Statistical Understanding: An Investigation in Two Countries

The Language of Statistical Understanding: An Investigation in Two Countries

Jonathan B. Moritz
University of Tasmania

Jane M. Watson
University of Tasmania

Lionel Pereira-Mendoza
Nanyang Technological University, Singapore


Abstract

Statistical understanding of students from Grades 3 to 9 in Australia and Singapore was assessed using a 20-item survey which involved items based on past research into probability and statistics. Responses were analysed using the language analysis software NUD•IST, and were classified according to the SOLO Model with Multimodal Functioning developed by Biggs and Collis. Results from the two countries are discussed in the light of curriculum and teaching practices.

© Jonathan B Moritz, Jane M Watson, and Lionel Pereira-Mendoza
Do not quote without permission

Statistical Understanding: An Investigation in Two Countries
A sound grasp of concepts in the areas of chance, data handling and statistical inference is critical for the levels of numeracy appropriate for informed participation in society today. (Australian Educational Council [AEC], 1991, p. 163)

Recent curriculum documents have acknowledged the increasing role that statistics should play in the mathematics curriculum, as a reflection of the increasing demands of data processing skills and statistical literacy (Watson, 1995). Statistics, including studies related to chance and data, are essential for the critical evaluation of claims made on the basis of statistical data in all aspects of society. The revolutions in education associated with constructivism and the push for reform to authentic assessment practices have emphasised the link between mathematical concepts and usage of mathematical language as applied to everyday situations. Constructivist approaches stress how meaning for terms is constructed individually, and often idiosyncratically. If understanding is valued rather than mechanical, algorithmic calculations, then fundamental statistical concepts must be constructed by students from their everyday understandings, and then applied to a variety of situations. The assessment of such understanding must also acknowledge the basic use of language and the contexts which may contribute to various usage of language.

Of particular interest in this study are the basic terms associated with understanding data handling and chance events: "average", "random", "sample", "impossible", "possible" and "certain". The following extracts from the National Council of Teachers of Mathematics (NCTM) Standards acknowledge the importance of understanding the concepts of "average", "random" and "sample" in relation to data collection and summarising data.

... students can formulate questions to determine the characteristics of an "average" student-age, height ... Sampling procedures are a critical issue in data collection. Which students should be surveyed to determine Mr. or Ms. Average? Must every student be questioned? If not, how can randomness in the sampling be assured and how many samples are needed to accumulate enough data to describe the average student? (NCTM, 1989, pp. 105-106)

Students must acquire intuitive notions of randomness, representativeness, and bias in sampling to enhance their ability to
evaluate statistical claims. These understandings would give students the appropriate tools for rejecting such television advertising claims as... (NCTM, 1989, p.169)

Note the recognition of the importance of basic language. In building an understanding of these concepts, the learning involves not just a rote-learned meaning for a term, but also an appreciation for when it is appropriately used or the domain over which the statistical tool may be applied.

Curriculum documents also stress the importance of understanding chance language at an early age. For example in Mathematics - A Curriculum Profile for Australian Schools, it is expected that primary school students

- Respond appropriately to and use 'possible' and 'impossible' for describing familiar events and actions.
- Identify possible outcomes for daily events ('After Daddy picks me up from school, we go to the shops, go straight home, or go and visit a friend'). (AEC, 1994, p. 46)

This reflects the necessity to use, with clarity, everyday language associated with chance events (AEC, 1991, p. 166), and

make statements about how likely are everyday experiences which involve some elements of chance and understand the terms 'chance' and 'probability' in common usage (AEC, 1991, p. 170).

The domains or contexts for statistical usage are important for three reasons. The first is the value of statistical education for authentic situations beyond the classroom, as noted above. By devoting particular attention to the contexts in which statistical techniques are applied, the relevance of the techniques is emphasised and students are encouraged to apply them with discernment for situations which are appropriate for their use.

Computing technology allows students to ... calculate statistical measures with remarkable precision using single computer keystrokes. What is missing-and what their study of statistics should provide-is an understanding of which measures are appropriate for a given problem and what such measures as mean, variance, and correlation can tell them about a problem. Furthermore, it is essential that students learn to interpret results intelligently. (NCTM, 1989, pp. 167-168)

The second reason for providing contexts for learning statistical concepts is to help to build upon students' intuitive notions. Understanding everyday chance language, such as "possible", "impossible" and "maybe" is an early stage in the development of chance measurement (Watson, Collis & Moritz, in press). Fischbein, Nello, and Marino (1991) report that many students have difficulty with these basic terms, in particular finding "certain events" more difficult to comprehend than "possible events" (p. 523). Students' intuitive notions for average, random and sample have also been documented
(Watson, Collis & Moritz, 1995b, 1996) and indicate that care must be taken in making assumptions about students' levels of understanding.

