Statistics Teachers as Scientific Lawyers

Joanne Woodward
The University of Auckland
<jwoo088@auckland.ac.nz>

Maxine Pfannkuch
The University of Auckland
<m.pfannkuch@auckland.ac.nz>

When Year 10 students are introduced to reasoning from box plots the type of classroom discourse that will lead them to understand statistical inferential argumentation is unknown. In this paper the discourse of one teacher and her class is analysed. Although the teacher required evidence for claims and introduced statistical vocabulary, she argued with the medians, lacked uncertainty, did not answer the original question or make sense of the conclusion. The implications for teaching are discussed.

According to Tukey (1977) statisticians are like detectives since they need to unlock stories in data. But they also need to be narrators of the stories they discover and excellent lawyers presenting reasoned arguments (Abelson, 1995). Little research has been done on the way teachers communicate inferences from data and yet this is critical for student learning (Rubin, Hammerman, & Konold, 2006). The teacher, as the master of statistical discourse, provides the accepted vocabulary, language structure, and behaviour, guiding and scaffolding the students to attend to the correct features of graphical representations and to build meanings recognised by the statistics community. This paper considers the language of one Year 10 teacher as she formulates inferences from box plots in the conclusion step of one investigation to discover how closely her argument models that of a statistician. Particular attention is given to whether the teacher displays her statistical thinking as reasoned arguments. In other words, is this teacher enculturating her students into a community of statistical practice with argumentative skills equal to a lawyer, or a community where statistical thinking is not present?

Background

Tukey's (1977) focus on discovery of stories in data using innovative visual representations was revolutionary. His quick pencil and paper methods of graph construction, such as the box plot, summarised the data in a more succinct way. The strength of the box plot, however, is also one of its weaknesses since students tend to reason solely using the five-number summary values, the cut-off points, rather than seeing the box plot as representing a distribution (Biehler, 2004). Such deficient inferential reasoning may result in *shaky* conclusions and give students the idea that statistics is deterministic (Ben-Zvi, 2006). In Year 10 students are expected to make informal inferences about populations by comparing samples displayed as box plots, that is, informally draw inferences by mainly looking at, comparing, and reasoning with box plots (Pfannkuch, 2006). Information represented in a box plot is dense, which makes it conceptually demanding (Bakker, 2004). There is limited research, however, on how students and teachers reason with and draw inferences from box plots and on describing how teachers guide and model to students the process of informal inference.

In the conclusion step of the investigative process, the need for inference is most obvious as in this step all the evidence must be presented and weighed. This involves interpretation of the results of analysis, bringing in new ideas, and communicating them in an appropriate format for the audience. Interrogation of the results is also present in the

Proceedings of the 30th annual conference of the Mathematics Education Research Group of Australasia
J. Watson & K. Beswick (Eds), © MERGA Inc. 2007

drawing of conclusions – possibilities for explanations are generated within the context, more information is sought, interpretations are critiqued, and unrealistic explanations are discarded whereas support is given for plausible ones (Wild & Pfannkuch, 1999). This interrogation process is very similar to a lawyer developing and presenting an argument.

Krummheuer (1995) considers classroom interaction as collective argumentation to develop working interims, which eventually become accepted knowledge. He describes argumentation as consisting of four elements: claims, grounds, warrants, and backings. In his scheme claims are usually the postulated solution to the problem, grounds are the facts to support the claim, warrants are the information joining the claims and grounds, whereas the backings are the global contexts, which give the warrants authority. Lampert (1990) describes a zig-zag process for the formation of conclusions that begins with conjecture, examines premises, and proposes counter arguments before agreement is reached, whereas Bakker, Derry, and Konold (2006) describe an inferential view of data as being a social exchange of questions, explanations, and justifications. Social interaction is also an important part of learning. Sfard (2000) likens discourse to playing a game. If the teacher is doing all the thinking then effectively the teacher is playing the game in her head. By verbalising her thoughts the teacher is inviting the students to play the game, which is important as in order to develop shared statistical meanings both the teacher and the students need to participate. Learning how to argue with data is the result of wanting to play the game, or in other words to communicate more effectively and recognise the superior discourse of the master, the teacher (Ben-Yehuda, Lavy, Linchevski, & Sfard, 2005).

Informal inference is a complex process, one in which researchers are still defining the rules of the game about how to talk about box plots (Pfannkuch, 2006). Even if all the rules of the game were understood, it would be impossible to make them all explicit. Instead the students need to experience thinking about data displayed in box plots and presenting inferences as reasoned arguments through interacting first with the teacher's thoughts.

