Journal of Maltese Education Research

Vol:2 No.1 2004 53-67
www.educ.um.edu.mt/jmer

Pupils' Understanding Of Probability & Statistics (14-15+) Difficulties And Insights For Instruction

Frank Bezzina

bezzinafrank@hotmail.com

Frank Bezzina holds a B.Ed (Hons) from the University of Malta and an MSc in Applied Statistics with Statistical Education from Sheffield Hallam University. He has recently submitted his Ph.D thesis titled 'Accounting for the Differential Performance of Boys and Girls in Secondary School Mathematics'. Mr. Bezzina teaches part-time with the Faculty of Education and the Faculty of Science.

Abstract:

This study highlights pupils' common errors and misuses of probability and statistics. Four hundred pupils, aged 14-15+, attending a number of Maltese non-selective (mixed-ability) schools were given a 'Probability & Statistics' test. Interviews were also conducted to inform and interpret the results of the written test.

The report highlights pupils' difficulties in the drawing of simple inferences from tables and statistical diagrams, in understanding 'equally likely' situations, in categorising data, in distinguishing between bar charts and histograms, in finding probability for simple and combined events, in summarising data and in choosing a suitable average. Faulty intuitions (e.g. the 'gambler's fallacy', the law of averages) and misconceptions (e.g. the 'sample space' misconception) were also evident.

The implications for teaching in the light of the present findings are discussed.

Introduction

An educated person must have an understanding of the basic statistical tools to function in a world that's becoming increasingly dependent on quantitative information (Sanders, 1995). Additionally, probability is an important topic in its own right as it allows us to understand, calculate and compare risks around us. This suggests that an understanding of probability and statistics can actually change the way we see the world and inform our judgements (Graham, 1994) and brings out the importance that the teaching of probability and statistics has (or should have) in the National Curriculum.

Although statistics has been an integral part of human life for a very long time, the main experience of statistics for most pupils was over the years obtained through the 'user subjects' such as Biology, Business Studies, Geography and Home Economics (Rangecroft, 1997). Statistics (and probability) appeared by name for the first time in

University of Malta, Msida, Malta MSD06
ISSN **1726-9725**

Malta in the early 1990s, as part of the 'Data Handling' section of the mathematics curriculum and in 1997, a statistics and operations research (OR) department was formally established at the University of Malta.

There has been a considerable amount of research into statistical education in foreign literature. A number of misconceptions that have been identified in probability and statistics include:

- *representativeness* - when pupils expect small samples to behave like the whole population (Shaughnessy, 1992),

- *availability* - when pupils estimate probability by the ease with which instances can be brought to mind rather than from evidence from data (Kahneman & Tversky, 1983)

- *gambler's fallacy* – after a long string of losses, the next bet is more likely to win (Fischbein, Nello & Marino, 1991)

- *effect of sample size* – pupils tend to neglect the sample size as if it is not relevant (Tversky & Kahneman, 1982)

- the *compound and simple events* misconception – if, say, two dice are rolled simultaneously, the tendency is to say that the pair 6-6 and the pair 5-6 both have the same chance of happening (Fischbein & Schnarch, 1997)

- the *sample*-space misconception – pupils do not recognise the possibility that all outcomes can occur (Jones , Langrall, Thornton & Mogill, 1999)

Research has also focused on the difficulties pupils have in the understanding of statistical diagrams (e.g. Rangecroft, 1994; Perry *et al.*, 1999) and on the understanding of statistical concepts such as the term 'average'. For instance, (i) serious difficulties are encountered in finding the mean of a combined group when given the means of two unequal sized groups (Pollatsek, Lima & Well, 1981), (ii) in understanding that the medians of two groups do not have an arithmetic relation to the median of combined groups (Mavarech, 1983) and (iii) that the median is sensitive to changes in only one or two observations in a data set (Jolliffe, 1999). It has also been highlighted that pupils have a tendency of learning the computations but lack clear intuition. Goodchild (1994) reports:

> it appears that pupils aged 13-14 years who demonstrated elsewhere in the interview that they were able to calculate the mean of a set of data, and, in addition, who are familiar with the word average, do not understand *average* as being a measure of a stochastic process.
>
> (p. 30)

There have been numerous disturbing aspects about the teaching of probability and statistics. One can mention the fact that teachers have acquired notoriety for not explaining the subject intelligibly and that, in the current teaching, the practical aspects are avoided or poorly presented (Sant, 1997). At the same time, there have been many recommendations in the foreign literature as to how probability and statistics should be taught. For instance, some researchers have suggested that teachers should:

a) present the subject in an interesting, timely (and occasionally humorous) way (Sanders, 1995)

b) provide ongoing experiences with experimental activities and random generators (Truran, 1994);

c) recognise and confront common errors in students' probabilistic thinking (Shaughnessy, 1992);

d) create situations requiring probabilistic reasoning that correspond to students' views of the world and try to arouse the feeling that statistics relates usefully to reality not just symbols, rules and conventions (Garfield & Ahlgren, 1988);

e) keep scrapbooks of the uses and abuses of statistics which they and their pupils encounter in their everyday life (Goodchild, 1994);

f) bridge the gap between teaching and learning by advocating the use of a general instructional model in which research-based knowledge of students' thinking is used to inform classroom instruction (Jones *et al.*, 1999).

