TEACHERS' UNDERSTANDING OF THE ARITHMETIC MEAN

ROSEMARY CALLINGHAM B.SC.(Hons.), Dip. Ed.

A research study undertaken as part requirement for the degree of Master of Educational Studies in the Centre for Education, University of Tasmania.

December 1993

This research study contains no material which has been accepted for the award of any other degree or diploma in any tertiary institution. To the best of my knowledge and belief, this study contains no material previously published or written by another person, except when due reference is made in the text of the research study.

R A Callinghan
....

ACKNOWLEDGEMENTS

I would like to acknowledge the help and support of many people throughout the period of this study. In particular, I wish to thank the teachers who willingly, and with interest, voluntarily completed questionnaires. The lecturers of the pre-service teachers involved both gave up their teaching time, and administered the survey, and for this I thank them.

Special thanks are extended to my supervisor, Dr. Jane Watson, who so generously listened to my dilemmas, discussed my ideas, and questioned my thinking. Her support and encouragement have been invaluable.

Finally, I must acknowledge particularly the support of my children, Laura and Tara, who have not only tolerated the *ad hoc* housekeeping, but have insisted that I get down to work when they felt I was prevaricating. Thank you.

CONTENTS

ABSTRACT

A review of the current literature about statistical and probabilistic ideas was undertaken. The concept of the arithmetic mean was chosen as the focus of an experimental study. A questionnaire was administered which required increasingly sophisticated understanding of this concept in the concrete-symbolic mode, and also allowed for ikonic mode processing. Responses gained from 136 pre- and in-service teachers were analysed from three perspectives. Qualitative analysis allowed some comparison of the responses from this sample to similar studies by other researchers. Differences between sub-groups of respondants were analysed quantitatively for significance. Consideration of these results from the viewpoint of cognitive development provided some possible explanations. Developmental cycles in both ikonic and concrete-symbolic mode were identified and described. A theoretical model is proposed to explain the interaction between these cycles. From this recommendations are made for teacher development, and further research.

CHAPTER 1

INTRODUCTION

Recent curriculum documents in several countries have emphasised the need to include more statistics and probability in the school curriculum (e.g. National Council of Teachers of Mathematics (NCTM), 1989; Australian Education Council (AEC), 1991). This is justified on the grounds that we need to become informed citizens. Many of the decisions being made today which affect our social and economic environment are based on the interpretation of data, and without an informed population the decisions may be misunderstood, or made on the basis of erroneous information. Ample evidence exists that there are major misconceptions about data interpretation, and that moves to change this have only been partly successful (e.g. Pollatsek, Lima & Well, 1981; Gal, 1992; Cox & Mouw, 1992).

Children bring to the classroom a number of intuitions, beliefs and misconceptions about probability and statistics. These will affect the learning outcomes of any teaching program, and for that reason need to be recognised and, where necessary, challenged. This necessity is familiar to the teacher of mathematics and mathematical problem solving. Where the teaching of statistical and probabilistic ideas differs from the teaching of "pure" mathematics, is that there are a number of beliefs about chance which are firmly adhered to, even when rational arguments are advanced against them. This accounts for the popularity of various forms of gambling, and makes them such spectacularly successful revenue raisers.

With the increased emphasis on chance and data as a strand of the "mainstream" mathematics curriculum, there has been a developing interest in the knowledge and skills of teachers required to teach these topics. If children's misconceptions and beliefs are to be challenged, teachers must both recognise them, and be able to supply experiences which will provide a scaffold for children to construct alternative understandings. If teachers' own knowledge and understanding has never been challenged, it is unlikely that they will recognise the misconceptions, let alone be capable of enabling reconstruction. Statistical reasoning has been a neglected area of the curriculum, both at school and teacher training level (Pereira-Mendoza, 1986). Thus, most of today's teachers are unlikely to have had opportunities to develop a pedagogy for this area of the curriculum.

Adult misconceptions about statistical and probabilistic ideas have been documented in a number of studies (e.g. Johnson, 1985; Mevarech, 1983; Kahneman, Slovic & Tversky, 1982; Russell & Mokros, 1991). While it would be hoped that teachers had better understanding, the lack of opportunities during teacher training to develop these concepts indicates that practising teachers are likely to be no better informed than any other group.

The implications of this are that the developing understandings of children may not be challenged to develop from the purely informal, gathered through everyday experiences, to the deeper knowledge about the concepts and ideas which are applied in data handling. In particular, teachers appear to have neither an understanding of the processes involved in statistical reasoning nor of the pedagogy needed to develop statistical understanding (Bright & Friel, 1993; Bright, Berenson & Friel, 1993).

Before recommendations can be made about increasing student understanding, it is necessary to be aware of what teachers know at present. Assumptions are made that the basic tools of descriptive and inferential statistics, calculations of various measures of central tendency and spread, are available to teachers, and that they, therefore, have the necessary knowledge required to develop understanding in students. Recent studies by Berenson, Friel and Bright (1993) and Mokros and Russell (1992) apparently contradict this belief.

This study has two main aims. The first of these is to review current understandings about chance and data concepts. This considers both the concepts held by children, and those of adults. It is in the context of children's understanding that teachers work, and thus they need to have a background about children's knowledge. The second aim is to focus on one aspect of the understanding of teachers in the light of the current research, in order to make some recommendations for professional development. The concept chosen for study was the arithmetic mean. Because this is a widely taught and used idea, assumptions are made that teachers have a well developed understanding of the concept. This study indicates that this may not be so.

The arithmetic mean is probably the most commonly taught summative statistical measure. As such it would be expected to be the best understood. The recent work of Russell and Mokros (1991) indicates that this may not be the case. The study presented here explores the understanding of Tasmanian teachers, both pre- and in-service, teaching from kindergarten to Grade 10, about some measures of centrality, in particular the arithmetic mean.

The study involved a number of teachers completing a questionnaire which aimed to explore their understanding of the arithmetic mean in different situations. Although this study did not attempt to replicate overseas studies, there was interest in seeing whether Tasmanian teachers had developed a fuller understanding of this statistical measure than appeared to be the case elsewhere (e.g. Pollatsek, Lima & Well, 1981; Russell & Mokros, 1991). Originally, primary and high school in-service teachers were involved, the intention being to compare these two groups. An initial survey, however, indicated that a number of other questions were emerging. These related to the apparently multi-modal thinking of the teachers concerned, and some unexpected differences between the two original groups. Accordingly the study was extended to include a greater number of teachers, and also included some pre-service teachers. This allowed more meaningful comparison between the groups of primary and high school teachers, and also led to comparisons between pre- and in-service teachers. Some quantitative analysis was also possible with the larger numbers. The results of the survey were considered qualitatively, with some classification of the types of response seen. As well, an explanation of the types of response was attempted, using a cognitive development perspective.

The framework for the qualitative, quantitative and cognitive developmental analysis was provided by a detailed examination of the related research. This included consideration of both children's and adults' understanding of statistical and probabilistic ideas. It is in the context of children's thinking that teachers work and, therefore, presumably refine their own knowledge. Some recent studies of children's and teachers' understanding of average in different situations (e.g. Russell & Mokros, 1991; Berenson, Friel & Bright, 1993) were examined in detail, as being particularly relevant to this study. This was considered against the background of current curriculum documents (AEC, 1991), and the place of the concept of average in the curriculum. Finally, attention was given to recommendations from the literature relating to the professional development needs of teachers, both in regard to statistics and probability, and from a more general point of view. This aspect of the related research review was seen as important since some of the findings of this study have implications for teacher development.

This review of the related research, the research questions addressed, together with the methodology used and results obtained are all presented in this report. A qualitative description of the responses obtained to the survey is given. Quantitative analysis of the results is carried out to establish independence or otherwise of different categories of response and groups of teachers. One possible explanation of some of the results is drawn from a cognitive development model proposed by Biggs and Collis (1991) and Collis and Biggs (1991). Interaction between ikonic and concrete-symbolic modes of thinking was seen in some responses to particular questions. A model for considering the multi-modal thinking of teachers is proposed. This considers parallel cycles of thinking in the ikonic and concrete-symbolic modes, and the interaction between them. Some recommendations for further research and the professional development of both pre- and in-service teachers are made following detailed considerations of the analysis of the survey results and the related research.

The research related to this study is reviewed in three sections: the development of concepts, understandings of average, and implications for professional development. Each of these is reported in a separate chapter. The first part of this literature survey is concerned with the development of probabilistic and statistical concepts, and this appears in chapter 2.

CHAPTER 2

Related Research

This chapter considers the related research in three sections. It provides the background for the experimental part of this study. The first section deals with the development of statistical and probabilistic concepts, and the misconceptions found in both adults and children. As the major focus of this study, the second section considers findings about the concept of average. Recent research studies from the United States (Mokros & Russell, 1992; Berenson, Friel & Bright, 1993) indicate that this may not be as well understood as might have been expected. Finally, implications from the related research of professional development for the implementation of chance and data as a strand of the "mainstream" mathematics curriculum are reviewed.

The Development of Probabilistic and Statistical Concepts

This review covers both the growth of children's understanding and the apparent misconceptions held by many adults in the areas of statistics and probability. It is necessary for teachers both to recognise misunderstandings and to know what intervention is necessary to overcome these. This applies to their own thinking as well as that of their students.

In order to make recommendations about the knowledge that teachers require about statistical and probabilistic concepts, it is necessary to have some understanding about the growth of these ideas in children. Inferential statistics in particular requires that the connection between chance events and data collection is made, so that the growth of both probabilistic cognition and statistical reasoning power need to be linked. There is then a responsibility for teachers to recognise children's developing thinking in both areas.

Developmental Perspective

The work of Piaget and Inhelder (1951) is still the major research into children's development of probability concepts. Through a number of experiments, they identified distinct stages, in accordance with Piaget's theories of development, in the acquisition of concepts of chance in children.

The earliest stage according to Piaget, is characterised by an inability to recognise possible from necessary actions. For this reason the child is unable to have any concept of chance. Unless the child is able to recognise that there is a possibility of some action, any notion of probability is at best very restricted, at worst cannot exist at all. Young children, under about 7 years, have no concepts of combination which would allow them to find all possibilities. Thus, they are unable to modify predictions and are constantly surprised by results other than those which they have confidently forecast. Teachers in the early childhood sector of education are not surprised when a child predicts that a yellow block will be drawn out of a bag that has been shown to contain only red and blue blocks.

Around 7 - 8 years, logical reasoning begins to appear and the child is able to start recognising different possibilities. This stage, according to Piaget, marks the earliest development of chance concepts. The child is able to modify predictions in the face of evidence, and to see the possibility of various combinations.

Only when the child reaches the third stage of development, around 11 - 12 years, does the system of probabilities become structured. This happens because the intuitive thinking relating to chance, and the logical thinking relating to an action, are synthesised into an integrated concept. Once this stage is reached it is possible to consider a quantification of probability (Piaget & Inhelder, 1951).

Other researchers, while broadly agreeing with this model, have found that children much younger than predicted by Piaget can in fact find all possible combinations when given a suitably structured task. Scardamalia (1977) gave children a combinatorial task based on types of cards. He related the task to a memory capacity model. When the task was within the predicted memory capacity, young children could develop a system which would allow them to predict all possible combinations. When the task was at or exceeded the predicted memory capacity, the children were more likely to fail.

Brainerd (1981) also researched a memory capacity model. Using variations on container tasks with primary age children, he found that younger age groups stored frequency data as well as older children, but failed to retrieve it efficiently. This failure he assigned to lack of working memory space rather than a lack of appropriate cognitive structures.

Other researchers, notably Noelting (1980) and Siegler (1981), postulated that children acquire probability concepts via a series of increasingly complex rules. Noelting identified stages within which were periods of "adaptive restructuring". He saw cognitive development as a discontinuous process with distinct steps forward. Siegler found that while pre-school children applied the straightforward choice of favourable elements consistently, children as young as eight years were able to take into account, at least partially, differences between favourable and unfavourable elements. His proposed sequence of rules does not appear to be completely hierarchical.

In contrast, some researchers have found that children already have intuitive conceptions about relative frequencies and probability before they come to school. In a number of different trials, children as young as 4 years were able to choose the correct "pay-off" colour when given suitable tasks (Falk *et al.* 1980). These intuitive ideas are gradually modified by instruction until they are integrated into a successful working concept of probability which does not occur until much later (Fischbein & Gazit , 1984).

The idea of a stage theory in the development of probability concepts seems to be well accepted. While there is some disagreement as to ages at which different concepts appear, that there are recognisable stages seems to be well founded.

The area of debate rather concerns the nature of the strategies used by children to cope with probability concepts. Researchers such as Piaget and Inhelder favour the growth of cognitive structures based on logical and combinatorial thinking. To some extent these ideas are challenged by the results of other researchers (e.g.Scardamalia, 1977; Brainerd, 1981; Siegler, 1981) which tend towards the idea that children use acquired strategies to cope with probability concepts.

The difficulty seems to be that probability can be seen as both objective, typified by the classical experiments using urns and other selection mechanisms, and subjective, indicated by the intuitive understanding referred to by Fischbein (1975) and Falk *et al* (1980). Both of these are valid perspectives.

Prediction of possibilities by young children does not necessarily mean that they can assign probabilities to events. Consider a task such as drawing red and blue marbles from containers such that there are four outcomes: red/red; red/blue; blue/red; blue/blue. Working concretely, relatively young children could find these outcomes subjectively. Predicting these theoretically requires the higher order thinking skills related to proportional reasoning (Piaget & Inhelder, 1951). Both intuitive and logical cognitive elements are utilised in the growth of the concept.

A recent study (Singer & Resnick, 1992) of children's representations of proportional relationships goes some way towards clarifying this. Children were given probability problems designed to test whether they used "part-part" or "part-whole" reasoning. In a collection of objects such as the red and blue marbles in the example given above, the number of any one colour may be related to either the number of the other colour, or to the total number in the collection. Where the number of, for example, red marbles is related to the number of blue marbles (7/5, red to blue), the response is classified as "part-part" reasoning. Relating the number of marbles to the total number (7/12, red to whole or 5/12, blue to whole) is classified as "part-whole" reasoning. In general a part-part strategy was used but indications were that some children were using the part-whole strategy some of the time. The part-part strategy may be a step, used intuitively, towards being able to relate the part to the whole. When the overall picture can be seen, children are able to reason objectively about it. If this is so, recognising the part-part reasoning could lead to strategies for teachers to use which would actively develop children's thinking about proportional reasoning and probability.

In their review of children's acquisition of probability concepts, Scholtz and Waller (1983) conclude that both intuitive conceptual knowledge and developing acquired strategies are required for successful completion of probabilistic problems, and that these mutually reinforce each other. Scholtz (in press) quoted in Shaughnessy (1992) has now proposed an information processing model, based on a theoretical system influenced by cognitive psychology, rather than computer operation. This is an attempt to bring together the intuitive and acquired reasoning, which are apparently disparate forms of thinking.

Understanding of statistical concepts

A number of researchers have considered specific areas of probability and statistics and the ideas which children bring to the classroom. These include randomness, comparison of odds, average and sampling.

In a longitudinal study, Green (1986) researched thinking about randomness and comparison of odds in children. He found that children as young as 7 years have a well developed sense of random events, such as raindrop patterns. In simulating coin tossing, children are very accurate in reflecting the equal probability of heads and tails to the point where the consistency is too good to be truly random. This thinking about randomness showed no significant change with age over the four years during which the study was carried out. There were, however, significant differences with age on comparison of odds questions, older children performing much better. As Green points out, the comparison of odds questions are based on ratio concepts which are explicitly taught in schools. Randomness, on the other hand, is not taught at all and this may have some bearing on the lack of improvement (Green, 1991). The clear developmental trend on comparison of odds questions reinforces the findings of Falk *et al* (1980).

A common fallacy in probability relates to how similar an event is to the population from which it is drawn. This has become known as "representativeness" and affects children's and adults' perceptions of sampling and outcomes of chance events. A particular variant of this is the "Gambler's Fallacy" in which people expect even short runs to reflect a theoretical probability such as 50:50 (Kahneman, Slovic & Tversky, 1982; Hope & Kelly, 1983). Adults generally recognise that the chances of a coin falling heads or tails, for example, are equal. This recognition is translated into an expectation that the number of heads in a run of, say, twenty tosses will be very close to, or exactly, half the number of tosses. A total of five heads and fifteen tails would be seen as unusual, and possibly biased.

A second common misconception is that of "availability". Subjects make judgements on the basis of their own experience. For example, an estimation of the probability of children under twelve years smoking regularly is likely to be higher if personal contact with young children who smoke has been experienced (Kahneman, Slovic & Tversky, 1982).

While people who are statistically naive could, perhaps, be expected to utilise such heuristics when faced with a difficult probability problem, they are surprisingly common among those who have a substantial training in probability and statistics (Shaughnessy, 1981). This may have an impact on teaching practice, especially if these misconceptions are held by teachers. It is difficult to challenge and reconstruct the ideas of learners if the thinking of the teacher is not clear.

