Is Statistical Literacy Relevant for Middle School Students?[1]

Jane M. Watson
University of Tasmania

In answering the question in the title with a resounding "Yes," this paper considers examples of the development of middle school students' understanding related to statistical literacy. The development spans six levels of increasing engagement with context, increasing flexibility of numeracy skills, increasing appreciation of variation, and increasing ability to describe terminology. The examples provided can be used as assessment items or as the basis of classroom learning activities.

Introduction

Several factors contribute to the importance of students developing statistical literacy skills at the school level. First is the expectation for participation as citizens in an information and data driven age where decision-making is likely to be based on critical skills from the realms of statistical literacy. Wallman (1993) foresaw this as both a public and private issue and Gal (2002) suggested two components required of adults. One component is the ability to interpret and evaluate critically statistical information in a variety of contexts and the other is the ability to communicate this understanding in a fashion that can have an impact on decision-making. With respect to these components of adult statistical literacy, Gal also noted the need for dispositions associated with attitudes and beliefs that would motivate citizens to be critical thinkers in this arena.

To have adults meeting the goals set by Wallman (1993) and Gal (2002) it is necessary to begin at the school level. Reform to the mathematics curriculum in many countries from around 1990 (e.g., Australian Education Council [AEC], 1991; National Council of Teachers of Mathematics [NCTM], 1989), which brought topics related to statistics and probability into the school curriculum, provided a starting point for consideration of statistical literacy. Although curriculum documents focus mainly on the traditional stages in statistical investigations, they at least acknowledge the importance of critical thinking. The NCTM (1989) for example states, "A knowledge of statistics is necessary if students are to become intelligent consumers who can make critical and informed decisions" (p. 105). The topics in the chance and data part of the school mathematics curriculum, for example, sampling, graphing, data reduction, and inference, form the foundation for building sophisticated thinking skills. Collectively these topics are another factor contributing to the development of statistical literacy at school and to its importance.

Out-of-school experiences, however, place basic statistical ideas in many and varied contexts. The traditional divisions of the school curriculum into different subjects, particularly at the secondary level, have worked against integrating the ideas of chance and data across the curriculum into subjects such as science, social science, or health where critical thinking can be developed in contexts that will be useful to students when they leave school. New curriculum reforms in some places, however, are moving to more holistic and integrated approaches to the school curriculum that encourage the inclusion of quantitative literacy across the curriculum as a tool for communication and critical thinking (Department of Education Tasmania, 2002). This factor provides further impetus to the development of statistical literacy as an important contributor to the more general area of quantitative literacy (Steen, 2001).

The research that resulted in the hierarchy and examples reported here began in 1993 with the aim to document school students' understanding of statistical topics in the new curriculum documents (e.g., AEC, 1991). The research was based on longitudinal surveys and interviews of students in grades 3 to 9, exploring foundational concepts, concepts applied in media contexts, and problem solving based on concepts (Watson, 1994). Analysis of data was based mainly on two hierarchical models. One arose from cognitive psychology (Biggs & Collis, 1982) and characterised responses based on their structural complexity, the number of components of the task employed, and the recognition of contradictions should they occur. The other model was a three-tiered hierarchy of statistical literacy (Watson, 1997), beginning with understanding of basic terminology, followed by understanding of terminology in context, and then by the ability to question claims made without appropriate statistical justification.

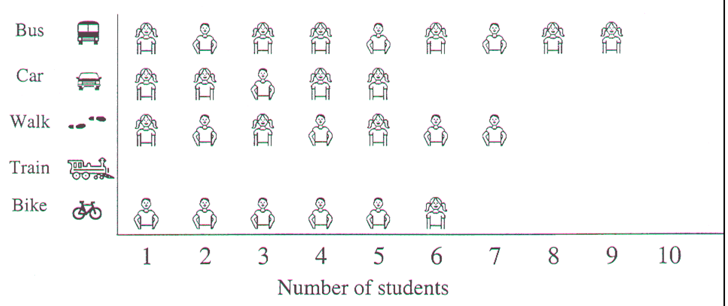
More recently Rasch (1980) analysis has been employed with survey items coded according to the above criteria to suggest levels of progression with respect to understanding of statistical literacy across the middle years of schooling (Watson & Callingham, 2003). The focus of this paper is the description of six levels of appreciation of aspects of statistical literacy approaching the goal of being able to communicate critical evaluations of statistical claims made in varying contexts. Selected examples of what students are likely to show they know and can do at each level are given.