A number of researchers have highlighted inconsistencies in the presentation of statistics (e.g. Li & Shen, 1994; Taverner, 1996), while others have provided some examples that teachers may use to help students understand and avoid misconceptions in probability and statistics (e.g. Fischbein & Schnarch, 1997; Graham, 1994).

Aim of Study

The aim of this study is to point out pupils' common misuses of probability and statistics and to help recognize and confront common errors in their thinking. In this way, secondary school teachers could become aware of what is likely to occur when presenting this topic of data handling. However, it is worth noting that this research was not based on classroom observation and so one cannot make pronounced statements on how the children were taught.

Method

Participants
400 Form 4 pupils (aged 14-15+) attending six Maltese non-selective schools, selected at random, participated in a study at the end of February 2003. These non-selective schools are mixed-ability institutions that guarantee primary and secondary schooling upon entry and operate in co-educational or single-sex settings. The modal age of the pupils was 14.

Instrumentation
A 'Probability and Statistics' test was designed for the study. The test items are representative of the 'Data Handling' syllabus for this particular age group. Some of the items were adapted from examples used by other researchers (e.g., Clegg, 1990;

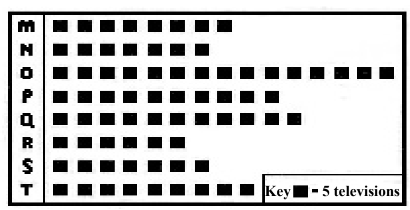
Graham, 1994; Monks, 1994; Truran, 1994) while others were purposely designed for the present study. Owing to the length of the test, a decision was taken to divide the test into two parts, each containing five tasks. Before conducting large scale testing, a pilot study was used to ensure that the timing of the test was appropriate and that the items were free from any technical words. The mathematics teacher administered the two parts of the test on two consecutive days. There was a time limit of 30 minutes for each part of the test and the teachers had specific instructions not to discuss any of the test items before all the interviews had been conducted. The test papers were then passed on to the researcher who examined the individual responses to the test items. Only pupils who participated in both parts of the test were included in the analysis.

Once all testing was conducted, thirty pupils from different schools were randomly selected and interviewed. These interviews were an integral part of the methodology used to inform and interpret the results of the written test items. It was evident that children from different schools made the same type of error and often used the same methods. This article will highlight those items that revealed widespread errors, before going to point out the implications for teaching.

Analysing Students' Responses

Items 1 and 2: skill required: reading, interpreting and drawing simple inferences from tables and statistical diagrams

1. The pictogram shows the total number of televisions in Year 7 pupils' homes.



a) How many televisions do 7P have?
b) Which class has the fewest televisions?

In item 1, pupils had to draw inferences from a pictogram. It was clear that most of the pupils were able to recognize that the length of the line of squares is a consistent measure of pupils it contains. In fact, parts (a) and (b) were answered correctly by 78% and 96% of respondents respectively. What a number of students failed to recognize was that each key (square) represented 5 televisions. This means that they must have ignored the small print that the pictogram contained. This also accounts for the discrepancy in the percentage of correct responses for parts (a) and (b), as in the latter, students still managed to obtain the correct answer without multiplying the number of squares by 5.

Teachers should stress the importance of the other vital, but less visible clues around the edges. In particular, the title, source or footnotes (as in this case) may provide essential information that may be required to qualify or refine conclusions. Graham (1994) warns:

> If people's eyes tend to blank out tables of figures, you can be darn sure that they will blank out the small writing that goes round them (p. 157).

2. The following table shows the results of the successful students (those obtaining a Grade ranging from 1 to 5) by *Gender* and T*ype of School* for the SEC examination in mathematics (May 2002)

Type of School	Successful Boys	Successful Girls	No. of Registered Candidates
Junior Lyceum	354	628	1445
Area Secondary	28	27	626
Church	222	489	854
Independent	102	50	257
Gozo Schools	94	127	350
Private Candidates	185	200	964
Post Secondary	38	71	501
Total	1023	1592	4997

Original Source: SEC Examinations 2002 Statistical Report, pg. 36, Table 3.3.

a) How many of students attending Area Secondary were unsuccessful?
b) What percentage of successful Junior Lyceum students were girls?
c) From the information given in this table, do you agree that the Junior Lyceum candidates were more successful than the Church School candidates in this examination? Why?