Other statistical misconceptions have also been identified, including notions of randomness, sampling, use and abuse of statistics, certainty/uncertainty and comparability of data (Shaughnessy, 1992). Most of these are ignored in the school curriculum, or at best, treated at a surface level only. Media reporting of "facts" based on incorrect inferences or insufficient data is widespread. Advertisers, knowingly or unknowingly, distort graphical representations frequently in order to sell products. One of the outcomes of a statistical education, Shaughnessy believes, should be the ability to recognise the misuse of statistics. Implicit in this is the notion of average, since it is one of the most commonly used terms.

It might be expected that apparently simpler concepts, such as the arithmetic mean, might be better understood by those persons responsible for teaching it. This does not appear to be the case. A number of studies (e.g. Pollatsek, Lima & Well, 1981; Mevarech, 1983; Mokros & Russell, 1992) have detailed adult difficulties with this concept. Few teachers have had the opportunity to experience conducting data-analysis projects with an emphasis on developing statistical inference (Gal, 1992). Consequently, their thinking about representative measures has been developed informally, leading to misconceptions such as those described in the studies referred to above, or theoretically, with little or no reference to practical implications. Further consideration of thinking about the arithmetic mean is given below.

It is worth noting here that most of the studies concerning adults, have been completed with college students. Some of these are also potential or practising mathematics teachers, but many others come from other disciplines. Statistics, by its nature, is cross-curricular. It may be that the statistical reasoning "taught" in other curriculum strands could, in fact, be even more problematical than in mathematics. In addition, many primary teachers lack confidence in their understanding of mathematics. This may be compounded in statistics because of the non-deterministic nature of the subject (Steinbring, 1986).

In summary, it seems that the evolution of stochastic understanding is not well understood. In addition some adult studies indicate that there are a number of well entrenched misconceptions which appear never to have been challenged during childhood. This compounds the problem for teachers who need both to recognise and overturn their own misunderstandings, as well as those of the children they teach.

Children's and Adults' Understanding of Average

The survey of the literature continues with a consideration of thinking about children's and adults' understanding of average. As the major focus of this study the relevant related research is reviewed in detail. The place of the concept of average in the current curriculum is also considered.

While there are a number of probabilistic and statistical ideas that are worthy of further study, probably the most common statistical idea which is taught in schools is the concept of the arithmetic mean. It is generally included in the syllabus at upper primary, lower secondary level. In *A National Statement on Mathematics for Australian Schools* (AEC, 1991) measures of central tendency, including the arithmetic mean, first appear in Band B, that is approximately Years 5 - 7, with an emphasis on interpretation of the data, rather than computation. Outcomes from the curriculum described in this document are defined in *Mathematics - The National Profile* (Curriculum Corporation, 1993). The relevant profile statement is found in Level 4.

> 4.26 Display frequency and measurement data using simple scales on axes and summarise data with simple fractions, highest, lowest and middle scores, and means.

Full understanding of different types of average is not expected until the child reaches Level 5, where the appropriate profile outcome reads

> 5.26 Display one-variable and two-variable data in plots and summarise data with fractions, percentages, means and medians.

Although Level 5 statements are considered to be appropriate to students in lower secondary school, there are considerable implications for teachers of much younger children. A profile statement is an outcome which is expected. In order to reach this outcome, the child will need to have appropriate experiences in earlier years for the correct notions to develop, so that the outcome can be met (Willis, 1993). Thus teachers in primary schools can no longer afford to leave the introduction of concepts to high school teachers, especially concepts such as those involved in probabilistic and statistical reasoning which appear to be, at least in part, developmental, and to require considerable intervention if intuitive misconceptions are to be corrected. Teachers will be expected to develop an understanding of statistical inference, which will include an appreciation of the nature of different interpretations of "average" and the value and appropriateness of use of the arithmetic mean.

The idea of "average" is widely met outside of mathematics classes. It has the dictionary definition:

> "Arithmetical mean; ordinary standard; generally prevailing degree etc"
>
> (New Oxford Illustrated Dictionary,1976)

followed by a detailed discussion of cricket terms! Roget's *Thesaurus* (1972) lists average under "mid-course" with the synonyms: "neutral, even, impartial, moderate, straight etc" and under "mean" listing "medium, intermedium, run of the mill, normal, balance, mediocrity, generality, rule, ordinary...". Nowhere, insofar as these sources are typical of common usage, is the idea of average or mean seen as "representative" in the mathematical sense of standing for a set of data.

The notion of average is referred to frequently in the media with phrases such as "the average wage" or "the average family" being typical of the usage. Thus, any thinking about average has a very strong informal background based on experience.

In the classroom, the idea is most often treated as an algorithm, "add them up and divide", with little relevance to data sets other than the most straightforward kind. The word "average" is generally used as a synonym for the arithmetic mean. Other measures of data summation, such as mode and median, are also dealt with in abstract contexts. Few attempts are made to develop an understanding of what these summative measures really do, that is summarise data in a particular way. Whether using the mean value is an appropriate approach to describing the data is rarely discussed. This implies that a change will be needed in teaching approaches if the emphasis on using summative measures to interpret data, as recommended in *A National Statement on Mathematics for Australian Schools* is to be fully implemented. Given this background, it is not surprising that this common concept appears to be widely misunderstood.

Studies by Mokros and Russell (1992) have indicated a number of ways in which the arithmetic mean is apparently understood by both children and adults. Five approaches were consistently noted. These were

- average as mode
- average as algorithm
- average as reasonable
- average as midpoint
- average as mathematical point of balance.

Each of these interpretations was fairly consistently adhered to by those children and adults using them. Each rationale also caused particular difficulties for its adherents when presented with atypical data sets.

Another study of children's thinking about the mean (Leon & Zawojewski, 1991) considered two statistical properties: the mean is located between the extreme values and the sum of the deviations is zero. Two representative properties were also studied. These were any value of zero must be taken into account when the mean is calculated, and the mean value is representative of the values that were averaged. This study concluded that the statistical properties of location and the sum of the deviations were better mastered than the representative properties.

In both of these studies, performance improved with age, but the Mokros and Russell study appeared to indicate that even very young children can have a sense of the mean as a representative value, in apparent disagreement with the second study. There were indications, however, that learning the algorithm for calculating the mean could interfere with the intuitive ideas that children were developing. Those students who had been taught the algorithm often seemed to have given up developing ideas of finding middle or balance points, and on some tasks actually performed worse than those who had no training in the concept of average. If this is the case, then it may explain why the older children studied by Leon and Zawojewski appeared to understand the representative aspects of the mean less well than might have been expected.

Although not analysed in the context of developmental psychology, the ideas about the mean identified by Russell and Mokros appear to be hierarchical to some extent. If so, introduction of the algorithm too early could interfere with the development of thinking about the mean as a point of balance, which seems to be a much more useful step towards understanding.

The work of Gal *et al* (1990) appears to confirm the developmental nature of children's thinking about the arithmetic mean. There were clear differences between age groups as to how the mean was used, with older students more likely than younger ones to appreciate the summary nature of the measure. A significant proportion of the children tested did not utilise their knowledge of the mean when presented with data analysis problems. The usage appeared to be related to context. School instruction does not seem to provide opportunities for children to refine environmental knowledge and resolve ambiguities.

When presented with problems requiring an understanding of the concept of a weighted mean, many college students were unable to calculate this (Pollatsek, Lima & Well, 1981). Their understanding of the mean was limited to the application of a straightforward algorithm, and no other strategy appeared to be available to them. Pollatsek *et al.* conclude that learning the method of calculation does not develop a full understanding of the concept.

As there is some suggestion that thinking about "average" is, at least in part, developmental, it is rather disconcerting that a number of teachers also held to some of the simpler interpretations of the concept. This would seem to indicate that their ideas had not been challenged and extended at any stage during their education, and points to a serious limitation of present approaches to teaching about the arithmetic mean. How similar the Australian situation is to the American one is not known. Early indications are that there are similarities when children are considered (Watson, 1993). If the thrust of the requirements of the Australian *Mathematics - The National Profile* is to be carried out, then considerable changes will need to be made in approaches to teaching about statistical measures.

Recent studies by Bright, Berenson and Friel (1993), indicate that elementary school teachers have little pedagogic knowledge about statistics. In addition, they seem to have only limited understanding of the relationships between statistical concepts (Bright & Friel, 1993). One of the areas of study considered the notions of "typical" and "middle of data" in the context of graph interpretation (Berenson, Friel & Bright, 1993). Teachers, they found, tended to "fixate" on particular features of a graph to explain these ideas. When line plots were used, teachers preferred to use the mode to describe what was "typical" of the data, and to use the centre of the range as a measure of the middle of the data. For a histogram representation, the interpretations of "typical" and centre of the data were more varied, but still tended to rely on one large feature of the graph rather than summarising all the data available. This finding reinforces those of Mokros and Russell (1992) and Pollatsek et al. (1981) that the concept of average does not appear to have been sufficiently well developed to be transferable to different contexts.

If these results are representative of teachers in Australia also, then considerable in-service work will be needed if teachers are to be able to teach statistics effectively enough for students to reach the required outcomes. There are also implications for the training of new teachers, since indications are that this is a neglected area of the teacher training curriculum (Pereira-Mendoza, 1986). Some of these implications for pre- and in-service courses are considered in the next section.

Implications for the Professional Development of In-service and Pre-service Teachers

Approaches to the professional development of pre-service and in-service teachers are reviewed in relation to the growth of stochastic ideas, and some general principles of good professional development. Strategies for overturning some of the fallacies and developing correct concepts have been suggested and these will now be considered. Many courses concerned with either teaching teachers statistics, or teaching teachers how to teach statistics have been described. Unfortunately, few of them appear to have been an unqualified success.

In the United Kingdom, a college level unit described by Goodall and Jolliffe (1986) was based on practical contexts, such as quality control, or genetics. The authors found that in these situations students responded enthusiastically claiming that the course enhanced their understanding. A major disadvantage was that of time. There is no doubt that to teach anything through practical experimentation does take longer. This has implications for teaching at all levels, not just pre-service teacher education. Concept development implies that individuals will need to revisit ideas at an increasingly sophisticated level throughout their education. Since the research reviewed in the previous chapters indicates that this does not appear to be the case for probabilistic and statistical ideas at present, educators at college level and beyond are faced with the need to alter well entrenched concepts in short periods of time.

Cox and Mouw (1992) reported on attempts to change the thinking about "representativeness". Graduate level students were given a short, sharp program of problem experiences, which required them to confront their faulty reasoning in several different contexts. The change recorded was not very great. No measure of the persistence of the change was made, the post test being given within one day of completing the instruction. What gains were made came from students being challenged with their faulty reasoning in a practical context, and direct involvement with practical contexts was recommended.

A study of college students reinforced the significance of context in relation to statistics and probability problems (Garfield & del Mas, 1991). Students appeared to be influenced by problem content, using different strategies for problems which were essentially the same but with changed background. Active intervention using coin tossing did not appear to improve thinking except to a small degree on coin tossing problems. The new knowledge was not transferred to applied problems. Similar results were noted by Jolliffe (1991).

The need for practical experimentation and involvement with data has been suggested by several researchers. Shaughnessy (1981) describes tasks which allow children to develop their understanding of probability through experiment. Russell (1991) and Garfield and Ahlgren (1986, 1988) similarly recommend practical involvement of children in learning statistical concepts. Teacher development courses should take a similar approach, since courses for adults which have been most successful appear to be those with a strong "hands on" bias (e.g. Jowett, 1991; Glencross, 1986). Fischbein (1990) goes further. He contends that since many of the notions about statistics and probability are developed through an intuitive process, attempts to educate teachers should provide similar experiences. Data should be related to real situations, and not just the urns and dice frequently used in practical courses. The idea of a mathematical model is very important, and the ability to translate a real world situation into a model should be explicitly developed. Fischbein also recommends that teachers should be faced with common misconceptions and given opportunities to analyse these and so develop an understanding of how these misunderstandings come about. There is evidence that teachers find teaching probability and statistics different from other areas of mathematics. The notion of uncertainty is one with which some teachers are uncomfortable (Russell, 1990; Pereira-Mendoza, 1990).

Teachers, it would appear, need to develop understanding of stochastic concepts. The most effective way of doing this seems to be practically based professional development in which teachers are able to develop concepts through the same sorts of methods which they would be using with their students.

The latest material being produced for schools in the area of chance and data relies on this practical approach. In the United States, the Quantitative Literacy project, a joint project between the American Statistical Association (ASA) and

NCTM, has published several booklets and included in-service training as part of the dissemination package. Another American project, Used Numbers, has produced material suitable for developing statistical concepts in primary age children. This also has been supported by in-service programs. Australian produced material has come from the Curriculum Corporation, with several publications and supporting computer material called *Chance and Data: Investigations*. These projects have provided teachers with well-structured resources from which to develop a statistical teaching program in their schools and classes. Professional development of teachers is implicit in the way in which the material is presented, allowing teachers to experience learning statistical concepts alongside their students. These materials are detailed in Appendix 1.

A warning about the value of this type of professional development teaching material is given by Pegg (1989). Unless teachers have the skills needed to analyse and then generalise the key elements of the approaches suggested, they will only ever be able to take a technical view. The material will be used, but any difficulties will not be foreseen, and may be dealt with inappropriately. At best the lessons will be used as one-off specials. The implications are that providers of professional development or in-service training, as well as mathematics educators involved in pre-service provision, need to utilise the existing knowledge about the development of concepts to build a framework for teachers to refine and extend their knowledge base. This may well have particular relevance to the teaching of statistics since teachers do not have a well developed statistics pedagogy (Bright, Berenson & Friel, 1993).

The recent work of Bright, Berenson and Friel (1993) describes approaches to in-service education which appear to hold the possibility of some success. Elementary school teachers entered the study because of their desire to learn how to teach statistics more effectively, following a mandated curriculum change. Thus, the group was already motivated to change ideas. The project involved both local support and a summer school workshop. There appeared to be gains made by the participants at the end of the three week intensive workshop although the persistence of these gains is not yet known, nor the affect these will have on classroom practice. The study is not yet complete.

What does seem to be unquestioned, is that teachers both need and want good quality professional development in this area. The general features of effective professional development have been identified by a number of people. Fullan (1991) indicates that both specific instructional change and organisational change within the school or institution are required. Both of these can be targeted through a planned implementation program of staff development. The conversion of the school culture to one in which workplace learning and the associated growth of skills and understanding are valued, fosters the changes being made by individuals. Good professional development should be seen in the context of the improved quality of teachers and increased student outcomes (Johnson, 1991). A number of conditions are required for this to occur including the need for the school climate to nurture the growth of ideas in teaching and learning. One-off courses involving teachers from many different schools who have no other support are unlikely to result in long term change unless the school organisation and leadership actively supports this.

In the context of improving teachers' understanding of both the basic concepts of statistics and probability, and bringing about changes in the way these are taught in the classroom, there are then some implications for professional development. It would appear that courses should be school or area based so that appropriate support structures for participants can be provided. The need for changes in this area of the curriculum must be acknowledged, and the explicit support of the school organisation given. From experience, the need for improvement in these areas is certainly acknowledged, but the recognition of the organisational support is not always in place. These issues are bigger than this study can deal with, but are worthy of further consideration.

A number of in-service courses for teachers have been described (e.g. Kepner & Burrill, 1990; Dunkels, 1990). These courses are based on helping teachers implement appropriate material in their classrooms, and may take several forms including school based professional development, extended courses and a sandwich model. In general teachers are enthusiastic about the practical approach taken and keen to utilise the methods suggested. There are recognised classroom management difficulties for some teachers, and networking or some other support system helps to overcome these. Understanding of pedagogy has been shown to be instrumental in improving learning outcomes for students (Peterson, Fennema, Carpenter & Loef, 1989). These practically based courses implicitly recognise this, as well as building in the support structures known to be necessary for change to take place (Murphy, 1992).

In summary, there are considerable implications from the research for the development of statistical and probability concepts in both children and adults. There is growing international interest in this area, with several projects currently considering different aspects and how these relate to classroom practice. With this in mind, it is necessary to establish the current understandings of teachers. The research questions framed in this study, and the method by which these were addressed are reported in chapter 3.

CHAPTER 3

METHODOLOGY

The manner in which the experimental part of this study was carried out is described in three major sections. Firstly, the research questions posed initially, and those which arose as the study proceeded are considered. Secondly, the survey design and implementation is discussed, including the choice of sample subjects and the practical restrictions imposed. Finally, the methods used to analyse the results are described, and consideration given to the limitations on interpretation arising from the design and administration of the study.

Research Questions

The aim of this study was to consider the understanding teachers have about the arithmetic mean. Several studies previously cited (e.g. Pollatsek, Lima & Well, 1981; Mokros & Russell, 1992; Berenson, Friel & Bright, 1993) indicated that understanding of the concept of average was less than might have been expected given that this is such a widely taught and commonly encountered idea. Little research appears to have been carried out in Australia on this topic.

Teachers come from a variety of training backgrounds. High school mathematics teachers are generally assumed to have an in-depth knowledge of their subject area. Most of them have specialist degree and Diploma of Education qualifications. Some primary teachers have similar qualifications, but the majority have Bachelor of Education qualifications, with generalist subject backgrounds but a deeper knowledge of learning and teaching.