A third reason for addressing students' intuitive notions is to provide contexts which help to correct misconceptions by direct confrontation. A major finding of educational research into students' understanding of probability is the resilience of intuitive misconceptions despite instruction (Fischbein & Gazit, 1984; Garfield & Ahlgren, 1988). In relation to "average", "sample", and "random", Gal, Mahoney and Moore (1992) reported that while many children have useful intuitions for these terms, others "associate with such terms meanings which are distant from, or in conflict with, formal notions" (p. 160). A suggested strategy for teaching is first to get students to use their own concepts to make predictions, and then to provide experiences which create cognitive conflict with misconceptions and a resulting impetus for change in understanding.

While there is general acknowledgment of the desirability of appreciating the context within which statistical terminology is used, building valid intuitions, and avoiding misconceptions, relatively less research has looked into students' initial understandings of terminology than has looked at sophisticated understandings such as weighted averages or conditional probability. Not only is there now interest in early intuitions, but also how these intuitions develop over time. The purpose of this investigation is to explore the contexts within which students apply chance language, and to look at the level of sophistication of basic understanding of the stochastic terms "average", "random", and "sample".

The importance of statistical literacy as a basis for deeper statistical understanding is recognised widely in the curriculum documents from several countries. The data from students in Australia and Singapore provide the opportunity to explore similarities and differences in student understanding. If students are displaying similar types of understanding, it may be possible to share curriculum initiatives between the cultures. If differences occur, then this is of interest for our appreciation of how students construct statistical understanding, and it may influence the contexts and examples used to improve teaching in the classroom.

Theoretical Model

The theoretical stance chosen for this research is that of the Structure of Observed Learning Outcomes (SOLO) model with multimodal functioning (Biggs & Collis, 1989, 1991; Collis & Biggs, 1991), as illustrated in Figure 1. Modes of functioning are similar to Piaget's developmental stages, and include sensori-motor, ikonic, concrete symbolic, formal and post formal (Biggs & Collis, 1982). For the years of formal schooling, interest focuses on how students build concepts in the concrete symbolic mode, using symbols and procedures to address concrete situations, such as quantitative reasoning in chance and data.
Often intuitions and imaging based on prior experiences may be observed in responses expressed in the ikonic mode, and in responses which involve ikonic support for a concept in the concrete symbolic mode. Ikonic support for the concrete symbolic mode is one form of multimodal functioning, where there is an interaction of modes (Biggs & Collis, 1989, 1991; Collis & Biggs, 1991). Ikonic functioning often influences decisions related to chance and gambling even into adult life.

Figure 1. The SOLO model (adapted from Biggs & Collis, 1989).

Within each mode, developmental levels may be observed according to the structure of how the information used is combined to produce a response. Prestructural responses (P) do not use any elements relevant to the mode required by the task. Unistructural responses (U) use only one relevant aspect of the mode. Multistructural responses (M) include several disjoint relevant aspects, usually in sequence. Relational responses (R) demonstrate an integrated understanding of the relationships between the different relevant aspects.

The SOLO model has been helpful in identifying cognitive levels and U-M-R cycles for topics in chance and data such as luck (Watson, Collis & Moritz, 1995a), chance measurement (Watson, Collis & Moritz, in press), average (Watson, Collis & Moritz, 1996), sampling (Watson, Collis & Moritz, 1995b), and looking for associations among variables (Watson, Collis, Callingham & Moritz, 1995). It forms the theoretical framework for identifying stages in statistical understanding.

Research Questions

This study aims to investigate the understandings of fundamental statistical concepts of students in Australia and Singapore. Two particular research questions are to be addressed:

•What contexts do students from Australia and Singapore apply to the interpretation of chance language and statistical terms?
•What SOLO levels do students from Australia and Singapore achieve in response to the statistical terms "average", "random" and "sample"?

Attention will be given to the role of ikonic support for the concrete symbolic concepts.

Method

Items

Four short answer items, shown in Figure 1, were the first four items in a 20-item questionnaire on chance and data (Watson, 1994; Watson, Collis & Moritz, 1994). Item 1 explores exposure to the basic terms "certain", "impossible", and "might possibly", and also the contexts in which students consider these terms to be applicable. Items 2, 3 and 4 are intended to permit students to respond either with the meaning they attach to the terms "average", "random" and "sample", or with an example of a context in which the term might be used.

