. L158 (Fig. 5.103, T1, Gr 7) gave the favourite colours with totals for each colour in a picture that is in essence a list, but the first impression is that of a picture.

Figure 5.102: U_2 response (T1, L27, Gr 4)

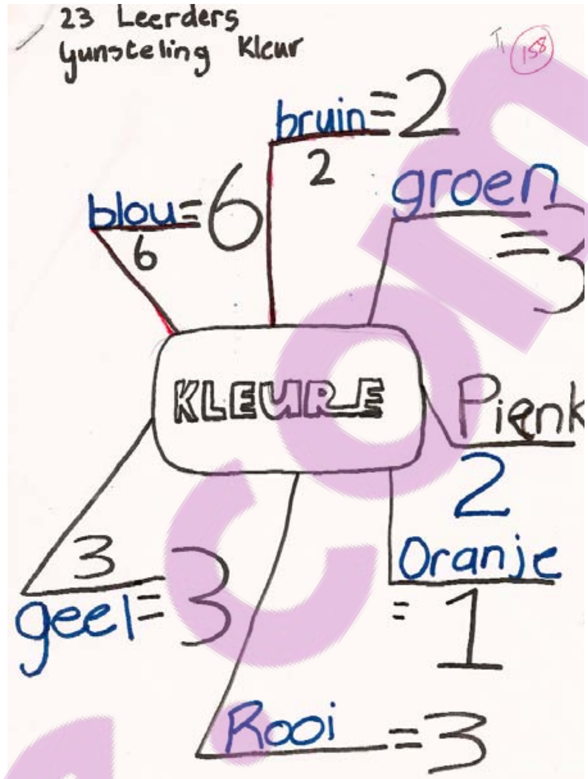


Figure 5.103: U_2 response
(L158, T1, Gr 7)

- Other examples of U_2 responses are summative lists. The number of learners preferring each colour is given ikonically and numerically by L123 (Fig. 5.104, T1, Gr 6), while L22 and L129 give the summary numerically only (Fig. 5.105, T1, Gr 4, and Fig. 5.106, T2, Gr 6).

Figure 5.104: U_2 response
(L123, T1, Gr 6)

Figure 5.105, U_2 response
(L22, T1, Gr 4)

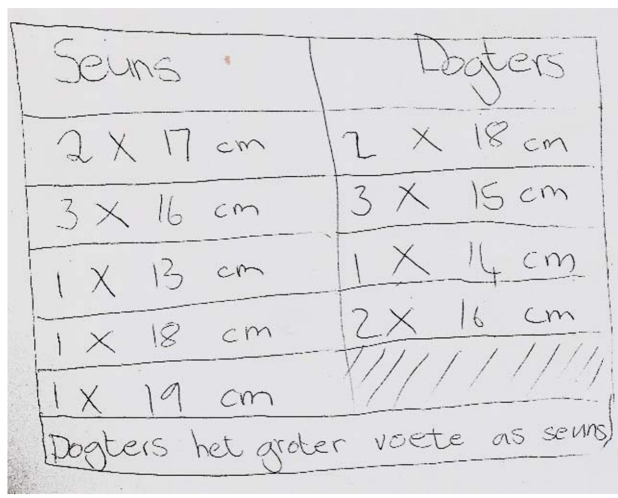


Fig. 5.106: U_2 response
(L129, T2, Gr 6)

- L127 (Fig. 5.107, T1, Gr 6) gave a summative table.

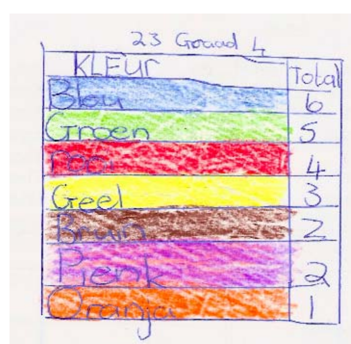
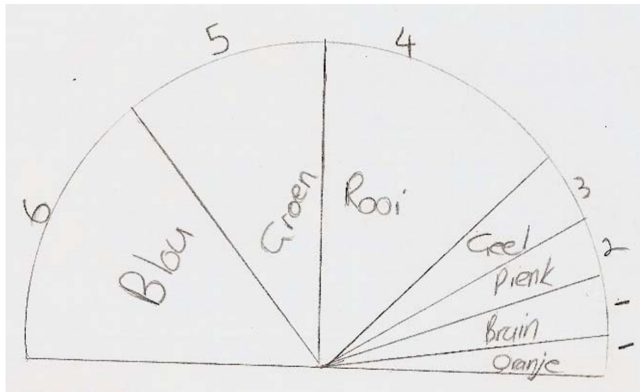


Figure 5.107: U_2 response
(L127, T1, Gr 6)

- L170 (Fig. 5.108, T1, Gr 7) created an incorrect pie graph using only half a circle. The proportional division of the half circle is not completely correct according to the numbers given above each segment, but gives an approximate idea of the number of children preferring each colour.

Figure 5.108: U_2 response (L170, T1, Gr 7)



An unexpected bulge of U_2 responses was found in Grade 4 (45% for T1 and 47% for T2) and Grade 6 (32% for T1 and 40% for T2). Very few Grade 5 learners responded on the U2 level in Task 1 and 2. As mentioned in 5.4.2 the only explanation for this fact may be that Grade 4's were exposed to data handling activities earlier in the year that the research was done while the others were not and that the Grade 6 learners participated in a series of well-planned data handling activities during the previous year.

Grade	Number of U_2 responses	% of U_2 responses	% of all possible responses in grade
4	14	45	39
5	1	3	2
6	10	32	23
7	6	19	26
Total	31	≈100	22

Table 5.35: Analysis of U_2 response across grades (T1)

Grade	Number of U_2 responses	% of U_2 responses	% of all possible responses in grade
4	7	47	19
5	1	7	2
6	6	40	14
7	1	7	4
Total	15	≈100	10

Table 5.36: Analysis of U_2 response across grades (T2)

5.4.4.2 Multistructural level, second cycle (M_2)

At the M_2 level appropriate bar graphs are used to organise and display data.

- Figures 5.109 and 5.110 show examples of bar graphs produced for Task 1. The first example shows an appropriate bar graph (Fig. 5.109, T1, L87, Gr 5). L135 (Fig. 5.110, T1, Gr 6) sequenced the favourite colours in descending order, but used incorrect percentages. It is not clear whether the learner produced the graph after calculating or estimating the percentages or if the percentages were estimated and added after the graph was completed.

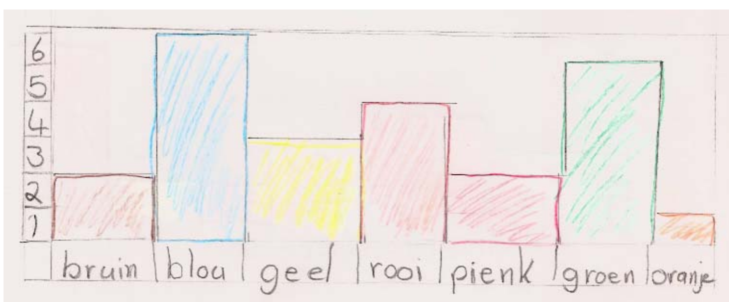


Figure 5.109: M_2 response (T1, L87, Gr 5)

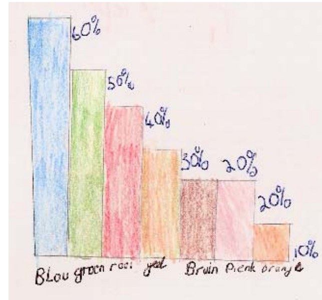


Figure 5.110: M_2 response (T1, L135, Gr 6)

- L164 (Fig. 5.111, T1, Gr 7) joined the vertical lines of the bars that should have been extended and used to form a bar graph, to produce an inappropriate broken line graph.

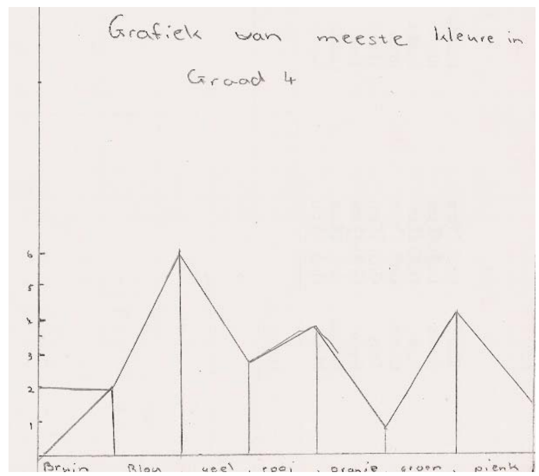


Figure 5.111: M_2 response (T1, L164, Gr 7)

- Two examples of M_2 responses for T2 are presented. L125 (Figure 5.112, T2, Gr 6) created a bar graph with strong ikonic support for T2, filling the bars with different coloured drawings. The bar graph in Fig. 5.113 (T2, L141, Gr 6) was drawn as a frequency table, but lines around the crosses that were added later turned it into a bar graph. Note that the line incorrectly includes the one cross for 13cm in the bar for 18cm and that feet length were not given in sequential order.

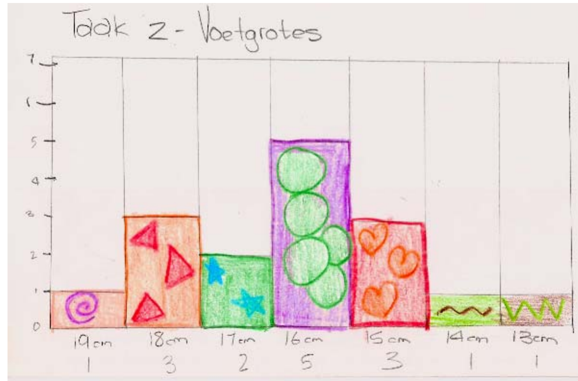


Figure 5.112: M_2 response (T2, L125, Gr 6)

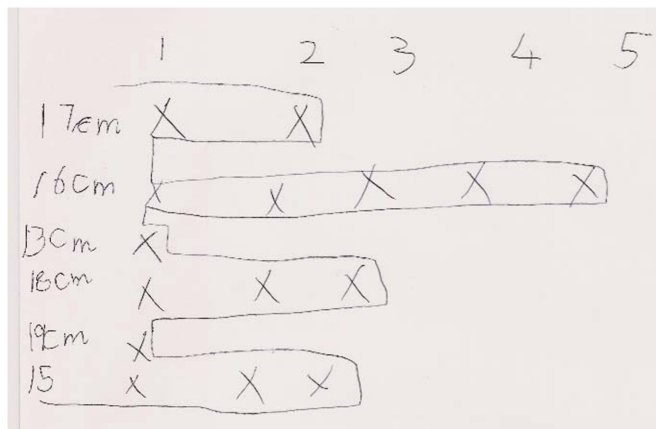


Figure 5.113: M_2 response (T2, L141, Gr 6)

As with U_2 responses, Grade 6 learners created most of the responses on this level, which constitutes half of all M_2 responses for T1 (almost half of all Gr 6 responses for this task) and 63% of responses for T2 (11% of all Grade 6 responses for T2). Almost half of the Grade 7 responses for Task 1 were on the M_2 level, but a most unexpected result was that not even one Grade 7 learner responded on this level for Task 2. Only 6% of all responses for Task 2 were on the M_2 level and came from Grade 4 – 6 learners in descending number according to age. Grade 7 learners are quite familiar with bar graphs, having drawn pictograms and bar graphs since Grade 4 and also pie graphs since Grade 5, but still no learner chose to use this kind of representation for Task 2. For the reader, this interesting result brings afore questions about the number of responses in each grade in each of the two U-M-R cycles. A comparative table showing these statistics will be presented and discussed after a few closing remarks about the relational level in the second cycle.