Examples of Tasks

Several tasks are presented in order to illustrate the typical progression of student responses through the six levels of the statistical literacy hierarchy. These tasks represent several strands of the school curriculum that contribute to the foundations required for the statistical literacy understanding that is desired when students leave school. The levels of response for these items are distributed across the six levels of statistical literacy as observed in the Rasch analysis (Watson & Callingham, 2003) and are discussed in the next section.

The tasks in Figure 1 are based on a pictograph on how children travel to school. Six questions are asked in relation to the graph and four of these (1., 2., 3., and 6.) are discussed here. Figure 2 contains a task based on finding the average of nine measurements, one of which is an outlier. The task and associated rubric allow for any of the measures of mean, median, or mode to be used as long as they are adequately justified. The item on sampling considered here asked, “What does ‘Sample’ mean? Give an example of a ‘sample’.”

How children came to school one day



1. How many children walked to school?
2. How many more children come by bus than by car?
3. Would the graph look the same everyday? Why or why not?
4. A new student came to school by car. Is the new student a boy or a girl? How do you know?
5. What does the row with the Train tell about how the children get to school?
6. Tom is not at school today. How do you think he will get to school tomorrow? Why?

Figure 1. Pictograph tasks

A small object was weighed on the same scales separately by nine students in a science class. The weights (in grams) recorded by each student are shown below.

6.3 6.0 6.0 15.3 6.1 6.3 6.2 6.15 6.3

The "average" value could be calculated in several ways.

1. How would you find the average? ______________________________
2. The average weight is _________ grams. [Show your working in the box below.]

Figure 2. Average task

A media item based on sampling is shown in Figure 3. The questions within the item are scaffolded to provide assistance if students do not see initially the difficulty of claiming a sample in Chicago represents the population of the United States. The item is coded based on answers to both parts (i.e., whether students require the hint in Question 2 or not). Another item based on a newspaper article is given in Figure 4. In this case students are asked to do two things: draw a graph illustrating a claimed cause-effect relationship and question the researcher about his claims. These two questions are coded separately.

ABOUT six in 10 United States high school students	**Family car is killing us, says Tasmanian researcher**

say they could get a handgun if they wanted one, a third of them within an hour, a survey shows. The poll of 2508 junior and senior high school students in Chicago also found 15 per cent had actually carried a handgun within the past 30 days, with 4 per cent taking one to school.

1. Would you make any criticisms of the claims in this article?
2. If you were a high school teacher, would this report make you refuse a job offer somewhere else in the United States, say Colorado or Arizona? Why or why not?

Figure 3. Sampling task

Twenty years of research has convinced Mr Robinson that motoring is a health hazard. Mr Robinson has graphs which show quite dramatically an almost perfect relationship between the increase in heart deaths and the increase in use of motor vehicles. Similar relationships are shown to exist between lung cancer, leukaemia, stroke and diabetes.

1. Draw and label a sketch of what one of Mr. Robinson's graphs might look like.
2. What questions would you ask about his research?

Figure 4. Association task

Two other tasks based on newspaper headlines address the Chance part of the curriculum. In the task in Figure 5 students are asked to order seven of eight newspaper headlines and in Figure 6, students are asked for an interpretation of an odds statement.

Here are eight chance words or phrases from headlines.
A. 58 per cent success at SkillShare
B. Impossible
C. It's a sure thing
D. Jack looking good for big one
E. Holden an unlikely American hero
F. No worries
G. Smith in doubt to play
H. There's a 50-50 chance

Please mark on the scale below the likelihood expressed by each of the seven phrases A to G. H is done as an example.

Likelihood

(low) (high)

|_________________________|_________________________|

0 H 1

Figure 5. Headline ordering task

North at 7-2
But we can still win match, says coach

1. What does "7-2" mean in this headline about the North against South football match? Give as much detail as you can.
2. From the numbers, who would be expected to win the game?