Method

Two teachers and their Year 10 classes of students participated in a case study, which considered the language used during six classroom lessons on informal inference. Both teachers taught at the same urban girls' school and both classes were in the average ability stream. The majority of the girls were of Pacific Island ethnicity, many of whom speak English as their second or third language.

The first researcher wrote three class activity outlines, which included student worksheets, overhead transparencies, and teachers' notes. The teachers participating in the study and those teaching at their school were consulted in the development of the resources. The activities encouraged students to act as statisticians unlocking the story in the data and learning in the spirit of Tukey (1977) through exploratory data analysis. Wild and Pfannkuch (1999) found practicing statisticians use an investigative cycle of defining the problem, planning, data management, analysis, and formulating conclusions. To emphasise this cycle the steps were used as section headings on the teacher's notes. The conclusion step was written by completing two statements, *I notice…* and *I wonder…,* as these were found to provide a useful structure to overcome the initial inertia students experience in writing conclusions (Pfannkuch & Horring, 2005).

The lessons were videoed and transcribed. To illustrate key findings about the language used by both teachers when formulating conclusions in investigations, this paper uses one of the teachers, a female with 6 years teaching experience, and one lesson, the sixth and last lesson on an activity called *Big Foot*.

Analysis

Informal inference has only recently been recognised as an important step to develop more formal inference concepts. The definition and concepts of informal inference are still being developed and so the tools to analyse the language of inference are also being created. A micro-analysis of the language used was based on an adaptation of Cadzen's (2001) initiation, response, and evaluation model for discourse analysis. Krummheuer's (1995) argumentation categories and the question categories of explanation, justification, noticing, wondering, and closed were added to her model to reflect the data captured better.

The fictional context for the *Big Foot* activity was provided as a story in the teacher's notes. The teacher read the story to the class about Alice and her twin brother going to their cousin's farm for a holiday. Normally Alice fits into bigger gumboots than her twin brother, but this year she notices that he has bigger gumboots. She wonders whether he has thick socks on or whether his feet are actually bigger than her feet. The problem to investigate is: Who have bigger feet, girls or boys? To answer this question a sample of real data from New Zealand CensusAtSchool was provided on the right foot length of 9, 11, and 13 year old male and female students. The sample size for each of the six groups was 24. Each group of students received data on one of the three age groups. Figure 1 shows box plots of the data but note that the students' graphs did not show outliers.

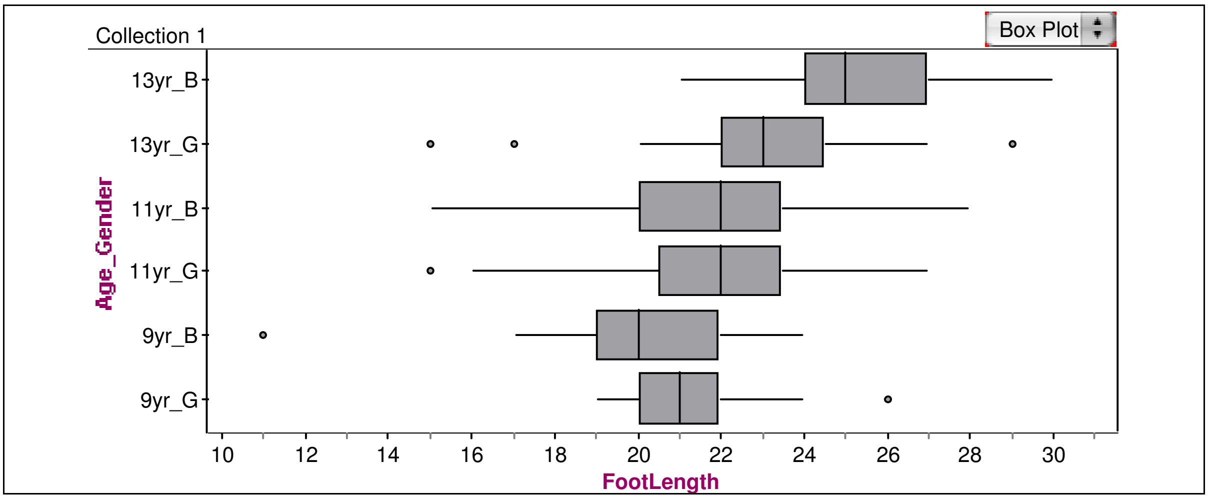


Figure 1. Box plots of the data provided for the *Big Foot* activity.

After the students drew their box plots and formulated their conclusions the teacher had a class discussion. From a detailed analysis of her argumentation language four main themes emerged, explaining the evidence, justifying the evidence, drawing conclusions from the evidence, and making sense of the conclusion.