Grade	Number of M_2 responses	% of M_2 responses	% of all possible responses in grade
4	4	10	11
5	6	14	15
6	21	50	48
7	11	26	48
Total	42	100	29

Table 5.37: Analysis of M_2 response across grades (T1)

Grade	Number of M_2 responses	% of M_2 responses	% of all possible responses in grade
4	1	13	3
5	2	25	5
6	5	63	11
7	0	0	0
Total	8	≈100	6

Table 5.38: Analysis of M_2 response across grades (T2)

5.4.4.3 Relational level, second cycle (R_2)

No second cycle relational responses were found for Task 1 or 2. At the R_2 level, learners should as in M_2 have re-organised data (feet lengths) into intervals that will make sense to manufacturers of beach sandals, displaying it in a bar graph, but interpretation and discussion of the data and graph should have been added. Appropriate summary statistics such as the mode, differences and similarities in the summative grouping and so forth should have been discussed. This level of functioning shows that a learner is in a transitional stage from the concrete symbolic mode to the formal mode in which formal statistical reasoning is used to deal with data. No learner however responded on this level in either of the tasks. Learners' exposure to summary statistics and analytical thinking in statistics (data handling) is at this stage (ages 10 -13) not of such a nature that they could have responded on the R_2 level. The Revised National Curriculum states that learners in Grade 7 should be able to determine measures of central tendency such as mean, median and mode and be able to distinguish between them, but the traditional way of teaching does not promote an integrated understanding of these measures to enable learners to use them in a meaningful way in different contexts. Most teachers' limited knowledge of statistics also hampers the "teaching for understanding" (Carpenter, Fennema & Franke 1996; Fennema, Carpenter & Peterson: c.a.), which enables learners to use their knowledge and skills in different contexts and to appreciate and utilise interrelationships between pieces of knowledge.

5.4.5 Summary of the two U-M-R cycles

Table 5.40 summarises the U-M-R responses in both tasks, giving the number of responses on each level as a percentage of the total number of responses on that level. As the results for each of the levels were discussed separately, only summarizing remarks will be presented here. Note that percentages do not add up to 100% due to rounding and the exclusion of the no-response and prestructural response categories from this specific table.

Task	Grade	Percentages of observed responses in the two U-M-R cycles						
		First U-M-R cycle				Second U-M-R cycle		
		Unistructural level (U_1)	Multistructural level (M_1)	Relational level (R_1)	Total percentage of first cycle responses in the grade	Unistructural level (U_2)	Multistructural level (M_2)	Total percentage second cycle responses in the grade
1	**4**	16	42	25	**27**	45	10	**25**
	5	65	19	25	**46**	3	14	**10**
	6	16	23	25	**18**	32	50	**43**
	7	3	15	25	**9**	19	26	**23**
2	**4**	6	23	9	**18**	47	13	**35**
	5	10	26	21	**26**	7	25	**13**
	6	0	32	38	**32**	40	63	**48**
	7	0	19	32	**23**	7	0	**4**

Table 5.39: Summary of observed responses in each U-M-R cycle

Results presented in Table 5.39 will first be discussed for Task 1. Grade 4 learners created approximately a quarter of the U_1 M_1 R_1 and of the U_2 M_2 R_2 responses for this task (Fig. 5.114 and Fig. 5.115). Grade 5 learners developed almost half of the first cycle responses, but they contributed only 10% of second cycle responses. Grade 6 and 7 learners created more second cycle than first cycle responses, with the largest percentage of second cycle responses coming from the Grade 6's (43%).

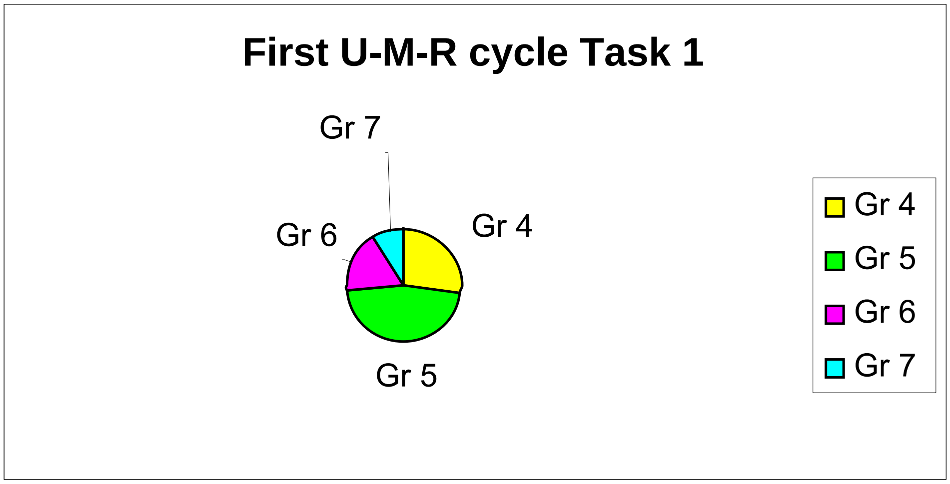


Figure 5.114: First U-M-R cycle per grade (T1)

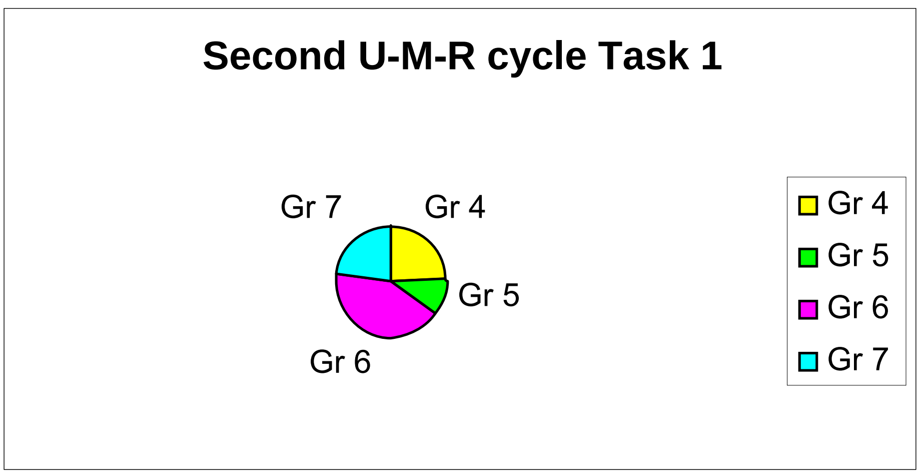


Figure 5.115: Second U-M-R cycle per grade (T1)

In Task 2, the Grade 6 learners again created almost half of all second cycle responses (48%) while Grade 7 learners produced very few responses on this level (4%) (Fig 5.116 and Fig. 5.117). Many Grade 6 and 7 learners responded on the R_1 level, feeling the need to give a quantitative summary of the data, but despite the fact that they may be familiar with the mean, median and mode they could not use their knowledge meaningfully in this task. As discussed in 5.4.4.3 the reason could be that the traditional way of teaching these concepts might not have equipped them to apply their knowledge in other contexts, pointing to a limited understanding thereof.

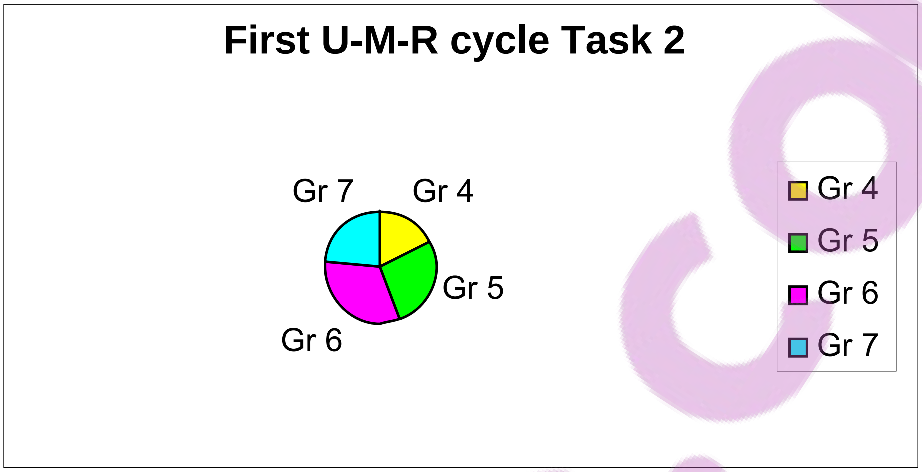


Figure 5.116: First U-M-R cycle per grade (T2)

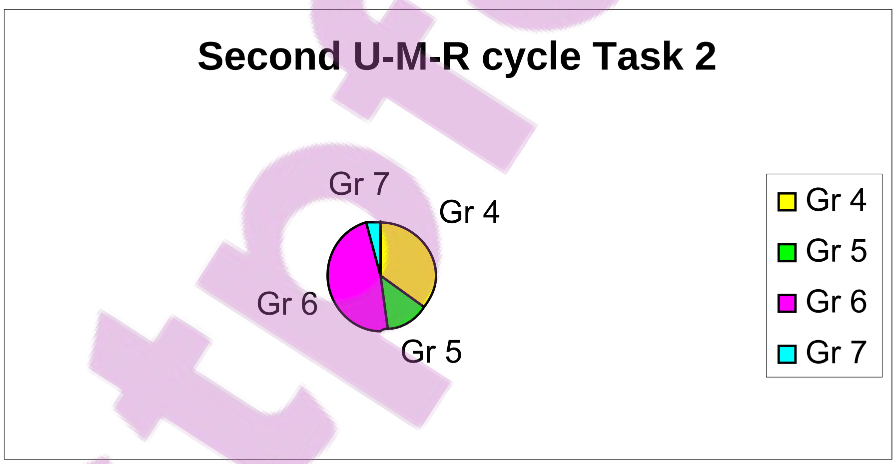


Figure 5.117: Second U-M-R cycle per grade (T2)

Table 5.40 summarises the U-M-R responses in both tasks, distinguishing between the two different U-M-R cycles and giving the number of responses on each level as a percentage of the total number of responses in the grade. As was the case in Table 5.39, percentages do not add up to 100% due to rounding and the exclusion of the no-response and prestructural response categories from Table 5.40.

In Task 1 the Grade 4 responses are equally divided between the two U-M-R-cycles. Contrary to expectations more Grade 5 responses were on a lower than a higher level (73% U1 M1 R1 and 17% U2 M2 R2 responses). Grade 6 and 7 learners responded as expected with larger bulges of responses in the second cycle: in Grade 6 30% of all responses were in the first U-M-R cycle and 71% on the second cycle while in Grade 7 25% of all responses were categorized in the first U-M-R cycle and 74% in the second cycle.

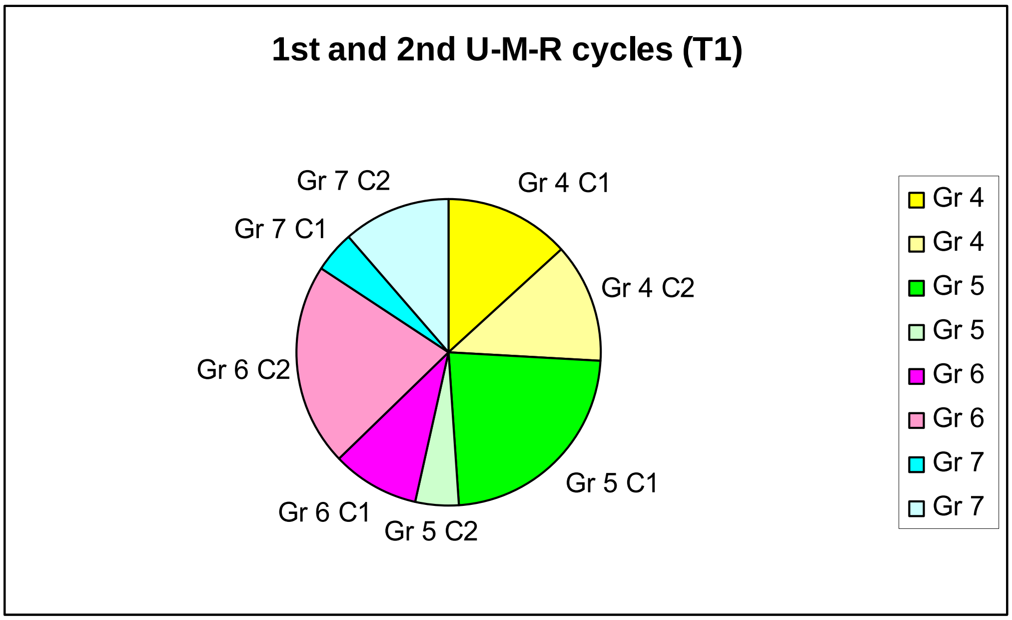


Figure 5.118: First and second U-M-R cycles per grade (T1)
(C1 is Cycle 1 and C2 is Cycle 2)

Grade 5 and 7 produced more responses in the first cycle than the second cycle in Task 2 (7% for Grade 5 and 4% for Grade 7), while in Grade 4 and 6 more responses were in the second cycle than in the first (22 % 2nd cycle responses for Grade 4 and 25% for Grade 6) (Fig. 5.119).

