STUDENTS' INDIVIDUAL AND COLLECTIVE STATISTICAL THINKING ®

Graham A. Jones, Edward S. Mooney, Cynthia W. Langrall and Carol A. Thornton,
Illinois State University
USA

Our paper describes a suite of studies involving students' statistical thinking in Grades 1 through 8. In our key studies (Jones et al., 2000, Mooney, in press), we validated Frameworks that characterised students' thinking on four processes: describing, organizing, representing, and analyzing *and interpreting data. These studies showed that the students' thinking was consistent with the four cognitive levels postulated in a general developmental model. We also report on two teaching experiments, with primary students (Jones et al., 2001; Wares et al., 2000) that used the Framework to inform instruction. Teaching experiment results showed that children produced fewer idiosyncratic descriptions of data, possessed intuitive knowledge of center and spread and were constrained in analysis and interpretation by knowledge of data context.*

OVERVIEW

In response to the critical role that data plays in our technological world, there have been widespread calls for reform in statistical education at all grade levels (e.g., National Council of Teachers of Mathematics, 2000; Australian Education Council, 1994). These reforms have advocated a more pervasive approach to data exploration, one that includes describing, organizing, representing and interpreting data. This expanded perspective has created the need for further research on the learning and teaching of statistics, especially in the elementary and middle grades, where instruction has tended to focus on graphing rather than data exploration (Shaughnessy, Garfield, & Greer, 1996).

In response to these calls for research, there have been an increasing number of studies on elementary and middle school students' individual statistical thinking (Curcio, 1987; Gal & Garfield, 1997; Strauss & Bichler, 1988; Mokros & Russell, 1995; Watson & Moritz, 2000), but relatively little research on students' collective thinking during instruction (Ben-Zvi, 2000; Cobb, 1999; Lehrer & Schauble, 2000). Existing research on students' statistical thinking has certainly not developed the kind of cognitive models of students' statistical thinking that researchers like Fennema et al. (1996) deem necessary to guide the design and implementation of instruction.

In this paper we will discuss how our research has developed and used cognitive frameworks to address these instructional issues. More specifically, the paper will: (a) discuss the formulation and validation of two related frameworks, one for elementary and one for middle school, that characterize students' statistical thinking; and (b) describe teaching experiments with Grades 1 and 2 children that were informed by the framework.

STATISITICAL THINKING FRAMEWORKS

In generating the frameworks, we identified four key statistical processes: describing data, organizing and reducing data, representing data, and analyzing and interpreting data. These processes which will be described below were modifications of similar processes identified by Shaughnessy et al. (1996). Based on our earlier work with *number sense* (Jones, Thornton, & Putt, 1994) and *probability* (Jones, Langrall, Thornton, and Mogill, 1997), the frameworks were formulated on the assumption that elementary and middle school students would exhibit four levels of statistical thinking in accord with Biggs and Collis's (1991) general development model. These levels of statistical thinking were described as idiosyncratic, transitional, quantitative and analytical, and in subsequent validation studies we confirmed the existence of these four levels and refined the descriptors of students' thinking in the frameworks (Jones et al., 2000; Mooney, in press).

Key Processes. The first process, *describing data*, incorporates what Curcio (1987) calls "reading the data." Curcio notes that reading the data means extracting information explicitly stated in the data display, recognizing graphical conventions, and making connections between context and data. Based on Curcio's definition, we generated tasks to assess students' thinking on this process. A sample of a middle school task is shown in Figure 1 (see question (D)). *Organizing and reducing data* incorporates mental actions such as ordering, grouping, and summarizing data (Moore, 1997). As such, it also involves using notions of center and spread. A sample of one of the questions used to assess this process is shown in Figure 1 (see question (O)). Our third process, *representing data*, incorporates constructing visual displays that sometimes require different organizations of data. A sample of a question used to assess this process with middle school students is shown in Figure 1 (see question (R)). The final process *analyzing and interpreting data* involves recognizing patterns and trends in the data and making inferences and predictions from the data. It incorporates what Curcio (1987) refers to as "reading between the data" and "reading beyond the data." The former involves using mathematical operations to combine or conpare data, while the latter requires students to predict from the data by tapping their existing schema for information that is not explicitly stated in the data. Wainer (1992) provides a similar perspective on analysis and interpretation. We used questions like (A) in Figure 1 to assess students' thinking on this process. Research from a number of studies (Beaton, et al., 1996; Curcio, 1987; Friel, Curcio & Bright, 2001; Mokros & Russell, 1995; Padilla, McKenzie & Shaw, 1986; Reading & Pegg, 1996; Watson & Moritz, 2000; Zawojewski & Heckman, 1997) was helpful in designing the questions for assessing the four key processes.