1(a). One thing that will certainly happen today is
1(b). One thing that is impossible today is
1(c). One thing that might possibly happen today is
2. If someone said you were "average", what would it mean?
3. What things happen in a "random" way?
4. If you were given a "sample", what would you have?

Figure 1. Items used from the 20-item survey

Samples
A total of 1014 Australian students were surveyed from Grades 3, 6
and-9. The students attended 13 schools distributed throughout
Tasmania, Australia. A subsequent sample of 114 students from
Singapore included students from Grade 2, 4 and 5. Table 1 shows a
summary of the number of students from each grade level who responded
to the survey items.

Table 1
Number of students in the study by Grade

<table>
<thead>
<tr>
<th></th>
<th>Tasmania - Grades</th>
<th>Singapore - Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>No. of students</td>
<td>369</td>
<td>245</td>
</tr>
</tbody>
</table>

Procedure
The 20-item chance and data questionnaire was administered to class
groups of Grade 4 level or above. A shortened survey including just
the first 10 items was administered students at Grade 3 level or below.
Responses to all items were analysed using the language analysis
software NUD•IST (Non-numerical Unstructured Data • Indexing Searching
and Theorizing) (Qualitative Solutions and Research, 1992). Command
files automatically assigned responses to categories, which were
manually checked for accuracy (Moritz, 1995). In this process,
responses were indexed at nodes in a tree data structure by searching
for words and character patterns. A search of this index system
permitted logical operations on nodes, resulting in each response set
being partitioned into disjoint categories. Categories of response for
Items-2, 3 and 4 were subsequently allocated SOLO modes and levels.
Responses to Item-1 were categorised by the context of the example
given.

Responses
Item 1a. One thing that will certainly happen today is
Item 1b. One thing that is impossible today is
Item 1c. One thing that might possibly happen today is
The first item focuses on language and personal experience of chance
events. The exact degrees of certainty that students attach to the
words are difficult to assess from these items, because while the items
invite students to give examples, the researchers cannot assess how
certain these examples are to occur. This means that SOLO levels of
sophistication are not appropriate for analysing responses to this
item. The contexts of responses are observed to fall into three broad
categories: school-based (S), personal (P), and general/wider-world
Responses categorised in the school-based context involve reference to things in the school schedule such as classes for specific subjects, having lunch, and going home. They also often use the term "we" in referring to the class.

**S:** We will go home. [Q1(a), Singapore, Grade 2]

**S:** Getting out of school work. [Q1(b), Tasmania, Grade-6]

**S:** We might have music lessons. [Q1(c), Singapore, Grade 4]

Responses in the personal context include "I", "My", "Mum" or other words related to the student's personal experience outside of school.

**P:** My mother will buy me a toy car. [Q1(a), Singapore, Grade-2]

**P:** I will move house. [Q1(b), Singapore, Grade 4]

**P:** I might get hurt. [Q1(c), Tasmania, Grade-6]

Responses in the general/world context include weather, world events, and highly imaginative comments.

**W:** The sun will rise in the daytime. [Q1(a), Singapore, Grade-4]

**W:** Human flying. [Q1(b), Singapore, Grade 4]

**W:** That it will rain today. [Q1(c), Singapore, Grade 5]

Also within the general/world context are some responses which offer a meaning rather than any specification of a contextualised example.

**W:** (a) Sure; (b) False; (c) Doubtful. [Singapore, Grade 4]

**W:** (a) 100%; (b) 0%; (c) 50%. [Singapore, Grade 5]

Many responses fall into more than one of the above contexts, and are categorised as multiple context (M).