For reasons described below, the study was broadened to include some pre-service teachers as well as the in-service teachers originally intended. There is interest in finding out whether new entrants to the teaching profession have a greater understanding than existing teachers. There are obvious implications for pre-service courses if this does not appear to be the case.

Evidence from other studies also indicated that the context of the question or the way in which it was presented, could influence the response (e.g. Berenson, Friel & Bright, 1993; Garfield & del Mas, 1991; Jolliffe, 1991). Without completely replicating overseas studies, there was interest in whether or not this could be confirmed, and the survey was designed with this in mind.

As the study proceeded, it became evident that there were apparent differences between responses to questions by sub-groups within the categories of pre-service and in-service teachers. This raised the question of whether or not these differences were significant. Furthermore, there were unanticipated differences within sub-groups in responses to questions having the same type of presentation. With this background in mind, the following questions were addressed specifically.

1. Are similar misconceptions to those described elsewhere present in Australian teachers?

2. Are there differences in response between primary and high school teachers?

3. Are there differences in response between pre- and in-service teachers?

4. Are there differences in response between sub-groups of the pre-service teachers, such as Diploma of Education students and final year Bachelor of Education students?

5. How does the context of the question alter response?

The results were analysed in a variety of ways to provide an answer to these questions, and some possible explanations. Details of this analysis are given later in this chapter.

Survey Design and Implementation

It was decided that teachers would be administered a questionnaire which considered various applications of the arithmetic mean. The questions were broadly based on approaches taken by other researchers. These are described in more detail below. Because of the time constraints under which teachers operate, it was considered important to keep the questionnaire as short as possible. It was therefore designed to fit on one A4 backed sheet, including space for the answers. This limited the number of questions to four. A short questionnaire covering the background and experience of participants was stapled to the survey sheet.

This study was carried out as part of the Chance and Data Project at the University of Tasmania. Permission to approach teachers in government schools, and clearance by the University Ethics Committee, was obtained as part of the larger project. Participation of all subjects was voluntary, and a permission form giving a brief outline of the project was signed by each person who completed a questionnaire. These were stored separately from the responses, so that participants could not be identified from their questionnaires. The permission form, background questionnaire and average questionnaire, are included in Appendix 2.

The Survey

All four questions related to the arithmetic mean, although the contexts did allow other interpretations of average to be used. Information was presented in ways which, from experience, teachers would recognise, utilising text, table and graphs as appropriate.

Question 1 involved the straightforward calculation of the average mass of an object given a set of measurements. This was not unlike a typical text book problem. The context was practical - a set of results from an experiment - and one outlier was included.

Questions 2 and 3 presented the information in graphical form. These questions were adapted from those used by Mokros and Russell (1992), and were related to the spelling scores of groups of students. Participants were asked to identify the group having the better spellers. Calculation of the arithmetic mean, using results from the graph was the expected response, although it was recognised that other answers were possible. Question 2 was seen as straightforward, involving two groups of only ten students, one of which was clearly better than the other from the graphs. Question 3 was essentially the same problem, but with two larger groups having different numbers of students. These two spelling groups were very similar in performance.

The final question was a weighted average question, similar to those used by Pollatsek, Lima and Well (1981). The context was finding the average number of babies born from data about small and large hospitals.

The questions were designed to be hierarchical to some extent, as well as placing the mean in different contexts with respect to modes of thinking. Question 1 was a relatively trivial calculation of the arithmetic mean, which called upon concrete-symbolic reasoning only. The second and third questions allowed the use of ikonic and concrete-symbolic modes of thought. Question 4 required concrete-symbolic thought at a higher level of operation than question 1. Further consideration of the structure of the questions and the modes of thinking will be made in the section concerning the treatment of the results.

Minor changes to the wording of the questionnaire were made after Stage 1 of the survey had been carried out. These made no noticeable difference to the responses. The questionnaire is included in Appendix 2.

Implementation Stage 1

The study was undertaken in two stages. Initially 27 teachers, teaching all grades from Kindergarten to Grade 10 were identified. Included in this group were some who had a considerable background in mathematics, others with a weaker mathematical background who were nevertheless teaching high school mathematics, and those with little or no background in mathematics. There was also a range of teaching experience from a pre-service teacher on school practice to several teachers having 10 or more years classroom experience.

The teachers were all members of school staffs which were undertaking a professional development workshop in mathematics. As such they were not a random group of teachers, but from experience appeared to be typical of school staffs, having a wide range of background experiences. The workshop topics did not include development of concepts of average.

At a scheduled workshop, after agreeing to take part in the survey, these teachers were given the questionnaires. On all occasions the questionnaire was administered in a group situation. Calculators were freely available. Generally, the questionnaire took about 20 minutes to complete. Although the questionnaires were administered to a group there was no discussion between subjects about the questions, although some did ask for the answers afterwards. These were given and explained where requested.

Implementation Stage 2

The data received from the initial survey of 27 teachers demonstrated a wide range of often creative responses to the questions. It appeared to indicate that the arithmetic mean was not as well understood as might have been expected. Following consideration of these responses, a second survey was carried out with a greater number of subjects. This second survey aimed to discover whether the initial responses were typical. It also included pre-service teachers to see whether the understanding of new members of the profession was different from that of in-service teachers. These pre-service teachers were all fourth year Bachelor of Education or Diploma of Education students who would be fully qualified to teach unsupervised in schools in 1994. A larger sample of high school teachers was also desired, since there had been only ten in the first stage.

The groups identified for the further survey were:

> Diploma of Education students studying primary method;
>
> Diploma of Education students studying mathematics and/or computing method;
>
> 4th year Bachelor of Education students studying both primary and secondary method;
>
> Another group of high school teachers of mathematics.

Minor changes were made to the wording to eliminate any possible misunderstanding of the questions. The results from the Stage 2 survey indicated that wording was not affecting response, and for analysis the two surveys were treated together.

The Stage 2 survey was administered in a similar manner to that of Stage 1, in some instances by the researcher, in others by lecturers known to the pre-service teachers. Participation was voluntary, and permission forms were received. As such the sample of teachers surveyed cannot be considered as random.

Treatment of Results

The results were considered from three different viewpoints: a qualitative analysis, a quantitative analysis and a developmental analysis.

Qualitative Analysis

The responses from Stage 1 were originally analysed descriptively, with consideration being given to the types of answers and the way in which they were communicated. Some attempt was made to classify different approaches to the problems, and errors were identified.

A similar approach was taken with the results from Stages 1 and 2 combined. Descriptive measures, such as percentage were used to summarise the results. These were tabulated and are presented in the next chapter.

The qualitative analysis was used particularly to address the first research question, relating to the understanding about the mean of Australian teachers. It was also used to select the sub-groups for further quantitative analysis as described below.

Quantitative Analysis

When the results were considered descriptively, it appeared that there were differences between some sub-groups in the study. These were analysed for significance using contingency tables and applying a X^2 test. This approach was applied particularly to question 4 in the survey, the weighted average question. This was generally found to be the most difficult question but had essentially only a right or wrong response. It was thus a suitable question for whether there were significant differences between the sub-groups identified in the research question, primary and high school teachers, pre- and in-service teachers and sub-groups within the pre-service teachers.

Different questions also seemed to provoke a different response from some groups. This related to the research question concerning the context. Some unexpected responses were seen in particular to survey questions 2 and 3, which were presented in graph form. There appeared to be some differences between primary and high school teachers in their approaches to these questions. These two groups were compared quantitatively to see whether the apparent differences were significant. It seemed also that there was a change in the type of response seen as subjects moved from the easier survey question 2 to the more difficult question 3. This change in response was also considered quantitatively for both high school and primary school teachers.

Developmental Analysis

The responses obtained were often unexpectedly creative, and varied from an unsophisticated calculation of the mean in a concrete-symbolic mode, to those which appeared to be utilising multi-modal functioning. The specific research questions addressed were those relating to differences between high and primary school teachers, and the influence of a change of context.

These were analysed using the framework provided by the SOLO taxonomy. The SOLO (Structure of the Observed Learning Outcome) Taxonomy was developed by Biggs and Collis (1982). They claim that increasing complexity of response can be recognised by the way in which information is utilised to answer a question. Five levels of response have been identified.

- Prestructural - there is no use or completely irrelevant use of the available information

- Unistructural - only one piece of the relevant information is utilised
- Multistructural - a number of bits of information are strung together, usually sequentially
- Relational - the various bits of information available are integrated into a coherent understanding of the situation
- Extended Abstract - over-arching principles are called upon, and the thinking is thus moved into a new level. This response provides the unistructural response for the next higher mode of thinking.

These categories form a cycle through which the level of response to any learning situation may be monitored (Campbell *et al.,* 1992). These cycles operate within the various modes of functioning.

Current neo-Piagetian thinking about the nature of intelligence postulates modes of thinking in which information is handled. As an individual progresses or matures, these modes become increasingly abstract. Thus at an early stage the mode is entirely sensori-motor, progressing through ikonic and concrete-symbolic to formal and post-formal thought. Contrary to traditional Piagetian thinking, the earlier modes of operating are not abandoned when the next mode of operating becomes available to an individual. Rather the earlier modes continue to be available and may be utilised to handle specific types of information (Biggs & Collis, 1991; Collis & Biggs, 1991).

Ikonic thinking is characterised by imagery and intuition, and is usually expressed in oral language in the form of "stories". It is used by adults as a powerful problem solving tool, and many significant advances in knowledge have been made by individuals first realising them intuitively, and only later establishing them by evidence and argument in other modes.

Concrete - symbolic thinking is significantly more abstract. It relies on symbol systems that apply to the experienced world. Thus it is expressed through written language and systems specific to particular fields of knowledge such as mathematics. More importantly in the context presented here, it is the mode in which we present justification of ideas and thinking to the world. Frequently arguments presented in this mode are seen as more objective than the stories which arise from the ikonic mode.

Beyond concrete symbolic thinking are Formal and Post-Formal modes of thought. These higher levels are those in which the thinking processes are increasingly abstract. While many people are capable of this level of thinking, it is not necessary for day-to-day functioning.

Within all of these modes of thinking are cycles referred to earlier, of increasing complexity. It has been further hypothesised (Biggs & Collis, 1991; Collis & Biggs, 1991) that within each mode there may be more than one cycle. Recent work involving volume measurement (Campbell, Watson & Collis, 1992) points to this possibility. These may be identified using the SOLO Taxonomy as described above. This model was used to identify and describe modes and levels of thinking in responses obtained to the survey questionnaire.

The questions were considered at two levels. At a macro level, they allowed examination of the unistructural, multistructural, relational sequence in the concrete-symbolic mode.with respect to the concept of average. At this level, question 1 required a unistructural response in the concrete-symbolic mode. While there are a number of prerequisite skills in the concrete-symbolic mode utilised to calculate the mean as presented in this question, with regard to the actual concept, it was only necessary to apply the algorithm to obtain an answer. Questions 2 and 3 needed at least a multistructural response. The necessary information had to be accessed from the graph, and then operated upon to find the mean. Because of the nature of the presentation of information, an ikonic mode of operation was also available. Finally, question 4, which involved calculation of a weighted average in the concrete-symbolic mode, required relational level thinking in the cycle which began with average. Not only was it necessary to understand the concept of average, but also to recognise the meaning of the data presented and integrate this into the overall picture. Thus, the questions gave an indication of a macro-level of response of the subjects to the concept of the arithmetic mean.

At a micro-level, the mathematical processes applied to each question was considered. The ways in which respondents carried out the necessary calculations was analysed, also within the framework provided by the SOLO Taxonomy. This provided a second level of consideration of the responses.

Because all the teachers involved in this study participated voluntarily, the sample was not random. This limitation must be borne in mind when the results are considered, since the sample may not be representative of all teachers. It is reasonable, however, to consider the teachers, both pre- and in-service, as showing many of the characteristics of typical teachers. Informal conversation with many of the in-service teachers at workshops or professional development sessions, indicated similar attitudes to those found in other groups of teachers, not involved in this study. The results obtained then from this study, while not being able to be said to be representative of teachers as a whole, can be considered to give a reasonable indication of what is typical among teachers, at least in Tasmania. The extent to which Tasmanian teachers are representative of teachers in other parts of Australia may be conjectured. It should also be noted that all the in-service teachers were employed in Tasmanian Government schools. For practical reasons, no attempts were made to survey teachers in private sector education.

The results obtained, and an analysis of these are reported in the following chapter. These consider the responses to the questionnaire on average from a qualitative, quantitative and cognitive developmental perspective.

CHAPTER 4

RESULTS

In this chapter, results from both the first and second stages of the study are considered. The three different approaches to the analysis of the results are applied. These are reported under the section headings of qualitative analysis, quantitative analysis and cognitive development perspective.

Qualitative Analysis

The responses to the questions are analysed descriptively with common errors being identified, and the types of response being classified. Descriptive measures, such as percentage are used to summarise these classifications. Where appropriate the results are presented in a table format. The results are presented in two sections which relate to stage 1 and stage 2 of the study.

Stage 1 results

The initial survey consisted of 27 teachers, 10 high school teachers including one pre-service teacher who was practice teaching at the time, and 17 primary teachers. These teachers taught all grades from kindergarten through to grade 10. Results seemed to indicate that misunderstandings about the arithmetic mean were more widespread than had been expected. A wide range of responses was received to the questions making classification of these very difficult. These initial results are briefly discussed question by question.

Question 1.

You want to do an experiment with your class as an exercise in using a weighing scale to find the average mass of a single block. You try the experiment yourself first and get the following results (all readings in grams.):

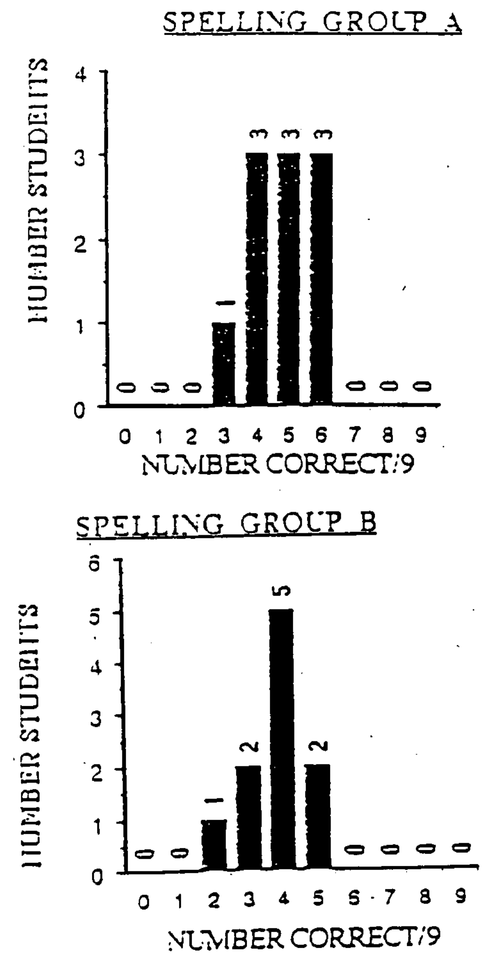
3.2 4.0 3.25 3.2 3.2 3.1 3.0 3.2

Use these results to find the average mass of the block.

This question required a straightforward calculation of the mean with one outlier. In the given context (finding the mean mass of a block) it would have been sensible to discard the outlier as having been a likely measurement error. No response showed this thinking, and in fact the presence of the outlier went totally unremarked. All teachers surveyed could calculate the mean of all results. The majority used the algorithm in two steps, i.e. first adding up the results and then dividing by the number of results as two separate processes. Although calculators were available, not all teachers used them. This was particularly so for primary teachers. Generally the processes used were relatively unsophisticated, and in SOLO terms were multi-structural rather than relational.

Out of the sample, four teachers (approximately 15%) answered with a technically incorrect response. These came equally from primary and high school teachers. In all cases the mistake was due to incorrect rounding of the result to two decimal places. That experienced teachers should be making this type of error gives some cause for concern.

Question 2.



The graphs show the results on the same spelling test for two different groups of equal size. Show how you would tell which group had the better spellers or did they spell equally well? Please explain your answer.

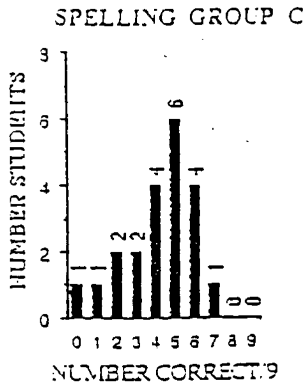
This question used a graphical rather than a tabulated representation of the data. Of the responses seen, fifty percent of all teachers used the arithmetic mean as the basis for their decision.

