MANUFACTURING LICORICE: MODELING WITH DATA IN THIRD GRADE

Lyn D. English
Queensland University of Technology
l.english@qut.edu.au

This paper reports on a study of 3rd-grade students' modeling with data, which involves comprehensive investigations that draw upon STEM-based concepts, contexts, and questions, and generate products supported by evidence and open to informal inferential thinking. Within a real-world STEM-based context of licorice manufacturing, students experienced the "creation of variation" as they compared and represented the masses of "licorice sticks" they made by hand (using Play-Doh) and those using a Play-Doh extruder. By generating their own statistical measures, students could observe the features of data distributions including center, range, typical, and middle, at a much younger age than usual. They could draw inferences from the models they created, with awareness of how variation limits the certainty with which predictions can be made. The study supports a potential route for advancing early statistical learning.

Keywords: Data Analysis and Statistics, Modeling

Introduction

This paper aligns with the first conference theme of potential routes for mathematics education for the future, specifically, pertaining to the statistical capabilities of young learners. Despite research revealing how young elementary students are more competent in dealing with statistical problems than is acknowledged (e.g., Lehrer & Schauble, 2012; Lehrer & English, in press), many curricula continue to delay core statistical experiences until the middle and early secondary school years (e.g., *Common Core State Standards: Mathematics*; http://www.corestandards.org/Math). Yet young children are exposed to a vast array of statistical information that can, at times, misinform, rather than inform their receptive minds. The ability to reason effectively with data, including entertaining uncertainty and risk, is integral to making meaningful, informed decisions across all spectrums of life. One cannot participate effectively in debates about community issues such as the environment, health care, and education, without this reasoning ability (English & Watson, 2015; Franklin, Kader, Mewborn, Moreno, Peck, Perry, & Scheaffer, 2007). Foundational statistical experiences need to begin early. By undertaking their own investigations, elementary school students can learn to make critical decisions with data, where variation and uncertainty are ever present.

This paper reports on third-grade students' modeling with data, which involves comprehensive investigations that draw upon STEM-based concepts, contexts, and questions, and generate products that are supported by evidence and are open to informal inferential thinking (Lehrer & English, in press). In the present study, students experienced the "creation of variation" as they compared the masses of "licorice sticks" they made by hand (using Play-Doh) with those made using a Play-Doh extruder kit ("factory-made"; adapted from Watson, Skalicky, Fitzallen, & Wright, 2009). Students chose their own forms of representation in displaying their models for the two forms of licorice production, and identified, compared, and explained the features of their data distributions.

Modeling with Data

There are various interpretations of modeling and modeling with data, as reported by English, Arleback, and Mousoulides (2016). As defined here, modeling with data includes: (a) an appreciation of how and why investigative questions are posed and refined within a STEM context; (b) competence in generating, selecting, and measuring attributes; (c) skills in organizing, structuring, and representing data; (d) an ability to interpret evidence-based models including features of data

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

distributions; and (e) making informal inferences while acknowledging variation in the data, and the uncertainty with which any conclusions can be drawn (cf. Makar, Bakker, & Ben-Zvi, 2011; Lehrer & English, in press). Figure 1 displays these core features of modeling with data. Consideration is given to a selection of these modeling components.