**M:** I am going to school. [Q1(a), Singapore, Grade 5]

**M:** The school to fall down on just me. [Q1(b), Tasmania, Grade-3]

**M:** For me to fly to the moon. [Q1(b), Tasmania, Grade-6]

**M:** I will get a scolding from my teacher. [Q1(c), Singapore, Grade 5]

**M:** The school fire alarm could go off. [Q1(c), Tasmania, Grade-9]

The multiple context category highlights some of the strengths and limitations of the classification automated by key-words. Watson, Collis and Moritz (1993) preferred to categorise responses by the dominant context, and so no acknowledgment was made of responses in multiple contexts. Responses in multiple contexts are well-defined by inclusion of key-words in responses, and yet, as the examples above show, can include incidental use of other contexts, such as personal pronouns in referring to other events. While this weakness must be acknowledged in interpreting results of the multiple category, there are two major benefits of key-word search automation. The first is that it ensures consistent categorisation of over 1000 responses to each of the three parts to the item. The second benefit is that it succeeds in identifying use of a context, as the category is defined according to specified key-words; an example is the categorisation of comments on subjects and school events within the school context. A dominant feature of the results is the majority of responses which are offered in single contexts, which highlights part of the value of the categorisation.
Table 2 presents the percentage of responses from each grade in the three categories. The world context dominates results, partly because it includes more general responses, but partly because it can easily be applied to all three items (e.g., sun will rise, human flying, it might rain). The school context dominates Item 1(a) results for Tasmanian students, who respond with certainty about time-table issues. The responses of Singapore students are similar to those of older Tasmania students who seem more concerned with the wider-world context.

Table 2
Percentage of Responses to Item Q1 by Context and by Grade

<table>
<thead>
<tr>
<th></th>
<th>Tasmania Grades</th>
<th>Singapore Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Item 1(a)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>&quot;certain&quot;</td>
<td>369</td>
<td>2</td>
</tr>
<tr>
<td>No context</td>
<td>441</td>
<td>15</td>
</tr>
<tr>
<td>World</td>
<td>172032</td>
<td>43</td>
</tr>
<tr>
<td>School</td>
<td>484422</td>
<td>13</td>
</tr>
<tr>
<td>Personal</td>
<td>7710</td>
<td>13</td>
</tr>
<tr>
<td>Multiple</td>
<td>232435</td>
<td>18</td>
</tr>
<tr>
<td>(N)</td>
<td>(322)</td>
<td>(310)</td>
</tr>
</tbody>
</table>

| Item 1(b)        |                 |                  |
| "impossible"     | 369             | 2                |
| No context       | 653             | 15               |
| World            | 395053          | 5372             |
| School           | 25198           | 10               |
| Personal         | 111013          | 10               |
| Multiple         | 191523          | 13               |
| (N)              | (322)           | (310)            |

| Item 1(c)        |                 |                  |
| "might possibly"| 369             | 2                |
| No context       | 642             | 18               |
| World            | 354251          | 2569             |
| School           | 25159           | 235              |
| Personal         | 111212          | 1313             |
| Multiple         | 222727          | 2313             |
| (N)              | (322)           | (310)            |

Item 2. If someone said you were "average", what would it mean? This item asks particularly for the meaning of "average", but places it in a situated context which invites specific examples. At the prestructural level are those who do not engage the item, but nevertheless feel it necessary to respond in some fashion. P: I have heard of it but I don't know what it means. [Tasmania, Grade-3] Students' initial appreciations for average derives from experience,
hearing others use the word in conversation. Use of these experiences as the sole basis for responding represents ikonic functioning. The second response of “good” indicates that average is in some sense an evaluative word.

IK: You were an average size. [Tasmania, Grade-3]
IK: It means I am good at my work. [Singapore, Grade-2]

Movement into the concrete symbolic mode occurs when students can give a single general idea associated with average when asked about being average themselves. The following are classified unistructural responses.

U: It would mean that you are OK. [Tasmania, Grade-3]
U: It means I am just right. [Singapore, Grade-4]
U: It would mean that I'm a normal person. [Singapore, Grade-5]
U: You are in between. [Tasmania, Grade-6]
U: It means that I am in the middle. [Singapore, Grade-4]
U: It means I am just a child who is just the same as others. [Singapore, Grade-4]
U: You are the same as everybody else. [Tasmania, Grade-9]

At the multistructural level are responses which appear to go beyond the single idea of centre to describe how the average relates to a data set or the method for obtaining the average from the data set. The following are examples.

M: It would mean I was not very good nor very bad. I was in the middle. [Singapore, Grade 4]
M: I'm not terrific or terrible, but I'm O.K. [Tasmania, Grade-6]
M: You weren't good or bad, just in between. [Tasmania, Grade-9]
M: It means that I'm the total divided by the number. [Singapore, Grade-4]
M: It would mean I was same as the most children he or she saw. [Singapore, Grade-4]
M: It would mean that I am the same as the majority of people. [Tasmania, Grade-9]

Notice that the final two multistructural examples are more sophisticated than the final two unistructural examples, as they acknowledge being similar to (the same as) the majority rather than everyone. It would appear that variability, or lack of total uniformity, in the population is being recognised.