Salaries of 15 Top Actors and Actresses (in millions of dollars)		Questions by Process
Actors	Actresses	(D) What does the table tell you?
$17.5	$12.5	
15.0	9.0	
20.0	11.0	
20.0	9.5	(O) What is the typical salary for the actresses?
20.0	2.5	
19,0	12.0	
20.0	3.0	
18.0	4.0	(R) Construct a graph that will allow you to compare the salaries of actors and actresses. Explain.
5.5	4.0	
6.0	2.5	
10.0	6.0	
16.5	8.5	(A) How do the actors' salaries compare to the actresses salaries?
12.5	4.5	
10.0	3.0	
7.0	10.0	

(D): Describing data; (O): Organizing and reducing data; (R): Representing data;
(A): Analyzing and interpreting data

Figure 1. Sample Middle School Protocol Task.

Thinking Levels. The validation process confirmed the existence of four levels of statistical thinking as postulated on the basis of the Biggs and Collis' (1991) developmental model. *Level 1* thinkers were consistently limited to idiosyncratic reasoning that was often unrelated to the given data and frequently focused on their own personal data banks. *Level 2* thinkers were beginning to recognize the importance of quantitative thinking and even used numbers to invent measures, albeit not always valid, for center and spread. Their perspective on data was generally single-minded and they seldom connected representations or analyses of the data to its context. Students exhibiting *Level 3* thinking consistently used quantitative reasoning as the basis for statistical judgments and had begun to form valid conceptions of center and spread. These students were cognizant of both the context and the data but they seldom made connections between the two. *Level 4* students used a more analytical approach in exploring data and showed evidence of being able to make connections between context and the data. They were able to look at the data both globally and locally; that is, to adopt both a macro and micro view of the data. In Figure 2 we present exemplars of middle school students' responses at each thinking level on the four statistical processes. The questions refer to those in Figure 1.

Question	Level 1	Level 2	Level 3	Level 4
(D)	It's about actors and actresses.	The salaries of actors and actresses. Like the first one made 12.5.	It's about 15 actor' and actresses' salaries. The first actor got 17.5 million.	It's about the salaries of 15 of the top actors and actresses. It lists the salaries in millions of dollars.
(O)	About 8 or 9 dollars.	I'd say 3 or 4 million dollars. (2 of 3 modal values)	About six million. It's in the middle.	6.8 million dollars. I found the average.
(R)	Represents just two data values; the first from each category (See Figure 3).	Represents just the first five values from each category (See Figure 3).	Constructs separate line plots for each data category (See Figure 3).	Integrates the data in a single display that uses ranks (See Figure 3).
(A)	[From the graph]The actor made 20 dollar and the actress made 12 dollars and 50 cents. I'm not exactly sure what this is showing.	The actors are normally higher than the actresses. Look at the top five [from the graph].	[From the graph]That for the actors most of the them earn in the... eight of them earn more than 15 and none of the actresses do.	If you look at the graph, the actors were always ahead of the actresses at each rank. Also, more than half of the actors make more than all of the actresses.

Figure 2. Exemplars of Thinking at each Level of the Framework.