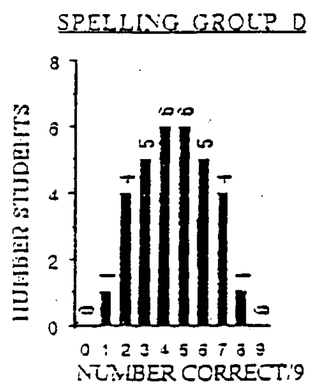
A surprisingly high percentage (approximately 27%) used an arbitrary pass mark to decide which group had more students who had "passed" the test. This was especially so among high school teachers (4 out of 10) probably reflecting a particular "mind set" with regard to assessment. This could be seen as similar to the findings of Mokros and Russell (1992) who identified a group who saw average as a "reasonable" value. Some teachers attempted to use the median, but were using the term inappropriately, identifying the middle value of the range, rather than the middle value of the results. This might reflect an unfamiliarity with the idea of the median, but is also in line with the Mokros and Russell (1992) finding of a group who saw average as a middle value.

A visual comparison of the graphs to compare the groups was used by two teachers, one from each category of high school and primary school. The difference was apparent from looking at the graphs, and given this, it was surprising that more teachers did not use this technique. It may indicate that, given the context of the question, teachers wanted to rely on what they perceived to be a more objective measure rather than on the purely ikonic visual comparison.

An incorrect response to this question was obtained from three teachers, one high school and two primary school teachers. The errors included misinterpretation of the graph (high school) and an attempt to use the range as a measure.

Question 3.





The graphs show the results on the same spelling test for two larger classes of different sizes. Show how you would tell which class had the better spellers or did they spell equally well? Please explain your answer.

This question was similar to question 2 but referred to two groups of students of different sizes, and with bigger groups than for the previous question.

Approximately 30% of teachers, 5/17 primary and 3/10 high school, were unable to respond correctly to this question. These ranged from no response to those indicating a complete lack of understanding, and one high school teacher who consistently misread the graphs. About 15% of the sample, two high school and two primary school teachers, used a visual estimation to make a comparison or to validate their calculation, even though this process was more difficult with this question.

Only two primary teachers used the arbitrary pass mark to make a comparison. It was interesting that no high school teachers used this strategy with this question, even though it was the preferred response to the previous question.

The arithmetic mean was used by approximately 42% of the respondents to obtain a correct result, and one teacher calculated the means incorrectly and then misinterpreted these results to draw a correct conclusion.

This question was a more complicated version of question 2 and it is interesting that the frequency of correct responses went down, and that teachers in some cases either did not recognise that this was a version of the same basic problem, or went to different strategies when faced with this more complex version. This bears out research about the importance of context (Garfield & del Mas, 1991; Jolliffe 1991).

Question 4.

Ten hospitals in Victoria took part in a survey of births.

4 large hospitals had an average of 70 babies born each month.

6 small hospitals had an average of 30 babies born each month.

What was the average number of babies born per month in the ten hospitals?

Please show your reasoning.

This question was a weighted average question. In general answers fell into only two categories, correct and incorrect, whereas for the graphical problems there were a variety of responses. The algorithm for this problem would be expressed as

$$\frac{(4 \times 70) + (6 \times 30)}{10} = 46.$$

Only one high school teacher used the algorithm to find the answer. Just over half of the teachers calculated the answer correctly but using a series of steps. Commonly this was to find the total births and then divide by the number of hospitals. A number of teachers (approximately 45%) could not get a correct response. These findings are in line with those of Pollatsek *et al.* (1981).

Even though the results indicated a relatively high proportion of correct answers, it was noticeable from the responses that teachers were very unsure of their answers. There were a number of question marks and alternative answers proposed so that even those people who could calculate the result correctly appeared to doubt their answers. It was this question also which was most frequently brought up during informal discussion after completion of the questionnaire.

Summary of the stage 1 analysis

The variety of responses to relatively straightforward questions indicated that, as in overseas studies, Australian teachers appeared to have a limited understanding of the concept of average. Because of the limited numbers involved, and the unexpectedly unsophisticated understanding demonstrated by the respondents, it was decided to widen the study in order to clarify this situation. Accordingly, the survey was extended to include groups of pre-service teachers and another group of high school teachers as described previously. The results from Stage 2 of the study are combined with the original Stage 1 results in all further analysis.

Total Survey Results

Included in the sample were teachers working in all grades from kindergarten to grade 10, both in the in-service and the pre-service groups. All the in-service teachers taught in Tasmanian Government schools. The pre-service teachers were completing a variety of methods courses, including early childhood, primary and secondary method among the Bachelor of Education group. The Diploma of Education group was taking either mathematics and/or computing method, or primary method. The high school in-service teachers came from three different high school, and the primary school in-service teachers came from three different primary schools. All Diploma of Education pre-service teachers were from the Hobart campus of the University of Tasmania. The Bachelor of Education pre-service teachers were from the two different campuses of the University of Tasmania, with 18 coming from the Hobart campus and 52 from the Launceston campus. Unless specifically stated, all teachers, pre- and in-service, of early childhood children, i.e. those children in kindergarten to grade 2, were included in the primary sample. This is justified on the grounds that teachers of primary school age children are increasingly being required to teach across the range of grades.

The total number sampled was 136 teachers, 36 in-service teachers and 100 pre-service teachers. Due to printing errors, a few questionnaires did not have questions 3 and 4 on them. Thus the sample number varied slightly for these questions. A summary of the composition of the sample surveyed is shown in table 1.

Table 1 Breakdown of the survey sample

	In-service Teachers	Pre-service Teachers
High School	19	17*
Primary Schools	17	
Dip. Ed		21
B.Ed.		62

* Includes 8 B.Ed. secondary method.

The results will be reported question by question, together with a description of the categories of response found. Where appropriate a table is used to summarise the results obtained. The questions appeared to be hierachical to some extent, with the numbers of incorrect responses increasing from questions 1 to 4. It should be noted that the "incorrect" category only included those who had made major errors in computation, totally misinterpreted the question in some way or had not made a response. Where calculations were essentially correct with only minor mistakes, such as rounding errors, they were counted as correct.

Question 1

The results are summarised in table 2.

TABLE 2 Summary of responses to question 1

Question 1 N=136

	IN-SERVICE TEACHERS				PRE-SERVICE TEACHERS													
	High School		Primary /ECE		Dip.Ed. Maths		Dip. Ed. Primary		B.Ed. High School		B.Ed. Primary		B.Ed. ECE Hobart		B.Ed. ECE Laun.		Total sample	
	No..	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Incorrect	1	5.3	0	0	1	11.1	2	9.5	0	0	0	0	0	0	1	4	5	4.2
Mean	17	89.5	17	100	8	88.9	19	90.5	8	100	19	100	18	100	24	96	130	94.9
Mean ex. outlier	1	5.3	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	< 1

Despite the high percentage of subjects correctly calculating the mean, the understanding demonstrated some lack of depth. Where the results were correctly calculated, it was done in a relatively unsophisticated way, using a multi-step approach. A typical response to question 1 is shown in Figure 1.

3.2 + 4.0 + 3.25 + 3.2 + 3.2 + 3.1
+ 3.0 + 3.2 = 26.15 ÷ 8 = 3.26 or 3.27

Figure 1

In only one instance out of the 136 responses was there a comment on the outlier. This came from a high school teacher who had a background in mathematics of First Year University level, and an employment background of work in both factory and medical research laboratories.

> *"As the weight of the block clearly doesn't change, I would discuss the 4.0 result which is clearly out for reasons differing from the other results and would discuss eliminating it, just as they do in diving competitions where they eliminate the highest and lowest score before averaging it."*
>
> High School Teacher

The comment about diving was made by other subjects, at least two using the method of eliminating highest and lowest scores before averaging the rest. If this strategy had been used in conjunction with a discussion about the effects of outliers it might have been justified, but it was used as the preferred algorithm regardless of the values involved.

A mean value lying outside the range was obtained by four subjects, approximately three percent of the sample. In all instances, there was no attempt to correct this, or to comment on it in any way. This suggests that these participants had little understanding of the meaning of their calculation, but applied the procedure without any analysis of either the problem or the answer.

Two high school teachers of some experience gave answers which, while correct in the sense of using the algorithm, indicated considerable lack of mathematical understanding of the process of division. These responses, appending a fraction to the end of a decimal representation, are shown in figure 2. That any teacher of high school mathematics should be showing such elementary mistakes is surely a matter for some concern. Both of these experienced teachers did not have a strong mathematical background; one was a science teacher who had majored in natural science, the other was a technical subjects teacher who came from a boat building background. Both are currently teaching mathematics. There were also many "rounding" errors which were not marked as incorrect responses but which deserve comment because of the frequency with which they occurred. All categories of teacher produced this type of error. These arithmetical errors were very unexpected, and rather disconcerting.

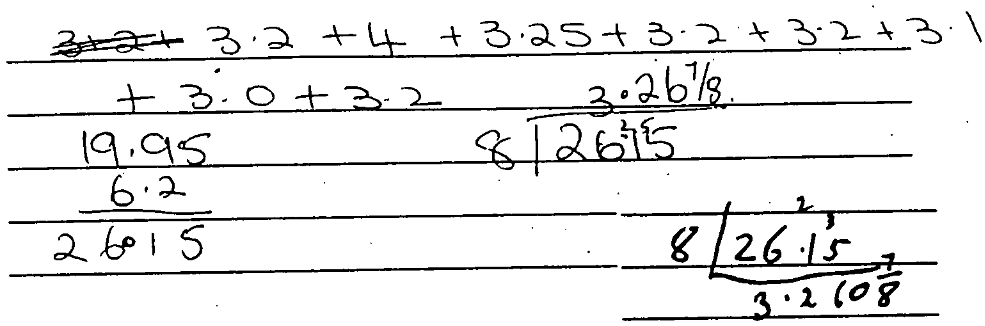


Figure 2

In general the vast majority (94.9%) of subjects could correctly calculate the mean by applying the algorithm albeit in a fairly unsophisticated manner, without considering the outlier. The lack of appreciation of the meaning of the one measurement which was apparently inaccurate may be related to a lack of experience with practical situations of the type indicated in the question. This has implications for professional development of teachers, and will be discussed further in a later chapter.

Questions 2 and 3

These questions are considered together since they are essentially the same problem, but at differing levels of complexity. The results are summarised in Tables 3 and 4.

These questions were presented graphically and provoked the widest range of response. It would seem that using the mean value to summarise graphically presented data is not widely experienced. Instead a variety of approaches was used, and this seemed to be related to the context to some extent. Because of the wide range of types of response the following categories were used in tables 3 and 4.

- The category "visual strategy only" refers to those answers which either referred directly or indirectly to respondents "seeing" the answer, or which gave no evidence of any calculation strategy being applied.
- Those responses grouped as "counting strategy" included references to the range of the results, the total score for the group or any other numerically based strategy other than stating a pass mark.
- Using an arbitrary "pass mark" was such a widespread strategy among some groups of teachers that it deserved a separate category. The better spelling group was seen as that which had more students scoring at or higher than some arbitrary pass mark. Interestingly, the pass mark was designated as "4" correct by some teachers and "5" correct by others indicating that even the value of a pass is subject to interpretation, even though the teachers who used this method were comfortable with it.
- the categories of "median" and "mean" indicate that the responses obtained showed correct use of these strategies. Only one subject utilised the median, although it was a reasonable approach to the question.

Some participants attempted to use a middle value but were not actually using a median. These responses were placed in the "Counting" category. Attempts to use a middle value are in line with the findings of Mokros and Russell (1992) and Berenson et al. (1993).

There appeared to be some differences between responses to the two questions. High school in-service teachers in particular seemed to change their approaches. In the simpler question 2 the preferred strategy of this group was an arbitrary pass mark, used by 31.6% of these respondents. In contrast, only small numbers of primary or early childhood teachers went to this approach, although it seemed that in-service teachers were marginally more likely to use it than pre-service teachers. In question 3 though, different strategies were adopted, especially by high school in-service teachers. The percentage using a pass mark was only 10.5% but the percentage using the mean had more than doubled to 63.2%. Pre-service teachers seemed to be more consistent in their strategies. Further discussion of these results will be undertaken when the developmental analysis is considered.

TABLE 3 Summary of responses to question 2

Question 2 N=136

	IN-SERVICE TEACHERS				PRE-SERVICE TEACHERS													
	High School		Primary/ ECE		Dip.Ed. Maths		Dip. Ed. Primary		B.Ed. High School		B.Ed. Primary		B.Ed. ECE Hobart		B.Ed. ECE Laun.		Total sample	
	No	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Incorrect	3	15.8	1	5.9	1	11.1	2	9.5	3	37.5	5	26.3	4	22.2	9	36	28	20.6
Visual Strategy Only	1	5.3	1	5.9	3	33.3	0	0	2	25.0	4	21.1	2	11.1	3	12	16	11.8
Counting Strategy	4	21.0	2	11.8	1	11.1	1	4.8	0	0	3	15.8	1	5.5	6	24	18	13.2
Pass Mark	6	31.6	3	17.6	0	0	2	9.5	2	25.0	1	5.3	1	5.5	2	8	17	12.9
Median	0	0	0	0	0	0	1	4.8	0	0	0	0	0	0	0	0	1	< 1
Mean	5	26.3	10	58.9	4	44.4	15	71.4	1	12.5	6	31.6	10	55.6	5	20	56	41.2

TABLE 4. Summary of responses to question 3

Question 3 N=131

	IN-SERVICE TEACHERS				PRE-SERVICE TEACHERS													
	High School		Primary/ ECE		Dip.Ed. Maths		Dip. Ed. Primary		B.Ed. High School		B.Ed. Primary		B.Ed. ECE Hobart		B.Ed. ECE Laun.		Total sample	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No	%	No.	%
Incorrect	3	15.8	4	23.5	1	11.1	2	9.5	6	75.0	8	53.3	6	33.3	10	41.7	40	29.4
Visual Strategy Only	2	10.5	2	11.8	2	22.2	0	0	2	25.0	2	13.3	1	5.5	6	25.0	17	13.0
Counting Strategy	0	0	1	5.9	0	0	2	9.5	0	0	2	13.3	1	5.5	2	8.3	8	6.1
Arbitrary Pass Mark	2	10.5	2	11.8	1	11.1	1	4.8	0	0	0	0	0	0	1	4.2	7	5.3
Median	0	0	0	0	0	0	1	4.8	0	0	0	0	0	0	0	0	1	< 1
Mean	12	63.2	8	47.1	5	55.6	15	71.4	0	0	3	20.0	10	55.6	5	20.8	58	44.3

Question 4

Question 4 was the weighted average question. A number of subjects commented on this question and it was the question that caused the most uncertainty amongst respondents, even when they had calculated the answer correctly. The results are summarised in table 5.

TABLE 5. Summary of responses to question 4

Question 4 N=131

	IN-SERVICE TEACHERS				PRE-SERVICE TEACHERS													
	High School		Primary/ ECE		Dip.Ed. Maths		Dip. Ed. Primary		B.Ed. High School		B.Ed. Primary		B.Ed. ECE Hobart		B.Ed. ECE Laun.		Total sample	
	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%	No.	%
Incorrect	3	15.8	8	47.1	4	44.4	3	14.3	3	37.5	10	66.7	5	22.2	19	79.2	55	42.0
Calc. Correct	16	84.2	9	52.9	5	55.6	18	85.7	5	62.5	5	33.3	13	77.8	5	20.8	76	58.0

A typical answer to this question is given below, demonstrating the uncertainty which many subjects showed about this question.

70+70+70+70+30+30+30+30+30+30 = 280 +180 = 460 ÷ 10 = 46.
But I don't really think you can compare them like this at [sic] 46 seems odd. Its way ↓70 + ↑30!!

(Primary Dip. Ed. Student. Psychology major)

Among the incorrect responses the commonest was an average of 10, obtained by adding the average numbers in each type of hospital (70 + 30) and dividing by the total number of hospitals (10). Altogether eleven subjects (8.4%) gave this response. The other common error was to obtain an average of 50 by adding the average numbers given for each type of hospital and dividing by the number of categories of hospital (2 - large and small). This answer was given by eight subjects (6.1%). These mistakes are typical of the erroneous answers obtained by Pollatsek, Lima and Well (1981). Overall, 58.1% of all responses showed correct reasoning, although much of the mathematics demonstrated was very simplistic, typified by the response shown above.

Summary of the qualitative analysis of the results.

A qualitative analysis of the results appeared to give clear indications that there are widespread misunderstandings about the arithmetic mean, and its uses. These were demonstrated by teachers from all groups, primary and high school, pre- and in-service teachers. The survey questions presented were graded in difficulty. Almost all teachers could respond correctly to question 1 but this had dropped to just over half for the weighted average question, question 4. Some of the types of response seen, especially to the two graphical questions, were similar to those reported in other studies (e.g. Mokros & Russell, 1992; Berenson, Friel & Bright, 1993).

Some variation in response was apparent between different groups of teachers. This was particularly so for the graph questions, where the use of an arbitrary "pass mark" strategy was more common among high school teachers in question 2, but not used by them at all in question 3. Some groups of pre-service teachers, in particular Bachelor of Education candidates, had particular difficulties with question 4, the weighted average calculation. These variations were analysed quantitatively, and this is reported in the next section.