At the relational level responses would reflect both the central measure aspect of an average and incorporate a firm representative aspect used for comparison. This level of response observed in interviews (Watson, Collis & Moritz, 1996), was not observed in responses to the survey item.

Table 3 shows the percentage of responses to Item Q2 by SOLO response level and by grade. Many Tasmanian Grade 3 students and Singapore Grade 2 students appear not to have any appreciation for the term “average”, although the Singapore students were more fluent in offering an attempt. The response distributions are similar for Tasmanian Grade 6 students as for Singapore Grade 4 and 5 students, with about 40
percent at each of unistructural and multistructural levels. With respect to the actual terms used, however, there were some differences. Tasmanian students commonly responded "okay" or "between", which were notably absent from Singapore responses, which tended to prefer using "just right" and "middle". With respect to the contexts of responses, Singapore responses tended to mention evaluation of school grades, while Tasmanian responses were more concerned with heights and weights.

Table 3
Percentage of Responses to Item Q2 by SOLO Response Level and by Grade

<table>
<thead>
<tr>
<th>Tasmania Grades</th>
<th>Singapore Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>Level</td>
<td>Responses</td>
</tr>
<tr>
<td>NRNo response</td>
<td></td>
</tr>
<tr>
<td>PIrelevance</td>
<td></td>
</tr>
<tr>
<td>IKExample only</td>
<td></td>
</tr>
<tr>
<td>U Idea of centre</td>
<td></td>
</tr>
<tr>
<td>M Ideas get ctre d st</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Item 3. What things happen in a "random" way?
This item asks students to give examples of random processes, although some students prefer to respond with a definition which talks about the meaning associated with random, or a whole class or set of examples. Students approach random as either applying to the random process which leads to a chance, unpredictable outcome, or applying chance to the outcome distribution, which may be an ordered, unordered, or equally likely distribution. Responses offer examples from game situations, natural phenomena, or human designed specific situations. Some students associate "random" with irrelevant concepts, and in particular with the word of "ransom". These responses are prestructural as they do not provide an example of the concrete situation for the concept. The last example given below might imply a thorough understanding of random, although this understanding is not clearly evident in the structure of the observed response.

P:A person that kidnap a child and ask for money is called random. [Singapore, Grade 5]
P:People hold you for money. [Tasmania, Grade-9]
P:Fighting. [Tasmania, Grade-3]
P:Girlfriends. [Tasmania, Grade-9]

Some students appear to have heard the word in the expression "random order", and so respond that the meaning is "ordered", which may be a precursor to views of random as "unordered" or "mixed". These responses express an idea about random which is strongly linked to experience without clearly establishing links between the term and exactly what it refers to, and so are considered to have ikonic support for a unistructural concept.

U+IK:By smaller to bigger number. [Singapore, Grade 4]
U+IK:Things happen one after another, sort of in order. [Tasmania, Grade-9]
Some students identify random with a specific, humanly constructed event or process, such as a random survey, a random breath test for drink-driving, or a CD-player with a feature for randomly selecting the order of songs.

U: Surveys are done randomly. [Tasmania, Grade-9]
U: In sports you are randomly drug tested. [Tasmania, Grade-6]
U: Random breath testing by police. [Tasmania, Grade-9]
U: On the C.D. when you press RANDOM PLAY it plays the tracks at random. [Tasmania, Grade-6]

Some responses include examples from natural, unpredictable phenomena,

such as the weather, death, or the sex of a baby.

U: It will rain. [Singapore, Grade 2]
U: Accidents happen in a random way. [Singapore, Grade 4]
U: Leaves falling off a tree. [Tasmania, Grade-9]
U: Birth of a newborn child, or death of an old man/woman. [Tasmania, Grade-9]
U: The weather; the depth of the sea; the positions of clouds. [Tasmania, Grade-9]

Note that the final response offers a number of examples, but that they are all drawn from the single context of natural phenomena and hence constitute a repeated unistructural response.

Some responses provide an example of random in the context of games or competitions, often with a defining characteristic of selection or choice. Notice that these models embody the meanings commonly associated with random, such as unpredictability, and equal likelihood.

U: Picking a name out of a hat. [Tasmania, Grade-9]
U: Competitions and tattslotto and Keno and that kind of stuff. [Tasmania, Grade-9]
U: Getting picked for a prize. "You're randomly selected". [Tasmania, Grade-6]

Some students attempt to define random related to chance, unpredictability and lack of a pattern. Again, students refer to either random processes or random outcomes.

