THE DEVELOPMENT OF A FRAMEWORK CHARACTERIZING MIDDLE SCHOOL STUDENTS' STATISTICAL THINKING ®

Cynthia W. Langrall and Edward S. Mooney
Illinois State University, USA

Given the importance of instruction in promoting students' statistical literacy, a cohesive picture of the development of students' statistical thinking is needed to better inform classroom teachers and curriculum developers. With this in mind, one of the authors developed a framework to characterize middle school students' statistical thinking within four statistical processes, across four levels of thinking. A subsequent study (Langrall, Mooney, Hofbauer, & Johnson, 2001) addressed gaps in Mooney's framework. This paper describes how the findings of the Langrall et al. study were merged with the framework and reports on resulting modifications to the entire framework.

INTRODUCTION

It is widely acknowledged that proficiency in statistical skills enables people to become productive, participating citizens in an information society (e.g., Australian Education Council, 1994; National Council of Teachers of Mathematics [NCTM], 2000; Secretary's Commission on Achieving Necessary Skills [SCANS], 1991; School Curriculum and Assessment Authority for Wales, 1996). In the United States, SCANS (1991) has recommended that benchmarks be established to inform statistics education at the secondary level to better prepare students for the workforce. In response to this recommendation, calls for reform in mathematics education have advocated a more pervasive approach to statistics instruction at all levels (NCTM, 2000). More specifically, the NCTM has recommended that middle school students (ages 12 to 14) have experiences collecting, organizing, representing, and interpreting data. However, research on middle school students' statistical thinking is sparse. A cohesive picture of middle school students' statistical thinking is needed to better inform classroom teachers and curriculum developers.

Researchers like Cobb et al. (1991) have identified the need for cognitive models of students' thinking to guide the planning and development of mathematics curriculum and instruction. There is evidence that research-based knowledge of students' thinking can assist teachers in providing meaningful instruction (Fennema & Franke, 1992). With this in mind, Mooney (2002) developed and validated the Middle School Students' Statistical Thinking (M3ST) framework based on a synthesis of the literature and observations and analyses of students' thinking in interview settings. The M3ST framework incorporated four statistical processes (describing, organizing and reducing, representing, and analyzing and interpreting data) and included descriptors that characterize four levels of students' statistical thinking within each process.

In validating the framework, Mooney (2002) found that two aspects of students' statistical thinking were not adequately addressed in his study: categorizing and grouping data and students' use of proportional reasoning. In a recent study (Langrall, Mooney, Hofbauer, & Johnson, 2001), descriptors for these two subprocesses were generated for inclusion in the M3ST framework. This paper will report on the merging of these descriptors and the subsequent modification of the entire framework.

COMPONENTS OF THE M3ST FRAMEWORK

The M3ST framework was constructed with the perspective that students' statistical thinking is represented by the cognitive actions in which they engage during the data-handling processes of describing, organizing and reducing, representing, and analyzing and interpreting data (Reber, 1995; Shaughnessy, Garfield & Greer, 1996). The framework characterizes students' statistical thinking for four processes, across four levels of thinking.

Statistical Processes

Describing data. Describing data entails the explicit reading of data presented in tables, charts, or graphical representations. Curcio (1987) considers "reading the data" as the initial stage of interpreting and analyzing data. Therefore, ability to read data displays can be considered the basis for students to begin making predictions and discovering trends. In initially developing the framework, Mooney (2002) identified four *subprocesses* related to describing data: (a) showing awareness of display features, (b) recognizing the same data in different data displays, (c) evaluating the effectiveness of data displays of representing data, and (d) identifying units of data values.

Organizing and reducing data. Organizing and reducing data involves arranging, categorizing, or consolidating data into a summary form. As with the ability to describe data displays, the ability to organize and reduce data is critical for learning how to analyze and interpret data. Measures of center and dispersion are important in analyzing and interpreting data, however, research indicates that middle school students rarely employ these statistics (Reading & Pegg, 1996; Friel & Bright, 1996). Mooney (2002) initially identified three subprocesses for organizing and reducing data: (a) grouping or ordering data, (b) describing data using measures of center, and (c) describing the spread of data.