Quantitative Treatment of Results

The initial treatment of the results appeared to indicate some differences in response by different groups. To investigate this further the questions were considered separately and differences tested for significance using contingency table and X^2 tests. The results of the various tests are reported below. Question 1 was not analysed quantitatively any further, since there was no apparent difference between different groups of respondents. The graphical questions appeared to show variation in response as subjects moved from question 2 to question 3, and there were also some differences apparent between high school and primary school teachers. The weighted average question was used particularly to compare sub-groups of respondents. These included high school and primary school teachers and different categories among the pre-service group of teachers. Details of the manner in which each quantitative test was conducted are given below, under the appropriate question.

Questions 2 and 3.

Because these questions were similar in type, they are considered together, although the types of response appeared to be different for each question. Because the large number of different categories of response provided numbers too small for meaningful contingency table analysis, they were combined into three classes

- incorrect, which included no response
- strategies other than the mean used to get a correct answer
- use of the arithmetic mean to obtain a correct answer.

These classifications were used for all comparative tests of questions 2 and 3.

Consideration of the initial results indicated that there appeared to be some variation in response between high school and primary school teachers. In order to obtain numbers large enough to perform a X^2 test of independence, pre- and in-service teachers in the different types of school were combined. Early childhood teachers were included in the primary category. The null hypothesis was that the type of response to the question was independent of the teaching sector, i.e. high school or primary school sector. The results for question 2 are presented in table 6 in Appendix 3.

From the X^2 test, the two categories, type of response and teaching sector, are not dependent ($X^2 = 4.87$, df = 2, p = 0.0874). The X^2 value obtained, however, is sufficiently high to indicate that there might be some relationship between the two groups. In other words, whether a respondent teaches in primary school or high school does not appear to affect the type of response made to question 2 to a great extent, but may play some part.

A similar test, using the same row and column categories, was carried out on responses to question 3. The results are summarised in table 7 in Appendix 3. In this case, the value of the chi-square indicates that the null hypothesis is supported ($X^2 = 0.22$, df = 2, p = 0.895). The type of response to question 3 is clearly independent of the teaching sector.

These results seem to indicate that the type of teacher may play some part in affecting the response seen to question 2 although this is not significant at the 0.05 level. The response to question 3 is independent of the type of teacher. When the relative percentages of each type of teacher are considered some interesting differences appear. In question 2 just over half (~53%) of high school teachers used strategies other than finding the mean, and just over one quarter (~ 28%) used the mean as a suitable measure for this question. In question 3 the figures are reversed with only 25% using strategies other than the mean and nearly half (~47%) using the mean to answer the question. In contrast, the proportion of primary teachers using the mean remained fairly constant (Q2,46%; Q3,43%). The percentage of teachers from each category making an incorrect response was very much the same for each question - approximately 20% in question 2 and approximately 30% in question 3. This would seem to indicate that the major factor in the apparent difference between question 2 and question 3 is the change in strategies used by high school teachers.

To investigate this further, a comparison was made of responses to each question by all teachers. The same categories of response were used as in the previous analyses. The types of response obtained by all teachers to question 2 and 3 are reported in table 8 in Appendix 3. The null hypothesis was that the type of response made was independent of the question being answered.

The chi-square test indicates that there appears to be some influence by the question on the type of response ($X^2 = 6.308$, df = 2, $p = 0.043$). The null hypothesis of independence between the categories is not supported. If the column percents are considered, the use of the mean, row 3, remains fairly constant from question 2 to question 3. The percentage of respondents using strategies other than the mean, row 2, drops however from 38.2% for question 2 to 25.2% for question 3. The percentage obtaining an incorrect result, row 1, rises from around 20% to around 30%. The most obvious inference is that those who used strategies other than the mean were not able to apply these strategies to the more complex question 3, and made more errors. The original qualitative analysis, though, indicated that high school teachers in particular were changing strategies from question 2 to question 3. In particular, they used a "pass mark" approach for question 2 but went to the mean in question 3. In order to resolve these apparently conflicting views, the responses were separated into the two categories of high school and primary school teachers. The types of response obtained by each of these groups were then tested for independence from the question by applying a chi-square test. Table 9 indicates the responses of primary teachers to the two questions, and table 10 shows the result for the same test applied to the types of response obtained by high school teachers. Both of these are included in Appendix 3.

The chi-square result obtained, supports the null hypothesis that for primary teachers the type of response obtained was independent of the question (X^2 = 3.17, df = 2, p = 0.205). For high school teachers, however, the response verges on statistical significance at the 0.05 level (X^2 = 5.92, df = 3, p = 0.052). In other words, high school teachers may be more likely to change their type of response in relation to the question asked. This variation becomes clearer when the column percents are considered. The greatest change in response from question 2 to question 3 is from the use of strategies other than the mean, row 2, to the use of the mean, row 3. Over 50% of high school teachers used an approach other than the mean for question 2, but only 25% used these strategies for question 3. In contrast, the use of the mean rose from about 28% in response to question 2, to nearly 50% for question 3. In contrast, primary teachers showed little change in the use of the mean, row 3, (Q2, 46%, Q3, ~43%). Rather, those using strategies other than the mean, row 2, dropped from 33% for question 2 to about 25% for question 3, and the percentage obtaining an incorrect answer rose from 21% for question 2 to nearly 32% for question 3.

For primary teachers, the results did not appear to depend on the question, whereas for high school teachers it did. Since the context of the questions, the spelling test results for two groups, and the presentation, in graph form, were the same, it would appear that the increased complexity of question 3 caused a shift in approach in high school teachers in particular. One possible explanation is considered in the developmental section of the results.

No statistical comparison of response between pre- and in-service teachers could be made on question 2 because of the low numbers in some categories of response. From the original results, summarised in Table 3, there did not appear to be any marked difference between these two groups on this question, so no further analysis was undertaken. Although a comparison could be made on question 3, it did not show any highly significant difference between these two groups, and was not explored further.

Question 4.

This question was one which involved calculation of a weighted average. If the number of incorrect responses is considered, this question was the one which was the most difficult for all groups of teachers, with just over 40% obtaining an incorrect answer, or no result at all. There appeared to be some marked differences between groups. Various contingency table analyses were undertaken and the results are detailed below.

For all analyses of types of response to this question, only two categories were used. These were

- incorrect, which included no response
- correct, including all responses which obtained the correct answer by some form of calculation.

Since some differences had been noted between the responses from high school and primary school teachers to earlier questions, these were the first groups considered here. The null hypothesis proposed that the response obtained was independent of whether the respondent was a high school or primary school teacher. The results of the chi-square test applied are reported in Table 11, Appendix 3.

The null hypothesis is not supported by the X^2 figure ($X^2 = 4.11$, df = 1, p = 0.043), and it appears that there is some relationship between the type of teacher and the response obtained to question 4. When the column percents are considered, the results for primary teachers, column 2, show that about half can calculate the weighted mean correctly. Nearly three-quarters of high school teachers, in contrast are able to obtain a correct result. It seems that high school teachers are significantly better at calculating a weighted average in the context of the question posed in the survey.

Because the responses were more clear cut than those for questions 2 or 3, this question was used to provide a basis for comparing other groups within the sample. The first of these comparisons was between pre- and in-service teachers. The same classification of response was used (row 1, correct, row 2, incorrect) as previously, and the test was performed in the same way. The chi-square test results for the comparison between pre- and in-service teachers are shown in table 12, Appendix 3. No significant difference was found between the responses of pre- and in-service teachers to question 4 ($X^2 = 2.33$, df = 1, p = 0.126). The null hypothesis that the response was independent of the category of teacher was supported.

Within the pre-service group of teachers, however, there appeared to be some differences. The results summarised in table 5 appeared to show that Diploma of Education pre-service teachers were more able to obtain a correct response than Bachelor of Education students. These two groups were, therefore, compared statistically, using the same techniques as for the high and primary school teachers, and the pre- and in-service teachers. It was postulated that the type of response to the question was independent of the type of pre-service training being undertaken. Results are summarised in table 13, Appendix 3.

This X^2 result provides very strong evidence against the hypothesis that the categories are independent ($X^2 = 9.314$, df = 1, p = 0.002). There is a highly significant difference between Diploma of Education and Bachelor of Education pre-service teachers in their response to this question. The Diploma group was smaller, and contained those pre-service teachers who had a background of a high level of mathematics study. As well, many of the primary method students had majored in psychology in their first degree, and had had some statistical training as part of their study. This may well have contributed to the difference.

Two groups which might have been expected to respond in similar ways were the early childhood pre-service teachers studying at Hobart and Launceston. The results summarised in table 5, however, seemed to indicate that there was some variance in the responses from these two groups to question 4. These two groups were relatively similar in size, Hobart having 18 and Launceston 24. It was hypothesised that the responses of these two groups would be independent of the place of study. A chi-square test was carried out and the results are shown in table 14, Appendix 3.

This result gives very clear evidence against the null hypothesis ($X^2 = 11.09$, df $= 1$, $p = 0.0009$). The comparison of these two apparently similar groups indicates that the place of education was highly significant in determining response to a weighted average question in this survey. When the column percentages are considered, pre-service students educated in Hobart appeared to be significantly more likely to calculate the answer correctly, with ~72% correct. In contrast, in the Launceston early childhood group, only ~ 20% could calculate this correctly. It should be borne in mind, that this result was only on one question of the four posed, and it was the question which was the most problematical for all groups. Because of small numbers, it was not possible to test the responses of these groups to other questions. The implications of this result will be considered in a later chapter.

<u>Summary of the quantitative analysis</u>

The statistical analysis provided quantitative evidence to support or refute some of the thinking about the research questions originally posed. It was applied particularly to the questions relating to similarities or differences between various sub-groups of teachers.

In summary, the findings were as follows.

- There was no significant difference in the responses of pre- and in-service teachers on question 4.
- There was a significant difference in the responses of high school and primary school teachers to question 4.
- There was a highly significant difference between responses of Diploma of Education and Bachelor of Education pre-service teachers to question 4.
- There was a highly significant difference between responses of early childhood pre-service teachers in training in Hobart and Launceston to question 4.

The implications of these findings will be discussed in chapter 5.

Some evidence that the context of the question influences response was found in relation to the graphing questions on the survey, questions 2 and 3. While no significant difference was found when the responses of these two groups were compared for each question, when the responses of each group to the two different questions were compared, a significant difference was found. Further analysis found that for high school teachers in particular, the responses obtained to these questions was dependent on the question. One possible explanation of this is presented in the next section, the cognitive development perspective.

Analysis of the Results From a Cognitive Development Perspective

The responses to the survey questions were considered in relation to a cognitive development model. They were classified using the framework provided by the SOLO Taxonomy, detailed in chapter 3. As in previous sections, the questions were considered individually, as the type of response differed considerably from question to question. The classification of the responses followed the unistructural, multistructural, relational model proposed by Watson *et al.*(1993). Thus a unistructural response was one in which only a single bit of information was used, a multistructural response was one in which several bits of information were used sequentially and a relational response was one in which all the information was taken into account and utilised in a coherent and systematic manner. Both ikonic and concrete-symbolic modes of thinking were identified, and a model is proposed for interaction between the two. The results are reported below, question by question.

Question 1.

This question was framed to access concrete-symbolic thinking only. Given the context of the question, a reasonable relational response with respect to the concept of average would have been to discuss the inclusion of the outlier, and probably calculate the results both with and without this value. No response showed this thinking, the most complex being the result quoted previously in the section concerning qualitative results. At this macro level, the typical response was seen as unistructural with a fixation on using the "add them up and divide" algorithm. All responses demonstrated this approach, even those who obtained an incorrect result. Altogether, four teachers (3%) obtained a result which lay outside the range, but all of these had attempted to use the algorithm. The error made was to leave out one of the repeated values of 3.2.

In two cases, this was shown to be one of the two values which occurred in juxtaposition. This could be an indication of some deeper misunderstanding. It was surprising that more high school teachers did not comment on the outlier, since most of them also taught science in which this type of thinking is important. This may indicate that the practice of using a number of measurements obtained from an experiment, critically examining them and then using the appropriate summative measure is not widespread. The context in which the survey questions were given might also be important - this was a "mathematics" survey rather than a "science" survey.

When the processes used were considered, at a micro-level the responses were overwhelmingly multistructural in terms of the mathematical methods used, typified by the example in Figure 1. One primary teacher also indicated values of the median and mode, demonstrating a growing understanding of the idea of average.

Question 4

This question also accessed the concrete-symbolic mode, but required a higher level of thinking than the response to question 1. The answers to this question were also relatively clear cut. Again the successful strategy was multistructural in approach at the micro-level, considering the processes used. The calculation was completed in a stepwise manner, with the total number of babies being calculated first and then the average obtained from this. A typical response is shown in figure 3.

$$4 \times 70 + 6 \times 30 = 280 + 180 = \frac{460}{10}$$

$$= 46 \text{ babies.}$$

Figure 3

Questions 2 and 3.

These two questions were different from the other two in that a visual component was present. A number of alternative solutions was demonstrated. Some of these appeared to rely on an ikonic component. Others seemed to use a combination of ikonic and symbolic reasoning. Very few relied totally on visual reasoning.

Some teachers used, or attempted to use, strategies based on counting. These included comments on the range of the results, sometimes including evidence that the shape of the graph had been a factor in their reasoning. Others calculated the total number of words correct and compared these for the two spelling groups, rather than calculating the mean. This strategy worked well for question 2 where the groups were the same size, but not for question 3 where the group size was markedly different. Those subjects who used this strategy for questions 2 and 3 could be seen to be concentrating on the number of correct answers, but neglecting the sample size. This may also be an indicator of deeper misconceptions about sampling and sample size.

A number of responses made use of an arbitrary pass mark. This seems to reflect the context of the question, and the teacher, since high school teachers appeared more likely to use this strategy. It would seem that teachers of older children are more likely to think in terms of a "pass" and to allocate an appropriate mark against which to measure this. High school teachers in particular seemed to use this strategy for question 2 but did not use it for question 3. This changed response from question 2 to question 3 was interesting, and will now be analysed further.

Question 2 was designed so that the difference between the two groups was immediately obvious visually. The groups were small, and of the same size, so that a direct comparison was easy and valid. Only sixteen teachers out of the sample of 136 (11.8%) used a visual comparison only. A much greater number (35 teachers, 25.7%) used some sort of numerical system other than recognised statistical measures such as the median or mean. These strategies included a consideration of the range of the data, the total number of questions that the group got correct or the numbers of students getting more correct than some arbitrary pass mark. This latter strategy was more prevalent among high school teachers.

Question 3 was more complex. The difference between the groups was very small and not obvious visually. In addition, group sizes were much bigger than those in the previous question and differed considerably. The proportion of teachers responding visually did not differ greatly from that in question 2 (17 teachers out of 131, 13.0%). Numerical methods, though, other than recognised statistical measures were preferred by only fifteen teachers (11.5%), less than half of those using these strategies for question 2. Some possible reasons for this difference will now be considered.

Question 2 effectively invited a response in the ikonic mode. It provided a very clear visual picture of the information, which could be simply compared. Teachers are also familiar with the context - deciding which group has better spellers. This context, however, invites justification from teachers. They work within the necessity to explain and justify judgements about individuals and groups of students, and in their day to day operations spend considerable time collecting evidence and records in order to do this. For teachers, therefore, these questions could produce some conflict between the visual presentation accessing an ikonic mode of thinking, and the need to explain and justify judgements of this nature in the concrete - symbolic mode.

Responses to question 2 seem to indicate that this could be the case. A relational ikonic response would take into account the "look" of the data, the shape and spread shown on the graph, the number of students in the group and the level of their response, compare the two groups, and integrate this into a judgement of which group is better. The necessity to justify this would then lead to translation of this judgement into concrete-symbolic mode. To base a justification simply on the appearance of the data would not generally be accepted in the context of the question. Teachers, therefore, appeared to prefer to use some numerical basis for this. Many went to the mean for corroboration of their intuitive findings, but others, in particular experienced high school teachers, used an arbitrary "pass mark" to support their findings. This approach is consistent with the every day working tools which teachers use. As previously indicated, in SOLO terms the concrete - symbolic responses were rarely more than multi-structural. This again is consistent with the context, since teachers are looking only for an acceptable explanation of their intuitive understanding, rather than a consideration of the "best" method for comparing the groups.

Question 3 was more problematic for teachers. It accessed the ikonic mode through its presentation in graphical form, but was considerably more complex than the earlier question. Nevertheless, a number of teachers responded with a visual justification only. The biggest difference was in the type of numerical justification utilised. Many of those teachers who had used other numerical methods, in particular the arbitrary pass mark, went straight to the mean when faced with this more difficult problem. While the methods used to calculate the mean were generally multi-structural at the micro-level, at a macro-level using the mean as opposed to less complex numerical methods could be seen as a higher level concrete-symbolic response. Teachers appeared to prefer a concrete

- symbolic mode of operating when faced with a problem which was not easily solved in the ikonic mode.

There is some evidence for this from some comments written in response to the questions. In answers to question 2 there were a number of responses of the following type.

> *I would look for the average of the score although visually it would suggest Group A are the consistently better spellers.*
>
> Primary Teacher L10

This teacher then went on to calculate the means of each group correctly.