U: When you get things mixed. [Singapore, Grade 5]
U: Quickly, roughly, any old way, etc. [Tasmania, Grade-9]
U: They happen in no pattern or order. [Tasmania, Grade-9]
U: Things that happen with no system involved. [Tasmania, Grade-9]

Responses which offer examples from more than one of the contexts mentioned above are classified at the multistructural level, as they show evidence of a multi-faceted appreciation for "random". These responses are typical of the multistructural level by providing a list of possibilities, with no mention of a generalizing characteristic.

M: Wind direction; computer number choice; coin tossing. [Tasmania, Grade-9]
M: The rolling of a dice; when rain comes from the sky; what the weather is like. [Tasmania, Grade-9]
M: Tattslotto numbers; numbers of birds flying in groups; wind gusts. [Tasmania, Grade-9]
Other multistructural responses provide a description which refers to multiple events or a class of events, but without clearly relating examples as required in the question. These responses often involve an idea about selection or being picked, with the added specification that the selection is being made without a system or reason.

M: If something happens in a random way it just comes about by chance. It means it wasn't particularly meant one way or another. [Tasmania, Grade-6]
M: Something is chosen without thinking i.e a number is picked out of 10. [Tasmania, Grade-9]
M: Someone or something is picked without looking for your name or things like that (picked out of a hat). [Tasmania, Grade-6]
Other responses at the multistructural level provide an example with a simple defining characteristic.
M: Death because you never know if your going to be next you just have to wait and see. [Tasmania, Grade-9]
M: Rain. Thunder. As weather is unpredictable. [Tasmania, Grade-9]
M: Random is anonymous, people are checked at random for their blood/alcohol level. [Tasmania, Grade-9]

Responses at the relational level typically provide a number of examples from different contexts, and then attempt to relate them by a generalised description. Sometimes responses provide the generalised description first, then use examples.

R: Random things happen out of the ordinary, it is unpredictable like the lotto and weather. [Tasmania, Grade-9]
R: Anything that does not "occur" on a regular basis or in a pattern, i.e., lotto numbers. [Tasmania, Grade-9]
R: Picking jellybeans from a packet. You choose the colours randomly & don't know which one you'll get. Nothing influences your choice. [Tasmania, Grade-9]

Table 4 shows the percentage of responses to Item Q3 by SOLO response level and by grade. Students from the two countries differed in the meaning they gave to the words, although not in the level of their responses. Students from Singapore tended to use "natural" examples of random in preference to "lottery" examples, although some of these students mentioned random "numbers", which were classified U+IK, because they did not make clear reference to a form of lottery or selection.

Table 4
Percentage of Responses to Item Q3 by SOLO Response Level and by Grade

<table>
<thead>
<tr>
<th>Level</th>
<th>Tasmania Grades</th>
<th>Singapore Grades</th>
</tr>
</thead>
<tbody>
<tr>
<td>NR No response</td>
<td>784016</td>
<td>3320</td>
</tr>
<tr>
<td>PI Irrelevant</td>
<td>14196</td>
<td>654634</td>
</tr>
<tr>
<td>U+IK Order/Numbers</td>
<td>14 2</td>
<td>02120</td>
</tr>
<tr>
<td>U Human</td>
<td>024</td>
<td>0 0</td>
</tr>
<tr>
<td>UN Natural</td>
<td>4613</td>
<td>25 2311</td>
</tr>
</tbody>
</table>
Item 4. If you were given a "sample", what would you have?
Responses to this item range from those totally unfamiliar with the term "sample", to those that can explain adequately what a sample is and the purpose of having a sample.
Responses which do not show any clear evidence of appropriate usage of the term "sample" are classified as prestructural.

P: Something that's too hot. [Tasmania, Grade 3]
P: Something that is a sample of something. [Tasmania, Grade-6]
P: I would have the answers. [Singapore, Grade 2]

Many responses include no clearly defined characteristic of a sample, but rather suggest an appropriate example. These responses seem to refer to everyday experiences to support a concrete idea about samples which is not clearly expressed. As such, these responses are classified as having iconic support for a unistructural idea in the concrete symbolic mode.

U+IK: You might get a sample of blood. [Tasmania, Grade 3]
U+IK: I would have an ice-cream sample. [Singapore, Grade 4]
U+IK: A sample of food from the supermarket. [Tasmania, Grade-9]
U+IK: I would have the product for free. [Singapore, Grade 5]

Responses which offer a single idea associated with the concept are classified as unistructural in the concrete symbolic mode. Responses include either a noun indicating that a sample is a bit or part, or an expression that the purpose of a sample is to test.

