Statistical Reasoning in the Middle School

2013 NCTM Annual Meeting & Exposition – Denver

Raymond Johnson [1] Susan Thomas [2]

[1]University of Colorado Boulder
Freudenthal Institute US
raymond.johnson@colorado.edu
http://mathed.net

[2]University of Colorado Boulder
susan.r.thomas@colorado.edu

April 18, 2013

University of Colorado Boulder

Outline

About Us

Raymond:

- PhD student in Curriculum & Instruction, Mathematics Education
- Learned to teach statistics on the job as a reaction to standards
- Instructor of *Basic Statistical Methods*

Susan:

- PhD student in Research and Evaluation Methodology
- Undergraduate degrees in mathematics and statistics

Together:

- Research on middle school teachers' perceptions of statistics in the CCSSM
  - A need for content knowledge (common, content, horizon) (Ball, Thames, & Phelps, 2008)
  - A need for curriculum and tasks

Why More Statistics?

Demand in standards

- NCTM 1989, 2000
- GAISE Report, CCSSM

Evolution of the discipline

- A new and rapidly evolving field
- Are you older than a box plot?

Big data and little data

- Google's Eric Schmidt: "There was 5 exabytes of information created between the dawn of civilization through 2003, but that much information is now created every 2 days, and the pace is increasing."(Kilpatrick, 2010)
- The "Quantified Self"

Stats vs. Math

"Statistics is a science in my opinion, and it is no more a branch of mathematics than are physics, chemistry and economics; for if its methods fail the test of experience – not the test of logic – they are discarded."

– John Tukey (1962, pp. 6-7)

"The twin sister of the 'certainty' in mathematics is the 'uncertainty' in statistics. We must prepare our students to deal with both types of quantitative reasoning as they grow in the mathematical sciences."

– Michael Shaughnessy (2010)

Statistical Thinking vs. Reasoning (DelMas, 2004)

Statistical *Thinking*: Knowing when and how to apply statistical knowledge and procedures

Statistical *Reasoning*: Explaining why results were produced or why a conclusion is justified

Examples of statistical reasoning:

- Stating implications
- Justifying conclusions
- Making inferences

Grade 6: A focus on variability and distribution

6.SP.A.1 Recognize a statistical question as one that anticipates variability...

6.SP.A.2 ...has a distribution which can be described by its center, spread, and overall shape.

6.SP.A.3 ...while a measure of variation describes how its values vary with a single number.

6.SP.B.5c Giving quantitative measures of ... variability(interquartile range and/or mean absolute deviation)...

6.SP.B.5d Relating the choice of measures of center and variability to the shape of the data distribution...

Research on Student Reasoning About Variability

- "...practically no research on students' conceptions of variability was reported prior to 1999" (Shaughnessy, 2007, pp. 972)
- "An underlying problem is that middle-grade students generally do not see 'five feet' as a value of the variable 'height,' but as a personal characteristic of, say, Katie." (Bakker & Gravemeijer, 2004, pp. 147-148)

Table 1. Between data and distribution

distribution (conceptual entity)			
center mean, median, midrange, …	**spread** range, standard deviation, inter-quartile range, …	**density** (relative) frequency, majority, quartiles	**skewness** position majority of data
data (individual values)			

(Bakker & Gravemeijer, 2004, p. 148)

Horizon Content Knowledge for Variability

- Standard Deviation (HSS-ID.A.2)
- Margin of Error (HSS-ID.B.4)
- Compare Two Treatments (HSS-ID.B.5)
- Evaluate reports (HSS-ID.B.6, everyday applications)
- Analysis of Variance/Multiple Comparisons (college statistics)

Selected Tasks

- Mean, Median, Mode, and Range
  MARS - `http://map.mathshell.org/materials/lessons.php?taskid=486`
- College Athletes
  Illustrative Mathematics - `http://www.illustrativemathematics.org/illustrations/1340`
- How Long Are Our Shoes?
  Bridging the Gap Between Common Core State Standards and Teaching Statistics (Investigation 3.4, pp. 98-110)

Discussion About Variability

- What kind of student thinking would you expect to see on this task?
- How might the task elicit reasoning about variability?
- Where in a sequence of tasks or lessons would you place this?

Grade 7: A focus on sampling and inference

7.SP.A.1 Understand that statistics can be used to gain information about a population by examining a sample of the population; generalizations about a population from a sample are valid only if the sample is representative of that population. Understand that random sampling tends to produce representative samples and support valid inferences.

7.SP.A.2 Use data from a random sample to draw inferences about a population...

7.SP.B.4 Use measures of center and measures of variability for numerical data from random samples to draw informal comparative inferences about two populations.