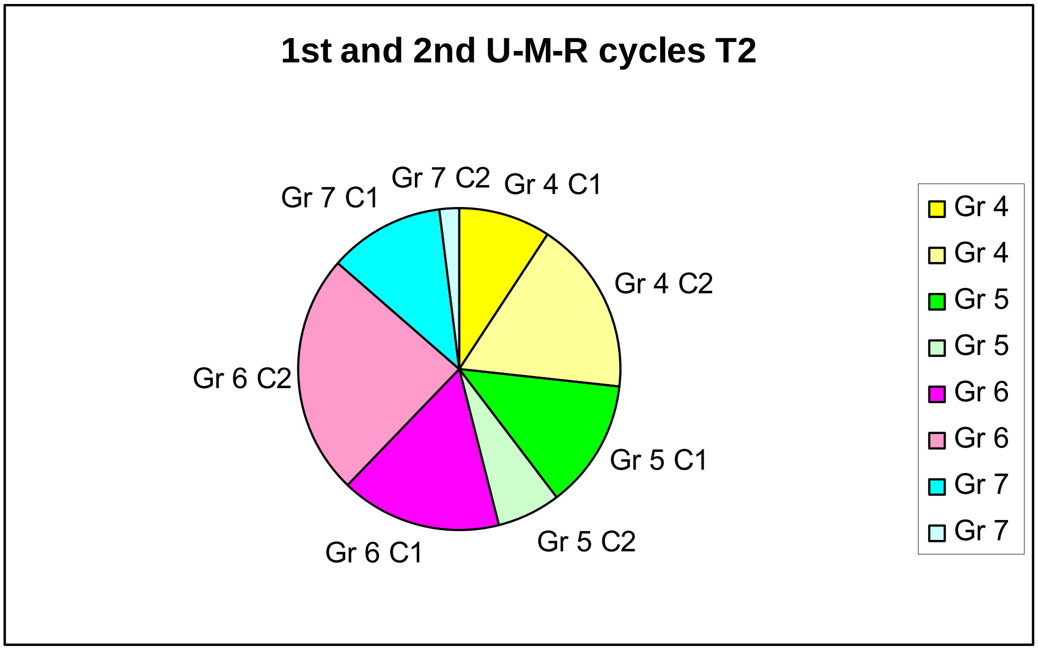


Figure 5.119: First and second U-M-R cycles per grade (T2)
(C1 is Cycle 1 and C2 is Cycle 2)

Table 5.40 summarises the percentage of all possible responses in each grade in for the two U-M-R cycles.

Task	Grade	Percentages of all possible responses in each grade in the two U-M-R cycles						
		First U-M-R cycle				Second U-M-R cycle		
		Unistructural level (U_1)	Multistructural level (M_1)	Relational level (R_1)	Percentage of possible responses in grade	Unistructural level (U_2)	Multistructural level (M_2)	Percentage of possible responses in grade
1	**4**	17	31	3	51	39	11	50
	5	59	12	2	73	2	15	17
	6	14	14	2	30	23	48	71
	7	4	17	4	25	26	48	74
2	**4**	6	31	8	45	19	3	22
	5	10	29	17	56	2	5	7
	6	0	34	30	64	14	11	25
	7	0	39	48	87	4	0	4

Table 5.40: Summary of all possible responses per grade in each U-M-R cycle

5.4.6 Summary of SOLO levels of responses

Table 5.41 gives a summary of the percentage of responses in all categories and shows that all learners responded to Task 1 and only 4% of responses for this Task were on the prestructural level. For Task 2 there were 8% no responses and 15% prestructural responses. As discussed in 5.4.5 more responses were on a higher level in Task 1 than in Task 2, indicating that learners found it more difficult to interpret and represent the data in Task 2 than that of Task 1.

Task	Grade	Percentages of responses: SOLO Levels							Number of possible responses
		No response/Incomplete (cannot tell)	Prestructural (P)	Unistructural first level (U_1)	Multistructural first level (M_1)	Relational first level (R_1)	Unistructural second level (U_2)	Multistructural second level (M_2)	
		0	**1**	**2**	**3**	**4**	**5**	**6**	
1	**4**	0	0	17	31	3	39	11	**36**
	5	0	10	59	12	2	2	15	**41**
	6	0	2	11	14	2	23	48	**44**
	7	0	0	4	17	4	26	48	**23**
	Total	0	4	25	18	3	22	29	144
2	**4**	19	14	6	31	8	19	3	**36**
	5	10	27	10	29	17	2	5	**41**
	6	2	9	0	34	30	14	11	**44**
	7	0	9	0	39	48	4	0	**23**
	Total	8	15	4	33	24	10	6	144

Table 5.41: Summary of SOLO levels

Twenty learners created multiple representations for Task 1 and ten for Task 2. Almost all of these learners created two representations. The only learner who chose to represent either task in more than two ways gave four different representations of Task 1 (Fig. 5.114, L107, T1, Gr 6). L107 initially displayed a summative list and then gave the favourite colours in descending order in three different ways. He first wrote the summary of the first four of the favourite colours in words and then concluded with two drawings of trophies and a rostrum, each time giving the favourite colours in descending order.

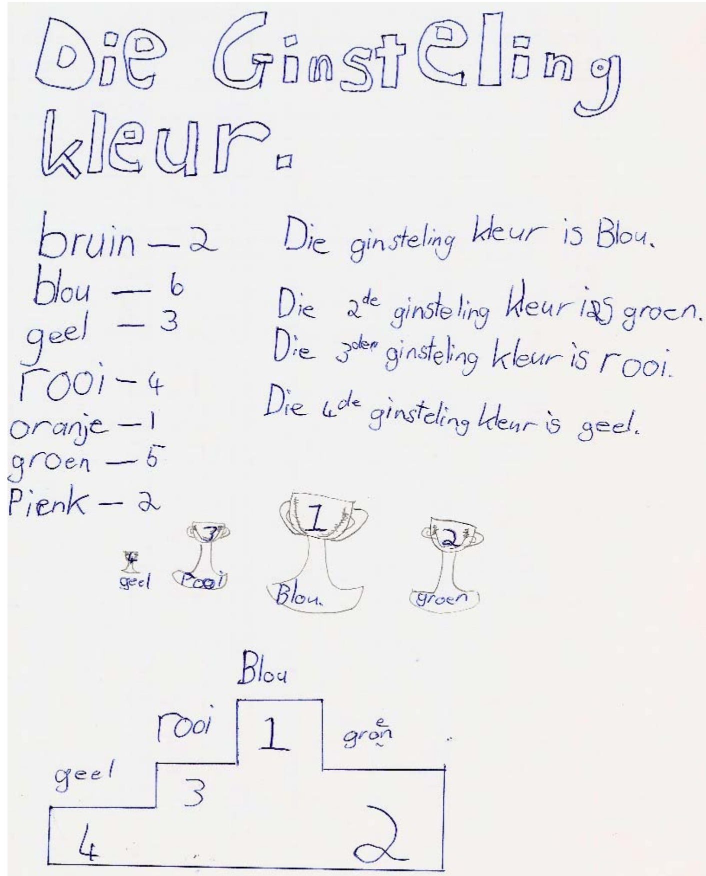


Figure 5.120: Multiple representations

(L107, T1, Gr 6)

The number of multiple representations in each grade is given in Table 5.43. No Grade 5 learners used multiple representations while most of the multiple representations were produced by Grade 6 learners (67%). Grade 4 and 7 learners produced an equal number of multiple representations (13,5% each).

Grade	Number of multiple responses T1	Number of multiple responses T2	Total number of multiple responses T1 and T2
4	4	1	5
5	0	0	0
6	13	7	20
7	3	2	5
Total	20	10	30

Table 5.42: Analysis of multiple responses across grades

All the second representations of these multiple representational sets of Task 1 were on a higher level than that of the first and six of the ten for Task 2 also had a second representation on a higher level. Multiple representations in which the second representation is on a higher level than the first are examples of successful modeling. Cox and Brna (1995:259) report that multiple representations are effective in problem solving and that the use of multiple representations was associated with good performance. For Task 2 four of the ten sets of multiple representations consisted of two inappropriate representations each, pointing to unsuccessful modeling.

The SOLO levels of response in the two tasks show that learners had more difficulty to interpret and represent the quantitative (numerical) data of Task 2 than the qualitative (categorical) data of Task 1. Different factors may have contributed to the difficulty experienced in the transnumeration or interpretation and representation of the data. These factors include the quantitative versus qualitative nature of the data, the contexts of the two tasks and the exposure of learners to data handling activities and different ways of representing data.

The two data tasks comprised of small data sets, but elicited a variety of responses and response levels, indicating that it is not the size or complexity of the data set that produced the rich variety of responses, but rather the nature of the questions asked about the data.

5.5 DATA ARRANGEMENT AND SPATIAL REPRESENTATION IN THE SOLO CONTEXT

When reflecting on the results of arrangement types and spatial representational types against the background of the SOLO Taxonomy framework, interesting facts come to light. Learners needed to arrange data appropriately to be able to represent it as required in the two tasks. The analysis yielded different combinations of clustered, sequential, summative and regrouped arrangement types. When regarding these arrangement types in the context of the SOLO Taxonomy framework, the hierarchical nature of arrangement types becomes apparent. No arrangement or inappropriate arrangement is typical of prestructural representational responses. Clustered and sequential clustered arrangements are typical of responses in the first cycle of the concrete symbolic mode, while summative, sequential summative, regrouped summative and sequential regrouped summative arrangement strategies are found in the second cycle of the concrete symbolic mode.

When considering spatial representations from the perspective of the SOLO Taxonomy framework, the representations showed an overt dissimilarity in mode, some responses representing data pictorially, clearly indicating the ikonic mode, while others were indicative of the concrete-symbolic mode with more abstract representations. Some of the more sophisticated responses however also showed ikonic support, for example filling the bars of a bar graph with different pictures as shown in 5.3.2. This kind of response is another indication that learners chose different problem solving paths when engaging in the tasks, as outlined in 3.7.5. Learners' experience and bias will influence the way in which the interaction between ikonic and concrete symbolic functioning takes place (Watson et al 1995:254). Some learners chose a straight concrete symbolic course of action in their modeling of the task (Fig. 5.115, L139, T1, Gr 6) while others followed an ikonic path throughout, even in multiple modeling cycles (Fig. 5.116 and 5.117, L27, T1, Gr 4).

Figure 5.121: Concrete symbolic problem solving path (L139, T1, Gr 6)

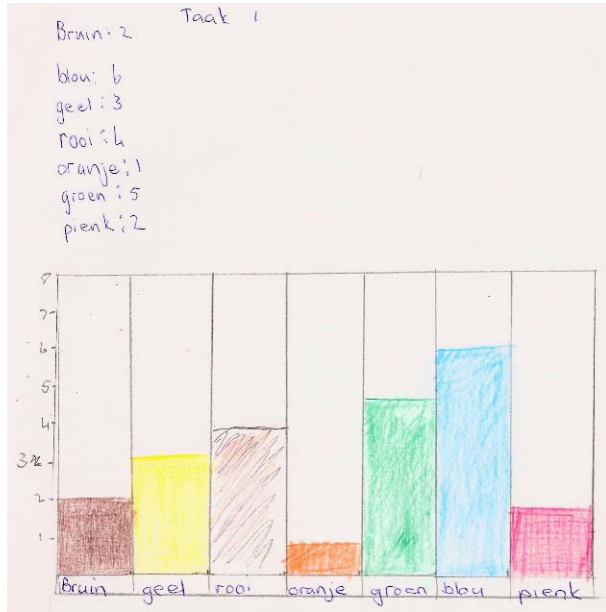


Fig. 5.122: Ikonic problem solving path – 1st of two representations (L27, T1, Gr 4)

Fig. 5.123: Ikonic problem solving path – 2nd of two representations (L27, T1, Gr 4)

5.6 SUMMARY

In Chapter 5 the analysis of the data was described and results considered. The data were analysed in three parts: the arrangement and spatial representation were analysed separately and then the level of representations were categorised according to the SOLO Taxonomy framework described in 3.7. Results for each of the three sections were discussed separately. Relationships between the different sections were discussed, regarding arrangement and spatial representation from a SOLO perspective. Examples of successful modeling as evident in multiple representations were given and the different problem solving paths chosen elucidated.

In Chapter 6 conclusions proceeding from the study will be given and recommendations regarding teaching will be offered. Limitations of the study and recommendations for future research will be discussed.