In item 2, the pupils were presented with 'real' data adapted from the SEC Examinations 2002 Statistical Report (Matsec Examinations Board, 2002). The table refers to the number of boys and girls within each school type (church school, independent schools, Junior Lyceum, etc) that obtained a Grade ranging from 1 to 5 (necessary for entry into Sixth Form) in the SEC 2002 examination in mathematics. Whereas the vast majority of pupils (86%) were able to answer item 2a (the number of unsuccessful Area Secondary candidates), only 56% answered item 2b correctly. From the denominator used, it was clear that a number of students were not able to distinguish between the percentages within a group and the percentage contributed by the group to the whole population. "What percentage of successful Junior Lyceum students were girls?" was misinterpreted as either "What percentage of successful candidates were Junior Lyceum girls?" or else as "What percentage of all candidates were successful Junior Lyceum girls?" Although these sentences seem to include the same elements (type of school, gender group and examination), careful reading will show that they do not represent identical populations. The general point that this finding illustrates is that every entry in a table can be expressed as a sentence, and indeed every table could be expressed in a series of such sentences, even though this would be extremely laborious in practice. The ability to move to and fro between 'sentences' and 'tables' lies at the heart of table reading and plays a crucial role in analysing and/ or constructing arguments (Davies & Straker, 1997).

In item 2c, 64% argued that the Junior Lyceum candidates were more successful that the Church School candidates. This is clearly WRONG. This is because they compared the absolute number of successful candidates, with no reference to the distribution of candidates in each category. It is true that the successful Junior Lyceum candidates amounted to 982, whereas the successful Church School girls amounted to 711 (i.e. 271 less). However, further analysis (based on the given information) reveals that 83.3% of the Church school candidates sitting for the examination were successful, whereas only 70% of the Junior Lyceum students sitting for this examination were successful. In an actual study, such a finding would warrant further investigation (by chasing up further sources), before reaching any conclusions. One would have to take into account several other factors such as the ability profile of the school's intake, the percentage of students who apply for the examination in each school and the socio-economic context of the school. However this was beyond the scope of this question, as no additional information was made available.

The errors in the students' responses here implicate that teachers should not assume that pupils possess very sophisticated table reading skills, in spite of the fact that table reading appears in a number of secondary school subjects. The point of a table is to communicate information effectively to the person who is reading it, but if this person fails to derive correct information, then it is a waste of space on the page (Graham, 1994). We should provide opportunities by which students can draw on the use of a range of detection skills in order to gain some better understanding of what lies behind the figures. This should help them make balanced and critical judgments about the way statistical inferences are used and presented to the public. Gorard, Rees & Salisbury (1999) argue that such methodological errors are also very common in published studies. They point out that if qualitative analysis may be attempting to explain a situation based on incomplete quantitative reasoning, then it follows that

> if a picture was found to be false, then even the most convincing explanations of that false picture would be in doubt, and the implications for remedial action inappropriate.
>
> (Gorard *et al*., 1999, p. 451)

Items 3 to 5: skill required – to understand the meaning of impossible, certain and equally likely events

3. a) What is the probability that the year 2006 will be after the year 2007?
   b) What is the probability that an apple will fall downwards if you drop it on this planet?
   c) What is the probability that I select a blue bead from a bag containing 4 red and 3 green beads?

In general, item 3 showed clearly that the majority of the pupils (70%) were aware that probabilities are usually measured as numbers between 0 and 1 and that a probability of zero (p = 0) is assigned when the outcome is impossible (e.g. "What is the probability that the year 2006 will be after the year 2007?"), and a probability of one (p = 1) refers to an outcome of certainty (e.g. "What is the probability that an apple will fall downwards if you drop it on this planet?"). It is worth noting that a number of pupils (4%) presented a 100 instead of a 1, and in the interviews a particular student insisted that he meant 100% not just 100.

4. State whether the following statements are true or false:
   a) If you throw a six-faced die, a 'six' is the most difficult outcome to get.
   b) If a fair coin was tossed twelve times and each time it showed 'tails' you would expect the eleventh toss to show 'heads'.