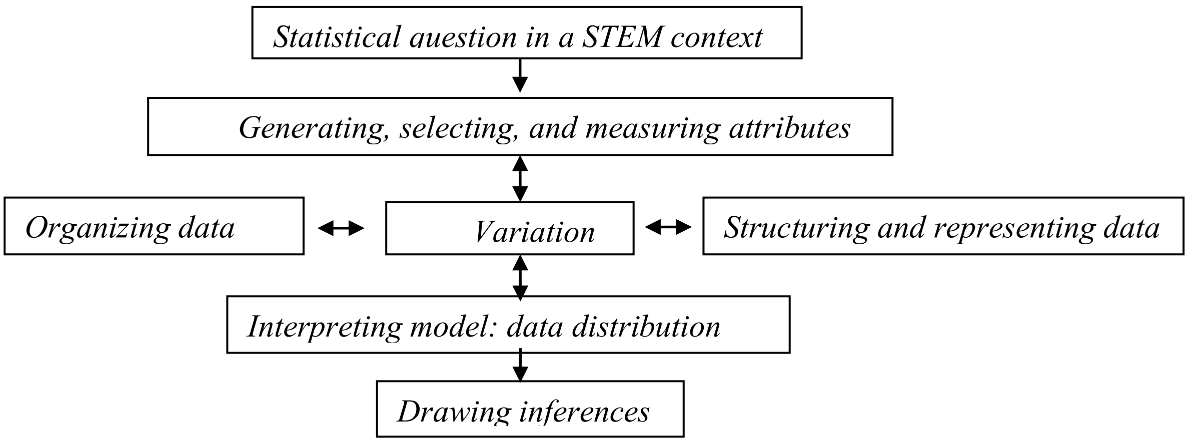

Figure 1. Modeling with data.

STEM Contexts

A statistical question is the starting point for any investigation and immediately raises the issue of problem context and cross-curriculum links. Because data are numbers in context (Moore, 1990), there is no statistics without a problematic situation from another field. Understanding the contextualized nature of data is crucial in developing a facility with statistics. Yet, elementary school curricula tend to give superficial or limited attention to the role of context, especially with respect to whether inferences drawn from the inquiry process align with both the question and context (Lavigne & Lajoie, 2007). With the increased focus on STEM education, including STEM integration (English, 2016) numerous rich contexts arise for undertaking statistical investigations. In the present study, engineering formed the statistical context where students explored the manufacture of licorice and the roles of various engineers (industrial, manufacturing, chemical) in the production process. Such a context highlights the need for quality control in the manufacturing process to reduce product variation.

Variation

Variation is the underlying concept linking all aspects of a statistical investigation; without variation, there would be no need for statistics (Cobb & Moore, 1997; Franklin et al., 2005; Garfield & Ben-Zvi, 2008; Konold & Pollatsek 2002; Moore, 1990; Watson, 2006). In simple terms, variation is “the quality of an entity (a variable) to vary, including variation due to uncertainty” (Makar & Confrey, 2005, p.28). As Watson (2006) highlighted, the reason data are collected and analyzed is to manage variation and draw conclusions and inferences about phenomena that vary. Although there is considerable research on older students’/adults’ awareness of variation there is less so on how this understanding can be developed with young students. This is a major concern especially given that secondary school and university students frequently apply statistical techniques without appreciating or understanding why, when, or how these are applied sensibly to a range of contexts (Garfield & Ben-Zvi, 2008).

Data Distribution

Developing the concept of variation necessitates some understanding of distribution, where

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

patterns in the variability of the data are of interest and are displayed visually (Makar & Confrey, 2005). Exploring learners' concept of variation provides a window into their understanding of distribution. Research has shown that younger students can come to recognize statistics as ways of measuring characteristics of distribution, which guide inferences about the questions posed (Lehrer & English, in press; Makar, 2014). One way to support and advance this early development is to provide children with opportunities to generate statistical measures of center and spread, and to observe how these indicate a distribution's characteristics (Bakker & Gravemeijer, 2004; Konold & Pollatsek, 2002; Lehrer & English, in press). Activities that involve repeated measures yield data distributions that display "signals" (measures of center) and "noise" (measures of variability), which can help students make sense of statistics as measures (Konold & Pollatesk, 2002). It has been argued that when students develop this understanding, they are viewing data through the lens of distribution, in contrast to just a set of data values (Bakker & Gravemeijer, 2004). In today's increasingly data-driven world, young learners deserve access to these core statistical foundations. Future directions in mathematics education need to consider increasing this access in the elementary mathematics curriculum.