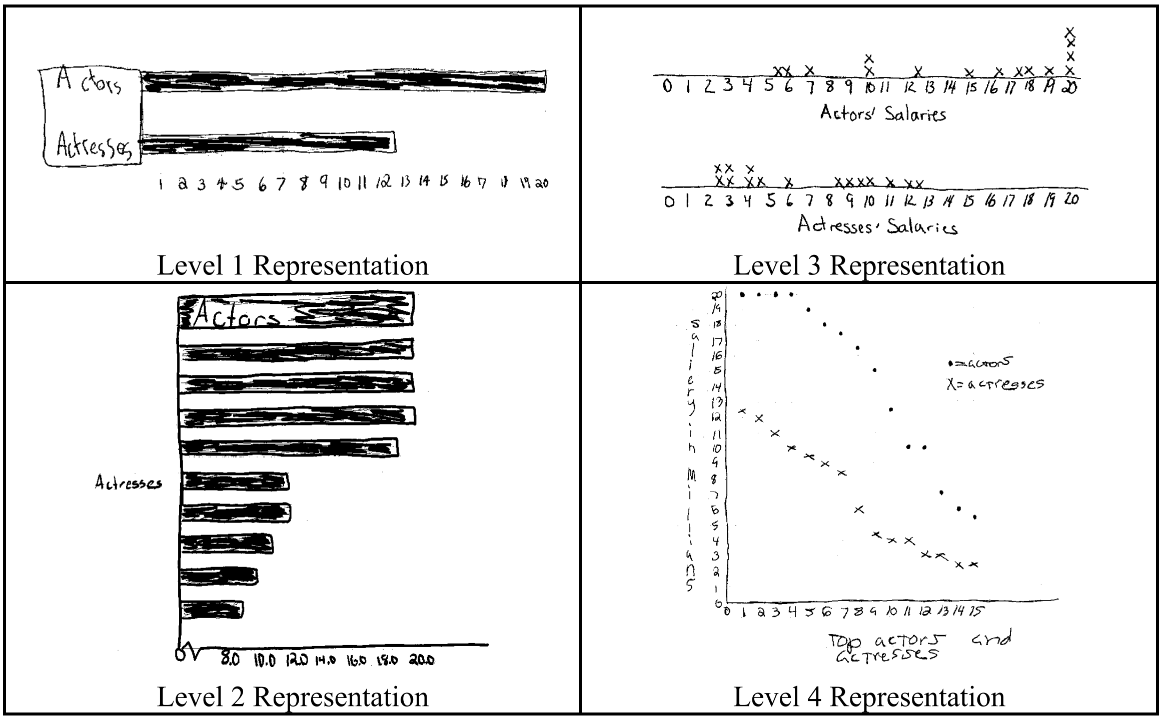

Figure 3. Exemplars of Representing Data

THE TEACHING EXPERIMENTS

A teaching experiment has been defined as a methodology that is aimed at capturing and documenting students' thinking over time (Steffe & Thompson, 2000). During a teaching experiment, researchers develop sequences of instructional activities or learning trajectories (Simon, 1995) and analyze students' individual and collective mathematical learning as it occurs in the social situation of a classroom or a small group (Cobb, 1999). In our teaching experiments, the learning trajectories (goals, tasks, and expected learning outcomes) were based on the elementary Framework, which was also used as a lens to trace changes in students' learning during the intervention.

Our Grade 2 teaching experiment (n=19) comprised 9 sessions each of 40 minutes. In the teaching sessions, the class's Butterfly Garden Project served as the context and provided both categorical and numerical data. The Grade 1 teaching experiment involved two classes and comprised 5 sessions of 40 minutes each. The data exploration tasks for Grade 1 teaching experiment were based on a data set generated from the "number of teeth" lost by the children in one class. The class that collected the data was referred to as the Collection Group (n=20) and the class that merely used the data was referred to as the NonCollection Group (n=18). All children in the Grade 2 class and both Grade 1 classes were assessed prior to and immediately following the teaching experiments using the same protocol that had been used to validate the Framework.

Effects of the teaching experiments: Quantitative Analysis. For the *Grade 2* teaching experiment, a Wilcoxon Signed Ranks Test (Siegel & Castellan, 1988) revealed significant growth between the pre and postintervention thinking levels of the students on each of the four statistical processes: *describing data* (p<.001); *organizing and reducing data* (p<.001); *representing data* (p<.002); and *analyzing and interpreting data* (p<.004). The most salient feature of the data following intervention was the increase in the number of students exhibiting statistical thinking beyond Level 2 and the decrease in the number students exhibiting Level 1. This pattern was most noticeable in the first three processes; in the case of analyzing and interpreting data the decrease in the number of students exhibiting Level 1 was similar but the increase in the number of students exhibiting statistical thinking beyond Level 2 was not as pronounced.