On the contrary, serious difficulties were encountered in the 'equally likely' situations. In fact only 36% of the pupils obtained a correct answer in both items. Item 4a revealed that a large number of pupils falsely believe that a 'six' is the most difficult score to obtain when a die is tossed at random. In the interviews, some pupils argued that "once you want to get a number it becomes hard, especially if you want a six" or "it is difficult, unless you are an expert on how far you throw it up and how much to spin it". This finding is in line with Truran (1994) who highlighted students' belief that a coin or die is subject to physical powers or else fail to appreciate that easier and harder are complementary. He claimed that this is a lack of understanding of physical symmetry. In fact, students who revealed a correct appreciation of symmetry argued that no number is easier because they're all on a similar plane surface, except they've got different dots on them, and the final outcome ultimately depends on how much it rolls and the way it lands. Perhaps this misconception associated with the 'six' might be tied to the fact that (a) many games with dice require a 'six' to be thrown before a player may start, (b) is the highest score or (c) they correctly feel that tossing a 'six' is less likely than not tossing a 'six'. However, this is quite a different belief from thinking that the six outcomes on the die are not equally likely. To improve student's understanding of symmetry, it has been suggested that games are almost certain to be very effective. The use of asymmetrical equipment should also help students appreciate the special features of symmetry. Freudenthal (1972) warns against the use of straightforward experiments, like asking students to tossing a die or a coin for a number of times to reinforce the 'equally likely' concept. Tossing a coin 100 times will only produce a relative frequency of 0.4 to 0.6 with 95% probability. It will be necessary to have 2500 to produce a relative frequency between 0.48 and 0.52 with 95% probability. In other words, "such an experiment might not guarantee results sufficiently to encourage a change in opinion" (Truran, 1994, p.51).

It was clear from item 4b that the ill-defined 'law of averages' was at play - that an unexpected row of tails will necessarily be corrected for a head on the next throw, as if the coin has a memory of past outcomes. Graham (1994) highlights that this misconception is sometimes known as the 'gambler's fallacy' as "it leads punters to believe that, after a long string of losses, their next bet is more likely to win" (p. 213). It also reveals that some misconceptions occur because of lack of belief in and understanding of the property of independence.

5. 100 drawing pins are spilt. 68 land pointing 'up' and 32 land pointing 'down'. When the experiment is repeated, which of the following is most likely to happen?

   a) 36 'up' and 64 'down'
   b) 63 'up' and 37 'down'
   c) 51 'up' and 49 'down'
   d) 84 'up' and 16 'down'
   e) or are all the events equally likely?

In item 5, a widespread misconception was that the underlying probabilities of a drawing pin landing 'up' or 'down' are each half. In fact, only 20% made the correct sort of inference and chose response (b). As for the incorrect responses, 14% opted for response (a), probably expecting (from symmetrical considerations) an immediate regression to the supposed mean; 22% opted for response (c), thus ignoring the experimental evidence in favour of their *a priori* belief; and 42% opted the 'equally likely' response, thus ignoring the asymmetric shape of the drawing pin. Monks (1994) highlights that "bright children are significantly better at this exercise, even prior to tuition" (p. 39). Although it is understood that some misconceptions prove especially difficult to remove, Shaughnessy (1983) recommends 'stimulation by simulation' in such a situation. The exercise could start off manually for a small number of simulations and then the number can be increased by using computer simulation, thus allowing the class to explore with larger data sets. Such opportunities help learners assimilate new concepts by providing 'thinking time' and the opportunity to test hypothesis, even if incorrect. Monks (1994) revealed that the use of (i) probe questioning, (ii) simulation, and (iii) general class discussion was a very effective teaching method to eradicate similar misconceptions in probabilistic thinking and always generated a greater awareness of the need for careful appraisal in questions concerning random processes. The inclusion of such activities may be time consuming, but "once the fundamental concepts are firmly understood, subsequent learning occurs rapidly" (Grant & Searl, 1997, p. 27).

Items 6 and 7: skills required – ability to categorise data (discrete or continuous) and to choose a suitable graph

6. State in each case whether the data is discrete or continuous.

a. The heights of pupils in your class
b. The number of brothers and sisters you have.
c. The colours of the rainbow
d. The running times of 6 athletes in the 100m race

It was surprising to note that the vast majority of the pupils (92%) were not aware of the important distinction between discrete and continuous data. In item 6, a response was considered to be correct only if a student answered all the given statements correctly. This should have minimised the chance of getting the correct answer through guessing. In the interviews, pupils insisted that these terms were never introduced in the mathematics lesson. If this is the case, it indicates a serious inconsistency in the presentation of statistics. It is important that the distinction between discrete and continuous data is given the attention it deserves in our teaching of the subject. A dictionary definition of the word discrete will read something like 'separate, detached from others, individually distinct, discontinuous'. Thus, discrete data can either take the form of word categories (e.g. onion, acid rain, Malta) or numbers where decimals make no sense (e.g. number of pencils). Continuous data are numerical (e.g. babies' birth weight, temperature, distance traveled, time taken) and are not restricted in the number of different values they can take (i.e. decimals make sense).