Explaining the Evidence

During the discussion of the conclusion, the teacher required the students to explain the claims they made and she emphasised selected evidence to support her arguments. Features of the teacher's language were requesting explanations, using statistical terms, using gestures, and using quantitative measures as she supported and guided the students to develop more complete arguments in their conclusions. The following excerpt, which illustrates some of these language features, occurred during the formulation of the conclusion from the foot length data for 11-year-olds.

T: These two sets of data here now are 11 year olds. The groups have done their own “I notice” “I wonder” but what do the rest of you think about 11 year old boys’ and girls’ foot size?
S: They're very similar.
T: Thank you, okay, they're very similar. What tells you that? Because. They're very similar because what?
S: They're both like.
S: The range.
T: The spread’s similar, similar range.
S: The boxes is similar.
T: The boxes are similar, the interquartile range is similar, real similar (measures them with her fingers on the overheads) the interquartile range.
S: The medians.
T: The medians are the same now. Do you know, did you notice that before they were one centimetre apart. Now at 11 years old what's happened?
S: They're the same.
T: They're the same. Great. Okay. At 11 year old, at 11 years old the girls from these data values have the same foot lengths.

The teacher often asked students to explain their observations either by asking directly “what are you looking at?” or by rewording the student’s answer as a clarification type question. A feature of her language that can be noted in the excerpt above is her use of closed questioning, which was often used to ask the students to explain their claims. She supplies another word such as “because” using a raised intonation and by revoicing the student’s response with the additional words, “What tells you that? Because. They're very similar because what?” Another feature of the teacher’s language was to revoice the students’ responses using statistical vocabulary so that rather than using the term *boxes*, the teacher uses the term *interquartile range* and then reinforces this substitute term, thereby implying that these are right words with which to argue. Further reinforcement is through hand gestures, pointing to or measuring the differences between the interquartile ranges. Another strong feature of her language for explaining the evidence, which is not illustrated in the excerpt, was her requirement for quantitative measures. A student would say, “the box is bigger” to which she would reply, “by how much?” To the student’s response “bigger by 5” she would revoice and typically add the measure of centimetres to the student’s response, thereby referencing and reinforcing the context. These quantitative measures, however, were not used as evidence for her argument but rather were observations.

The excerpt above also demonstrates how she typically guided the students to see the median as the most important feature. The students offered a variety of views to support the claim that the foot sizes for 11-year-old boys and girls are similar. They suggest the range, boxes, and the medians are the same. Although the teacher provided a visual explanation by drawing along the length of the boxes to highlight the range and the boxes, the median received the most attention from the teacher and appeared to be the answer she was requiring. In an earlier lesson she explicitly stated that the median was the most important feature of a box plot. So although the teacher did require explanations, these tended to focus on only one feature of the box plot, the median. The only time she used numbers to support the argument was when she reasoned with the median.

Justifying the Evidence

Although the teacher provided the backing for the claim that “girls have a larger foot length than the boys” using the values of the medians, the teacher did not provide any warrants for using the median. For example: “so what we’re saying girls is the median value is 21 and here it’s 20. So for year 9 [9-year-old] girls the average, typical value of their right foot is one centimetre bigger than for the boys”. The concept of using the median as a representation of the data set was not used to support the teacher’s arguments in the conclusion, instead the use of the medians was presented in a way that suggested representativeness was a ground, an understood and accepted concept in the classroom.

Drawing Conclusions from the Evidence

Several features of the teacher’s language may have conveyed a sense of certainty about the conclusion that was drawn from the evidence. One feature was the way the conclusion was reached. The teacher focused on a single statistic, the median to compare the data sets. This may have communicated to the students that only the median should be attended to and the rest of the information in the graph could be ignored. The teacher described the boxes as being at the median, as if the whole box was just a single line: “See this little box for boys is at 25 and this one’s at 23”. Reasoning with only one feature of the data also suggested there was a single procedure to follow to formulate a conclusion, rather than a weighing of the evidence. If the boys’ median foot length was numerically larger than the girls’ median foot length then the boys had a larger foot length, and vice versa. A sense of finality may have been communicated to the students, which could have prevented them from exploring the data any further.

Definite language was often used rather than expressing uncertainty by using phrases such as *tends to* or *could show* and so students may have assumed there was a single right answer. The teacher did occasionally use informal variable language; for example, she described the 13-year-old boys’ foot length as being “on average, two centimetres bigger than girls’ foot length”. Usually, however, the teacher used exact language such as, “at 11 years old the girls from these data values *have* the same foot lengths”.

The *Big Foot* activity was introduced by a story about Alice and her twin brother. Although the teacher did answer the statistical question, about whether boys or girls had bigger feet, the purpose for the investigation, to tell Alice whether her brother’s feet had grown or whether he was just wearing thicker socks was not part of the conclusion drawn. Therefore the hallmarks of the teacher’s drawing of conclusions from data were certainty and a lack of reference to the original problem under consideration.