For the *Grade 1* teaching experiment, the Wilcoxon Signed Ranks Test revealed differences between pre and postintervention thinking levels that were significant for some statistical processes and not for others. For *describing data*, only the NonCollection Group showed a significant difference (Collection Group, p< .08; NonCollection Group, p< .01); for *organizing and reducing data*, both Groups showed significant differences (Collection Group, p< .04; NonCollection Group, p< .01); for *representing data*, neither group showed significant differences (Collection Group, p<.17; NonCollection Group, p< .42); and for *analyzing and interpreting data* only the Collection Group showed a significant difference (Collection Group, p< .01; NonCollection Group, p< .65). While the statistical thinking of the children in the two groups changed in slightly different ways, the evidence does not support a stronger overall growth in favor of the Collection Group. When the two classes were combined the differences between children's pre and postintervention thinking levels were significant for all statistical processes except representing data. For the three significant statistical processes, the most salient feature of the data was that the number of children exhibiting Level 3 increased following the intervention and this was accompanied by a decrease in the number exhibiting Level 1 thinking.

Effects of the teaching experiments: Qualitative Analysis. Several learning patterns emerged from the analysis of instruction and in particular the detailed case-study analysis of 4 target students in each of the Grades 1 and 2 classes. These learning patterns are described by statistical process. With respect to *describing data*, children brought varying degrees of prior knowledge about meanings and conventions associated with contextual data displays. Experiences with different kinds of data during instruction seemed to focus their thinking and produced less idiosyncratic descriptions. Categorical data was more troublesome for these children than numerical data. Children's intuitive thinking with respect to *organizing and reducing data* was problematic. Although they were reluctant to use paper and pencil to reorganize data (especially categorical data), technology proved very helpful in stimulating their organizing strategies. Our results also show that collectively children revealed conceptual knowledge of center and spread that was multifaceted (Watson & Moritz, 2000) and useful in informing instruction. The difficulty for the teacher lay in deciding how and when to use children's different representations of center (e.g. median or mode). Children's prior knowledge in *representing data* appears to be constrained by limited accessibility to pervasive sorting and organizing schemas. However, instruction that incorporated technology or the use of unfinished graphs showed potential in stimulating children's sorting schema and ipso facto their capability for constructing representations. With respect to *analyzing and interpreting data*, children's thinking, prior to the intervention, was more normative on tasks that involved reading between the data than on tasks that involved reading beyond the data. The intervention revealed some unanticipated problems with tasks that focused on reading between the data, especially those that involved identifying and comparing two subsets of the data. Our analysis also highlighted the importance of children's knowledge of the data context in relation to tasks that involved reading beyond the data.

CONCLUSIONS

Given the prior knowledge and growth that elementary and middle school students demonstrated on all four statistical processes, there is evidence that they can accommodate a broader approach to data exploration. However, if instruction on data exploration is to reach its full potential in the elementary and middle grades, there is a need for further research to build learning trajectories that link the different levels of children's statistical thinking identified in the Framework.

REFERENCES

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

Beaton, A.E., Mullis, I.V.S., Martine, M.O., Gonzalez, E.J., Kelly, D.L., & Smith, T.A. (1996). *Mathematics achievement in the middle school years: IEA'S third international mathematics and science study (TIMSS)*. Chestnut Hill, MA: Boston College.

Ben-Zvi, D. (2000). Toward understanding the role of technological tools in statistical learning. *Mathematical Thinking and Learning, 2*, 127-155.

Biggs, J.B., & Collis, K.F. (1991). Multimodal learning and intelligent behavior. In H. Rowe (Ed.), *Intelligence: Reconceptualization and measurement* (pp. 57-76). Hillsdale, NJ: Erlbaum.

Cobb, P. (1999). Individual and collective mathematical development: The case of statistical analysis. *Mathematical Thinking and Learning, 1*, 5-43.