7. This frequency table gives the lengths, t (in minutes), of 20 telephone calls in an office one day.

Length of Call, *t* (in minutes)	$0 \le t < 5$	$5 \le t < 10$	$10 \le t < 15$	$15 \le t < 20$
Frequency	9	7	3	1

a) Which class would 5.00 belong to?
b) Why is it essential to group continuous data before you display them?
c) Choose a suitable graph that might be used to represent this data?

In item 7, the data was grouped into class intervals of five minutes. The results obtained here are rather worrying. Whereas 35% managed to recognize that 5.00 belongs to the $5 \le t < 10$ class (with 65% indicating a lack of understanding of inequality signs), only 4% could explain why it is essential to group continuous data before displaying them. In theory, there is no limit to the number of possible times a telephone call can take. Essentially, the only restriction here lies in the degree of accuracy with which measurements are taken and this is a crucial property of continuous data. Pupils need to have a clear understanding of the distinctive *nature of the data* being represented. If not, this might have serious implications. For instance, how will they be able to decide which graph is most suitable in a particular situation?

In item 7c, the vast majority of the pupils had difficulty in identifying a graph that describes the 'shape' of data that has been measured on a continuous number scale. In fact, only 5% answered correctly (i.e. 'histogram'). The others pointed out either 'bar chart' or a 'pie chart''. These forms of data representation are suitable for separate 'discrete' categories. An additional condition for using a pie chart is that the complete pie needs to represent something that is complete and meaningful. The complete pie here, however, corresponds to the sum of all the lengths of telephone calls and this is not a particularly useful or interesting collection of things. The interviews revealed that students could not distinguish between a bar chart and a histogram. This reminded me of the problems encountered by students on Teaching Practice as well as by many experienced teachers when teaching this topic. Usually, when presented with a problem, it is relatively easy to confirm one's views by referring to a textbook. Unfortunately, secondary school textbooks often reveal an inconsistency in the method of presentation of bar charts and histograms. An inconsistency is also present in a number of questions set in the SEC examination in mathematics (e.g. September 2000, P1, No. 5(ii); September 1997, P2B, No. 15).

Such conflicting evidence indicates that much clearer explanations need to be given as to when, where and why bar charts and histograms should be used and how they should be labelled. Taverner (1996), for example, offers the following definitions:

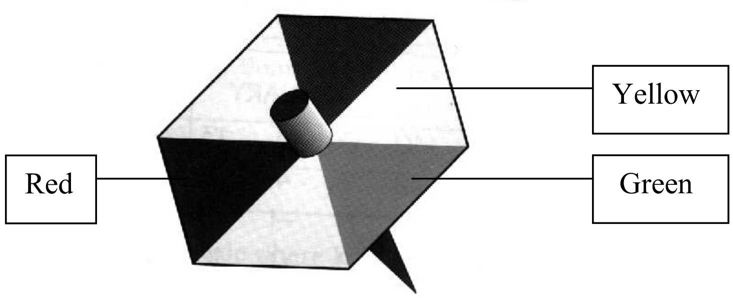
> A bar chart should have horizontal or vertical bars, of equal width, that do not touch. It should be used when representing discrete data, regardless of whether it is grouped or not. If it is grouped the groups should be of equal size. The length of the bar is proportional to the frequency.
>
> A histogram should have vertical bars. It should be used for continuous data, and discrete data that has been put into groups of different sizes. The horizontal axis should be labelled at the join of each pair of bars. The area of the bar is proportional to the frequency (p. 9).

It is important to note that histograms are much more difficult conceptually than bar charts and do cause major problems for many pupils. This difficulty comes from the fact that the use of area instead of height is used to represent frequency. Rangecroft (1994) suggests the "use of square tiles to represent unit frequencies; these can be counted and more importantly moved around to show the effect of varying class size" (p. 12).

Apart from bar charts, histograms and pie charts, teachers could also present pupils with other types of data illustration diagrams (e.g. dot plots, box plots, stem plots, scatter plots and time graphs). Box plots were recently introduced in the SEC syllabus for mathematics as it was felt that they present pupils with a powerful means of comparing two or more distributions. Hopefully, in the near future, other forms of data representation will be introduced in the mathematics curriculum, to aid the understanding of data sets.

Items 8 and 9: skills required - finding probability of a single event; use of sample space to calculate probabilities with two independent events.

8. When this spinner is spun, the side that lands on it is the outcome.



How many possible outcomes are there?
What is the probability of getting (i) green (ii) yellow (iii) not red.