CHAPTER 6
CONCLUSIONS AND RECOMMENDATIONS

6.1 SUMMARY AND FINDINGS OF THE STUDY

The research focus of this study is to understand the role of representation in mathematical and statistical modeling and problem solving as evident in learners' arrangement and representation of statistical data. The areas of research in the study were modeling and problem solving in mathematics and statistics; the nature and roles of representation; types of data arrangement and data representation; levels of representation of Grade 4-7 learners and the SOLO categorisation of learner representations.

In the literature study (Chapters 2 and 3), the first five objectives of the study were addressed, namely an investigation of the nature of problem solving and modeling, representation in mathematics, data arrangement and representational types in data representation, as well as the SOLO Taxonomy as evaluative tool. The research design (sixth objective) was expounded upon in Chapter 4. Analysis of the data, including the categorising of the statistical thinking level evident from learner responses, as well as the detailed discussion of the findings (objectives 7 and 8) were described in Chapter 5. The synthesis of the empirical investigation and the research question (objective 8) were discussed in Chapter 5 and will again be addressed in this chapter. The ninth objective, namely the detailing of implications of the study for classroom practice and teacher training will also be detailed in this chapter.

Findings of the study will be discussed in three parts, corresponding with the three focal points of the study and differences between this study and other studies in the field will be pointed out. The three sections are data arrangement, data representation and the SOLO statistical thinking levels evident from the representations. The crucial role of representation in mathematics and statistics constitute an essential component in teachers' understanding of how learners think and develop mathematically. A close relationship between modeling, problem solving and representation exists: modeling goes through multiple cycles in which representations play an all-important role. When modeling a problem, learners form internal representations to make sense of the problem, which are then expressed in external form and changed in the phases of the modeling process. Learners had to arrange and represent the data in the process of transnumeration during the modeling of the two tasks of the study. Arrangement and representation strategies occur on different levels of statistical thinking which were analysed using the Solo model.

6.1.1 Data arrangement

The first focus area in the study that will be discussed is the ability to arrange data. This ability to organise data is regarded as critical in the analysis and interpretation of data (Mooney 2002:26; Chick 2003:207, 208). The importance of data arrangement and the paucity of literature on the subject led to the need to analyse arrangement and representational types separately to obtain an insight into learners' intuitive arrangement strategies. The representation tasks in the investigation required learners to arrange data in the process of transnumeration (see 3.6.2; 3.6.4 and 5.2). Categories of arrangement described in the literature had to be extended to make provision for combinations of the different arrangement types found in learner responses.

The findings regarding arrangement correlate with conclusions in other research on classification of data, showing that arrangement types increased in sophistication with increased grade level (Lehrer & Schauble 2000; Mooney 2002). Exposure to more sophisticated types of arrangement in class activities and in the media, as well as better developed higher order thinking skills contribute to this increased sophistication. Learners had difficulty with the transnumeration of Task 2, resulting in an unexpected bump of inappropriate responses in this task. Unfamiliarity with numerical data may be one of the contributing factors, as teachers in the Intermediate and early Senior Phase tend to concentrate on categorical data. Of interest was that arrangement types are often used in combination with each other, for example sequential arrangement per se was not used on its own, but arrangement strategies such as sequential summative arrangement or regrouped sequential summative arrangement were observed. The hierarchical nature of arrangement types becomes apparent when regarded in the context of the SOLO Taxonomy framework (see 5.5). A higher level arrangement strategy points to a higher SOLO level of statistical thinking. No arrangement or inappropriate arrangement is typical of prestructural representational responses. Clustered and sequential clustered arrangements are typical of responses in the first cycle of the concrete symbolic mode, while summative, sequential summative, regrouped summative and sequential regrouped summative arrangement strategies are found in the second cycle of the concrete symbolic mode.

6.1.2 Data representation

The second area of research involved the representational types used by learners. Although types of arrangement and representation were analysed separately, they are inseparable parts of the process of transnumeration. The focus was on spontaneous representations, therefore there was no mention of the word *graph* in the tasks, nor was any representations specifically taught or shown to the learners as examples of possible representations, a fact that renders this study different from other studies in the field.

The findings concerning representation include facts regarding the range and types of representation and the number of representations produced by a learner. The ultimate success of a representation is dependent on successful transnumeration of the data, which includes identifying the message in the data, choosing a representation and arranging the data. Learners experienced more difficulty to transnumerate numerical data than categorical data and the context of the tasks also influenced the transnumeration. While all learners responded to the categorical task, more than 8% did not respond to the numerical task and almost 70% of responses in the numerical task were inappropriate compared to 30% in the categorical task. Learners in the two tasks produced a rich variety of representations which included idiosyncratic, unsophisticated responses as well as standard statistical representations. Lack of statistical tools and lack of exposure to learning activities in which different representational types are used in many cases led to these unsophisticated self-invented representations. An unexpected range of different representational types were found in learner responses. The different representational types found in learner responses were 'no representation', pictures, lists, tables, pictograms, bar graphs, pie graphs and line graphs (see 3.6.5; 4.7.2 and 5.3). An extra category, anomalous responses, had to be added for responses that did not fit into one of these categories, such as the pairing off of data or descriptions and cases where the data values were just added up with no other kind of representation. A number of learners did not respond to the second task. The number of learners not responding to this task decreased with increased grade level, which is an indication that learners in higher grades work faster and/or have more insight into the task and were therefore able to produce a representation. For the categorical task the most popular types were pictures, bar graphs and lists in descending order while the trend is just the reverse for the numerical task, namely lists, bar graphs and pictures also in descending order. Lists, pictures and bar graphs were the most popular representational types used in both tasks. A large percentage of all responses were pictures, which shows that although the target mode for the Grade 4 to 6 learners are the concrete symbolic mode, many learners tend to respond in the ikonic mode to tasks, or use ikonic support in the concrete symbolic mode (see 5.5). Ikonic support played an important role in both tasks, it was however more prominent in Task 1. Learners either chose an ikonic problem solving path from the start or used ikonic support in the concrete symbolic mode. The relationship between arrangement type and representational type also yielded unexpected results: more sophisticated types of arrangement did not necessarily turn out sophisticated representational types, showing that representational types are not hierarchical (see 5.3.10).

6.1.3 Statistical thinking levels as determined by using the SOLO Taxonomy framework

The third focus area was the level of statistical thinking evident in learner representations. The SOLO Taxonomy proved a very useful evaluative tool to determine the statistical thinking level of learners (2.5.5; 3.7; 5.4.5 and 5.5). Responses indicate that learners found it easier to transnumerate the data of the first task than that of the second. While there was evidence that the statistical thinking of more than half the learners has moved beyond the intuitive phase of the first UMR-cycle to consolidation of concepts in the second cycle, only few responded on a higher level in the second task. Factors that contributed to this state of affairs are the context of the two tasks, the quantitative versus qualitative nature of the data in the tasks and the statistical tools or representational skills learners have at their disposal. As grade level increased, sophistication in arrangement strategies increased, with accompanying increase in statistical thinking level. The unexpected big differences in grade level performance between the Grade 5 learners and others are partly explained by their exposure to rich data handling learning activities in the classroom. This fact confirms that the well-planned data handling activities some of the grades were exposed to have developed representational and higher order thinking skills.

Learners also chose different problem solving paths when engaging in the tasks, taking either an ikonic path or a concrete symbolic path, with possible interaction between ikonic and concrete symbolic functioning at different stages in the modeling activity. The modeling process became evident when a learner produced more than one representation. Learners' second representations typically were on a higher level than the first, indicating that the learner had rethought and refined the first response. This however does not imply that no modeling took place when only one representation was created. The format of the tasks did not provide for learners to be questioned about their solutions, so if a learner produced only one representation, it had to be analysed on face value.

The focus on multiple representations is significant difference between this study and others describing SOLO as evaluative tool in categorising data tasks. A significant difference between learner responses in the two tasks was the number of multiple representations found in each of the tasks. More evidence of successful modeling in learner responses was found in Task 1 than in Task 2, again indicating that learners were more comfortable representing categorical than numerical data. Successful modeling was evident from multiple representations in which the second representation was on a higher SOLO level than the first. This was true of all multiple representation sets in the first task and of most of those in the second task. Multiple representations were found to be effective in problem solving and the use of multiple representations was associated with good performance.

The variety of responses and response levels elicited in the two tasks indicate that the nature of the tasks rather than the size of the data set play a conclusive role in data tasks.

6.2 LIMITATIONS OF THE STUDY

The research was done within time constraints imposed by the Education Department and the school involved. A maximum of 60 minutes were available for doing the data tasks and a number of learners did not manage to complete both tasks within this time limit.

The current study cannot provide insight into reasons why learners used the specific representations and chose specific modeling and problem solving paths. Deductions about a learner's process of modeling and level of statistical thinking are limited when only written responses are analysed. Interview protocol tasks yield more insight onto the modeling process, as all the learners' attempts and thinking can be trailed in detail and a learner can be questioned during the process of representing the data. Another aspect that can be incorporated in interview protocol tasks is the introduction of cognitive conflict. When a learner is shown other appropriate representations after he or she has completed his or her own, it may result in reflection on the effectiveness of his or her own representation.

6.3 IMPLICATIONS OF THE RESEARCH

6.3.1 Teaching implications

Teachers need to keep in mind that teaching statistics is open-ended (Burrill 1990:17) and context dependent (Cobb & Moore 1997:801) when planning data handling activities. Learners start to realise the importance of mathematics and statistics in their lives when exposed to open-ended data handling activities set in real-life contexts. Real data can be messy, and learners need to be exposed to problems where they have to make sense of raw data. Mathematics textbooks typically provide learners with organised data and then require them to construct a particular graph. Learners should however be exposed to problem situations within meaningful contexts for them to be able to make the connection between school mathematics and the real world. It is critical that learners realise that some representations are more useful “telling the story of the data” than others (Chick 2003: 207). They should therefore represent the same data set in different ways to be able to see the different stories the same data set can tell and compare the effectiveness of various ways of arranging data for analysis or representation. Whenever possible, learners must be given the opportunity to make decisions about how to represent data verbally, numerically, graphically and symbolically, with ample opportunity for discussion of the special characteristics of each representation (Chick & Watson 2001:106; Burril 1990:17). Discussions about strengths and weaknesses of different representations are invaluable in developing good transnumerative skills. Furthermore, the creation of cognitive conflict as starting point for such discussions may better encourage learners to consider other possibilities except their own. Habits of reflection and speculative thinking are critical factors in representation of statistical data and should be fostered, but can only be developed over time if enough opportunities are created for group and class discussions. Facilitating such discussions on the part of the teacher also is a skill that develops with over time.

6.3.2 Implications for further research

This empirical study was conducted in only one school with a reasonably homogenous population. Future research concerning spontaneous representations could be extended to include a larger number of learners with a wider range of abilities and cultural backgrounds. In such studies context would however need careful consideration.

The role of school taught techniques on intramodal development should be investigated because responses in the second cycle ($U_2M_2R_2$) will to some extent depend on previous experiences and techniques to which learners have been exposed. Such a research project should include a large number of learners and also older learners who have had more data handling experiences.

The fact that learners from four consecutive grades completed the same two tasks in the study provided meaningful insights with regard to the development of representational skills of primary school learners. More research studies should be conducted where consecutive grades are involved in completing the same open-ended tasks in rich real-life contexts. This point is supported by Dr Rosemary Callingham from the University of New England in Australia when she commented on the usefulness of this study in personal communication with the researcher:

> It will be useful to have something developmental applying to representation and graphing particularly. It opens the possibility of addressing other kinds of graphing tasks too (not just statistical ones) (2005: e-mail).

Since the real world handling of data in the work place outside the school is inconceivable without the use of technology such as calculators or computers, there is a need for similar studies focusing on learners' data handling abilities with such technologies.

6.4 CONCLUSION

This study contributes to the research literature in the field of representation and statistical thinking. The analysis and results led to a more integrated picture of Grade 4-7 learners' representation of statistical data and of the statistical thinking levels evident in their representations.

The SOLO model which incorporates a structural approach as well as a multimodal component proved valuable in the analysis of responses. The acknowledgement of different problem solving paths and the contribution of ikonic support in the concrete symbolic mode possible with the use of this model promote a more in-depth analysis of responses.

BIBLIOGRAPHY

BAKKER, A. 2004 *Design research in statistics education on symbolizing and computer tools*. Utrecht: CD-ß Press

BATANERO, C. 2002 Opening address at ICOTS6, Preface ICOTS6 Proceedings CD.