Model Representation and Interpretation

Models created through working with data are usually defined as systems of representation, where structuring and displaying data are fundamental; the structure is created, not inherent (Lehrer & Schauble, 2007). Young learners' ability to create and work with a range of representations, including those that extend beyond traditionally accepted formats, is underestimated and needs to be given more recognition and nurturing. In particular, the explicit consideration of variation in relation to representations has not been a key feature of research in the elementary years. Yet, a major foundational component of young students' statistical growth is being able to interpret the meaning, within a given context, of a distribution that displays variation, clusters, modes, and unexpected values; this might not involve conventional text-book types of graphs. Early experiences with a range of representations that effectively display variation in data sets are important but have remained largely neglected in many elementary curricula until recently. Greater insights are needed into how young learners deal with variation in representations that they, themselves, create from their investigations, including how they respond to questions on comparing variation in different data sets, and how they identify and justify the sources of variation that they encounter. With the increasing impact of technology, young students are exposed to more complex and more varied representations that require careful interpretation and critical analysis rather than mere visual inspection.

Drawing Inferences

Informal inference, a precursor to formal inference, has been highlighted as a foundational component that also has not received the required attention especially in the elementary grades. Informal inference is the process of using the evidence provided by data to answer questions beyond the data, acknowledging the uncertainty associated with the conclusion reached (Makar, 2016). Variation is the key to accepting a conclusion with some degree of uncertainty (Franklin et al., 2007). The confidence with which one can form a decision, however, depends on creating a balance between variation and expectation/prediction (Watson, 2006). In the senior secondary courses of study, this balance is expressed in tests of significance or confidence intervals but learning to appreciate variation and its relationship to expectation/prediction needs to begin in the elementary grades with appropriate hands-on experiences and student/teacher questioning.

In addressing these foregoing components of modeling with data, this paper reports on three questions investigated: (1) *How did students represent the models generated for each licorice-*

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

making method? (2) *How did the students interpret variation and overall data distribution?* and (3) *What was the nature of the informal inferences children drew from their models?*

Methodology

Participants

The activity was implemented in two schools, one a private girls' school, and the other, a co-educational Catholic school. The data in this paper are confined to one class in the former school (mean age of 8.8 years), which was situated in a middle socioeconomic area.

Research Design

The activity was the first that was implemented in a 4-year longitudinal design-based study (Cobb, Jackson, & Dunlap, 2016). This research design caters for complex classroom situations that contain many variables and real-world constraints, supports learning and informs future learning experiences, and facilitates contributions to both theory and practice. Data collection included video-taping of three focus groups as they worked the activity, as well as all class discussions, which were subsequently transcribed for analysis. Focus groups comprised three students of mixed achievement levels selected in consultation with the class teacher.

Activity and Implementation

The activity was created in collaboration with the teacher and formed part of her regular mathematics program in the area of data and probability. The teacher implemented the activity across three school days. The researcher and research assistant were in attendance for the entire activity to observe the students' learning. A detailed lesson description was prepared for the teacher, as was a workbook for students. Students recorded their responses to a number of questions as they worked the activity. Although the students completed the activity in groups, they were to record their own answers and explanations in their workbooks.

The activity comprised several parts including: (a) Reviewing an earlier science activity where students made tubes of lip balm, and discussed variation in their products; (b) Learning about engineers and engineering involved in the manufacture of licorice and foods in general (students viewed a YouTube clip of the American Licorice Co.); (c) Experiencing the notion of variation through exploring packets of manufactured licorice; (d) Investigating questions regarding differences in making licorice sticks by hand (using Play-Doh) and with a Play-Doh extruder ("manufactured"). For each of the hand-made and "manufactured" methods, students identified, measured, compared, and recorded attributes including mass; within-group results were compared; (e) Collating group data on the masses, and representing the group data in a format of choice, for each method; (f) Sharing and interpreting resultant group models from each method with the whole class, including identifying the range and "typical" masses displayed in each group model; (g) Collating all group data and creating a class representation; interpreting the resultant whole-class model of the distribution of the licorice stick masses, for each method.