Research on Student Reasoning About Sampling and Inference

- "Over-reliance on sample representativeness is likely to lead to the notion that a sample tells us *everything* about a population; over-reliance on sample variability implies that a sample tells us *nothing*." (Rubin, Bruce, & Tenney, 1991, p. 315)
- Higher-performing students "developed a multi-tiered scheme of conceptual operations centered around the images of repeatedly sampling from a population, recording a statistic, and tracking the accumulation of statistics as they distribute themselves along a range of possibilities." (Saldanha & Thompson, 2003, p. 261)

Horizon Content Knowledge for Sampling and Inference

- Understand statistics as a process for making inferences about population parameters based on a random sample from that population (HSS-IC.A.1)
- Randomization related to sample surveys, experiments, and observatonial studies (HSS-IC.B.3)
- Sampling Distributions and Central Limit Theorem (college level statistics)

Selected Tasks

- What's Your Favorite Subject?
  http://www.illustrativemathematics.org/illustrations/973
- Counting Trees
  http://map.mathshell.org/materials/tasks.php?taskid=386&subpage=expert
- Candy Bars
  http://map.mathshell.org/materials/tasks.php?taskid=396&subpage=expert

Discussion About Sampling and Inference

- What kind of student thinking would you expect to see on this task?
- How might the task elicit reasoning about sampling and inference?
- Where in a sequence of tasks or lessons would you place this?

Grade 8: A focus on covariation

8.SP.A.1 Construct and interpret scatter plots for bivariate measurement data to investigate patterns of association between two quantities. Describe patterns such as clustering, outliers, positive or negative association, linear association, and nonlinear association.

8.SP.A.2 Know that straight lines are widely used to model relationships between two quantitative variables. For scatter plots that suggest a linear association...

8.SP.A.3 Use the equation of a linear model to solve problems in the context of bivariate measurement data, interpreting the slope and intercept.

Research on Student Reasoning About Covariation

- "It seems unwise, for example, to specify ... that by middle school, students will learn how to 'make conjectures about possible relationships' between two characteristics of a sample on the basis of scatterplots' (NCTM 2000, p. 248)." (Konold, 2002, p. 5)

Table 1. Characteristics of four levels of speculative data generation

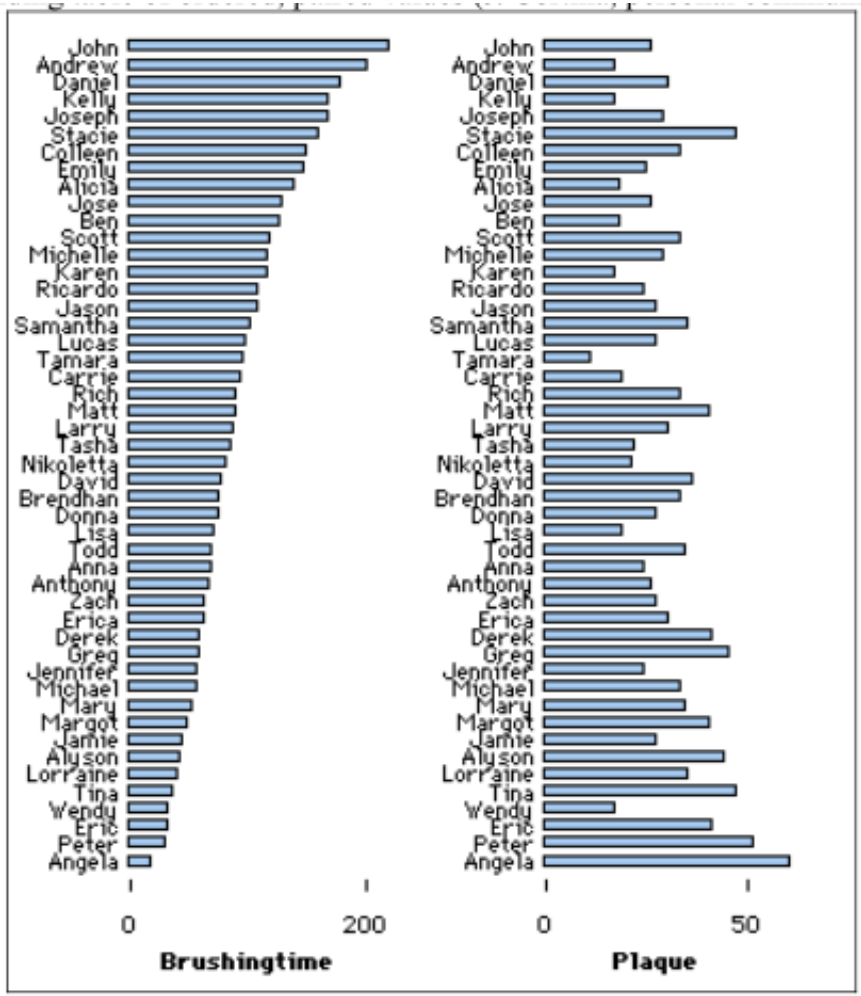
Level	Description
0. Nonstatistical	Responses represent either: (a) context in a narrative but without a data set of more than one value of one variable, or (b) graph axes or values, denoted by number or spatial position, but without a context indicating a data variable
1. Single Aspect	Responses represent either: (a) correspondence in a single bivariate case, or (b) variation of values for a single variable
2. Inadequate Covariation	Responses represent both variables but either: (a) correspondence is shown with inappropriate variation for at least one variable, such as one variable only has two distinct values (often categorical), or (b) variation is shown for each variable with inappropriate correspondence, such as not in the correct direction
3. Appropriate Covariation	Responses represent both variables with appropriate correspondence between the variation of values for each variable