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

Fennema, E., Carpenter, T.P., Franke, M.L., Levi, L., Jacobs, V.R., & Empson, S.B. (1996). A longitudinal study of research to use children's thinking in mathematics instruction. *Journal for Research in Mathematics Education, 27*, 403-434.

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.

Gal, I. & Garfield, J.B. (1997). *The assessment challenge in statistics education*. Amsterdam, The Netherlands: IOS Press.

Jones, G.A., Thornton, C.A., & Putt, I.J. (1994). A model for nurturing and assessing multidigit number sense among first grade children. *Educational Studies in Mathematics, 27*, 117-143.

Jones, G.A., Langrall, C.W., Thornton, CA., & Mogill, A.T. (1997). A framework for assessing and nurturing young children's thinking in probability. *Educational Studies in Mathematics, 32*, 101-125.

Jones, G.A., Thornton, C.A., Langrall, C.W., Mooney, E., Perry, B., & Putt, I. (2000). A framework for characterizing students' statistical thinking. *Mathematical Thinking and Learning, 2*, 269-308.

Jones, G.A., Langrall, C.W., Thornton, C.A., Mooney, E.S., Wares, A.S., Jones, M.R., Perry, B., Putt, I.J., & Nisbet, S. (2001). Using students' statistical thinking to inform instruction. *Journal of Mathematical Behavior, 20, 109-144*.

Lehrer, R., & Schauble, L. (2000). Inventing data structures for representational purposes: Elementary grade students' classification models. *Mathematical Thinking and Learning, 2*, 51-74.

Mokros, J., & Russell, S.J. (1995). Children's concepts of average and representativeness.

Journal for Research in Mathematics Education, 26, 20-39.
Mooney, E.S. (in press). A framework for characterizing middle school students' statistical thinking. *Mathematical Thinking and Learning*.
Moore, D.S. (1997). *Statistics: Concepts and controversies* (4th ed.). New York: Freeman.
National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: Author.
Padilla, M.J., Mckenzie, D.L., & Shaw, E.L. (1986). An examination of the line graphing ability of students in grades seven through twelve. *School Science and Mathematics, 86*, 20-26.
Reading, C., & Pegg, J. (1996). Exploring understanding of data reduction. In L. Puig and A. Gutiertez (Eds.), *Proceeding of the 20th Conference of the International Group for the Psychology of Mathematics Education* (Vol. 4, pp. 187-194). Spain: Universitat de Valencia.
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* (Part 1, pp. 205-238). Dordrecht, The Netherlands: Kluwer.
Siegel, S., & Castellan, N.J. (1988). *Nonparametric statistics for the behavioral sciences* (2nd edn.). New York. McGraw-Hill.
Simon, M. (1995). Reconstructing mathematics pedagogy from a constructivist perspective. *Journal for Research in Mathematics Education , 26,* 114-145.
Steffe, L.P., & Thompson, P.W. (2000). Teaching experiment methodology: Underlying principles and essential elements. In A.E. Kelly, and R.A. Lesh (Eds.), *Handbook of research design in mathematics and science education* (pp. 267-306). Mahwah, NJ: Laurence Erlbaum.
Strauss, S., & Bichler, E. (1988). The development of children's concepts of the arithmetic average. *Journal for Research in Mathematics Education, 19*, 64-80.
Wainer, H. (1992). Understanding graphs and tables. *Educational Researcher, 21* (1), 14-23.
Wares, A., Jones, G. A., Langrall, C. W., & Thornton, C. A. (2000). Young children's statistical thinking. A teaching experiment. In M.L. Fernandez (Ed.), *Proceedings of the 22nd Annual Meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education*, (Vol. 1, pp. 321-327). Columbus, OH: ERIC Center for Science. Mathematics, and Technology Education.
Watson, J.M., & Moritz, J.B. (2000). The longitudinal development of the understanding of average. *Mathematical Thinking and Learning, 2*, 11-50.
Zawojewski, J.S., & Heckman, D.S. (1997). What do students know about data analysis, statistics, and probability? In P.A. Kenney and E.A. Silver (Eds.), *Results from the sixth mathematics assessment of the national assessment of educational progress* (pp. 195-223). Reston, VA: National Council of Teachers of Mathematics.