Figure 6. Odds task

The tasks chosen for discussion here allow for a range of responses, some of which can display critical questioning and acknowledgement of uncertainty. They represent the kind of interpretations to be expected as students move from the mathematics class to other school subjects and contexts outside of school. Being able to address the issues in these types of tasks is relevant to the intellectual development of middle school students.

Six Levels of Statistical Literacy at School

Based on tasks such as those above and others (Watson &Callingham, 2003), analysis of responses across grades from grade 3 indicated that important contributors to improved performance across levels were increasing engagement with context, greater facility with numeracy skills, greater appreciation of variation, and deeper appreciation of the meaning of terminology associated with application of statistical concepts (e.g., sample, average, and random).

Idiosyncratic Understanding – Level 1. At this level students can do little that would satisfy a statistician. Generally their appreciation of context is non-existent, idiosyncratic, or based on personal experience, and this reveals itself particularly in relation to sampling, chance, and inference. The task to evaluate an article with a non-representative sample related to gun access in United States schools (Figure 3) is likely to be answered with beliefs such as "students shouldn't have guns." Chance outcomes are likely to be suggested based on favourite colours or numbers. The explanation of the odds statement in Figure 6 is likely to suggest the current score in the game, e.g., "North 7 goals to South 2 goals," or just refer generally to betting. For many tasks related to graphing and average there is little or no evidence of engagement. Definitions of terms are unlikely to display more than tautological appreciation, whereas work with tables and pictographs suggests students can successfully read from appropriate cells, count in a one-to-one fashion, and perform simple two-digit additions. Students, for example, are likely to answer the first two questions in Figure 1 correctly. Meaningful engagement is hence limited to straight-forward contexts with limited demands. On the one hand for the third question in Figure 1, students are likely, when directly asked if the graph will be the same every day, to say, "No, it could change," recognising the possibility of variation. On the other hand the response to the last question on Tom, is likely to avoid making a decision, for example saying, "the same way as yesterday" or "I can't tell."

Informal Understanding – Level 2. Although engagement with tasks at this level is still likely to be colloquial, with students distracted by irrelevant features, there are times when students demonstrate appreciation of single elements of the concepts that are relevant to tasks. This is shown for example in responses to definitions, where students are likely to focus on a single feature of "sample," such as "a test," or of "average," such as "normal." In responses to tasks requiring inferences, however, they are likely to focus on story-telling based on their own personal experience or to focus on inappropriate features of a pictograph, such as suggesting Tom (Q6 in Figure 1) comes to school by train because there is a gap in the graph. For chance predictions justification is likely to be based on "anything can happen" and variation in repeated trials is unlikely to be to be volunteered. When explicitly asked for a surprising chance outcome, however, many can provide an appropriate one. Numeracy skills are likely to be limited to one-step table and graph calculations, involving addition or subtraction, or to an appreciation of "half" as a chance when displayed on a spinner.

Appreciation of context is very limited; for example, in ordering in headlines in Figure 5, one or two headlines are placed in the inappropriate half of the number line.

Inconsistent Understanding – Level 3. Although students' responses at this level are more likely to engage with context than at the earlier levels, they are likely to be selective, often depending on a supportive format in the statement of the task. More than one feature is sometimes considered in responses but statistical ideas are likely to be expressed qualitatively rather than quantitatively. When attempts at quantitative approaches are made, for example for the item on average in Figure 2, although meaningful measures are selected, the calculations are not correct. In some contexts, for example surveying classmates in a school setting, students are likely to understand the purpose of sampling but not be able to detect bias in sampling methods. In other contexts, such as drawing a graph to represent Mr. Robinson's claim in Figure 4, students are likely to provide a formless graph with no labels or a graph with a single comparison of two values (e.g., one bar for car use and one bar for heart deaths). When provided with alternatives from which to choose, students are likely to *identify* correct interpretations of "15% chance of getting a rash" and "50 families having an average of 2.2 children." Qualitative success in chance contexts is likely to relate to identifying "more" as related to greater probability and to interpreting relative chances in relation to the conjunction of two events, again in terms of "greater" or "less" chance.