U: A little bit. [Tasmania, Grade 3]
U: I would have something to try. [Singapore, Grade 4]
U: I would have an example. [Singapore, Grade 2]

At the multistructural level, responses make reference to a sample either as a part or as a test of some larger "total", but do not include both ideas. Sometimes the total is included by way of an example.

M: A small piece of something. [Tasmania, Grade 6]
M: A little of something, not the whole thing but a little piece of it.

[ Tasmania, Grade 9]
M: Something free for you to try. [Singapore, Grade 4]
M: I would have a sample of dirt, something they have done tests on. [Tasmania, Grade 9]

Relational level responses integrate the "small part of something" aspect of a sample with the purpose of providing a test.

R: I would have a bit of something to taste, try or eat etc. [Singapore, Grade 4]
R: You would have a small amount of something, like a new washing up liquid to test in your own home. [Tasmania, Grade 9]
R:A thing taken out of a larger group, for you to view as a representative of that group. [Tasmania, Grade 9]
Relational responses demonstrate the acquired concrete symbolic concept of a sample as "a small part representing a whole (population)". This concept is the basis for assessing reliability of inferences based on samples in concrete situations with reference to sample size and sample selection.
Table 5 shows the percentage of students in each grade who gave each level of response. Older students generally responded at higher levels than younger students. In relation to the research questions, the responses of students from Australia and Singapore appear quite similar, with the exception of a high rate of irrelevant responses at Grade 2, which suggests that young students from Singapore who had not heard the term were more concerned to give a response while Tasmanian students who did not know preferred not to respond. It is interesting to note that the modal "active" response involved an example only. This suggests that experiences with the word "sample" in concrete contexts are important for younger students' early notions of sample.

Table 5
Percentage of Responses to Item Q4 by SOLO Response Level and by Grade

<table>
<thead>
<tr>
<th>Grade</th>
<th>NR No response</th>
<th>PI Irrelevant</th>
<th>US Small part; A test of tot</th>
<th>MS Small pt of tot; A test of</th>
<th>RS Small part repres tot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Tasmania Grades</td>
<td>42167</td>
<td>842</td>
<td>1116</td>
<td>163749</td>
<td>31525</td>
</tr>
<tr>
<td>Singapore Grades</td>
<td>3014</td>
<td>63829</td>
<td>183123</td>
<td>33120</td>
<td>030</td>
</tr>
</tbody>
</table>

Discussion
This study aimed to investigate the meanings and contexts of responses to statistical language of students from Australia and Singapore. While the results from the Singapore sample are tentative, due to a small sample size and grade levels not directly comparable to the Tasmanian sample, the conceptual development is parallel as viewed from a SOLO perspective. The fact that the conceptual development is parallel between students in Australia and Singapore at the basic levels of understanding of the language of chance and data would be encouraging to those who wish to share curriculum materials and teaching methodologies. It also may be that professional development programs for teachers could be shared between the two countries as well as among others in the region.
There were, however, a number of differences in the contexts that students used as illustrations that reflect their experiences, particularly in relation to the chance words "impossible", "possible", "certain", and "random". Students from Singapore were more likely to focus on events from the natural world, rather than consider human
designed structures such as school schedules or games and competitions. The reasons for this are not clear. While it would appear plausible that language and "society" play a role in students' conceptions, it is not possible to say whether the understandings of the students in this study were grounded in experiences from school or from the surrounding environment. It is at least worth noting that while the concept of average is introduced in both countries somewhere around Grade 4, in Singapore there is no formal teaching of random or probability in the primary school mathematics programme.

It is important for teachers to realise the variety of student response they are likely to receive when exploring terms associated with chance and data in the classroom. As has been pointed out by Fischbein, Nello, and Marino (1991) in drawing out the didactical implications of their research, it should not be taken for granted that children understand spontaneously the meaning of the terms "impossible", "possible", and "certain". Children have to be trained to distinguish between rare and impossible, and between highly frequent and certain (pp.-528-529). Knowledge such as this should assist in planning questions, feedback and experiences to help raise levels of response. The basic understanding of statistical language as investigated in the current study has received little previous research attention, and yet is very relevant to contemporary school curricula. The responses of younger children indicate that experiences in concrete contexts, such as the food sample in the supermarket or the water sample to test in science, can be used to build the concept of a sample as a part to represent a total. In relation to "random", some responses provide an example of random in the context of games or competitions. These are the common experimental situations which may be used in the classroom. Notice also that these models embody the meanings commonly associated with random, such as unpredictability and equal likelihood, other key concepts in the development of statistical and probabilistic ideas.