Data Analysis

For the present paper, data are drawn from the students' workbooks, together with the recorded and transcribed group work and whole class discussions. In conjunction with an experienced research assistant, content analysis (Patton, 2002) was applied in initially identifying, coding, and categorizing the data recorded in the students' workbooks. A further round of refined coding was undertaken to ensure meaningfulness and accuracy. Iterative refinement cycles for video-tape analyses of conceptual change (Lesh & Lehrer, 2000) were applied in reviewing the transcribed focus group and

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

whole class discussions to ascertain the students' learning pertaining to variation, distributions, and inferential reasoning.

Results

This section considers how students represented their models, their interpretations of the variation and data distribution displayed, and the nature of the inferences they drew.

Models Represented

Students represented in various ways their models displaying the licorice-making results. Although forms of bar graphs were most popular, these differed in students' approaches to organizing and structuring their data. For example, many students (78%, N=23) structured their data according to each group member's results (e.g., Monica, Kate, Sarah), while some (13%) ordered the data differently, such as from the "biggest licorice" to "second biggest", to "second smallest", to "smallest licorice" as illustrated in Figure 2. One student displayed each member's heaviest licorice stick only, while another used both tallies and a 3-way table.

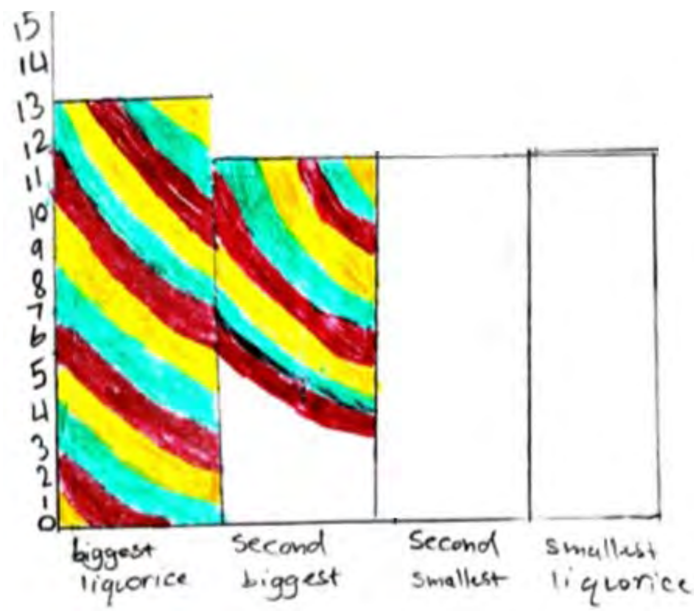

Figure 2. An example of one student's representation for her group's data

Interpreting Variation and Data Distribution

Students were readily able to identify the variation in masses for both the hand-made and factory-made methods, with 83% (N=23) identifying the variation in the former and all students (N=24) for the latter.

Likewise, the students had few difficulties in giving an initial reason for this variation in the hand-made sticks (87%, N=23) with explanations including reference to some sticks being "fatter" or "too thin" or "thicker". They found it more difficult, however, when asked to provide more than one reason.

Sixty-one percent (N=23) were able to offer two acceptable reasons for the hand-made variation, while less than half (48%) were able to give an appropriate third reason (e.g., they simply stated that the sticks "are all different weights"). Over half of the students (63%, N=24) were able to give three appropriate reasons for the factory-made method (58% offering a first reason, 71% a second, and 58% a third).