Consistent Non-critical Understanding – Level 4. At this level students are likely to display a consolidation of appreciation of various contexts without the ability to question claims of a suspicious nature. They are hence more successful with tasks that do not require finding errors in reasoning. Multiple features of concepts are likely to be shown in giving definitions, for example describing a sample as "a small part of something bigger," an average in terms of the mean algorithm or a descriptive middle, and variation as "change like the weather varying over the period of a few days." In suggesting methods for sampling classmates in a school setting, they are likely to suggest representative methods such as "I'd pick 10 from each grade," rather than random methods. In judging the article on guns in schools in the United States in Figure 3, students are likely to suggest that "people could be lying" or "you need to ask everyone in the United States." Graphing skills improve in a technical sense in that students are likely to identify highest values and ranges, and show partial success in attempting to draw an association of two variables, for example for the article about Mr. Robinson in Figure 4 indicating an increase in one of the variables with time or plotting two values only for each of car use and heart deaths. In questioning Mr. Robinson's claim they are likely to ask about his sample size or where he collected the data. For the average task in Figure 2, students are likely to calculate the mean correctly but without considering the outlier. They are likely to interpret straight-forward conditional statements and order the likelihood of newspaper headlines in Figure 5 successfully. For the odds question in Figure 6, suggestions are likely to relate qualitatively to percent chance of winning, odds, or a prediction of the final score. In interpreting a pictograph, students are likely to draw inferences based on frequency but not to include an element of uncertainty, for example in suggesting Tom (in Figure 1, Q6) will come to school by bus tomorrow because most children do.

Critical Understanding – Level 5. At this level students are likely to engage in critical analysis in both familiar and unfamiliar contexts that do not require sophisticated mathematical reasoning. They are likely to use terminology appropriately and give integrated definitions, for example of "sample" as a "a small part of a whole to test and see what it is like." In suggesting methods of selecting a sample of students from a school, students are likely to include both random and representative methods or two different random methods, as well as identifying bias in the methods suggested by others. For the Chicago/United States sampling dilemma in Figure 3, students are likely to recognise the difficulty in relation to considering the situation in another state, such as Arizona. Graphs drawn to show the association of two variables are likely to be acceptable, for example for the task in Figure 4. For the odds headline in Figure 6, students are likely to provide fractions or ratios to justify the chances of North winning, but these are incorrect, usually in the wrong direction. In relation to variation students are likely to volunteer words like "about" when making chance predictions and to notice change over time or mention variation explicitly when looking for unusual features of a bar chart from the media.

Critical Mathematical Understanding – Level 6. Students at this level not only display critical questioning skills in all contexts but also can employ proportional reasoning and understanding of independent events when calculating probabilities. They further realise the importance of including an aspect of uncertainty in drawing inferences related to tasks, for example in suggesting how Tom (Q6 in Figure 1) will get to school tomorrow, they say, "He might come by bike because it looks like more boys come by bike." The language issue is illustrated with the newspaper article in Figure 3 about access to guns in United States schools. At this level students recognise the non-representative sample without a hint and without specific mention of "sample" or "population" in the article. Students are likely to question the presence of the outlier in the data set in Figure 2 and eliminate it when calculating the mean. When asked to question Mr. Robinson's methods in the article in Figure

4, students at this level question a cause-effect claim and often suggest the presence of a lurking variable. They are also likely to be able to quantify the chance of South winning as 7/9 in the game referred to in the headline in Figure 6.

Implications

Clearly Level 6 is the goal by the time students leave school but without an appreciation of the preceding levels of likely progression, it is not possible to plan experiences that will assist students to the higher levels of understanding. It is important to know the kinds of misapprehensions and partial comprehensions that develop and change over the middle years of schooling. Although it is not possible to place grade levels definitively within the levels of development as there is considerable variation observed (e.g., Watson & Callingham, in press), it should be noted that by the end of grade 10, the end of compulsory schooling, many students are not performing at the highest level described above.

One of the important findings in this research has been the influence of context in determining levels of understanding. At the first two levels students struggle with interpreting all but the most straight-forward contexts. At Levels 3 and 4, they cope with tasks that require an understanding of concepts in various settings but not the criticism of questionable claims. At the top two levels students appreciate the subtleties of context and show a propensity to question most claims that are made without proper statistical justification. Other findings with respect to specific topics in the school chance and data curriculum tend to confirm development found in other studies.