The spinner in item 8 had repeated outcomes. In finding the number of possible outcomes, a number of students did take into consideration the repetitions of each colour. As a result, only 52% answered item 8a correctly. Surprisingly, 77% of the pupils answered correctly the three parts of item 8b. Although this might sound contradicting, it can be explained by the fact that students are usually trained by means of drill exercises similar to this. In fact, item 8b happens to be a typical textbook question and the results obtained here clearly reflect this kind of preparation. In marking item 8b, a correct answer was accepted only if a student gave a correct response to all the three parts, as some pupils managed to obtain part of the question right through the wrong reasoning. For instance, a particular student argued the probability of getting 'not red' was 2 out of 3 as yellow and green are favourable. In this case, the number of outcomes on the spinner was ignored, and it so happened that 4 out of 6 is equivalent to 2 out of 3. This implies a lack of understanding of the term 'possible outcomes'. Here, the total number of possibilities on the spinner is 6 (as 2 are red, 3 are yellow and 1 is green). This student thought that the spinner had 3 outcomes as it had 3 distinct colours (red, green and yellow), thus revealing the sample space misconception.

9. a) Copy and complete this possibility space, showing the possible outcomes when you throw a die and a coin together.

DIE

COIN	1	2	3	4	5	6
Head		(H, 2)				
Tail				(T, 4)		

a) Use this sample space to find the probability of getting
i) an even number on the die
ii) a head and an even number
iii) a tail and a number greater than 4

In item 9, there was a correct response from 86% of the students in drawing the possibility space. A number of students did not follow the same pattern (thus ignoring the order) while others left this item unattempted. Only 46% managed to use the sample space correctly. A number of common errors made were: (a) the counting of each space as 2 outcomes instead of 1; (b) failure to distinguish between 'greater than 4' and 'at least 4'; and (c) failure to present probability as fraction or decimal, but giving answer in wholes (e.g. 4). The same incorrect reasoning in identifying the number of possible outcomes highlighted in the spinner problem was again evident. For instance, in finding the probability of getting an even number, a pupil claimed that the die had 3 even numbers out of 6, so this is equivalent to a ½. No reference was made to the fact that {(H,2), (H,4), (H,6), (T,2), (T,4), (T,6)} were the favourable outcomes and these constitute 6 out of 12 or ½ of the sample space. Alternatively, pupils could have argued that the probability of tossing 'heads' or 'tails' is 1 and the probability of tossing an even number is ½. These are independent events and so the probability of being successful at both events is 1 x ½ = ½. This particular student ignored totally the possibility space outcomes and still obtained the right answer through the wrong reasoning. Thus, a word of warning is required. Teachers need to be aware that even when a student answers correctly, it does not necessarily imply that the concept has been grasped. Rangecroft (1997) argues that pupils might apply techniques blindly, with no understanding of the basic statistical concepts involved, especially when they are presented with fairly artificial examples.

Jones *et al*. (1999) revealed that overcoming a misconception in sample space (in the form of subjective judgements that are extremely persuasive for students), and dealing with probability situations that incorporate the need for quantitative reasoning (e.g., "one out of three") were key patterns in producing growth in probabilistic thinking (a term used to describe children's thinking in response to a probability situation). This study shows that ongoing experiences with experimental activities involving continuous and discrete generators enable students to recognize that no one outcome is certain in probability situations and that students benefit from working in pairs with an adult mentor, since they receive individualized instruction that is not generally possible in a classroom. These results also support the claim that the probabilistic thinking framework can be used to develop an effective instructional program in probability. Further research in this direction is needed, as such programs enhance student learning. In fact, they should provide opportunities to "build on students' prior experience, foster their thinking through problem-focused experiences and monitor their understanding" (Jones *et al.*, 1999, p. 517).

Item 10: ability to compute the mean, mode, median and range for raw data; choosing a suitable average.

> A crucial human skill is to be selective about the data we choose to analyse and, where possible, to summarise the information as briefly and usefully as possible (Graham, 1994, p.64).

10. The amount of wear suffered by tyres made from 2 different types of rubber was measured. For each type of rubber, 7 tyres were tested (by subjecting them to the equivalent of 1000 miles driving). The results (mm x 10) were:

    Rubber A: 16, 16, 16, 18, 14, 16, 16
    Rubber B: 15, 13, 13, 22, 23, 13, 13

    a) Compute the mean, mode, median and range for each type of rubber
    b) Which, in your opinion, is the most reliable type of tyre? Why?

In item 10a, students were asked to compute the three measures of central tendency (mean, mode and median) and a measure of spread (the range) for the amount of wear suffered by tyres made from two different types of rubber. Whereas 56% of the students gave a correct answer, others either interchanged the mean and the median, forgot to place the data in ascending order before finding the median or else presented incomplete computations (e.g. range = "95 – 27" or even "27-95").