Representing data. Representing data involves displaying data in a graphical form. According to Friel, Curcio and Bright (2001), the *graphical sense* involved in representing data "includes a consideration of what is involved in constructing graphs as tools for structuring data and, more important, what is the optimal choice for a graph in a given situation" (p. 145). Representing data, like the previous two processes, is important in analyzing and interpreting data. How data are represented will determine the trends and predictions that can be made. Also, different data displays can communicate different ideas about the same data. Research has shown that middle school students are unclear of the necessary features of a graph (e.g., Berg & Phillips, 1994) and find making graphical representations of situations difficult (e.g., Mevarech & Kramarsky, 1997). Mooney (2002) concluded that three subprocesses underlie representing data: (a) constructing a data display for a given data set, (b) completing a partially constructed atypical data display (a data display not usually presented at the middle school level), and (c) constructing an alternate data display for data presented in a given display.

Analyzing and interpreting data. Analyzing and interpreting data consists of identifying trends and making inferences or predictions about the data. Curcio (1987) has identified two levels beyond "reading the data" that form the basis for analyzing and interpreting data. "reading between the data" involves making comparisons within the data. "Reading beyond the data" entails the ability to make extensions, predictions or inferences from the data. These levels were used in development of the initial M3ST framework resulting in three subprocesses for analyzing and interpreting data: (a) making comparisons within data sets or data displays, (b) making comparisons between data sets or data displays, and (c) making inferences from a given data set or data display. The ability to analyze and interpret data builds upon the ability to read data displays, organize and reduce data, and represent data.

Cognitive Levels

The levels of students' statistical thinking in the M3ST framework were based on the general developmental model of Biggs and Collis (1991). Their model incorporates five modes of functioning: sensormotor (from birth), ikonic (from around 18 months), concrete symbolic (from around 6 years), formal (from around 14 years), and post formal (from around 20 years). Within each mode, three cognitive levels (unistructural, multistructural, and relational) recycle and represent shifts in the complexity of students' reasoning. According to Biggs and Collis, each of the five modes of functioning emerges and develops in a way that incorporates the continuing development of earlier modes. Thus, they also recognize two other cognitive levels: the prestructural which is related to the previous mode and the extended abstract which is related to the next mode. We consider the ikonic, and concrete symbolic modes to be most applicable to middle school students. Following this model, Mooney (2002) concluded that students in his study exhibited four levels of statistical thinking: *idiosyncratic*, (associated with the prestructural level and representing thinking in the ikonic mode), and *transitional*, *quantitative* and *analytical* (associated respectively with the unistructural, multistructural and relational levels; representing thinking in the concrete symbolic mode).

Descriptors

For each statistical process in the framework, the four levels of students' thinking are characterized by descriptors pertaining to each of the subprocesses. Descriptors were developed with the assumption that each subsequent descriptor subsumes the characteristics of the previous levels. For example, the process describing data includes four descriptors for the subprocesses *identifying units of data values*, one for each level of thinking: Level 1, misinterprets or does not identify units of data values; Level 2, identifies the units of data values incompletely; Level 3, identifies the units of specific data values; and Level 4, identifies the units of general data values. It is our intention that teachers will be able to use these descriptors as benchmarks to assess or characterize a student's thinking over a series of tasks. Thus, descriptors are not task specific and we expect variability in the level of thinking exhibited by a student's response across tasks.

OVERVIEW OF THE DEVELOPMENT OF FRAMEWORK DESCRIPTORS

The studies conducted by Mooney, (2002) and Langrall et al. (2001) used the same method to generate and refine framework descriptors. This method comprised six components: (a) developing initial descriptors for the subprocesses based on research; (b) constructing a protocol to assess students' thinking of the subprocesses; (c) interviewing students using the protocol; (d) analyzing students' responses; (e) refining the initial descriptors; and (f) consolidating the refined descriptors into the framework.