BENDICK, J. & LEVIN, M. 1973 *Mathematics illustrated dictionary*. London: Kaye & Ward Ltd.

BEVER, T.G. 1986 The aesthetic basis for cognitive structures. In M. Brand & R.M. Harnish (eds), *The representation of knowledge and belief*, 314-356. Tucson: University of Arizona 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. 1991 Multimodal learning and the quality of intelligent behaviour. In Rowe, H.A.H. (ed), *Intelligence: reconceptualization and measurement*. New Jersey: Lawrence Erlbaum Associates.

BOONE, W. & ROGAN, J. 2005 Rigour in quantitative analysis: The promise of Rasch anlaysis techniques. *African Journal of Research in Mathematics, Science and Technology Education* 9(1) 25-38

BRANCA, N. & KILPATRICK, J. 1972 The consistency of strategies in the learning of mathematical structures. *Journal for Research in Mathematics Education*, 3, 132-140.

BRIGHT, G.W. & FRIEL, S.N. 1998 Graphical representations: helping students interpret data. In S.P.Lajoie, *Reflections on Statistics:learning, teaching and assessment in Grades K-12*, 63-88. Mahwah, New Jersey: Lawrence Erlbaum Associates.

BURRILL, G. 1990 Implementing the Standards: Statistics and probability. *Mathematics Teacher* 83(2) 113-118

BURRILL, G.& ROMBERG, T. 1998 Statistics and probability for the middle grades: Examples from Mathematics in Context. In S.P. Lajoie (ed), *Reflections on statistics: Learning, teaching and assessment in Grades K-12*, 33-62. Mahwah, New Jersey: Lawrence Erlbaum Associates.

BYRNES, J.P. 1992 The conceptual basis of procedural learning. *Cognitive development*, 7, 235-237.

BYRNES, J.P. 1999 The nature and development of representation: forging a synthesis of competing approaches. In I.E. Sigel (ed), *Development of mental representation: theories and applications*, 273-293. Mahwah, New Jersey: Lawrence Erlbaum Associates.

CAMPBELL, K.; WATSON, J. & COLLIS, K. 1992 Volume measurement and intellectual development. *Journal of Structural Learning and Intelligent Systems* 11, 279-298.

CARPENTER, T.P. & MOSER, M. 1984 The acquisition of addition and subtraction concepts in grades one through three. *Journal for Research in Mathematics Education* 15, 179-202.

CARPENTER, T.; FENNEMA, E.; FRANKE, M.L.; LEVI, L. & EMPSON, S.B. 1996 *Children's mathematics: Cognitively guided instruction*. Portsmouth, NH: Heinemann

CASSIM, I; GEJU, N.; NEL, I.; WESSELS, D.C.J. & WESSELS, H.M. 2006 (in press) *Mathematics through issues. Learners' book Grade 8.* Pretoria: Clever Books

CHARLES, C.M. & MERTLER, C.A. 2002 *Introduction to educational research.* (4th ed) Boston: Allyn and Bacon.

CHICK, H.L. 2000 Young adults making sense of data. In J. Bana & A Chapman (eds): *Mathematics education beyond 2000*. Proceedings of the 23rd annual conference of the Mathematics Education Research Group of Australasia (MERGA), 157-164.

CHICK, H.L. 2003 Transnumeration and the art of data representation. In L. Bragg; C. Campbell; G. Herbert & J. Mousley (eds): *Mathematics Education Research: Innovation, Networking, Opportunity* (MERINO). Proceedings of the 26th annual conference of the Mathematics Education Research Group of Australasia (MERGA), 207-214.

CHICK, H.L. & WATSON, J.M. 1998 Showing and telling: primary students' outcomes in data representation and interpretation. Proceedings of the 20th International Conference of the Mathematics Education Research Group of Australasia (MERGA), Vol 1, 153-160.

CHICK, H.L. & WATSON, J.M. 2001 Data representation and interpretation by primary school students working in groups. *Mathematical Education Research Journal* 13(2), 91-111.

CHRISTIANS, C.G. 2000 Ethics and politics in qualitative research. In N.K.Denzin & A.S. Lincoln (eds), Handbook of qualitative research (2nd ed). Thousand oaks, CA: Sage Publications.

CIFARELLI, V.V. 1998 The development of mental representations as a problem solving activity. *Journal of Mathematical Behavior* 17(2), 239-264.

COBB, G.W. & MOORE, D.S. 1997 Mathematics, Statistics and teaching. *The American Mathematical Monthly* 104, 801-824.

COBB, P; WOOD, T., YACKEL, E., NICHOLLS, J., WHEATLEY, G., TRIGATTI, D. & PERLWITZ, M. 1991 Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education* 22(1), 3-29.

COBB, P; YACKEL, E & MCCLAIN, K (eds) 2000 *Symbolizing and communicating in mathematics classrooms. Perspectives on discourse, tools and instructional design.* Mahwah, NJ: Hillsdale: Lawrence Erlbaum.

COCKING, R.R. 1999 Introduction. In I.E. Sigel (ed), *Development of mental representation*. vii-xiii. Mahwah, NJ: Lawrence Erlbaum Associates.

COLLIS, K.F. & BIGGS, J.B. 1986 Using the SOLO taxonomy. *SET: Research information for teachers* 1(1). Australian Council for Educational Research.

COLLIS, K.F. & ROMBERG, T.A. 1991 Assessment of mathematical performance: an analysis of open-ended test items. In M.C. Whitrock & E. Baker (Eds.), *Testing and cognition* 82-130. Englewood Cliffs, NJ: Prentice Hall.

CONFREY, J. 2000 Leveraging constructivism to apply to systemic reform. *Nordisk Mathematik Didaktik (Nordic Studies in Mathematics Education)* 8(3), 7-30.

COX, R & BRNA, P. 1995 Supporting the use of external representations in problem solving: the need for flexible learning environments. *Journal of Artificial Intelligence in Education* 6(2), 239-302.

CRESWELL, J.W. 2002 *Educational research. Planning, conducting and evaluating quantitative and qualitative research*. Upper Saddle River: Merrill Prentice Hall.

DAVIS, R.B. 1984 *Learning mathematics: The cognitive science approach to mathematics education.* Norwood, NJ: Ablex.

DeBELLIS, V.A. & GOLDIN, G. 1993 analysis of interactions between affect and cognition in elementary school children during problem solving. In J.R. Becker & B.J.Pence (eds) Proceedings of the 15th International Conference of PME-NA, Vol 2, 56-62.

DeBELLIS, V.A. & GOLDIN, G. 1997 In E. Pehkonen (ed), Proceedings of the 21st International Conference of PME, Vol 2, 209-216.

DENIS, M. 1994 *Image and cognition* (2nd ed). Paris: Presses Universitaires de France.

DENZIN, N.K. & LINCOLN, Y.S. (eds) 2000 *Handbook of qualitative research* (2nd ed). Thousand oaks, CA: Sage Publications.

DEPARTMENT OF EDUCATION (DoE) 2002 *Revised National Curriculum Statement Grades R-9 (Schools), Mathematics.* Pretoria: Department of Education.

DEPARTMENT OF EDUCATION (DoE) 2003 *National Curriculum Statement Grades 10-12 (General), Mathematics.* Pretoria: Department of Education.

DEWEY, J. 1926 *Democracy and education*. New York: MacMillan.

DEWEY, J. 1929 *The quest for certainty*. New York: Minton, Balch & Company.

DIENES, Z.P. & JEEVES, M.A. 1965 *Thinking in structures*. London: Hutchinson Educational.

DIENES, Z.P. & JEEVES, M.A. 1970 *The effects of structural relations on transfer*. London: Hutchinson Educational.

Di SESSA, A. 1983 Phenomenology and the development of intuition. In D. Gentner & A. Stevens (eds), *Mental models*, 15-33 Hillsdale, NJ: Erlbaum.

DOERR, H.M. & ENGLISH, L.D. 2001 A modeling perspective on students' learning through data analysis. The Proceedings of the 25th International Conference of PME, Vol 2, 361-368.

DOERR, H.M. & ENGLISH, L.D. 2003 A modeling perspective on students' mathematical reasoning about data. *Journal for Research in Mathematics Education* 34(2), 110-136.

DOSSEY, J.A.; McCRONE, S.; GIORDANO, F.R. & WEIR, M.D. 2002 Mathematics methods and modeling for today's mathematics classroom: A contemporary approach to teaching Grades 7-12. Spain: Brooks/Cole.

DREYFUSS, T. 1991 Advanced mathematical thinking processes. In D.Tall (ed), *Advanced mathematical thinking*. Dordrecht: Kluwer.

EISENHART, M.; BORKO, H.; UNDERHILL, R.; BROWN, C.; JONES, D. & AGARD, P. 1993 Conceptual knowledge falls through the cracks: complexities of learning to teach mathematics for understanding. *Journal for Research in Mathematics Education* 24(1), 8-40.

ENGLISH, L. 1995 Cognitive Psychology and Mathematics Education. In English, LD & Halford, GS 1995 *Mathematics Education: Models and processes* 1-20. Mahwah, New Jersey: Lawrence Erlbaum.

ENGLISH, L.D. & HALFORD, G.S. 1995 *Mathematics Education: Models and processes*. Mahwah, New Jersey: Lawrence Erlbaum.

ERNEST, P. 1991 *The philosophy of Mathematics Education.* Basingstoke, England: Falmer Press.

FENNEMA, E., CARPENTER, T.P., FRANKE, M.L., LEVI, L., JACOBS, V.R. & EMPSON, S.B. 1996 A longitudinal study of research to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education* 27, 403-434.

FILSTEAD, W.J. 1990 Q*ualitative methodology*. Chicago: Markham.

FRIEL, S.N. & BRIGHT, G.W. 1998 Teach-Statistic: A model for professional development in data analysis and statistics for teachers K-6. In S.P.Lajoie (ed), *Reflections on Statistics: learning, teaching and assessment in Grades K-12*, 89-120. Mahwah, New Jersey: Lawrence Erlbaum Associates.

FRIEL, S.N.; CURCIO, F.R. & BRIGHT, G.W. 2001Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education* 32, 124-158.

FREUDENTHAL, H. 1973 *Mathematics as an educational task*. Dordrecht: D Reidel

FRIEL, S.N.; CURCIO, F.R. & BRIGHT, G.W. 2001 Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education* 32(2), 124-158.

FUSON, K.C. 1986 Teaching children to subtract by counting up. *Journal for Research in Mathematics Education* 17, 172-189

GAGNé, R. 1970 *The conditions of learning (second edition).* New York: Holt, Rinehart & Winston.

GAL, I. 1998 Assessing statistical knowledge as it relates to students' interpretation of data. In S.P. Lajoie (ed), *Reflections on statistics: Learning, teaching and assessment in Grades K-12*, 275-295. Mahwah, New Jersey: Lawrence Erlbaum Associates.

GAL, I. & GARFIELD, J. 1997 Curricular goals and assessment challenges in Statistics Education. In I. Gal & J. Garfield (eds), The assessment challenge in Statistics Education. Amsterdam: IOS Press

GARDNER, H. 1991*The unschooled mind: how children think and how schools should teach.* New York; Basic Books.

GERBER, R., BOULTON-LEWIS, G. & BRUCE, C. 1995 Children's understanding of graphic representations of quantitative data. *Learning and Instruction* 5, 77-100.

GOLDIN, G.A. 1983 Levels of language in mathematical problem solving. In J.C. Bergeron& N. Herscovics (eds), Proceedings of the 5th International Conference of PME-NA, Vol 2, 112-120.

GOLDIN, G.A. 1984 The structure of heuristic processes. In J.M. Moser (ed), Proceedings of the 6th International Conference of PME-NA, 183-188.

GOLDIN, G.A. 1987 Cognitive representational systems for mathematical problem solving. In C. Janvier (ed), *Problems of representation in the teaching and learning of mathematics*. Hillsdale, NJ: Erlbaum.

GOLDIN, G.A. 1988 The development of a model for competence in mathematical problem solving. In A. Borbás (ed), Proceedings of the 12th international Conference of PME, Vol 2, 358-365.

GOLDIN, G.A. 1996 Theory of mathematics education: The contributions of constructivism. In L.P. Steffe; P. Nesher; P. Cobb; G.A. Goldin & Greer, B. (eds), *Theories of Mathematical learning* 303-306. Mahwah, New Jersey: Lawrence Erlbaum Associates.