(Moritz, 2005, p. 239)

Example 1 - Bivariate Table

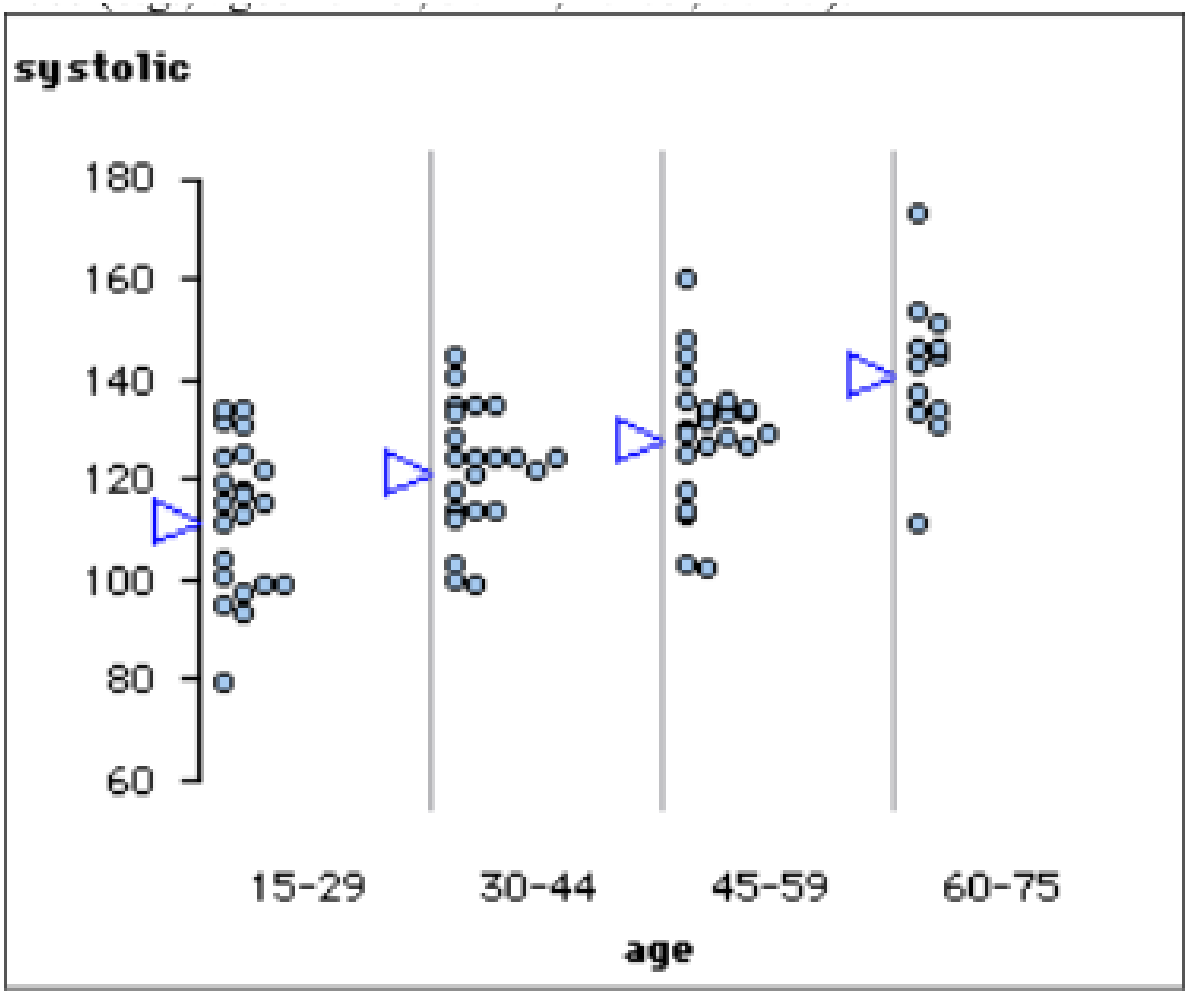
(Moritz, 2005, p. 244)