Both studies used students in grades six through eight (ages 12 to 14) from a Midwestern U.S. school. In each case, twelve students, four from each grade level, were selected for case-study analysis. Using an interview protocol, each student was individually interviewed during a 60-minute, audio-taped session. In generating protocols for each study, tasks were developed with a series of questions designed to assess students' thinking across the four levels of the framework. Questions were designed so students could respond orally or by generating tables or data displays. In most cases, a task assessed more than one statistical process. In both studies, data sources consisted of the transcribed interviews, students' written work and data displays, researcher field notes, and summaries generated during the analysis. A double-coding procedure (Miles & Huberman, 1994) was used to analyze students' responses. The authors independently coded each student's response to each question. Responses were coded by levels based on: (a) the initial descriptors for each level and (b) descriptors generated from the data analysis that characterized students' responses, yet were not present in the initial descriptors. This occurred in the following manner. After all students' responses to a question were read, the authors compared the responses to the corresponding descriptors to describe the levels of students' statistical thinking. If descriptors did not adequately characterize students' responses, the responses were examined as a whole to discern patterns of thinking. These patterns were used to revise the corresponding descriptors and the students' responses were then recoded using the revised descriptors to characterize students' levels of statistical thinking. If few or no students demonstrated thinking at a particular level of a subprocess, we interpolated the descriptor for that level based on students' thinking at other levels. Throughout this process, differences in coding were discussed and agreement was negotiated.

FRAMEWORK MODIFICATIONS

In modifying the M3ST framework, descriptors were added, altered, moved, or removed from the framework to clarify the meaning of each process and to promote consistency in language. These changes are reported below for each statistical process. The current version of the M3ST framework is presented in Figures 1 through 4.

Describing data. The subprocess *recognizing the same data in different data displays* was removed from the framework because we do not believe it reflects a general characteristic of describing data. Rather, it is task specific in the sense that the situation would have to warrant examining multiple displays of the same data. The subprocess *evaluating the effectiveness of data displays in representing data* was moved since it was more appropriately related to the process representing data. These revisions have resulted in the following subprocesses for describing data: *showing awareness for display features* (D.1) and *identifying units of data values* (D.2).

Describing Data In general, student performing or explaining questions, tasks or activities which involve describing data will . . .		
	Level 1 - Idiosyncratic **D.1.1** Demonstrate little awareness of display features of a table, chart or graphical representation. **D.2.1** Misinterpret or not identify units of data values.	**Level 2 - Transitional** **D.1.2** Demonstrate some awareness of display features a table, chart or graphical representation. **D.2.2** Identify the units of data values incompletely.
	Level 3 - Quantitative **D.1.3** Demonstrate complete awareness of display features a table, chart or graphical representation, **D.2.3** Identify the units of specific data values.	**Level 4 - Analytical** **D.1.4** Demonstrate complete awareness of display features a table, chart or graphical representation including features that are irrelevant or cosmetic. **D.2.4** Identify the units of general data values.

Figure 1. Descriptors for Describing Data.

Organizing and reducing data. Descriptors related to the subprocess *ordering data* were removed because the ways students might order data do not vary enough to distinguish different levels of statistical thinking. Based on the Langrall et al. (2001) study, descriptors for *grouping data* were changed to distinguish summative forms of grouping instead of representative forms of grouping. As a result of these revision, the subprocesses for Organizing and Reducing Data are: *grouping data* (O.1), *describing data using measures of center* (O.2), and *describing the spread of data* (O.3).

Organizing and Reducing Data In general, students performing or explaining questions, tasks, or activities which involve organizing and reducing data will . . .		
	Level 1 - Idiosyncratic **O.1.1** Not attempt to group data. **O.2.1** Not be able to describe data in terms of representativeness or typicalness. **O.3.1** Not be able to describe the spread of the data in terms representative of the spread.	**Level 2 - Transitional** **O.1.2** Group data but not in a summative form. **O.2.2** Describe the typicalness of data using invented measures that are partially valid. **O.3.2** Describe the spread of the data using invented measures that are partially valid.
	Level 3 - Quantitative **O.1.3** Group data in a summative form or group data by creating new categories or clusters. **O.2.3** Describe the typicalness of data using a measure of center from a flawed procedure or a valid and correct invented measure. **O.3.3** Describe spread of data using a measure from a flawed procedure or a valid and correct invented measure.	**Level 4 - Analytical** **O.1.4** Group data in a summative form by creating new categories or clusters. **O.2.4** Describe data using a valid and correct measure of center. **O.3.4** Describe spread of data using valid and correct measure.