GOLDIN, G.A. 1998 Representational systems, learning, and problem solving in mathematics. *Journal of Mathematical Behaviour* 17(2), 137-165.

GOLDIN, G.A. 2002 Representation in mathematical learning and problem solving. In L.D. English (ed), *Handbook of International research in Mathematics Education*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

GOLDIN, G.A & GRAMICK, J. 1980 Algorithmic structure and children's error patterns in arithmetic. In R. Karplus (ed), The Proceedings of the 4th International Conference of PME, 14-23

GOLDIN, G.A. & JANVIER, C. (eds) 1998 Representations and the psychology of mathematics education: Part 1. *Journal of Mathematical Bahavior* 17(1), Special Issue.

GOLDIN, G.A. & KAPUT, J.J. 1996. A joint perspective on the idea of representation in learning and doing mathematics. In L.P. Steffe; P. Nesher; P. Cobb; G.A. Goldin & Greer, B. (eds), *Theories of Mathematical learning* 397-430. Mahwah, New Jersey: Lawrence Erlbaum Associates.

GOLDIN, G.A. & SHTEINGOLD, N. 2001 Systems of representations and the development of mathematical concepts. In A.A. CUOCO & F.R. CURCIO (eds), *The roles of representation in school mathematics*. Reston, VA: NCTM.

GRAVEMEIJER, K.P.E. 1994 *Developing realistic mathematics education.* Utrecht: CD Bèta Press.

GRAVEMEIJER, K.P.E. 1998 Emergent models as an instructional-design heuristic. Paper presented at the International Conference on Symbolizing and Modeling in Mathematics Education, Freudenthal Institute, Utrecht, The Netherlands.

GRAVEMEIJER, K.P.E. 1999 How emergent models may foster the constitution of formal mathematics. *Mathematical Thinking and Learning* 1, 155-177.

GRAVEMEIJER, K.P.E.; COBB, P.; BOWERS, J. & WHITENACK, J. 2000

Symbolizing, modeling and instructional design. In P. Cobb, E. Yackel & K. McClain (eds), *Symbolizing and communicating in mathematics classrooms*. 225-274. Mahwah, New Jersey: Lawrence Erlbaum Associates.

HALFORD, G.S. 1993 *Children's understanding: the development of mental models.* Hillsdale, NJ: Lawrence Erlbaum

HALFORD, G.S. 1995 Cognition and cognitive development. In English, LD & Halford, GS 1995 *Mathematics Education: Models and processes* 21-56. Mahwah, New Jersey: Lawrence Erlbaum.

HALFORD, G.S. 1999 The properties of representations used in higher cognitive processes: developmental implications. In I.E. Sigel (ed), *Development of mental representation: theories and applications,* 147-168. Mahwah, New Jersey: Lawrence Erlbaum Associates.

HERSCOVICS, N. 1989 The description and analysis of mathematical processes. In C.A. Maher; G.A. Goldin & R.B. Davis (eds), Proceedings of the 11th International Conference of PME-NA, Vol 2, 1-28.

HERSCOVICS, N. & BERGERON, J.C. 1983 Models of understanding. *Zentralblatt für Didaktik der Mathematik* 8312, 75-83.

HERSCOVICS, N. & BERGERON, J.C. 1984 A constructivist versus a formalist approach in the teaching of mathematics. In B. Southwell et al (eds), Psoceedings of the 7th International Conference of PME, 190-196 Sydney, Australia: University of Sydney.

HERSCOVICS, N. & BERGERON, J.C. 1988 An extended model of understanding. In M.J. Behr; C.B. Lacampagne & M.M. Wheeler (eds), Proceedings of the 10th International Conference of PME-NA, 15-22.

HIEBERT, J & CARPENTER, TP 1992 Learning and teaching with understanding. In DA Grouws (ed), *Handbook of Research on mathematics teaching and learning,* 65-97 New York: McMillan.

HIEBERT, J.; CARPENTER, T.P.; FENNEMA, E.; FUSON, K.C.; WEARNE, D.; MURRAY, H.; OLIVIER, A. & HUMAN, P. 1997 Making sense: Teaching and learning mathematics with understanding. Portsmouth: Heinemann.

HITTLEMAN, D.R. & SIMON, A.J. 1992. *Interpreting educational research*. (3rd ed) Upper Saddle River, New Jersey: Merrill Prentice Hall.

HOLMES E.E. 1995 New directions in elementary school mathematics. Interactive teaching and learning. Englewood Cliffs, NJ: Merrill

HORVATH, J. & LEHRER, R. 1998 A model-based perspective on the development of children's understanding of chance and uncertainty. In S.P. Lajoie (ed), *Reflections on statistics: Learning, teaching and assessment in Grades K-12*, 121-148. Mahwah, New Jersey: Lawrence Erlbaum Associates.

INHELDER, R. & PIAGET. J. 1964 *The early growth of logic in the child*. New York: Basic Books

JANVIER, C. (ed) 1987 *Problems of representation in the teaching and learning of mathematics.* Hillsdale, NJ: Erlbaum.

JONES, G.A.; MOONEY, E.S.; LANGRALL, C.W. & THORNTON, C.A. 2002 Students' individual and collective statistical thinking. Proceedings of the 6th International Conference on Teaching Statistics (ICOTS 6), Cape Town (electronic copy – CD).

JOHNSON, K. & HERR, T. 2001 *Problem solving strategies*. Emeryville: Key Curriculum Press.

JOHNSON, Y. & HOFBAUER, P. 2002 Describing middle school students' organization of statistical data. In D.S. Mewborn; P. Sztajn; D.Y. White; H.G. Wiegel; R.L. Bryant & K Nooney, The Proceedings of the 24th International Conference of PME-NA, Vol 3, 1281-1291.

JOHNSON-LAIRD, P.N. 1983 *Mental models: Towards a cognitive science of language, inference, and consciousness.* Cambridge, M.A.: Harvard University Press.

JOHNSON-LAIRD, P.N. & BYRNE, R.M.J. 1991 *Deduction. Essays in cognitive psychology.* Lawrence Erlbaum: Hove.

KAPUT, J.J. 1987a Toward a theory of symbol use in mathematics. In C. Janvier (ed), *Problems of representation in the teaching and learning of thematics*. 159-195 Hillsdale, NJ: Erlbaum.

KAPUT, J.J. 1987b Representation systems and mathematics. In C. Janvier (ed), *Problems of representation in the teaching and learning of mathematics*, 19-26. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

KAPUT, J.J. 1991 Notations and representation as mediators of constructive processes. In E. von Glasersfeld (ed), *Radical constructivism in mathematics education*. 53-74 Dordrecht, Holland: Kluwer Academic Publishers.

KAPUT, J.J. 1998 Representations, inscriptions, descriptions and learning: a kaleidoscope of windows. *Journal of Mathematical Behaviour* 17(2), 265-281.

KILLEN, R. 2004 Writing outcomes, performance indicators and assessment criteria. In J.G. Maree & W.J. Fraser (eds), *Outcomes-based assessment*, 65-86. Sandown: Heinemann.

KOSSLYN, S.M. 1980 *Image and mind*. Cambridge, MA: Harvard University Press.

KRUTETSKII, V.A. 1976 The psychology of mathematical abilities in schoolchildren. Chicago:University of Chicago Press.

LAJOIE, S.P. 1998 *Reflections on statistics: learning, teaching and assessment in Grades K-12.* Mahwah, New Jersey: Lawrence Erlbaum & Associates.

LAJOIE, S.P.; LAVIGN, N.C.; MUNSIE, S.D.; & WILKIE, T.V. 1998 Monitoring student progress in statistics. In I. Gal & J.B. Garfield (eds), *The assessment challenge in statistics education*, 199-231. Amsterdam: IOS Press.

LAKOFF, G. 1987 *Women, fire and dangerous things*. Chicago, IL: University of Chicago Press.

LEHRER, R. & ROMBERG, T. 1996 Exploring children's' data modeling. *Cognition and Instruction* 14(1), 69-108.

LEHRER, R. & SCHAUBLE, L. 2000. Inventing data structures for representational purposes: Elementary grade students' classification models. *Mathematical Thinking and Learning* 2, 51-74.

LESH, R. 1981 Applied mathematical problem solving. *Educational Studies in Mathematics*. 12, 235-264.

LESH, R. 1999 The development of representational abilities in middle school mathematics. In I.E. Sigel (ed), *Development of mental representation: theories and applications,* 323-349. Mahwah, New Jersey: Lawrence Erlbaum Associates.

LESH, R.; CARMONA, G. & POST, T. 2002 Models and modeling. In D.S. Mewborn; P. Sztajn; D.Y. White; H.G. Wiegel; R.L. Bryant & K Nooney, The Proceedings of the 24th International Conference of PME-NA, Vol 1, 89-98.

LESH, R. & DOERR, H.M. 2000 Symbolizing, communicating and mathematizing: Key components of models and modeling. In P. Cobb, E. Yackel & K. McClain (eds), *Symbolizing and communicating in mathematics classrooms*. 361-384. Mahwah, New Jersey: Lawrence Erlbaum Associates.

LESH, R.; LANDAU, M. & HAMILTON, E. 1983 Conceptual models and applied mathematical problem solving research. In R. Lesh & M. Landau (eds) *Acquisition of mathematics concepts and processes,* 263-343. New York: Academic Press.

LESH, R.; POST, T. & BEHR, M. 1987 Representations and translations among representations in mathematics learning and problem solving. In C. Janvier (ed), *Problems of representation in the teaching and learning of mathematics*, 33-40. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

LESTER, F.K. & GAROFALO, J. (eds) 1982 Mathematical problem solving: Issues in research. Philadelphia, PA: Franklin Institute Press.

LEVINS, L. & PEGG,J. 1993 Students' understanding of concepts related to plant growth. *Research in Science Education* 23, 165-173.

LI, K.Y. & SHEN, S.M. 1992 Students' weaknesses in statistical projects. *Teaching Statistics*, 24(1), 2-8.

MARTINEZ, M.E. 1999 Cognitive representations: distinctions, implications and elaborations. In I.E. Sigel (ed), *Development of mental representation: theories and applications,* 13-32. Mahwah, New Jersey: Lawrence Erlbaum Associates.

McKENZIE, D.L. & PADILLA, M.J. 1986 The construction and validation of the test of graphing in science (TOGS). *Journal of Research in Science Teaching* 23(7), 571-579.

McKLINTOCK, C.E. 1984 Heuristic processes as task variables. In G.A. Goldin & C.E. McKlintock (eds), *Task variables in mathematical problem solving*. 171-234. Philadelphia, PA: Franklin Institute Press.

McLEOD, D.B. & ADAMS, V.M. (eds) 1989 *Affect and mathematical problem solving: a new perspective.* New York, NY: Springer Verlag.

MOONEY, E.S. 2002 A framework for characterising middle school students' statistical thinking. *Mathematical Thinking and Learning*, 4(1), 23-63.

MOONEY, E.S.; LANGRALL, C.W.; HOFBAUER, P.S. & JOHNSON, Y.A. 2001 Refining a framework on middle school students' statistical thinking. Proceedings of the 23rd International Conference of PME-NA, Vol 1, 437-447.

MOORE, D.S. 1990 Statistics for all: Why, what and how? In D. Vere-Jones (ed), Proceedings of the 3rd International Conference on Teaching Statistics (ICOTS3), Vol 1, 423-428

MOORE, D.S. 1997 New pedagogy and new content: The case for statistics. *International Statistical Review,* 65, 123-165.

MORITZ, J. 2002 Study times and test scores: what students' graphs show. *Australian Primary Mathematics Classroom* 7(1), 24-31.

MUDALY, V 2004 *The role and use of Sketchpad as a modeling tool in secondary schools*. Unpublished D.Ed Thesis. Durban, University of KwaZulu-Natal.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS 1989 *Curriculum and evaluation standards for school mathematics*. NCTM: Reston, Virginia.

NATIONAL COUNCIL OF TEACHERS OF MATHEMATICS 2000 *Principles and Standards for school mathematics*. NCTM: Reston, Virginia.

NESHER,P. & KILPATRICK, J. 1991 *Mathematics and cognition*. Cambridge, UK: Cambridge University Press.

NISBETT, R.E. et al 1983 The use of statistical heuristics in everyday inductive reasoning. *Psychological Review* 90, 339-363.