There are many possibilities for future research in association with students' development of understanding of statistical literacy while at school. Extending the survey work to include interviews will provide further evidence of students' constructed understandings. Work with sampling suggests this will be a useful approach (Watson & Moritz, 2000). As well classroom-based research including instruction specifically aimed at improving statistical literacy will be useful in determining which types of activities are most useful in this regard. Suggestions of activities are made by Watson (2002). Employing technology, perhaps to access media contexts or to develop familiarity with software, may prove to be an important part of instruction trialled in the classroom. Finally, it will be important to carry out research with teachers themselves to explore their understandings and help them develop the classroom strategies required to improve the levels of performance of their students.

In the meantime there are implications for teaching the Chance and Data curriculum as outlined in the Victorian *Curriculum Standards Framework* (CSF) (Board of Studies, 2000). This research points to the importance of acknowledging partial progress toward goals when assessing outcomes and yet the CSF goals are stated in positivist "complete success" terminology. Teachers need to be aware that whereas some students will achieve complete goals, others will get partway there and should not be relegated as failures. In terms of statistical literacy as illustrated in the examples in this paper, the CSF does not address the critical questioning that should be developing over the middle school years, or the skills of predicting while acknowledging uncertainty when considering data. "Variation" is also a topic that can and should be recognised as appropriate for middle school students.

Is statistical literacy relevant for middle school students? Not only is the answer "Yes" in terms of the overall goals of critical thinking and quantitative literacy before students leave school, but also the means of achieving success are available if teachers will take a bit of time to include appropriate tasks to motivate student thinking across the middle school years.

REFERENCES

Australian Education Council. (1991). *A national statement on mathematics for Australian schools.* Carlton, Vic.: Author.

Biggs, J.B., & Collis, K.F. (1982). *Evaluating the quality of learning: The SOLO taxonomy.* New York: Academic Press.

Board of Studies. (2000). *Mathematics: Curriculum Standards Framework II*. Carlton: Author.

Department of Education Tasmania. (2002). *Essential learnings framework 1*. Hobart: Author.

Gal, I. (2002). Adults' statistical literacy: Meanings, components, responsibilities. *International Statistical Review, 70*, 1-51.

National Council of Teachers of Mathematics. (1989). *Curriculum and evaluation standards for school mathematics*. Reston, VA: Author.

Steen, L.A. (Ed.). (2001). *Mathematics and democracy: The case for quantitative literacy.* Washington, DC: Woodrow Wilson National Fellowship Foundation.

Wallman, K.K. (1993). Enhancing statistical literacy: Enriching our society. *Journal of the American Statistical Association, 88*, No. 421, 1-8.

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* (Vol. 1, pp. 73-80). Rabat, Morocco: National Institute of Statistics and Applied Economics.

Watson, J.M. (1997). Assessing statistical literacy using the media. In I. Gal & J.B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 107-121). Amsterdam: IOS Press and The International Statistical Institute.

Watson, J.M. (2002). Lessons from variation research II: For the classroom. In M. Goos & T. Spencer (Eds.), *Mathematics - making waves*. (Proceedings of the Nineteenth Biennial Conference of the Australian Association of Mathematics Teachers Inc., Brisbane, pp. 424-432). Adelaide, SA: AAMT, Inc.

Watson, J.M., & Callingham, R.A. (2003). Statistical literacy: A complex hierarchical construct. *Statistics Education Research Journal, 2*(2), 3-46.

Watson, J.M., & Callingham, R.A. (in press). *Statistical literacy: From idiosyncratic to critical thinking*. Paper accepted for the International Association for Statistics Education Roundtable on "Curricular Development in Statistics Education," Lund, Sweden.

Watson, J.M., & Moritz, J.B. (2000). Developing concepts of sampling. *Journal for Research in Mathematics Education*, *31*, 44-70.

---

[1] Some parts of this paper appeared under the title "Statistical Literacy at the School Level: What Should Students Know and Do?" in the *Bulletin of the International Statistical Institute 54th Session* (Volume LX, Book 2, Invited Papers, pp. 68-71). Berlin: ISI, 2003.