> *Group A's graph seems to show they are achieving better spelling results for the greater part of the class.*
>
> Primary Teacher B1

> 1. *You could just measure the total score for each to get a quick idea. Gp A = 48 Gp B = 38*
> 2. *Visually Gp A looks better though I'd explain visually* [sic] *representations can be misleading.*
> 3. *I'd talk about spread (which is greater in Gp. B). Clearly if we are discussing individuals then Gp B has the best individual performances.*
>
> High School Teacher 11

This teacher was the same one who had noted the outlier in the first question. The response by this teacher to question 3 indicated a very good grasp of the idea of average.

I would work out the mean for each plus the standard deviation to see if there is any obvious difference. If not then you could do some statistical tests on them if you had more advanced students. Do a chi squared or t test or something.

High School Teacher 11

In question 2 this teacher had commented on the visual aspects of the graph, albeit with the qualification about misleading graphs. In question 3, there was no evidence of ikonic thinking. Rather, the appropriate concrete - symbolic method had been used. The subject also had considered the questions in the context of teaching students, which might have made a difference to the way in which the questions were regarded. From experience, most teachers when teaching are concerned to make students aware of the limitations of particular methods.

Another response to question 3 was the following.

Spelling group D are the better spellers as on average they score 4.5 correct per child as group C on average score 4.2 per child.

Note if we consider number of students getting 5 or more correct the results are reversed! Criteria for selection will dictate the answer you want.

High School Teacher 14

For question 2, this teacher had used the strategy of finding the total number correct for each group and comparing these. The arbitrary pass mark strategy had also been applied to question 3 but abandoned in favour or the mean as a more suitable measure, in the absence of any other direction. Only concrete-symbolic thinking was indicated.

Several other high school teachers changed strategies from question 2 to question 3, and these were generally from some sort of numerical justification to using the mean. No reasons were given for the changed approaches. Teachers, rather, seemed to be choosing the method which seemed most appropriate to them under the circumstances.

Responses to these two questions could be considered at two levels. At a macro level respondents were choosing to answer in a way which was appropriate to the context. This led to a relational ikonic response for question 2. At this level, the choice of the mean as the preferred measure for question 3 could be interpreted as a relational response in the concrete-symbolic mode. Teachers had considered the question, recognised the increased complexity over question 2 and had chosen a response appropriately. Recognising that the groups had greater numbers than those in question 2, noting the different numbers from each other so that a direct comparison of the graphs was difficult, integrating this with existing knowledge and choosing calculation of the mean as a suitable strategy is at worst a very good multi-structural response.

It seems possible that the changeover seen from ikonic to concrete-symbolic thinking demonstrates a parallel structure of modes of thinking, hypothesised by Watson *et al* (1993). Multi-modal thinking, particularly when faced with a novel problem, has been reported by a number of researchers (e.g. Collis & Romberg, 1990; Watson, Campbell & Collis, 1993). Essentially, when faced with a problem, people select which mode of thinking is most appropriate for them in that situation. While sensori-motor functioning is available, in most school based circumstances the real choice is between ikonic and concrete-symbolic thought. The problem may then be followed through in the chosen mode, or there may be transfer from one mode to the other. This may be represented diagrammatically as shown in figure 4.

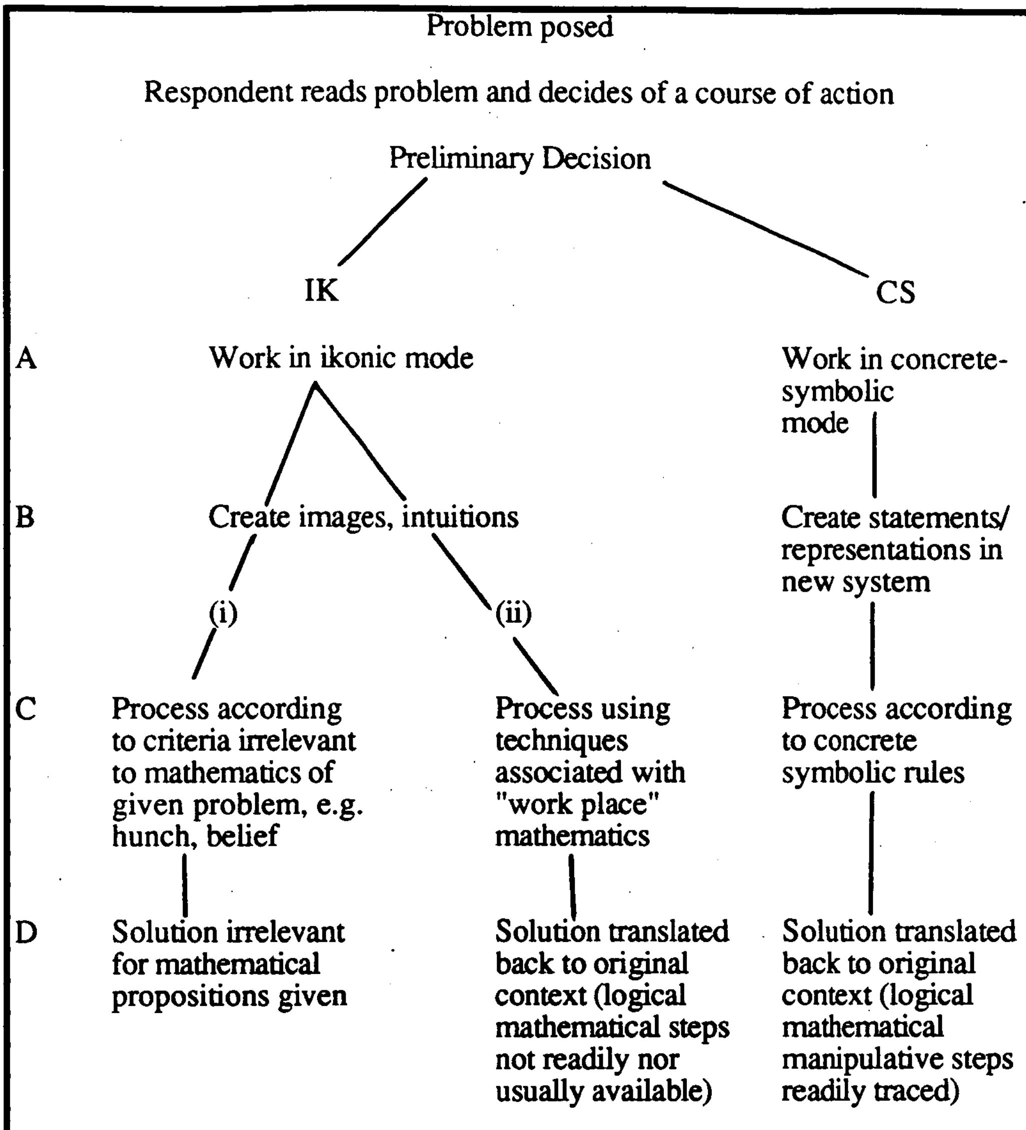


Figure 5. The Problem Solving Path (from Watson, Campbell & Collis, 1993)

Following this pathway, teachers faced with a straightforward graphical problem, such as that in question 2, could choose the ikonic (IK) route, or the concrete-symbolic (CS) path. An individual choosing the ikonic path, then has a choice of two possible paths, labelled (i) and (ii), for processing the problem. The choice of the IK(ii) route allows transfer into the concrete-symbolic mode at points B and C. Given the context of the question, spelling scores, the use of "work place" mathematics to process the problem, and transfer into the concrete symbolic mode to justify and communicate the answer seems very appropriate for teachers.

When faced with a more complex ikonic problem, teachers seemed to prefer to move down the concrete-symbolic path. High school teachers particularly changed their strategy from the easier ikonic problem of question 2 to the more complex problem of question 3. This was not so evident for primary school teachers who showed no significant difference in their responses to these questions.

Using this model it seems possible that the generalised categories of response to questions 2 and 3 could be mapped onto a problem solving path in either ikonic or concrete-symbolic modes. One possible mapping is shown in figure 6. The categories IK_A, IK_B and IK_C refer to the levels of ikonic response indicated in figure 5. Campbell et al (1992) hypothesised that developmental sequences in the concrete-symbolic mode could be described as cycles of increasingly complex responses. It could be that the responses obtained here also demonstrate a developmental sequence in the ikonic mode. Thus the levels IK_A, IK_B and IK_C might be considered as equivalent to a unistructural, multi-structural and relational SOLO Taxonomy level. The U, M and R categories under concrete-symbolic refer to the SOLO Taxonomy levels referred to in chapter 3.

Type of response	SOLO level of response Ikonic	Concrete-symbolic
Intuition leading to incorrect response	IK_A	Pre-structural
Visual strategy only	IK_A	U
"Counting" strategy	IK_B	M
"Pass mark" strategy	IK_C	M
Use of Mean	No ikonic support	R

Figure 6. Summary of types of response to survey questions 2 and 3

The first two types of responses, incorrect and visual only, were categorised as a lower level ikonic response, IK_A, since no attempt had been made in the ikonic mode to process the information by considering the various features of the graph. Incorrect responses were mainly those which used no concrete-symbolic reasoning, or made no attempt to solve the problem, typified by comments such as "too difficult". Those classified as using a visual strategy were those which referred only to visual features in the justification. This was classified as a unistructural concrete-symbolic response. The "counting" strategy followed the ikonic pathway to IK_B initially, since these responses mainly focussed on some particular feature of the graph, such as the spread or range of the data. This was seen as a multistructural response in the concrete-symbolic mode since this ikonic information was translated into a concrete-symbolic justification which included some sort of numerical processing. Using the "pass mark" was classified as a higher level ikonic response, IK_C, since work place mathematics was involved. In the concrete-symbolic mode it was classified as multistructural since the same level of reasoning was employed as the "counting" strategy. Use of the mean was seen as the top level of concrete-symbolic reasoning in this problem solving cycle. Teachers using this method, especially in question 3, did not appear to rely on ikonic support. Some teachers in response to question two, did appear to reason in both modes, using the mean to justify their initial ikonic response.

In the concrete-symbolic mode it seems that two developmental cycles are being seen with respect to the concept of average. A first cycle of this mode could be that described in figure 6. In order to process a straightforward calculation of the arithmetic mean, a number of skills are required as a pre-requisite. These include addition, division and some understanding of the order in which these need to be done. Thus, any calculation of the mean, however simplistically performed, needs relational thinking in this first cycle of the concrete-symbolic mode. This provides the top level of response in the first concrete-symbolic cycle, described in figure 6.

A second cycle, which starts with the notion of the mean, may be that accessed by the hierarchical sequence of the questions posed in this survey, as indicated in chapter 3. Question 1 required only a unistructural understanding of the mean in this cycle. The second and third questions needed at least multistructural understanding since a number of ideas had to be used to arrive at a response. Finally, question 4 was relational in nature in this cycle since understanding of the mean had to be placed into the context of different sized hospitals and then integrated into a single figure which summarised all the data available. Questions 2 and 3 were different from the others in that they also allowed an ikonic response. Thus, they provided more choice in the problem solving path than those questions framed only in the concrete-symbolic mode. The responses received from teachers indicated that many of them exercised this choice when it was appropriate to them.

The types of response seem then to form a hierarchy in the ikonic mode, which is paralleled by similar structures of increasingly complex response in the concrete-symbolic mode. It is interesting, however, that a particular group of teachers, namely high school teachers, appeared to function mainly in the ikonic mode when faced with an "easy" problem, but moved straight to relational level functioning in the lower level cycle of the concrete-symbolic mode when faced with a "difficult" problem. It may be that these teachers were exercising some level of conscious control over the choice of problem solving path.

One interesting feature of the responses is that all respondants indicated that they had reached this relational level of thinking in a lower level concrete-symbolic cycle in their answers to question 1. In question 2, however, as shown in table 3, nearly 40% of the respondents chose not to use this level of reasoning. Instead, less than relational level concrete-symbolic thinking was demonstrated in this earlier cycle. In question 3, less than 25% chose to use these lower level strategies and the higher level use of the mean was preferred by about 44% of respondents. It seems that in the context of the questions 2 and 3, where ikonic thinking could be used, the teachers in this study preferred to use high level ikonic thought, supported by low level concrete-symbolic justification for a simple problem, but moved to higher level concrete-symbolic thinking when the problem was less straightforward.

There are a number of implications from this finding. Amongst these is the question of whether it is more efficient to process information in a lower level mode, or a lower level developmental cycle within a mode. Further consideration will be given to this issue in chapter 5.

The qualitative, quantitative and cognitive development, perspectives used to analyse the results obtained, have produced a number of findings. These will now be explored further, and some conclusions drawn.

CHAPTER 5

DISCUSSION

The results obtained from this study will be discussed with reference to the original research questions, and the relevant related research. The research questions originally posed were as follows.

1. Are similar misconceptions to those described elsewhere present in Australian teachers?
2. Are there differences in response between primary and high school teachers?
3. Are there differences in response between pre- and in-service teachers?
4. Are there differences in response between sub-groups of the pre-service teachers such as Diploma of Education students and final year Bachelor of Education students?
5. How does the context of the question alter response?

Each of these questions will be addressed individually, together with the relevant results. The samples used for this study were not random selections and the results obtained must be considered with this limitation in mind. The teacher groups, however, while not truly random, were involved in the professional development experiences which provided the researcher access to them because of their school involvement, and not because of any great personal interest. In this they are typical of many teachers, so that the samples used were less atypical than a group which had chosen to be involved because they had a particular interest in mathematics education.

Research Question 1

Are similar misconceptions to those described elsewhere present in Australian teachers?

With the limitations in mind, it would appear that the situation among these Tasmanian teachers is not unlike that found in other countries. The understanding of notions of average was surprisingly limited. When the findings are compared in detail to other research, some similarities are seen.

When faced with a straightforward calculation of an arithmetic mean, Tasmanian teachers were generally able to calculate this. As with the findings of Pollatsek *et al.* (1981) this was at an unsophisticated level, using a step-wise approach. All teachers indicated that they knew how to calculate the mean, although some made errors in the computation, but knowledge of the calculation did not appear to have allowed any real development of the concept of average. Teachers with a well developed sense of average would notice when their result fell outside the range, and would recognise the importance of an outlier. This lack of depth in their understanding is in line with the findings of Pollatsek *et al.* previously cited.

When the calculation was placed in a different context, and presented in a graphical format, a number of teachers were unable to translate their knowledge of the computation to the new situation. This was even more marked when the situation was made more complex, as in question 3.

Further considerations about teachers responses to graphical representation come from the work of Berenson, Friel and Bright (1993). They researched a number of alternative conceptions and fixations of elementary teachers faced with different graphical representations. They found that teachers tended to "fix" on some feature of the data as represented on the graph, and the nature of the fixation was related to the type of graphical representation used. Bar charts are the simplest form of graph commonly used by teachers to present data, and the researchers suggested that some teachers have difficulty moving to more complex forms of representation such as histograms.

Although in the present study the graphical representations were of the same type, rather than different types as in the study by Berenson *et al.*, there were some similarities in response noted. Some subjects had idiosyncratic preferences about the graph interpretation, not unlike the "fixation" found by Berenson *et al.* Alternative conceptions in statistics, Berenson *et al* found, were linked to some large, concrete feature of the graph, commonly the mode or the horizontal scale. Often teachers attempted to use the horizontal scale to determine the centre or middle of the data. A typical value was often seen as the mode of the data. This is not dissimilar to some of the strategies demonstrated by teachers in response to questions 2 and 3. Some of the responses categorised as "counting" strategies used the horizontal scale to identify a range, for example. The difference, however, is that in this study different approaches were seen using the same type of data and representation which had increasing complexity, whereas in the American study, the data type and representation were different. Many of the recommendations for professional development made by Berenson *et al.* could probably be applied to teachers here, and some of these will be considered later.

The weighted average problem, question 4, was problematical for nearly half the teachers surveyed. As in the study by Pollatsek *et al.* (1981), successful strategies used were mainly multi-step, rather than integrated into an elegant mathematical solution. Some teachers who were able to succeed on what was essentially the same problem in question 3, were unable to translate this to the written context. The non-transferability of strategies is a recurring theme, and bears out the findings of other researchers about the importance of context (e.g. Gal *et al.*, 1990; Garfield & del Mas, 1991; Jolliffe, 1991).

The issue of non-transferability relates to the cognitive development perspective. If a concept has been developed over a period of time, through a variety of modes of thinking, with a clear target mode in mind, and with a determined effort to link this into existing cognitive structures, it is more likely that the necessary relational thinking will evolve which would allow the ideas to be drawn upon in a variety of situations. When the cognitive development model of Biggs and Collis (1982, 1991) is applied to the results, it provides an explanation for the simplistic approach referred to above. The teachers involved in the study do not appear to have had access to experiences which would allow their concept development to proceed beyond a multi-structural level, in the developmental cycle which starts with the notion of the mean. This level is sufficient to allow successful calculation of a straightforward arithmetic mean but not sufficient to allow for the other factors, such as outliers, to be taken into account.

In general, then, the teachers in this study demonstrated very much the same sorts of response to questions about average as has been shown in other studies. While one should beware of assuming that these Tasmanian teachers are representative of Australia's teachers as a whole, there are certainly grounds for questioning teachers' understanding of what is arguably the most commonly taught and used statistical measure.