In collating the group results to form a class plot for each licorice-making method and describing the data distributions, 62% of the students (N=24) offered at least one feature for the hand-made licorice. Of these students, 33% noted multiple characteristics, such as "…lots of spaces and humps and sections and a lot at the start." Those who noted just one aspect gave reasons such as, "very, very lumpy", or "zig-zag." In contrast, all but one student was able to describe the distributional features of the class plot for the second method, with 79% (N=24) describing multiple features, suggesting that their understanding of distribution was developing as they experimented with the licorice-making methods. An example of the class plot for each licorice-making form appears in Figure 3.

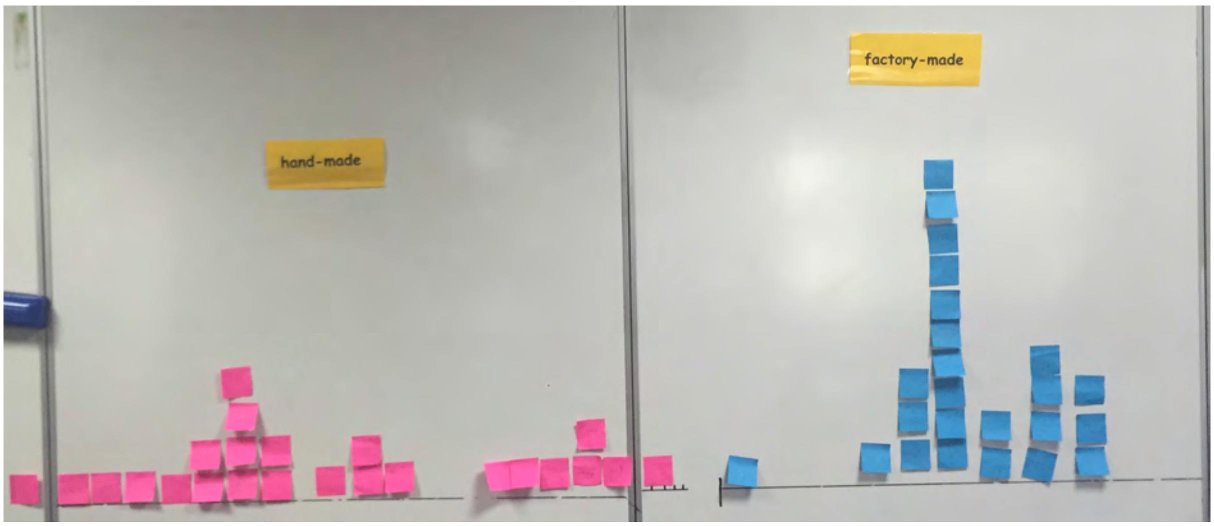

Figure 3. Class plots for each method.

In recording responses to questions about the variation in the masses and the typical mass of the licorice sticks for each class plot, the students again performed better in interpreting the plot for the factory-made licorice. For the hand-made, 58% of students (N=24) could describe the variation in the class plot (e.g., "There's more on 10 and less on 7; there are a lot of people between 8 and 16") and 62% for the factory-made class plot. Although a little over half the students (58%, N=24) could describe the typical mass for the hand-made class plot, 83% (N=24) could do so for the factory-made class plot. The students could readily recognize the difference in the two plots, with 88% (N=24) noting at least one difference. Of these, three students identified multiple differences such as, "Tuesday's (hand-made method) plot had a lot of variation. Thursday's plot (factory-made) had not that much variation. Thursday's plot was a lot taller than Tuesday's plot." The majority of students (71%, N=24) could explain that using the Play-Doh extruder was more accurate in producing sticks of a consistent mass (e.g., "Because it's a machine like, the machine makes them all about the same size and when you're doing them with your hands you can't really tell if they're going to be the same size or not").

Drawing Inferences

On completion of the class models created for each licorice-making method, students were asked, "If you made one more piece of licorice, what do you think (predict) its mass might be? How did you decide?" As part of a follow-up class discussion, the students were also asked, "If another student came into our class and made some licorice, what do you think hers would be (mass of licorice stick)?" Students were readily able to respond to the first question above with 88% identifying an appropriate mass range for the hand-made and 96% for the equipment-made (N=24).