The use of situated contexts is important not only to develop the statistical concepts, but also to make students confront their intuitive misconceptions. When teaching about sampling, or other aspects of chance and data in social contexts, it must be acknowledged that sometimes students will not focus on the desired concepts. This can provide the basis for class discussion to compare different meanings and contexts for statistical terms. The discussion permits students to encounter and confront everyday meanings for terms. For example, a common usage of average is to refer to something which is below standard, e.g., "the movie was pretty average". This out-of-school usage of the term average provides the teacher with an opportunity to discuss the mathematical definition of the term and how it differs from the general usage. This is not limited to statistical situations, but occurs in many branches of mathematics, such as the use of the word "similar" in geometry.

Sources of situated contexts involving data handling include practical
experiments with concrete materials, and also secondary sources such as news reports. For the concept of sample, experiences with experimental data collection might include sample products and sample selection from a bag of hidden objects, as well as the collection of meaningful class sample data. As noted earlier, for probabilistic language and in particular random, open discussion of situations suggested by students, such as games, natural events or personal experience, provide meaningful contexts, some of which lead to experimentation. The concept of average, following sample, can be illustrated if meaningful data are collected in classroom situations. The links among the various concepts and contexts may be established quite early, as is demonstrated by Watson and Pereira-Mendoza (1996) for bar charts, another fundamental concept in data handling.

The outcomes described in this study occur as students' responses are first appearing in the concrete symbolic mode. The questions asked did not elicit higher level responses. Other research (Watson, Collis, Callingham & Moritz, 1995; Watson, et al., 1995b, 1996, in press), however, has documented a second U-M-R cycle which applies the basic concepts studied here in problem solving situations. An understanding of both the concepts and the contexts in which they apply is a necessary foundation for later work. It is at the relational (R) level that these ideas are consolidated into usable constructs. In practice, it would be assumed that classroom discussion would build on the sort of written responses obtained in this study to achieve the higher levels of understanding required for solving problems in realistic contexts.

Acknowledgments

This research was funded by the Australian Research Council, Grant Number A79231392.

References


improve probabilistic intuitions? An exploratory research study.
probability judgements in children and adolescents. Educational Studies
in Mathematics, 22, 523-549.
statistical terms. In W. Geeslin & K. Graham (Eds.), Proceedings of
the Sixteenth PME Conference. Vol. 3 (p. 160). Durham, NH: Program
Committee of the 16th PME Conference, USA.
concepts in probability and statistics: Implications for research.
Contemporary Approaches to Research in Mathematics, Science, Health and
Environmental Education. Symposium conducted at Deakin University,
Melbourne.
National Council of Teachers of Mathematics. (1989). Curriculum and
Qualitative Solutions and Research. (1992). Non-numerical
Unstructured Data • Indexing Searching and Theorizing (NUD•IST) v3.0.
[Computer program]. Melbourne: LaTrobe University.
Watson, J. M. (1994). Instruments to assess statistical concepts in
the school curriculum. In National Organizing Committee (Ed.),
Proceedings of the Fourth International Conference on Teaching
Statistics. Volume 1 (pp. 73-80). Rabat, Morocco: National Institute
of Statistics and Applied Economics.
mathematics and society. In A. Richards, G. Gillman, K. Milton, & J.
Oliver (Eds.), Flair: Forging links and integrating resources (pp.

understanding chance and data concepts. In J. P. da Ponte & J. F. Matos
(Eds.), Proceedings of the Eighteenth International Conference for the
Psychology of Mathematics Education. Vol. 4 (pp. 369-376). Lisbon: PME.
Watson, J. M., Collis, K. F., Callingham, R. A., & Moritz, J. B.
Assessment of statistical understanding in Australian schools. Paper
presented at the Statistics '93 conference, Wollongong, NSW.
statistical understanding in Grades 3, 6 and 9 using a short answer
questionnaire. In G. Bell, B. Wright, N. Leeson, & G. Geake (Eds.),
Challenges in Mathematics Education: Constraints on Construction (pp.
675-682). Lismore, NSW: Mathematics Education Research Group of
Australasia.
understanding of luck. In B. Atweh & S. Flavel (Eds.), Proceedings of the Eighteenth Annual Conference of the Mathematics Education Research Group of Australasia (pp. 550-556). Darwin, NT: MERGA.