Research Questions Relating to Variations in Response Between Different Groups of Teachers

- .Pre- and in-service teachers.

There appeared to be no significant difference overall to the responses from pre- and in-service teachers. While this is not very surprising given the low priority given to chance and data in the curriculum until recently, it is rather disconcerting that given what is now known about children's learning, young people who have only recently left the school system do not appear to have any better understanding of common concepts than established teachers. The obvious conclusion to be drawn is that the understanding about learning which researchers have developed, and are continuing to evolve, are not being translated into classroom practice. The findings of Pollatsek *et al.* (1981), Garfield and Ahlgren (1989), and Gal *et al* (1990), amongst others need to be taken to teachers.

Much of the new material in the chance and data strand of the curriculum does just this. The Quantitative Literacy materials, the Used Numbers series and the Australian produced Chance and Data Investigations, detailed in Appendix 1, are all research based. As Pegg (1989) points out, however, unless the cognitive development of teachers is at a sufficiently high level to be able to generalise from these examples, the lessons will be little more than well run "one offs".

While mindful of the limitations of the study, it nevertheless seems clear that courses need to be developed both for in-service professional development, and for pre-service training which will allow teachers to address misunderstandings, and become fully confident in this area of the curriculum. Further comments concerning teacher development will be made later.

• High school and primary school teachers.

There were some significant differences between primary and high school teachers when answering question 4, the weighted average question. High school teachers were more likely to calculate the result correctly. This result is not surprising, given that high school mathematics teachers generally have studied a higher level of mathematics. It is rather disconcerting though since implicit in the Australian national curriculum documents (AEC, 1990; Curriculum Corporation, 1993) is the recognition that concepts need to be developed in children over a period of time. The appropriate preliminary experiences, therefore, for children to acquire concepts such as "average", which are needed before a weighted average can be fully understood, need to be provided by teachers in the primary sector. While significantly worse than high school teachers at calculating a weighted mean, primary teachers were no worse at finding averages in other situations. When the responses were considered within the framework provided by the SOLO taxonomy, they demonstrated multi-structural features, indicating that ideas about the mean had not been integrated into a cohesive whole. This also has implications for professional development.

• Pre-service teachers in Diploma of Education and Bachelor of Education courses.

There were highly significant differences between these two groups on question 4. While this is hardly surprising, given that some of the Diploma students came from the mathematics/computing method course, and had therefore recently completed a mathematics degree, it is grounds for some concern. Many Bachelor of Education trained teachers move into high school teaching. Some intend to do this, and are training in specialised areas such as physical education. From experience, however, it is just these teachers, physical education or technical studies teachers, that end up teaching high school mathematics to make up their load. It was a teacher with this sort of background, in this case boat building, that made one of the errors shown in figure 2. The concepts and skills considered in this study are not unusual, and are commonly taught. It is not unreasonable, therefore, to expect all pre-service teachers to have a basic grounding in this area.

- Early childhood pre-service teachers from Hobart and Launceston campuses.

With the previous comments in mind, it is very worrying that there was such a highly significant difference between the Hobart and Launceston groups of teachers. The clear difference between the performance of Hobart based students and Launceston based students cannot pass without comment. It seems that students in Launceston have not had opportunities to build their knowledge base of the concept of the arithmetic mean to a level which would successfully allow them to calculate a weighted average. It cannot be argued that the early childhood trained pre-service teachers will not need to teach this concept, and therefore do not need to know it. Given the trend towards generalist teachers, many early childhood trained teachers will find themselves teaching upper primary classes, and primary teachers are moving into middle school areas. This makes it even more imperative that recently trained teachers have a sound understanding of learning areas, as well as pedagogy. Further investigation may be indicated concerning the relatively poor performance of the Launceston pre-service teachers on a question relating to a weighted mean, compared to an apparently equivalent group from Hobart. There are a number of possible explanations which this study cannot address.

To summarise then, teachers, it seems from this albeit limited study, need improved content knowledge in the area of measures of centrality. In Tasmania, at least, both pre- and in-service teachers appear to lack specific content knowledge about the arithmetic mean. During the period of this study, it was possible to ask a very small number of teachers to respond to children's understanding of average, including mean, median and mode. While the numbers involved were too small for this to become part of this study, it was interesting to note that teachers did not appear to recognise misunderstandings, or correct reasoning, in children. The understanding of average demonstrated in this context was computation of the arithmetic mean. When a child used the mode as an average in an appropriate context it was marked "wrong". Further studies concerning teachers' reactions to children's work in this area of the curriculum would be worth undertaking.

<u>Research Question 5</u>

<u>How does the context of the question alter response?</u>

The qualitative analysis of the results indicated that subjects found the questions increasingly difficult as they moved from a straightforward calculation of the mean, through two problems where the information was presented graphically, to calculation of a weighted average. Of most interest were the two graph questions where some unexpected differences were found in responses to the questions. There was a significant difference between the responses of all teachers to questions 2 and 3. When this was broken down into high school and primary school teachers, it seemed that the varying responses were not significant for primary teachers but were for high school teachers. One possible explanation of this was provided by a cognitive development model.

There appeared to be a strong interplay between ikonic and concrete-symbolic thinking, with teachers accessing modes of thinking appropriate to the situation. Where the problem was more complex, the preferred mode was concrete-symbolic, reflecting the heavy emphasis placed on that mode in education. It is interesting to surmise whether teachers who had been able to build up the concept of average through kinaesthetic and ikonic modes would have shown the same behaviour.

This preference for accessing the concrete-symbolic mode when faced with a more complex problem, could also explain some of the differences between high and primary school teachers. In general the high school teachers had studied mathematics at a higher level than primary teachers. They could therefore feel more confident of their judgements and be less concerned about the justification than primary trained teachers. They were satisfied with low level justifications, such as a "pass mark". Primary teachers, in contrast, needed to use concrete-symbolic thinking to justify straightforward tasks, such as that in question 2. The behaviour of primary teachers in response to question 2 could be seen as equivalent to that of high school teachers on question 3. Teachers who have limited content knowledge, need to concentrate on the mathematical reasoning in order to obtain a correct response. Mathematics, by its nature, has a strong basis in concrete-symbolic thinking. Consequently, the primary teachers who were less well versed in mathematics tended to use strategies associated with the concrete-symbolic problem solving path described in chapter 4. Only on question 3, the more complex graph question, did high school teachers consistently use concrete-symbolic reasoning. Their preferred path seemed to be ikonic with a move into concrete-symbolic reasoning to justify their results.

A study of this size cannot be taken as necessarily representative of teachers as a whole, but there are indications that this multi-modal thinking in teachers is worthy of further research. A clear indication of parallel cycles of response in the ikonic and concrete-symbolic modes was seen when problems accessing ikonic thinking were presented. There was also some evidence for a second cycle in the concrete-symbolic mode. Cycles of this nature in children's thinking have been reported by other researchers (Watson, Campbell & Collis, 1993). It is a reasonable supposition that they also occur in adults.

This study, though limited, also has obvious concerns for pre- and in-service teacher education. What then does this study, given its limitations, mean for professional development? The implications of this report will now be considered in chapter 6.

CHAPTER 6

IMPLICATIONS FOR PROFESSIONAL DEVELOPMENT AND FURTHER RESEARCH

The implications of this study are considered both for pre- and in-service teacher development, and for further research. While limited in scope, from this study there appear to be some clear indications of future directions in both of these areas.

The teachers involved in this study did not demonstrate a relational understanding of the application of the arithmetic mean. It appears from this that improved pre- and in-service training is needed. If recommendations from research are considered, this should be practically based, have appropriate support structures for participants as an integral part of the course, and be conducted in such a way that participants are able to construct their understanding of the concepts involved through a variety of modes of thinking. Approaches such as those taken through use of the Quantitative Literacy Project materials (e.g. Kepner & Burrill, 1990) or Teach-Stat, a three year project, aimed at developing teaching skills in statistics among elementary school teachers (Bright, Berenson & Friel, 1993) appear to have the most promise of success. Short courses, even when intensive and highly focussed, do not appear to result in great gains in understanding (e.g. Cox & Mouw, 1992). The importance of continued teacher development programs is evident from the work of Pegg (1989) which indicates that teachers need opportunities to refine understanding of pedagogy in order to utilise fully research based teaching materials appropriate to their subject area.

A number of models for in-service professional development have been proposed. The Mathematics Curriculum and Teaching Program (MCTP) produced a set of guidelines for professional development to support the teaching materials (Owen *et al*, 1988). These suggested several models of teacher development based on research findings concerning effective in-service provision. These models include theory, modelling, opportunities for practice and feedback. They are similar in approach to those being utilised by "Teach-Stat" and the Quantitative Literacy projects previously mentioned. Specific training in chance and data concepts utilising these models is needed as a matter of urgency. While being wary of generalisations from this limited study, it seems evident that in-service courses will need to address basic ideas that might have been expected to be understood.

Some consideration also needs to be given to the learning area content of pre-service courses. The concept of the arithmetic mean is a widely taught and used idea and it seems that pre-service teachers demonstrate no better understanding than in-service teachers. Pre-service teachers also need opportunities to refine their thinking and increase skills in this area. Since there is a growing emphasis on chance and data in the mathematics curriculum, it seems reasonable that all teachers should have a well developed understanding of the essential everyday concept of average. From this study, it seems that courses for pre-service teachers may need to be reviewed, since it is evident that they have only a limited understanding.

The responses at the micro-level referred to in Chapter 3 may be part of the same lower level concrete-symbolic cycle as that described in Figure 6 in Chapter 4. Alternatively, it may be that the processes involved utilise different reasoning. No conclusions can be drawn from the evidence presented here, and further research is needed to clarify this issue.

There are also other implications for future research. If the multi-modal functioning which is apparent in this study is a common phenomenon, then attention should be paid to developing the skills of moving from one mode of thinking to another. The utilisation of the ikonic mode for low-level processing of skills, frees up the concrete-symbolic mode with its emphasis on the relational development of ideas.

One implication of this may be that the more content knowledge teachers have, the less they need to concentrate on the subject matter. They can use lower level ikonic thinking processes to deal with this, which frees the concrete-symbolic mode for different tasks. These tasks could include applying appropriate pedagogic knowledge to student learning. From the personal experience of the researcher, both in the classroom and in work with other teachers, experienced teachers commonly react to children's queries about content at an intuitive level. They can look at a piece of work presented by a child and recognise whether the answer is correct or not without actually processing it. This thinking may be part of a second cycle in the ikonic mode following the cycle identified in this study. Having considered the work in this way, the teacher can then pose suitable questions to help the child refine understanding. Where the teacher is unfamiliar with the content, the answer may need to be processed in the concrete-symbolic mode before a suitable response can be made to the child. These are only conjectures, but this study provides some evidence for multi-modal functioning in teachers which is worthy of further consideration.

If the multi-modal thinking identified here is confirmed in other areas, then there are huge implications for generalist teachers, and teacher education at all levels. Opportunities need to be provided for teachers to move from one mode of thinking to another, in the same way that multi-modal approaches are now being recommended for children's learning (Biggs & Collis, 1991). It would seem that teacher development courses, both pre- and in-service, should aim to develop learning area thinking and skills to a level such that they can be routinely processed in ikonic mode, allowing the concrete-symbolic mode to develop cognitive structures relating to other aspects of teachers' work. This is in line with the view of Pegg (1989) that teachers need to develop a relational understanding of the process of teaching. No generalisations can be made from a study of this size and scope, but further research into the psychology of teaching could provide more information about this area, and make recommendations for pre- and in-service teacher development.

"Numeracy" or mathematical literacy across the curriculum is currently gaining status. Statistical literacy is now being recognised as an integral part of numeracy, and built into the mainstream mathematics curriculum as the chance and data strand, It is evident from this albeit limited study, that the common concept of average is only understood in a simplistic way. This begs the question of teachers' understanding of less commonly taught, but equally important, stochastic concepts such as randomness or sampling. From the evidence presented here, even with the qualifications of sample size and selection, there is still a huge challenge waiting for mathematics educators.

REFERENCES

Australian Education Council (1991). A National Statement on Mathematics for Australian Schools. Canberra: Author.

Brainerd, C.J. (1981). Working memory and the developmental analysis of probability. Psychological Review, 88, 463-502.

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). Multi-modal learning and the quality of intelligent behaviour. In H. Rowe (Ed.). Intelligence: Reconceptualisation and measurement. Hillsdale, N.J.: Lawrence Erlbaum.

Bright, G.W.; Berenson, S.B. & Friel, S. (1993). Teachers' knowledge of statistics pedagogy. Paper presented at the annual meeting of the Research Council for Diagnostic and Presriptive Mathematics, Melbourne, FL.

Berenson, S.B., Friel S. & Bright, G. (1993). Elementary teachers' fixations on graphical features to interpret statistical data. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Bright, G.W. & Friel, S.N. (1993). Elementary teachers' representations of relationships among statistics concepts. Paper presented at the annual meeting of the American Educational Research Association, Atlanta, GA.

Campbell, K.J., Watson, J.M. & Collis, K.F. (1992). Volume measurement and intellectual development. Journal of Structural Learning, 11, 3, 279-298.

Collis, K.F. & Biggs, J.B. (1991). Developmental determinants of qualitative aspects of school learning. In G.T. Evans (Ed.). Learning and teaching cognitive skills. Melbourne:A.C.E.R.

Collis, K.F. & Romberg, T.A. (1990). "The standards": Theme and assessment. In K. Milton & H. McCann (Eds.), Mathematical Turning Points - Strategies for the 1990's. Hobart: A.A.M.T.

Cox, C., & Mouw, J.T. (1992). Disruption of the representativeness heuristic: Can we be perturbed into using correct probabilistic reasoning? Educational Studies in Mathematics, 23, 163-178.

Curriculum Corporation. (1993). Mathematics - the national profile. Carlton, Victoria: Curriculum Corporation.

Dunkels, A. (1990) Examples from the in service classroom (age group 7-12). In A. Hawkins (Ed.). Training teachers to teach statistics. Voorburg: International Statistical Institute.

Falk, R., Falk, R. & Levin, I. (1980). A potential for learning probability in young children. Educational Studies in Mathematics, 11, 181-204.

Fischbein, E. (1990). Training teachers for teaching statistics. In A. Hawkins (Ed.). Training teachers to teach statistics. Voorburg: International Statistical Institute.

Fischbein E. & Gazit, A. (1984). Does the teaching of probability improve probabilistic intuitions? Educational Studies in Mathematics, 15, 1-24.

Fullan, M. & Stiegelbauer, S.(1991) Professional development of educators. In The New Meaning of Educational Change. 2nd Edition. New York: Teachers College Press.

Gal, I., (1992). Reaching out: Some issues and dilemmas in expanding statistics education. Proceedings of the August 1992 International Statistics Institute Roundtable on Teaching Data-analysis, Bishop University, Canada.

Gal,I.,Rothchild, K., & Wagner, D.A. (1990,April). Statistical concepts and statistical reasoning in children: Convergence or divergence. Paper presented at the American Educational Research Association Conference, Boston.

Garfield, J.B., & Ahlgren, A. (1986). Difficulties in learning probability and statistics. In R. Davidson & J. Swift (Eds.) Proc. 2nd International Conference on teaching statistics. Victoria, B.C.: International Statistical Institute.

Garfield, J.B., & Ahlgren, A. (1988). Difficulties in learning basic concepts in probability and statistics: Implications for research. Journal for Research in Mathematics Education, 19, 44-63.

Garfield, J.B., & del Mas, R. (1991). Students conceptions of probability. In D. Vere-Jones (Ed.) Proc. 3rd International Conference on teaching statistics. Vol. 1: School and general issues (pp. 320-328). Voorburg: International Statistical Institute.

Glencross, M.J., (1986). Initial and inservice training of teachers of statistics: Inservice training to assist junior high school teachers to implement a new syllabus in statistics. In R. Davidson & J. Swift (Eds.) Proc. 2nd International Conference on teaching statistics. Victoria, B.C.: International Statistical Institute.

Goodall, G.W., & Jolliffe, F.R. (1986). The training of Brunel University undergraduates who intend to become statistics teachers. In R. Davidson & J. Swift (Eds.) Proc. 2nd International Conference on teaching statistics. Victoria, B.C.: International Statistical Institute.

Green, D. (1986). Children's understanding of randomness: Report of a survey of 1600 children aged 7 - 11 years. In R. Davidson & J. Swift (Eds.) Proc. 2nd International Conference on teaching statistics. Victoria, B.C.: International Statistical Institute.

Green, D. (1991). A longitudinal study of pupils' probability concepts. In D. Vere-Jones (Ed.) Proc. 3rd International Conference on teaching statistics. Vol. 1: School and general issues. Voorburg: International Statistical Institute.

Hope, J.A., Kelly, I.W. (1983). Common Difficulties with probabilistic reasoning. Mathematics Teacher, 76, 565 - 570.

Johnson, N. (1991). Effective professional development: the key to quality teaching and learning. Occasional Paper No. 23. Jolimont, Vic.: Incorporated Association of Registered Teachers of Victoria.