Example 2 - Paired Case-Value Plots



(Konold, 2002, p. 3)

Example 3 - Scatterplot Slices



(Konold, 2002, p. 2)

Horizon Content Knowledge for Covariation

- Two-way frequency tables and relative frequencies (HSS-ID.B.5)
- Fit a function (linear, quadratic, exponential) to the data (HSS-ID.B.6)
- Interpret linear models (slope, intercept) in the context of the data (HSS-ID.C.7)
- ANCOVA, Propensity Score Matching, Regression Discontinuity, Logistic and Nonparametric Regression (college statistics)

Selected Tasks

- Hand Span and Height
  `http://www.illustrativemathematics.org/illustrations/1097`
- US Airports
  `http://www.illustrativemathematics.org/illustrations/1370`
- Scatter Diagram
  `http://map.mathshell.org/materials/tasks.php?taskid=381`

Discussion About Covariation

- What kind of student thinking would you expect to see on this task?
- How might the task elicit reasoning about covariation?
- Where in a sequence of tasks or lessons would you place this?

Resources

- *Bridging the Gap Between Common Core State Standards and Teaching Statistics*: `http://www.amstat.org/education/btg/` & `http://www.nctm.org/catalog/product.aspx?id=14444`
- MARS: `http://map.mathshell.org/materials/index.php`
- Illustrative Mathematics: `http://www.illustrativemathematics.org/`
- AIMS Project: `http://www.tc.umn.edu/~aims/index.htm`

Contact

Raymond Johnson

- raymond.johnson@colorado.edu
- `http://mathed.net`
- @MathEdnet, +RaymondJohnson

Susan Thomas

- susan.r.thomas@colorado.edu

References I

Bakker, A., & Gravemeijer, K. (2004). Learning to reason about distribution. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 147–168). New York, NY: Kluwer.

Ball, D. L., Thames, M. H., & Phelps, G. (2008). Content knowledge for teaching: What makes it special? *Journal of Teacher Education*, *59*(5), 389–407. Retrieved from `http://jte.sagepub.com/content/59/5/389` doi: doi: 10.1177/0022487108324554

DelMas, R. C. (2004). A comparison of mathematical and statistical reasoning. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 79–95). New York, NY: Kluwer. doi: doi: 10.1007/1-4020-2278-6_4

References II

Kilpatrick, M. (2010, August). *Google CEO Schmidt: "People aren't ready for the technology revolution".* Retrieved from `http://readwrite.com/2010/08/04/google_ceo_schmidt_people_arent_ready_for_the_tech`

Konold, C. (2002). Alternatives to scatterplots. In *Proceedings of the sixth international conference on teaching statistics* (pp. 1–6). Cape Town, South Africa: International Association for Statistical Education. Retrieved from `http://www.stat.auckland.ac.nz/~iase/publications/1/7f5_kono.pdf`

Moritz, J. (2005). Reasoning about covariation. In D. Ben-Zvi & J. Garfield (Eds.), *The challenge of developing statistical literacy, reasoning and thinking* (pp. 227–255). New York, NY: Kluwer. doi: doi: 10.1007/1-4020-2278-6_10

References III

Rubin, A., Bruce, B., & Tenney, Y. (1991). Learning about sampling: Trouble at the core of statistics. In D. Vere-Jones (Ed.), *Proceedings of the third international conference on teaching statistics* (pp. 314–319). Voorberg, The Netherlands: International Statistics Institute.

Saldanha, L., & Thompson, P. (2003). Conceptions of sample and their relationship to statistical inference. *Educational Studies in Mathematics*, *51*, 257–270.

Shaughnessy, J. M. (2007). Research on statistics learning and reasoning. In F. K. Lester (Ed.), *Second handbook of research on mathematics teaching and learning: A project of the national council of teachers of mathematics* (pp. 957–1009). Reston, VA: National Council of Teachers of Mathematics.

Shaughnessy, J. M. (2010). *Statistics for all – the flip side of quantitative reasoning.* Retrieved from `http://www.nctm.org/about/content.aspx?id=26327`

References IV

Tukey, J. W. (1962). The future of data analysis. *The Annals of Mathematical Statistics*, *33*(1), 1–67.