Item 10b should have presented a good illustration of how one might go about selecting a suitable average for the given wear and tear problem, since the choice of average really depends on the context and on the purpose for calculating it in the first place. Although this item was attempted by 64% of the pupils, only 2% considered the median or the mode in deciding which is the most reliable type of rubber. Two apparently high-achieving pupils also considered the range before reaching a conclusion. The rest focused solely on the mean, claiming that there was no difference between the two types of rubber. Thus, knowledge of the mean seems to begin and end with an impoverished computational formula. Pollatsek *et al*. (1981) warn us that the learning a computational formula is a poor substitute for gaining an understanding of a concept. Perhaps the lesson to be learnt from this example is that the mean is a somewhat overworked average, and one needs to question whether it meets the main criterion of any well-chosen average; namely, 'Does it provide a useful and representative summary of the data set?' If it doesn't, then don't use it. In this case, it would have been more appropriate to use the median or the mode. Additionally, the range should give an idea of how widely spread the figures are.
Students should be warned that other peoples' motives when describing sets of numbers may not be so pure and they may even be a deliberate attempt to mislead. Clegg (1990) argues:

> We have all heard the saying 'lies, damned lies and statistics'; quoting the mean when the mode would be a more appropriate 'average' is a perfect illustration of the type of thing that gives statistics a bad name!
>
> (p. 19)

Could it be that we are presenting our pupils with fairly artificial examples - invented on the spur of the moment or taken directly from the school textbook? Teachers must not assume that "just because the word average is in common usage that it is properly understood" (Goodchild, 1994, p. 30). Rangecroft (1996) suggests that teachers should use primary data (collected on purpose for the exercise) and secondary data (taken from sources such as Sports Encyclopaedias, Social Trends, Building Societies, Europe in Figures, etc.) as a teaching resource, so that pupils "learn not only about statistics per se but also its value as a tool" (p. 4). Even if supported by a computer data base, a calculator or a statistical package, it is important that pupils gain a clear sense of what this information might be telling, what an average or measure of spread is, and why we might bother to calculate it.

Conclusion

The Schools Council Project on Statistical Education (POSE) suggested that the basic aims in teaching statistics to students aged 11-16 are:

> 1. Children should become aware of and appreciate the role of statistics. That is, they should know about the many and varied fields in which statistical ideas are used, including the place of statistical thinking in other academic subjects.
>
> 2. Children should become aware of, and appreciate, the scope of statistics. That is, they should know the sort of questions that an intelligent use of statistics can answer, and understand the power and limitations of statistical thought
>
> (Schools Council Project on Statistical Education, 1980, p. 27)

It is doubtful whether the teaching of statistics at present goes far enough at all towards achieving these aims. It is hoped that by reading this article, teachers think of the discrepancies that exist between what is intended and what is in fact received. For most of us, there is a tendency to keep on teaching in the way we have always taught. The teaching of probability and statistics is often seen as an initiation into rules and procedures, which although powerful and very attractive to teachers, are often seen as meaningless by pupils. "Life does not present mathematics under the heading 'Remember when you did this one'…" (Burrill, 1997, p. 605). Learning means getting involved in the struggle with the concepts that students are to learn. By getting students involved with problems (e.g., real-life situations involving probability), they are made to struggle with ideas and take part in dialogue. In this way, they become involved in trying to formulate a reasonable solution that is different from just applying a procedure.

Continuing methods of feedback should be devised and where the discrepancies indicate a failure, remedies need to be devised. Teachers should be encouraged to co-operate with other members of staff to meet the problems and difficulties they encounter in the teaching of statistics, whereas schools should provide teachers with in-service courses, where necessary. Mathematics co-ordinators and heads of departments can draw on research findings. Whilst there is a small body of knowledge at present on how children learn statistics, there is a great deal of research into learning generally which has implications for the learning of probability and statistics. Additionally, a number of studies, as highlighted earlier, have discussed misconceptions in probability and statistics and give some examples that may be used to help students understand and avoid these misconceptions.

Ultimately, what really matters is that our pupils understand the principles of what they are doing and are able to spot other people's statistical jiggery-pokery (Graham, 1994). However, to achieve this, we need to provide our pupils with a bedrock understanding of what statistics is all about.

References:

Burrill, G. (1997) President's Report: Choices and Challenges. *Journal for Research in Mathematics Education*, 28 (5), pp. 602-611.

Clegg, F. (1990) *Simple Statistics: A course book for the social sciences*. UK: Cambridge University Press.

Davies, G. & Straker, C. (1997) *Educational Statistics*. Sheffield: SHU.

Fischbein, E., Nello, M.S. & Marino, M.S. (1991) Factors affecting probabilistic judgements in children and adolescents. *Educational Studies in Mathematics*, 22, pp. 555-583.