Figure 2. Descriptors for Organizing and Reducing Data.

Representing data. The subprocess *evaluating the effectiveness of data displays in representing data* was moved from describing data to this statistical process since we believe this subprocess should be part of the decision-making strategy involved in determining the optimal choice of representation for a given set of data. The subprocess *constructing an alternate data display* was removed because we believe it is subsumed in the subprocess pertaining to evaluating data displays. The descriptors related to the subprocess *completing a partial data display*, were removed because we believed this subprocess was task specific in that it could only be assessed for a given rather than across a variety of tasks. These revisions have resulted in the following subprocesses for Representing Data: *constructing a data display* (R.1), and *evaluating the effectiveness of data displays* (R.2).

Representing Data In general, student students performing or explaining questions, tasks or activities which involve representing data will . . .	**Level 1 - Idiosyncratic** **R.1.1** Be unable to construct a display or constructs a display for that is both incomplete and unrepresentative of the data. **R.2.1** Evaluate the effectiveness of data display based on irrelevant features or reasons.	**Level 2 - Transitional** **R.1.2** Construct a display that is partially complete and representative of the data or complete and unrepresentative of the data. **R.2.2** Evaluate the effectiveness of data display based on relevant display features.
	Level 3 - Quantitative **R.1.3** Construct a complete and representative display. The display may have a few minor flaws. **R.2.3** Evaluate the effectiveness of a data display based on relevant display features with some reference to the context the data is presented.	**Level 4 - Analytical** **R.1.4** Construct a complete, representative and appropriate display. **R.2.4** Evaluate the effectiveness of a data display based on relevant display features and the context the data is presented.

Figure 3. Descriptors for Representing Data.

Analyzing and interpreting data. Descriptors for the subprocess *proportional reasoning* were added to the framework based on the results of the Langrall et al. (2001) study. The descriptors are based on the degree to which students reasonably use, and are able to quantify, relative thinking when analyzing data. Thus, the subprocesses for this process include: *making comparisons within data sets or data displays* (A.1), *making comparisons between data sets or data displays* (A.2), *making inferences from a given data set or data display* (A.3), and *using proportional reasoning* (A.4).

Analyzing and Interpreting Data In general, student performing or explaining questions, tasks or activities which involve analyzing and interpreting data will . . .	**Level 1 - Idiosyncratic** **A.1.1** Make no or incorrect comparisons within data displays or data sets. **A.2.1** Make no or incorrect comparisons between data displays or data sets. **A.3.1** Make inferences that are not based on the data or inferences are based on irrelevant issues. **A.4.1** Not use relative thinking.	**Level 2 - Transitional** **A.1.2** Make a single correct comparison or a set of partially correct comparisons within or between data displays or data sets. **A.2.2** Make a single correct comparison or a set of partially correct comparisons between data displays or data sets. **A.3.2** Make inferences that are partially based on the data. Some inferences may be only partially reasonable. **A.4.2** Use relative thinking qualitatively.
	Level 3 - Quantitative **A.1.3** Make local or global comparisons within data displays or data sets. **A.2.3** Make local or global comparisons between data displays or data sets. **A.3.3** Make inferences that are primarily based on the data. Some inferences may be only partially reasonable. **A.4.3** Uses relative thinking quantitatively but not in a reasonable manner.	**Level 4 - Analytical** **A.1.4** Make local and global comparisons within data displays or data sets. **A.2.4** Make local and global comparisons between data displays or data sets. **A.3.4** Make reasonable inferences based on data and the context. **A.4.4** Uses relative thinking quantitatively in a reasonable manner.

Figure 4. Descriptors for Analyzing and Interpreting Data.