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

The majority of students could also offer appropriate reasons for each decision, referring to either their own data (42% for hand-made and 33% for factory-made) or the whole class data (29% for hand-made and 46% for factory-made). Their reasons included, "I think because most of mine were around ten and mine were both exactly 1cm wide and 8cm long;" "because it is about the average;" and "I decided because 13g is the typical mass of sticks in the class."

During class discussions, students frequently referred to chance and uncertainty when explaining what the mass of a licorice stick made by a new student might be. One student explained that, "It might be 13 because most people got … 13 so maybe that's the typical number." Another student responded, "I think maybe 12, because if she came in, there's a chance, because the Fun Factory makes all of them um pretty similar and, and she could make it, but I decided on that [13g] because I think there's a more likely chance that she would because it won't always be bigger, she might get it a little smaller than some." The teacher asked a further question, namely, "Would you expect, say, if we did it again next week and we used the same Play-Doh, and we used the same Fun Factory, would you expect the same plots?" Alesha commented, "I think they might be different because like we could do something, we may have like cut it a bit further or because it's really hard to get everything exact, so it won't always be exact." Monica agreed, "…maybe or maybe not, I sort of agree … you actually don't know because … when you made three of them like last week they weren't all the same mass, they weren't all 15 or they weren't all 13…"

Discussion

This paper has illustrated how third-grade students can engage in modeling with data involving core statistical concepts and processes, when presented with a motivating context and a meaningful hands-on activity. Using a STEM-based context involving a licorice factory, students were able to explore the important roles of the various engineers responsible for manufacturing high-quality products. The importance of quality control in the real-world provided a valuable context for appreciating how the two forms of licorice making yielded different variations in the sticks produced. By generating and observing "variation in action", the students could see how variation is an important factor throughout a statistical investigation. They developed an understanding of the reasons behind the greater variation in hand-made sticks and hence, the difference in data distributions of their hand-made and factory-made sticks. Students could identify foundational data distributional features including center, range, typical, and middle, which are usually not introduced until the later grades and then frequently in a computational manner. In line with other research, activities in which students create their own statistical measures of center and spread enable them to observe and understand the features of a data distribution (e.g., Bakker & Gravemeijer, 2004; Konold & Pollatesk, 2002). As Franklin et al. (2007) emphasized, "Statistical education should be viewed as a developmental process" (p. 13) and, as such, these foundational experiences need to commence in the elementary grades.

From the individual and whole-class representations of the models generated, students could draw inferences including predicting the masses of further sticks that might be made. Some awareness of chance and uncertainty was present as the students realized that variation in both licorice-making methods meant that predictions could not be drawn with absolute certainty. Given students' realization of the uncertainty in drawing conclusions due to variation, the activity can provide foundations for chance explorations. For example, investigations involving the chances of selecting particular candies from factory produced packets can yield unexpected results due to variation in contents. Returning to the conference theme, the present results provide further support for moving mathematics education along a path that capitalizes on elementary students' early statistical talents. Given the research that has already revealed these talents, greater attention is needed to further advance the field.

Galindo, E., & Newton, J., (Eds.). (2017). *Proceedings of the 39th annual meeting of the North American Chapter of the International Group for the Psychology of Mathematics Education.* Indianapolis, IN: Hoosier Association of Mathematics Teacher Educators.

Acknowledgments

This study was supported by funding from the Australian Research Council (ARC; DP150100120). Views expressed in this paper are those of the author and not the ARC. Interstate collaborators (Jane Watson & Noleine Fitzallen) are acknowledged, together with the senior research assistant, Jo Macri, and the classroom teacher and her students. Feedback from the paper reviewers is also gratefully acknowledged.

References

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). Dordrecht: Springer.

Cobb, G. W. & Moore, D. S. (1997). Mathematics, statistics, and teaching. *The American Mathematical Monthly, 103* (9), 801-823.