Jolliffe, F. (1991). The loss of intuition - A lesson for the school teacher? In D. Vere-Jones (Ed.), Proc. 3rd International Conference on teaching statistics. Vol. 1: School and general issues (pp. 320-328). Voorburg: International Statistical Institute.

Jowett, G. (1991). Expanding statistical education: A New Zealand retrospect. In D. Vere-Jones (Ed.), Proc. 3rd International Conference on teaching statistics. Vol. 1: School and general issues (pp. 320-328). Voorburg: International Statistical Institute.

Kahneman, D., Slovic, P. & Tversky, A. (Eds.) (1982). Judgement under uncertainty: Heuristics and biases. Cambridge: University Press.

Kepner, H.S. Jr. & Burrill, G.F. (1990). In-service training for statistics in Wisconsin: A cooperative effoert between the professional organisation, local school districts, state department of education and universities. In A. Hawkins (Ed.). Training teachers to teach statistics. Voorburg: International Statistical Institute.

Leon, M. & Zawojewski, J. (1991). Use of the arithmetic mean: An investigation of four properties. Issues and preliminary results. In D. Vere-Jones (Ed.) Proc. 3rd International conference on teaching statistics. Vol. 1: School and general issues (pp. 320-328). Voorburg: International Statistical Institute.

Mokros, J.R., & Russell, S.J. (1992). Children's concepts of average and representativeness (Working Paper 4-92). Cambridge, MA: TERC.

Mevarech, Z.R., (1983). A deep structure model of students statistical misconceptions. Educational Studies in Mathematics, 14, 415-429.

Murphy, C. (1992). Study groups foster schoolwide learning. Educational Leadership, 50, 71-74.

Noelting, G. (1980). The development of proportional reasoning and the ratio concept part 1: Differentiation of stages. Educational Studies in Mathematics, 11,217 - 253.

Noelting, G. (1980). The development of proportional reasoning and the ratio concept part 2: Problem structure at successive stages; Problem solving strategies and the mechanism of adaptive restructuring. Educational Studies in Mathematics, 11, 331-363.

Owen, J., Johnson, N., Clarke, D., Lovitt, C. & Moroney, W. (1988). Guidelines for consultants and curriculum leaders. Canberra: Curriculum Development Centre.

Peard, R.F. (1991). Misconceptions in probability: A comparison of gamblers and non-gamblers within the Australian social context. Centre for Mathematics and Science Education Technical Report No: 3. Brisbane: Queensland University of Technology.

Pegg, J. (1989). Analyzing a mathematics lesson to provide a vehicle for improving teaching practice. Mathematics Education Research Journal, 1, 2, 18 - 33.

Pereira-Mendoza, L. (1986). A comparison of the statistics curriculum for children aged 5-11 in Britain, Canada and the U.S.A. In R. Davidson & J. Swift (Eds.) Proc. 2nd International Conference on teaching statistics. Victoria, B.C.: International Statistical Institute.

Piaget, J. & Inhelder, B. (1951). The origin of the idea of chance in children. Translated 1975. New York: Norton & Co. Inc.

Peterson,P.L., Fennema,E., Carpenter T.P. & Loef, M. (1989). Teachers' pedagogical content beliefs in mathematics. Cognition and Instruction. 6 (1), 1-40: Lawrence Erlbaum Associates, Inc.

Pollatsek, A.,Lima, S., & Well, A.D., (1981). Concept or computation: Students understanding of the mean. Educational Studies in Mathematics, 12, 191-204.

Russell, S.J. (1990). Issues in training teachers ro teach statistics in the elementary school; a world of uncertainty. In A. Hawkins (Ed.). Training teachers to teach statistics. Voorburg: International Statistical Institute.

Russell, S.J. (1991). Counting noses and scary things: Children construct their ideas about data. In D. Vere-Jones (Ed.) Proc. 3rd International Conference on teaching statistics. Vol. 1: School and general issues (pp. 307-313). Voorburg: International Statistical Institute.

Russell, S.J., & Mokros, J.R. (1991). What's typical? Children's and teachers' ideas about average. In D. Vere-Jones (Ed.) Proc. 3rd International Conference on teaching statistics. Vol. 1: School and general issues (pp. 307-313). Voorburg: International Statistical Institute.

Scardamalia, M. (1977). Information processing capacity and the problem of horizontal decalage. Child Development, 48.28 - 37.

Scholz, R. & Waller, M. (1983). Conceptual and theoretical issues in developmental research on the acquisition of the probability concept. In R. Scholz (Ed.) Decision Making Under Uncertainty. Amsterdam: Elsevier Science Publishing Co.

Siegler, R.S. (1981). Developmental sequences within and between concepts. Monographs of the Society for Research in Child Development, 46, 2.

Shaughnessy, J.M. (1981). Misconceptions of probability: From systematic errors to systematic experiments and decisions. In A. Schulte (Ed.) Teaching Statistics and Probability (Yearbook of the National Council of Teachers of Mathematics), pp.90-100. Reston, VA: NCTM.

Shaughnessy, J.M. (1992). Research in probability and statistics: Reflections and directions. In D.A. Grouws (Ed.) Handbook of research on mathematics teaching and learning (pp 465-494). New York: MacMillan.

Singer, J.A. & Resnick, L.B. (1992). Representations of proportional relationships: Are children part-part or part-whole reasoners? Educational Studies in Mathematics, 23, 231-246.

Steinbring, H. (1986). The interaction between teachinbg practice and theoretical conceptions - A cooperative model of in-service training in statistics for mathematics teachers (Grades 5 - 10). In R. Davidson & J. Swift (Eds.) Proc. 2nd International Conference on teaching statistics. Victoria, B.C.: International Statistical Institute.

Truran, J. (1990) From tossing coins to probabilistic understanding - How can research help classroom practice? In K. Milton & H. McCann (Eds.), Mathematical Turning Points - Strategies for the 1990's. Hobart: A.A.M.T.

Watson, J.M. (1992). What research is needed in probability and statistics education in Australia in the 1990's? In Southwell, B., Perry,B & Owens, K. (Eds.), Proceedings of the Fifteenth Annual Conference of the Mathematics Reasearch Group of Australasia, 556-567. Kingswood, NSW:MERGA

Watson, J.M. & Collis, K.F. (1993) Initial considerations concerning the understanding of probabilistic and statistical concepts in Australian students. In B. Atweh, C. Kanes, M. Carss & G. Booker (Eds.), Contexts in mathematics education, 575-580. Brisbane: Mathematics Education Research group of Australasia.

Watson, J.M.,Campbell, K.J. & Collis, K.F. (1993). Multimodal functioning in understanding fractions. Journal of Mathematical Behaviour, 12, 45-62.

Willis, S. (1993) Personal Communication.

Appendix 1

Research Based Teaching Materials

Lovitt, C. & Lowe, I. (1992) Chance and Data Investigations Carlton, Victoria: Curriculum Corporation.

Quantitative Literacy Series

Landwehr, J.M. & Watkins A.E. (1991). Exploring Data. Palo Alto, CA: Dale Seymour

Newman, C.M., Obremski, T.E. & Scheaffer R.L. (1991). Exploring Probability. Palo Alto, CA: Dale Seymour

Landwehr, J.M., Swift, J. & Watkins, A.E. (1991). Exploring Surveys and Information From Samples. Palo Alto,CA: Dale Seymour

Gnanadesikan, M., Scheaffer, R.L. & Swift, J. (1991). The Art and Techniques of Simulation. Palo Alto,CA: Dale Seymour

Used Numbers Series

Friel, S.N., Mokros, J.R. & Russell, S.J. (1992) Used Numbers: Real Data in the Classroom. Palo Alto,CA: Dale Seymour

Titles:

Counting: Ourselves and our families	Grades K - 1.
Sorting: Groups and Graphs	Grades 2 - 3
Measuring: From paces to feet	Grades 3 - 4
Statistics: The shape of data	Grades 4 -6
Statistics: Prediction and sampling	Grades 5 - 6
Statistics: Middles, means and in-betweens	Grades 5 - 6

Appendix 2

Survey documents

The University of Tasmania has recently commenced a study into the teaching and learning of chance and data concepts. This strand of the mathematics curriculum is developing in importance, and little basic research has been carried out. The aim is to make recommendations about approaches to teaching about chance and data handling in the classroom.

This survey is being carried out as part of the Chance and Data Project. Participation in the study is voluntary. I would be grateful if you would complete the permission form and the survey questions

----

I, ______________________________ am willing to participate in the Chance and Data Project.

Signed ________________________ Date _________________________

____

Please answer the following questions as accurately as you can.

1. Name ______________________________

2. School ______________________________

3. Years of teaching experience ______________________________

4. Do you have any work experience other than teaching? If "YES" please give details..______________________________

____

5. Which grades do you usually teach e.g. Early Childhood, Upper Primary, 7 - 10 etc. ______________________________

6. What level of mathematics did you complete e.g. HSC Level 2, 2nd year uni, B.Ed. maths etc ______________________________

7. Do you remember ever being taught about concepts of average? If "YES" please give details..______________________________

____

Thank you for your help with this survey.

SCHOOL:________________________

NO: ________________

TEACHER QUESTIONNAIRE

1. You want to do an experiment with your class as an exercise in using a weighing scale to find the average mass of a single block. You try the experiment yourself first and get the following results (all readings in grams):

3.2 4.0 3.25 3.2 3.2 3.1 3.0 3.2

Use these results to find the average mass of the block.

____

____

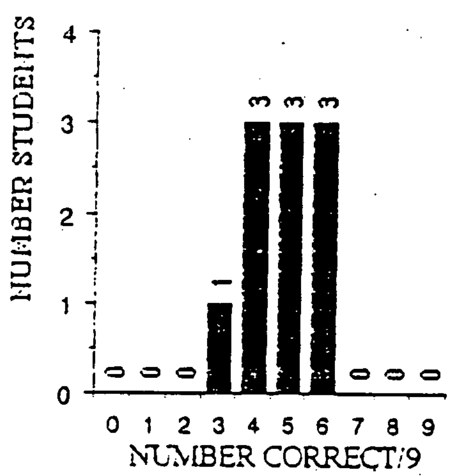
____

____

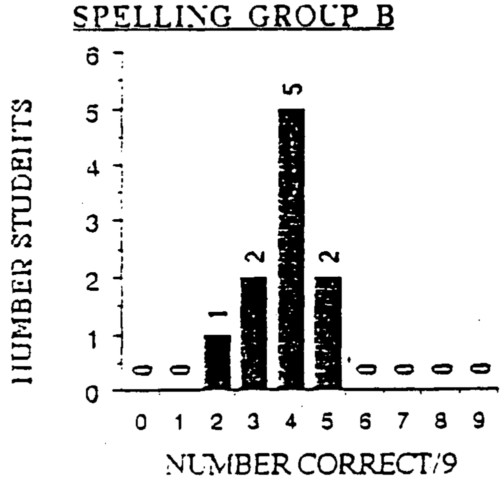
____

2.

SPELLING GROUP A



SPELLING GROUP B



The graphs show the results on the same spelling test for two different groups of equal size. Show how you would tell which group were the better spellers or did they spell equally well? Please explain your answer.

____

____

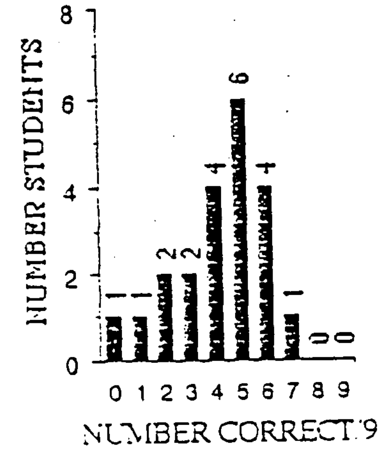
____

____

____

SPELLING GROUP C

3



The graphs show the results on the same spelling test for two larger classes of different sizes. Show how you would tell which class had the better spellers or did they spell equally well? Please explain your answer.

____

____

____

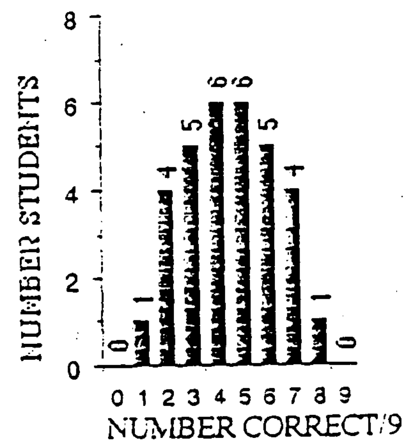
____

____

____

____

SPELLING GROUP D



4. Ten hospitals in Victoria took part in a survey of births.
4 large hospitals had an average of 70 babies born each month.
6 small hospitals had an average of 30 babies born each month.
What was the average number of babies born per month in the ten hospitals?
Please show your reasoning.

____

____

____

Appendix 3

Categories of responses to questions 2 and 3 used in the contingency tables

Row 1: incorrect, which included no response;

Row 2: strategies other than the mean used to get a correct answer;

Row 3: use of the arithmetic mean to obtain a correct answer.

Table 6 Comparison of Response Between High School and Primary Teachers on Question 2.

Column 1: High School Teachers   Column 2: Primary School Teachers

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	7	21	28
Row 2	19	33	52
Row 3	10	46	56
Totals:	36	100	136

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	19.444%	21%	20.588%
Row 2	52.778%	33%	38.235%
Row 3	27.778%	46%	41.176%
Totals:	100%	100%	100%

$X^2 = 4.87$ df = 2 $p = 0.0874$

Table 7 Comparison of Response Between High School and Primary Teachers on Question 3

Column 1 High School Teachers Column 2 Primary Teachers

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	10	30	40
Row 2	9	24	33
Row 3	17	41	58
Totals:	36	95	131

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	27.778%	31.579%	30.534%
Row 2	25%	25.263%	25.191%
Row 3	47.222%	43.158%	44.275%
Totals:	100%	100%	100%

$X^2 = 0.22$ df = 2 p = 0.895

Table 8 Comparison of all teachers responses to questions 2 and 3

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	28	40	68
Row 2	52	33	85
Row 3	56	58	114
Totals:	136	131	267

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	20.588%	30.534%	25.468%
Row 2	38.235%	25.191%	31.835%
Row 3	41.176%	44.275%	42.697%
Totals:	100%	100%	100%

$X^2 = 6.308$ df = 2 p = 0.043

Table 9 Comparison of Primary Teacher Response to Questions 2 and 3

Column 1: Q 2 Column 2: Q 3

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	21	30	51
Row 2	33	24	57
Row 3	46	41	87
Totals:	100	95	195

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	21%	31.579%	26.154%
Row 2	33%	25.263%	29.231%
Row 3	46%	43.158%	44.615%
Totals:	100%	100%	100%

$X^2 = 3.17$ df = 2 p = 0.205

Table 10 Comparison of High School Teachers Responses to Questions 2 and 3

Column 1: Q 2 Column 2: Q 3

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	7	10	17
Row 2	19	9	28
Row 3	10	17	27
Totals:	36	36	72

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	19.444%	27.778%	23.611%
Row 2	52.778%	25%	38.889%
Row 3	27.778%	47.222%	37.5%
Totals:	100%	100%	100%

$X^2 = 5.92$ df = 2 p = 0.052

Categories of responses to question 4 used in contingency tables

Row 1: incorrect, which included no response

Row 2: correct, including all responses which obtained the correct answer by some form of calculation.

Table 11 Comparison of responses from high school and primary school teachers to question 4

Column 1 High school teachers   Column 2 Primary teachers

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	10	45	55
Row 2	26	50	76
Totals:	36	95	131

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	27.778%	47.368%	41.985%
Row 2	72.222%	52.632%	58.015%
Totals:	100%	100%	100%

$X^2 = 4.11$ $df = 1$ $p = 0.043$

Table 12 Comparison between in-service and pre-service teachers on question 4

Column 1 Inservice Teachers   Column 2 Pre-service Teachers

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	11	43	54
Row 2	25	52	77
Totals:	36	95	131

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	30.556%	45.263%	41.221%
Row 2	69.444%	54.737%	58.779%
Totals:	100%	100%	100%

$X^2 = 2.33$ df = 1 $p = 0.126$

Table 13 Comparison of Diploma of Education and Bachelor of Education pre-service teachers' responses to question 4

Column 1 Dip. Ed. Column 2 B. Ed.

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	7	37	44
Row 2	23	28	51
Totals:	30	65	95

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	23.333%	56.923%	46.316%
Row 2	76.667%	43.077%	53.684%
Totals:	100%	100%	100%

$X^2 = 9.314$ $df = 1$ $p = 0.002$

Table 14 Comparison Between Early Childhood Trained Pre-service Teachers from Hobart and Launceston

Column 1 Hobart Column 2 Launceston

Observed Frequency Table

	Column 1	Column 2	Totals:
Row 1	5	19	24
Row 2	13	5	18
Totals:	18	24	42

Percents of Column Totals

	Column 1	Column 2	Totals:
Row 1	27.778%	79.167%	57.143%
Row 2	72.222%	20.833%	42.857%
Totals:	100%	100%	100%

$X^2 = 11.09$ df = 1 $p = 0.0009$