DISCUSSION

The M3ST framework was developed to provide a coherent picture of middle school student's thinking across four statistical processes. We believe that frameworks such as this one can play an important role in informing classroom instruction as well as guiding curriculum design. Teachers and curriculum developers can use the M3ST framework to construct tasks that are within the scope of students' statistical thinking yet able to promote the development of data handling processes called for in mathematics education reform documents. The modifications and refinements we have made to the M3ST framework were intended to make the framework more accessible to teachers. This refinement process was part of an extended research program that includes a large-scale validation of the framework followed by the implementation of a professional development program using the framework with middle teachers to guide instruction in statistics.

REFERENCES

Australian Education Council. (1994). *Mathematics: A curriculum profile for Australian schools.* Carlton, Victoria: Curriculum Corporation.

Berg, C.A., & Phillips, D.G. (1994). An investigation of the relationship between logical thinking structures and the ability to construct and interpret line graphs. *Journal of Research in Science Teaching, 31,* 323-344.

Biggs, J.B. & Collis, K.F. (1991). Multimodal learning and quality of intelligent behavior. In H.A.H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-66). Hillsdale, NJ: Erlbaum.

Cobb, P. et al. (1991). Assessment of a problem-centered second-grade mathematics project. *Journal for Research in Mathematics Education, 22*, 3-29.

Curcio, F.R. (1987). Comprehension of mathematical relationships expressed in graphs. *Journal for Research in Mathematics Education, 18,* 382-393.

Fennema, E., & Franke, M.L. (1992). Teacher's knowledge and its impact. In D.A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 147-164). New York, NY: Macmillan.

Friel, S.N., & Bright, G.W. (1996). *Building a theory of graphicacy: How do students read graphs?* Paper presented at the annual meeting of the American Educational Research Association, New York, NY. (ERIC Document Reproduction Service No. ED 395 277).

Friel, S.N., Curcio, F.R. & Bright, G.W. (2001). Making sense of graphs: Critical factors influencing comprehension and instructional implications. *Journal for Research in Mathematics Education, 32*, 124 – 158.

Langrall, C.W., Mooney, E.S., Hofbauer, P.S., & Johnson, Y.A. (2001). Refining a framework on middle school students' statistical thinking. In R. Speiser, C.A. Maher and C.N. Walter (Eds), *Proceedings of the twenty-third annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education* (Vol. 1, pp. 437–448). Columbus, OH: ERIC Clearinghouse for Science, Mathematics, and Environmental Education.

Mevarech, Z.A., & Kramarsky, B. (1997). From verbal descriptions to graphic representations: stability and change in students' alternative conceptions. *Educational Studies in Mathematics, 32*, 229-263.

Miles, M.B., & Huberman, A.M. (1994). *Qualitative data*. Thousand Oaks, CA: Sage.

Mooney, E.S. (2002). Development of a middle school statistical thinking framework. Submitted for publication, *Mathematical Thinking and Learning, 4*(1).

National Council of Teachers of Mathematics (2000). *Principles and standards for school mathematics.* Reston, VA: Author.

Reading, C., & Pegg, J. (1996). Exploring understanding of data reduction. In L. Puig and A. Gutiérrez (Eds.), *Proceedings of the 20th Conference of the International Group for the Psychology of Mathematics Education: Volume 4* (pp.187-194). Spain: Universitat de Valencia.

Reber, A.S. (1995). *Dictionary of psychology*. London: Penguin Books Limited.

School Curriculum and Assessment Authority for Wales. (1996). *A guide to the national curriculum*. London: Author.

Secretary's Commission on Achieving Necessary Skills (1991). *What work requires of schools: A SCANS report for America 2000*. Washington, DC: Department of Labor.

Shaughnessy, J.M., Garfield, J., & Greer, B. (1996). Data handling. In A.J. Bishop, K. Clements, C. Keitel, J. Kilpatrick, and C. Laborde (Eds.), *International handbook of mathematics education* (pp. 205 – 237). Dordrecht, The Netherlands: Kluwer Academic Publishers.