O'DAFFER, P.; CHARLES, R.; COONEY, T.; Dossey, J. & SCHIELAK, J. 2002 (2nd ed) *Mathematics for elementary school teachers*. Boston: Addison-Wesley.

OLSON, D & TORRANCE, N 1996 *The handbook of education and human development. New models of learning, teaching and schooling.* Cambridge: Blackwell Publishers.

PARKER, J. & WIDMER, C.C. 1992 Teaching mathematics with technology. *Arithmetic teacher* 39(8), 48-52

PEGG, J. 1992 Assessing students' understanding at the primary and secondary level in mathematical sciences. In J. Izard & M. Stephens (eds), *Reshaping* *assessment practice: Assessment in the mathematical sciences under challenge*, 368-385. Melbourne: Australian Council of Educational Research.

PEGG, J. & DAVEY, G. 1998 Interpreting student understanding in geometry: a synthesis of two models. In R. Lehrer & D. Chazan (eds), *Designing learning environments for developing understanding of geometry and space.* Mahwah, New Jersey: Lawrence Erlbaum & Associates.

PEREIRA-MENDOZA, L. 1995 Graphing in the primary school: Algorithm versus comprehension. *Teaching Statistics* 17(1), 2-6.

PIAGET, J. 1951 *Play, dreams and imitation*. Melbourne: Australia

PIAGET, J. 1965 *The child's conception of number*. New York, NY: W.W. Norton & Co (First published in French, 1941, under the title *La genèse du nombre chez l'enfant*).

PIAGET, J. 1967 *Six psychological studies.* New York, NY: Random House.

PIAGET, J. 1969 *Science of education and the psychology of the child.* New York, NY: Viking Press.

PIAGET, J. 1970 *Structuralism.* New York: Basic books.

PIAGET. J.& INHELDER, R. 1956 *The child's conception of space.* London: Routledge & Kegan Paul (Original work in French 1948).

PIAGET. J. & INHELDER, R. 1971 *Mental imagery in the child: A study of the development of imaginal representation.* London: Routledge & Kegan-Paul.

POLYA, G 1945 *How to solve it?* Princeton, NY: Princeton University Press.

POLYA, G. 1965 *Mathematical discovery: On understanding, learning and teaching and problem solving.* Vol I & II. New York, NY: John Wiley & Sons.

PRESMEG, N. 1986 Visualisation in high school mathematics. *For the learning of mathematics* 6(3), 42-46.

PRESMEG, N. 1992 Prototypes, metaphors, metonymies and imaginative rationality in high school mathematics. *Educational Studies in Mathematics.* 23, 595-610.

PYLYSHYN, Z. 1973 What the mind's eye tells the mind's brain: a critique of mental imagery. *Psychological Bulletin* 80, 1-24.

READER'S DIGEST ILLUSTRATED OXFORD DICTIONARY. 1998 Oxford: Oxford University Press.

READING, C & PEGG, J. 1995 Teaching statistics: Background and implications. In L. Grimison & J. Pegg, *Teaching secondary school mathematics: theory into practice.* Marrickville, Australia: Harcourt Brace & Company.

RESNICK, L.B. 1983 Toward a cognitive theory of instruction. In S. Paris, G.M. Olson & H.W. Stevenson (eds), *Learning and motivation in the classroom*, 5-38. Hillsdale, NJ: Lawrence Erlbaum Associates.

ROMBERG, T. 1999 Understanding reality mathematically. Unpublished manuscript, personally made available by the author (2005).

ROMBERG, T.A.; CARPENTER, T.P. & KWAKO, J. 2005 Standards-based reform and teaching for understanding. In T.A. Romberg & T.P. Carpenter & Dremock, F. (Eds*), Understanding mathematics and science matters*, 3-26. Mahwah, New Jersey: Lawrence Erlbaum Associates.

SCHAEFFER, R.L.; WATKINS, A.E. & LANDWEHR, J.M. 1998. In S.P. Lajoie (ed), *Reflections on statistics: Learning, teaching and assessment*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

SCHAEFFER, R.L. 2000. Statistics for a new century. In Burke, M.J. & Curcio, F.R. (eds), *Learning Mathematics for a new century*. Reston,VA: NCTM.

SCHOENFELD, A.H. 1985 *Mathematical problem solving.* New York: Academic Press.

SCHOENFELD, A.H. 1987 *Cognitive science and mathematics education.* Hillsdale, NJ: Erlbaum.

SCHOLNICK, E.K. 1999 Representing logic. In I.E. Sigel (Ed), *Development of mental representation: theories and applications.* 113-128. Mahwah, New Jersey: Lawrence Erlbaum Associates.

SCHROEDER, T.L. & LESTER, F.K. 1989 Developing understanding in mathematics via problem solving. In P.R. Trafton & A.P. Schulte (eds), *New directions in elementary school mathematics.* Reston, VA: NCTM.

SHAUGHNESSY, M. & WATSON, J. 2003 Stochastical thinking, learning and teaching. The Proceedings of the 27th International Conference of PME, Vol 1, p 192.

SHAUGHNESSY, JM, GARFIELD, J & GREER, B 1996 Data handling. In A.J. Bishop; K. Clements; C. Keitel; J. Kilpatrick & C. Laborde (eds), *International Handbook of Mathematics Education* 205-237 Dordrecht, The Netherlands: Kluwer.

SHULMAN, L. 1986 Those who understand: Knowledge growth in teaching. *Educational Researcher* 15(2), 4-14.

SIGEL, I. E. 1999. Approaches to Representation as a Psychological Construct: A Treatise in Diversity. In I.E. Sigel (Ed) *Development of mental representation: theories and applications.* 3-12. Mahwah NJ: Lawrence Erlbaum.

SILVER, E.A. (ed) 1985 *Teaching and learning mathematical problem solving: multiple research perspectives.* Hillsdale, NJ:Erlbaum.

SKINNER, B.F. 1953 *Science and human behavior*. New York: Free Press.

SKINNER, B.F. 1974 *About behaviorism.* New York: Alfred A Knopf.

SNEE, R.D. 1990 Statistical thinking and its contribution to quality. *The American Statistician* 44, 116-121.

STEYN, A.G.W.; SMIT, C.F.; DU TOIT, S.H.C. & STRASHEIM, C. 2000 *Moderne Statistiek vir die praktyk*. Pretoria: Van Schaik.

TREFFERS, A. 1993 Wiskobas and Freudenthal: Realistic mathematics education. *Educational Studies in Mathematics* 25, 89-108.

TRURAN, K. 1997 Beliefs about teaching Stochastics held by primary pre-service teaching students. In F. Biddulph & K. Carr (eds), Proceedings of the 20th International Conference of the Mathematics Education Research Group of Australasia (MERGA), 538-545.

TRURAN, K.; GREER, B. & TRURAN, J. 2001 Discussion Group: Stochastical thinking, learning and teaching. The Proceedings of the 25th Conference of PME, Vol 2, 258.

VAN DE WALLE, J.A. 1996 (4th ed) *Elementary and Middle School Mathematics: Teaching developmentally.* New York: Longman.

VAN DEN HEUVEL-PANHUIZEN, M.1 2001 *Children learn mathematics: A learning-teaching trajectory with intermediate attainment targets*. Utrecht: Freudenthal Institute.

VERGNAUD, G. 1987 Problems of representation in the teaching and learning of mathematics: Conclusion. In Janvier, C (ed), *Problems of representation in the teaching and learning of mathematics*. Hillsdale, New Jersey: Lawrence Erlbaum & Associates.

VERGNAUD, G. 1998 A comprehensive theory of representation for mathematics education. *Journal of mathematical behaviour*, 17(2), 167-181

VON GLASERSFELD, E. 1987a Learning as a constructive activity. In C. Janvier (ed), *Problems of representation in the teaching and learning of mathematics*. 3-18. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

VON GLASERSFELD, E. 1987b Preliminaries to any theory of representation. In C. Janvier (ed), *Problems of representation in the teaching and learning of mathematics*. 215-226. Hillsdale, New Jersey: Lawrence Erlbaum Associates.

VON GLASERSFELD, E. 1990 An exposition of constructivism: Why some like it radical. In R.B. Davis, C.A. Maher & N. Noddings (eds), *Journal for Research in Mathematics Education Monograph No 4: Constructivist views on the teaching and learning of mathematics*. Reston, VA: NCTM.

VON GLASERSFELD, E. 1996 Aspects of radical constructivism and its educational recommendations. In L.P. Steffe; P. Nesher; P. Cobb; G.A. Goldin & Greer, B. (eds), *Theories of Mathematical learning*. Mahwah, New Jersey: Lawrence Erlbaum Associates.

WARES, A; JONES, G.A.; LANGRALL, C.W. & THORNTON, C.A. 2000 Young children's statistical thinking: A teaching experiment. In Fernández, M.L. (ed) Proceedings of the 22nd International Conference of the PME-NA 22, Vol 1, 321-328.

WATKINS, A.E.; SCHAEFFER, R.L. & COBB, G.W. 2004 *Statistics in action*. Emeryville, CA: Key Curriculum Press.

WATSON, J. 1998. Professional development for teachers of probability and statistics. *International Statistical Review* 66(3), 271-289.

WATSON, J.M. & CALLINGHAM, R.A. 1997 Data cards: An introduction to higher order processes in data handling. *Teaching Statistics* 19, 12-16.

WATSON, J.M., COLLIS, K.F., CALLINGHAM, R.A. & MORITZ, J.B. 1995 A model for assessing higher order thinking in statistics. *Educational research and evaluation* 1(3), 247-275.

WATSON, J.M., COLLIS, K.F. & CAMPBELL, K.J. 1995 Developmental structure in the understanding of common and decimal fractions. *Focus on Learning Problems in Mathematics* 17, 1-24.

WATSON, J.M., & MORITZ, J.B. 2001 Development of reasoning with pictographs: representing, interpreting and predicting. *Educational Studies in Mathematics* 48, 47-81.

WESSELS, D.C.J. 2000 Konstruktivisme en Intuïsionisme: Ontwikkelingslyne in Wiskundeonderwys. *Tydskrif vir Christelike Wetenskap/Journal for Christian Scholarship* 36(1&2), 119-148.

WESSELS, D.C.J. & KWARI, R. 2003 The nature of problem solving in mathematics: Traditional vs contemporary views. *Tydskrif vir Christelike Wetenskap/ Journal for Christian Scholarship.* 39(1&2), 69-94.

WILD, C.J. & PFANNKUCH, M. 1999. Statistical thinking in empirical enquiry. *International Statistical Review* 67(3), 223-265.

WRIGHT, B.D. & MASTERS, G.N. 1982 *Rating scale analysis Rasch measurement.* Chicago, IL: MESA Press.

APPENDICES

A. APPLICATION TO THE DEPARTMENT OF EDUCATION TO CONDUCT RESEARCH

B. LETTER TO THE PRINCIPAL OF THE SCHOOL

C. LETTER TO PARENTS

Fax to 011-355 0512 Ms Nomvula Ubisi

For Official Use	
Ref. No.:	

GAUTENG DEPARTMENT OF EDUCATION

GAUTENG PROVINCIAL GOVERNMENT

RESEARCH REQUEST FORM

REQUEST TO CONDUCT RESEARCH IN INSTITUTIONS AND/OR OFFICES OF THE GAUTENG DEPARTMENT OF EDUCATION

1. PARTICULARS OF THE RESEARCHER

1.1	Details of the Researcher
Surname and Initials	MÜLLER, HM
First Name/s	HELENA
Title (Prof / Dr / Mr / Mrs / Ms)	MRS
Student Number (if relevant)	0449-364-8 (UNISA)
ID Number	5601150076081

1.2 Private Contact Details	
Home Address	*Postal Address (if different)*
Steekbaardstr 732	Posbus 38642
Garsfontein X10	Garsfontein Oos
Postal Code 0060	*Postal Code* 0060
Tel: (012) 993 2449	
Cell: 083 280 6603	
Fax: (012) 348 1305 (W)	
E-mail: helena.muller@absamail.co.za	

2. PURPOSE & DETAILS OF THE PROPOSED RESEARCH

2.1 Purpose of the Research (Place cross where appropriate)	
Undergraduate Study - Self	
Postgraduate Study - Self	X
Private Company – Commissioned by Provincial Government or Department	
Private Research by Independent Researcher	
Non-Governmental Organisation	
National Department of Education	
Commissions and Committees	
Independent Research Agencies	
Statutory Research Agencies	
Higher Education Institutions	

2.2 Full title of Thesis / Dissertation / Research Project
The relationship between spatial development and the teaching and learning of basic statistics