Cobb, P., Jackson, K., & Dunlap, C. (2016). Design research: An analysis and critique. In L. D. English & D. Kirshner (Eds.), *Handbook of international research in mathematics education* (3rd ed.) (pp. 481-503). New York, NY: Routledge.

English, L. D. (2016). STEM education K-12: Perspectives on integration. *International Journal of STEM Education*, 3(1), http://dx.doi.org/10.1186/s40594-016-0036-1.

English, L. D., & Watson, J. M. (2015). Exploring variation in measurement as a foundation for statistical thinking in the elementary school. *Journal of STEM Education, 2*(1), 1-20.

English, L. D., Arleback, J. B., & Mousoulides, N. (2016). Reflections on progress in mathematical modelling research. In A. Gutierrez, G. Leder, & P. Boero (Eds.), *The Second Handbook of Research on the Psychology of Mathematics Education* (pp. 383-413). Rotterdam: Sense Publishers.

Franklin, C., Kader, G., Mewborn, D., Moreno, J., Peck, R., Perry, M., & Scheaffer, R. (2007). *Guidelines for assessment and instruction in statistics education (GAISE) report: A preK-12 curriculum framework*. Alexandria, VA: American Statistical Association.

Garfield, J., & Ben-Zvi, D. (2008). *Developing students' statistical reasoning: Connecting research and teaching.* Dordrecht: Springer.

Johnson, C. C., Peters-Burton, E. E., & Moore, T. J. (2015). *STEM roadmap: A framework for integration.* London: Taylor & Francis.

Konold, C, & Pollatsek, A. (2002). Data analysis as the search for signals in noisy processes. *Journal for Research in Mathematics Education, 33*(4), 259–289.

Lavigne, N. C., & Lajoie, S. P. (2007). Statistical reasoning of middle school children engaged in survey inquiry. *Contemporary Educational Psychology, 32*, 630-666.

Lehrer, R., & English, L. D. (In press). Introducing children to modeling variability. In Ben-Zvi, D., Garfield, J., & Makar, K. (Eds.). *International handbook of research in statistics education*. Springer.

Lehrer, R., & Schauble, L. (2007). Contrasting emerging conceptions of distribution in contexts of error and natural variation. In M. C. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 149-176). NY: T&F.

Lehrer, R., & Schauble, L., (2012). Seeding evolutionary thinking by engaging children in modeling its foundations. *Science Education, 96*(4), 701-724.

Lesh, R. A., & Lehrer, R. (2000). Iterative refinement cycles for videotape analyses of conceptual change. In R. A. Lesh & A. E. Kelly (Eds.), *Handbook of research design in mathematics and science education* (pp. 665-708). Mahwah, NJ: Lawrence Erlbaum.

Makar, K. (2014). Young children's explorations of average through informal inferential reasoning. *Educational Studies in Mathematics Education, 86*, 61-78.

Makar, K., & Confrey, J. (2005). "Variation-talk": Articulating meaning in statistics. *Statistics Education Research Journal, 4*(1), 27-54, http://www.stat.auckland.ac.nz/serj

Makar, K., Bakker, A., & Ben-Zvi, D. (2011). The reasoning behind informal statistical inference. *Mathematical Thinking and Learning, 13*(1-2), 152-173.

Moore, D.S. (1990). Uncertainty. In L.S. Steen (Ed.), *On the shoulders of giants: New approaches to numeracy* (pp. 95-137). Washington, DC: National Academy Press.

Patton, M. Q. (2002). *Qualitative research & evaluation methods* (3rd ed.). Thousand Oak: Sage.

Watson, J. M. (2006). *Statistical literacy at school: Growth and goals*. New York: Routledge.

Watson, J., Skalicky, J., Fitzallen, N., & Wright, S. (2009). Licorice production and manufacturing: All-sorts of practical applications for statistics. *Australian primary Mathematics Curriculum, 14*(3), 4-13.