Fischbein, E. & Schnarch, D. (1997) The evolution with age of probabilistic, intuitively based misconceptions. *Journal for Research in Mathematics Education*, 28, pp. 96-105.

Freudenthal, H. (1972) The Empirical Law of Large Numbers or “The Stability of Frequencies”. *Educational Studies in Mathematics*, 4, pp. 484-490.

Garfield, J.B. & Ahlgren, A. (1988) Difficulties in Learning Basic Concepts in Probability and Statistics: Implications for Research. *Journal of Research in Mathematics Education*, 19 (1), pp. 44-63.

Goodchild, S. (1994) School Pupil’s Understanding of Average. In, D. Green (Ed.) *Teaching Statistics at its Best* (pp. 28-30). University of Sheffield: Teaching Statistics Trust.

Gorard, S., Rees, G. & Salisbury, J. (1999) Reappraising the apparent underachievement of boys at school. *Gender and Education*, 11 (4), pp. 451-454.

Graham, A. (1994) *Statistics: An Introduction*. London: Hodder & Stoughton.

Grant, F. & Searl, J. (1997) Practical Activities in the mathematics classroom. *Mathematics in School*, 26 (1), pp. 27-28.

Jolliffe, I.T. (1999) An Example of Instability in the Sample Median. *Teaching Statistics*, 21 (1), p. 29.

Jones, G.A., Langrall, C.W., Thornton, C.A. & Mogill, A.T. (1999) Students’ Probabilistic Thinking in Instruction. *Journal of Research in Mathematics Education*, 30 (5), pp. 487-519.

Kahneman, D. & Tversky, A. (1983) Extensional versus Intuitive Reasoning: The Conjunction Fallacy in Probability Judgement. *Psychological Review*, 90 (4), pp. 293-315.

Li, K.Y. & Shen, S.N. (1994) Students’ Weaknesses in Statistical Projects. In D. Green (Ed.) *Teaching Statistics at its Best* (pp. 42-48). University of Sheffield: The Teaching Statistics Trust.

MATSEC Examinations Board (2002) *SEC Examinations Statistical Report*. University of Malta: Matsec Support Unit.

Mavarech, Z.R. (1983) A Deep Structure Model of Students’ Statistical Misconceptions. *Educational Studies in Mathematics*, 14, pp. 415-429.

Monks, A.R. (1994) Equally likely. In, D. Green (Ed.) *Teaching Statistics at its Best* (pp. 38-39). University of Sheffield: Teaching Statistics Trust.

Perry, B., Jones, G.A., Thornton, C.A., Langrall, C.W., Putt, I.J. & Kraft, C. (1999) Exploring Visual Displays involving Beanie Baby Data. *Teaching Statistics*, 21 (1), pp. 11-13.

Pollatsek, A., Lima, S. and Well, A.D. (1981) Concept or Computation: Students’ understanding of the mean. *Educational Studies in Mathematics*, 12, pp. 191-204.

Rangecroft, M. (1994) Graphwork- Developing a Progression. In, D. Green (Ed.) *Teaching Statistics at its best* (pp. 7-12). Sheffield: The Teaching Statistics Trust.

Rangecroft, M. (1996) ‘Data Handling – in real life’. *Mathematics in Schools*, 25 (1), pp. 3-4.

Rangecroft, M. (1997) *Teaching Statistics*. Sheffield: SHU.

Sanders, D. (1995) *Statistics: A first course*. New York: McGraw Hill Inc.

Sant, L. (1997) The Department of Statistics and Operations Research. *Times of Malta*, 18th December, p. 27.

Schools Council Project on Statistical Education (1980) *Teaching Statistics 11-16*. Slough: Foulsham Educational.

Shaughnessy, J.M. (1992) Research in probability and Statistics: Reflections and directions. In, D.A. Grouws (Ed.) *Handbook on Research on Mathematics Teaching and Learning* (pp. 465-494). New York: Macmillan.

Shaughnessy, J.M. (1983) Misconceptions of probability, systematic or otherwise, teaching probability and statistics so as to overcome some misconceptions. *Proceedings of the First international Conference on Teaching Statistics Volume 2*. University of Sheffield.

Taverner, S. (1996) Inconsistencies in the presentation of Statistics or Mind the Gap. *Mathematics in School*, 25 (2), pp. 28-29.

Truran, J. (1994) Children's Understanding of Symmetry. In, D. Green (Ed.) *Teaching Statistics at its best* (pp. 49-51). Sheffield: The Teaching Statistics Trust.

Tversky, A. & Kahneman, D. (1982) Judgement of and by Representation. In, D. Kahneman, P. Slovic & A. Tversky (Eds.) *Judgement under Uncertainty: Heuristics and Biases* (pp. 84-100). Cambridge: Cambridge University Press.