2.3 Value of the Research to Education (Attach Research Proposal)
See attached Research Proposal

2.5 Student and Postgraduate Enrolment Particulars (if applicable)	
Name of institution where enrolled	Unisa
Degree / Qualification	D. Ed Didactics (Math Ed)
Faculty and Discipline / Area of Study	Education: Mathematics Ed
Name of Supervisor / Promoter	Prof DCJ Wessels

2.6	Employer (where applicable)	Helena Müller	
Name of Organisation		GDE	
Position in Organisation		Teacher	
Head of Organisation			
Street Address		Laerskool Lynnwood Rodericks Road Lynnwood	
Postal Code		0040	
Telephone Number (Code + Ext)		+27 012-348 1306	
Fax Number		+27 012- 348 1305	
E-mail		admin @ lynn-woodlaer.co.za	

2.7	PERSAL Number (where applicable) 14392941

1	4	3	9	2	9	4	1

3. PROPOSED RESEARCH METHOD/S

(Please indicate by placing a cross in the appropriate block whether the following modes would be adopted)

3.1 *Questionnaire/s (If Yes, supply copies of each to be used)*

YES		NO	

3.2 *Interview/s (If Yes, provide copies of each schedule)*

YES	X	NO	

3.3 *Use of official documents*

YES		NO	X
If Yes, please specify the document/s:			

3.4 *Workshop/s / Group Discussions (If Yes, Supply details)*

YES		NO	X

3.5 *Standardised Tests (e.g. Psychometric Tests)*

YES		NO	X
If Yes, please specify the test/s to be used and provide a copy/ies			

4. INSTITUTIONS TO BE INVOLVED IN THE RESEARCH

4.1 *Type of Institutions (Please indicate by placing a cross alongside all types of institutions to be researched)*

Primary Schools	X
Secondary Schools	
ABET Centres	
ECD Sites	
LSEN Schools	
Further Education & Training Institutions	
Other	

4.2 ***Number of institution/s involved in the study (Kindly place a sum and the total in the spaces provided)***

Type of Institution	Total
Primary Schools	3
Secondary Schools	
ABET Centres	
ECD Sites	
LSEN Schools	
Further Education & Training Institutions	
Other	
GRAND TOTAL	

4.3 ***Name/s of institutions to be researched (Please complete on a separate sheet if space is found to be insufficient)***

Name/s of Institution/s
LAERSKOOL LYNNWOOD
LAERSKOOL FLEUR
LAERSKOOL OOST-EIND

4.4 District/s and other GDE Offices where the study is to be conducted. *(Please indicate by placing a cross alongside on all districts to be canvassed)*

District	
Johannesburg East	
Johannesburg South	
Johannesburg West	
Johannesburg North	
Gauteng North	
Gauteng West	
Tshwane North	
Tshwane South	X
Ekhuruleni East	
Ekhuruleni West	
Sedibeng East	
Sedibeng West	

Office/s (Please indicate)
D 4

NOTE:

If you have not as yet identified your sample/s, a list of the names and addresses of all the institutions and districts under the jurisdiction of the GDE is available from the department at a small fee.

4.5 Number of pupils to be involved per school

Grade	1		2		3		4		5		6	
Gender	B	G	B	G	B	G	B	G	B	G	B	G
Number							20	20	20	20	20	20

Grade	7		8		9		10		11		12	
Gender	B	G	B	G	B	G	B	G	B	G	B	G
Number	20	20										

4.6 Number of educators/officials involved in the study

Type of staff	Teachers	HODs	Deputy Principals	Principal	Lecturers	Office Based Officials
Number	–	–	–	–	–	–

4.7 *Are the participants to be involved in groups or individually?*

Participation	
Groups	X
Individually	

4.8 *Average period of time each participant will be involved in the test or other research activities (Please indicate time in minutes)*

Participant/s	Activity	Time
All		± 60 min

4.9 *Time of day that you propose to conduct your test/research.*

School Hours	During Break	After School Hours
X		

4.10 *School term during which the research would be undertaken*

First Term	Second Term	Third Term
		X

DECLARATION BY THE RESEARCHER	
1. *I declare that all statements made by myself in this application are true and accurate.*	
2. *I have taken note of all the conditions associated with the granting of approval to conduct research and undertake to abide by them.*	
Signature:	[signature]
Date:	30/5/2002

DECLARATION BY SUPERVISOR / PROMOTER / LECTURER	
I declare that: -	
1. The applicant is enrolled at the institution / ~~employed by the organisation~~ to which the undersigned is attached.	
2. The questionnaires / structured interviews / tests meet the criteria of: ✓ • *Educational Accountability* ✓ • *Proper Research Design* ✓ • *Sensitivity towards Participants* ✓ • *Correct Content and Terminology* ✓ • *Acceptable Grammar* ✓ • *Absence of Non-essential / Superfluous items*	
Surname:	WESSELS
First Name/s:	DIRK CORNELIS JOHANNES
Institution / Organisation:	Unisa
Faculty / Department (where relevant):	Education DEPT. OF FURTHER TEACHER Educ
Telephone:	012 - 429 4474
Fax:	012 - 429 4922
E-mail:	wessedcj@unisa.ac.za
Signature:	[signature]
Date:	30/05/2002

N.B. This form (and all other relevant documentation where available) may be completed and forwarded electronically to either Ntombi Maswanganyi (violetm@gpg.gov.za) or Nomvula Ubisi (nomvulau@gpg.gov.za). The last 2 pages of this document must however contain the original signatures of both the researcher and his/her supervisor or promoter. These pages may therefore be faxed or hand delivered. Please mark fax - For Attention: Ntombi Maswanganyi at 011 355 0512 (fax) or hand deliver (in closed envelope) to Ntombi Maswanganyi (Room 910) or Nomvula Ubisi (Room 914), 111 Commissioner Street, Johannesburg.

Posbus 38642
Garsfontein-Oos
0060
6 Augustus 2002

Geagte Meneer Louw

Ek is tans besig met navorsing wat deel vorm van die Spatial Orientation and Spatial Insight (SOSI)-projek by die Departement Verdere Onderwys (Wiskunde Onderwys) by Unisa. Hierdie projek word befonds deur die National Research Foundation . Die navorsing handel oor die vlak van ruimtelike vaardighede en statistiese denke by leerders van Graad 4 – 7. Die projekleier is Prof DCJ Wessels.

Die GDE het toestemming verleen dat ek die navorsing by Laerskool Lynnwood mag doen. Ek wil daarom graag u toestemming vra om Graad 4 – 7 leerders van u skool by hierdie navorsing te betrek. Leerders wat aan die projek deelneem, sal 'n aantal ruimtelike vaardigheidstake sowel as datahanteringstake aflê. Individuele onderhoude, wat op videoband vasgelê sal word, sal daarna met sommige leerders gevoer word. Indien leerders geselekteer word vir 'n individuele onderhoud, sal daar per telefoon reëlings met hulle ouers getref word. Individuele resultate en onderhoude word vertroulik hanteer en leerders sal ten alle tye anoniem bly by die bespreking of terugrapportering van die resultate van die navorsing. Video-opnames sal indien ouers dit so verkies, nie leerders se gesigte wys nie.

Alle Graad 4 – 7 leerders sal briewe ontvang waarin inligting oor die projek en leerders se deelname deurgegee word. Deelname van leerders is heeltemaal vrywillig en afhanklik van skriftelik etoestemming van ouers deur middel van die skeurbriefie wat aan die inligtingsbrief geheg is. Ons hoop om omtrent 15 – 20 leerders uit elke graad by die projek te betrek. Aangesien die take in Engels en Afrikaans beskikbaar is, maak dit nie saak van watter taalgroep die leerders is nie.

Dit is belangrik dat leerders sal weet dat daar nie regte of verkeerde antwoorde of oplossings vir die take is nie, maar dat ons geïnteresseerd is in elkeen se unieke, interessante manier om die probleem aan te pak en op te los. Dit is juis in die verskeidenheid van unieke antwoorde waar die waarde van die navorsing lê. Hierdie inligting sal aan leerders gekommunikeer word voordat hulle met die take begin. Dit sal egter help as die registeronderwysers dit so aan hulle stel wanneer die inligtingsbriewe uitgedeel word.

Die navorsingsprojek by u skool sal die volgende behels:

- Een uursessie waarin die ruimtelike take gedoen sal word.
- Een uursessie waarin die datatake gedoen sal word.
- Individuele onderhoude met ongeveer 12-15 leerders wat na aflegging van die take geselekteer sal word.

Graad 4 – 7 leerders kan die take almal gelyk na skoolure in die saal (of verdeel in twee klaskamers) voltooi, soos dit die skool sal pas. Ons is bereid om die werkers wat die banke heen en weer sal moet dra, te vergoed. Die enigste wyse waarop onderwyspersoneel van u skool betrokke sal wees, is deur die uitdeel en inneem van briewe in die registerklasse van Graad 4 – 7.

Ons het vandag eers die Departement se goedkeuring vir die loodsing van die projek ontvang en aangesien ons slegs ongeveer tien skooldae tot ons beskikking het om die navorsing af te handel, sal ek dit waardeer indien ek u vandag of môre nog sal kan kom sien – ek sal die sekretaresse by u skool skakel om 'n afspraak te maak. Sal u asseblief ook so gaaf wees om solank datums te bepaal vir die aflê van die ruimtelike en datatake (indien moontlik Maandag, Dinsdag of Woensdagmiddag volgende week).

Ek wil u by voorbaat dank vir u skool se deelname aan die projek, waarmee ons poog om duidelikheid te kry oor Graad 4 – 7 leerders se ruimtelike en statistiese denke. Die navorsingsresultate behoort van belang te wees vir die uitbouing van die nuwe kurrikulum, asook vir onderwysersopleiding en die ontwikkeling van leerondersteunngsmateriaal. Ons sal graag die resultate van die projek met u deel sodra dit beskikbaar word. Dit behoort vir die Wiskunde onderwysers van die Intermediêre en Seniorfase insiggewend te wees.

Opregte groete

HMüller

Helena Müller
Sel no 083 280 6603

e-posadres: lynnwoodlaer@intekom.co.za of helena.muller@freemail.absa.co.za

Posbus 38642
Garsfontein-Oos
0060
7Augustus 2002

Geagte Ouers

Ek is tans besig met navorsing wat deel vorm van die Spatial Orientation and Spatial Insight (SOSI)-projek by die Departement Verdere Onderwys en Opleiding (Wiskunde onderwys) by Unisa. Hierdie projek word befonds deur die National Research Foundation. Die navorsing handel oor die vlak van ruimtelike vaardighede en statistiese denke by leerders van Gr 4 - 6.

Ek wil daarom graag u toestemming vra om u kind(ers) by hierdie navorsing te betrek. Leerders wat aan die projek deelneem, sal na skoolure 'n aantal ruimtelike vaardigheidstake sowel as datahanteringstake aflê. Individuele onderhoude, wat op videoband vasgelê sal word, sal daarna met sommige van die leerders gevoer word. Indien u kind geselekteer word vir 'n individuele onderhoud, sal daar per telefoon reëlings met u getref word. Individuele resultate en onderhoude word vertroulik hanteer en leerders sal ten alle tye anoniem bly by die bespreking of terugrapportering van die resultate van die navorsing. Video-opnames sal, indien u dit so verkies, nie u kind(ers) se gesigte wys nie.

Dit is belangrik dat leerders sal weet dat daar nie regte of verkeerde antwoorde of oplossings vir die take is nie, maar dat ons geïnteresseerd is in elkeen se unieke, interessante manier om die probleem aan te pak en op te los. Dit is juis in die verskeidenheid van unieke antwoorde waar die waarde van die navorsing lê.

Sal u asseblief die skeurstrokie voltooi en **môre** u kind se registeronderwyser inhandig. Indien u enige vrae oor die projek het, is u welkom om my by die onderstaande telefoonnommer of e-posadres te kontak.

Ek wil u by voorbaat bedank vir u kind(ers) se deelname aan die projek, waarmee ons poog om duidelikheid te kry oor Gr 4 – 6 leerders se ruimtelike en statistiese denke. Die navorsingsresultate behoort van belang te wees in die uitbouing van die nuwe kurrikulum, asook vir onderwysersopleiding en die ontwikkeling van leerondersteuningsmateriaal.

Opregte groete

Helena Müller

Tel no 012 993 2449 (na 19:00)
e-posadres: *lynnwoodlaer@intekom.co.za*