Making Sense of the Conclusion

For the conclusion the teacher did not discuss the significance of the differences in the box plots in terms of the context, that is, age and foot length. The difference between the boys’ and girls’ foot length was simply stated as “1 centimetre”, “the same”, and “2 centimetres”.

Often the *I wonders* came from students’ personal opinions or experience of the context rather than the data. This is evident when the students suggested growth spurts occurred when they were 14 years old, for which they did not have data, or for years the data showed the opposite.

T: Alright girls, very quickly, people over at this table were talking about growth spurts, lovely, I wonder if boys undergo a growth spurt between the ages of …
S: 9.
S: 13 and…
S: 14 (several students).
T: 11 and 13.
S: Between 9 and 13.
S: I reckon 10 and 14.
T: Shh, S10's saying she did the year 9s but she said she found out the boys had smaller feet, year 9 boys had smaller feet. She said actually that's probably not true because she knows that her Dad's got bigger feet than her Mum, probably.

As shown in the excerpt, the teacher engaged with these personally-based wonders but did nor challenge them and hence lost an opportunity to explore the difference between evidence-based statistical reasoning and personal experience. Although her words reflected the data she did not explicitly redirect students' attention back to the box plots and data.

Discussion and Conclusion

Informal inference is a recent introduction into the curriculum. In this study the teacher was learning a new way of teaching statistics but more importantly she had not experienced or been enculturated into the discourse of informal inference. To expect to see a perfect statistical discourse modelled in a real classroom is unrealistic. However, there are some issues that arise from the analysis that need to be considered.

Abelson (1995) identified two facets of argumentation: inference, which is the process of deriving logical conclusions from data, and providing persuasive arguments based on the analysis. Students enter a classroom expecting the teacher knows and will provide the correct answers. They also expect there is a single correct answer. Yet this is not the case in statistical investigations. Analysis of data usually provides a multiplicity of results rather than one clear answer and some are contradictory (Biehler, 1997). The teacher in this study presented only one interpretation of the data and did not request alternative interpretations from the students. The argument was one sided, with the teacher developing her stance only. Her conclusion was certain, resulting in a deterministic rather than a probabilistic stance. Perhaps the teacher focused on simplifying her process of reasoning with the data and so she removed the arguments she was having in her head and only verbalised the *winning* argument.

The teacher also tended to reduce the complex relationships in the box plots by only using the medians as evidence for the arguments and by not linking her observations back to the context or the problem under investigation. The teacher did not verbalise her thinking or reasoning process, nor did she justify the use of the median by providing a warrant, instead she used a series of questions to funnel the students to focus on the median. Although the students could answer the questions, they were not learning about how to think about the box plot or the reasoning process. If this is not occurring while the teacher is present then the students are unlikely to think for themselves when the teacher is gone (Mason, 2000). Wild and Pfannkuch (1999) call for statistical thinking to be articulated: in a classroom this call surely should be louder. The verbalisation of the inner dialogue alerts the student to its existence; noticing features on a graph becomes a process rather than a plucking of ideas from the air. Thinking is learnt in the same way knowledge is learnt, through interaction with a knowledgeable other (Perkins, Jay, & Tshman, 1993) or as Mason (2000, p. 97) states "a student learns to think mathematically by being in the presence of a relative expert who makes their thinking processes explicit". When students interact with the thinking of teachers they have a model for thinking and the experience of thinking. Modelling also provides a way for students to hear how the language and discourse is used in the context, and how it is structured.

Increasing the fluency of students' discourse will mirror an increased understanding of graphs (Ainley, Nardi, & Pratt, 2000). If understanding emerges in use (Bakker et al., 2006) then teachers need to invite the students to participate in a learning dialogue. The teacher in this study did ask the students to offer their opinions but rather than exploring the students' stances by asking them for the basis of their claims, as Bakker et al., (2006) suggest, the teacher evaluated them. The teacher did invite the students to support their views by explaining them, but the judgement still rested with the teacher rather than inviting the other students to agree or disagree.

Formulating thoughts into words helps clarify students' thinking. Words can act as a pump for statistical ideas that do not yet exist for students (Sfard, 2000). When the teacher introduced the terms spread, range, and interquartile range the students then had the words to argue with and new referents for exploration and elaboration. Technical knowledge, however, is not sufficient to interpret graphs to provide a meaningful answer in terms of the problem being investigated. To synthesise an answer the evidence needs to be weighed (Pfannkuch, 2006). The teacher in this study did not explicitly model this process although there were periods of silence where she may have been thinking through the evidence. In particular she did not challenge the students when they were attempting to make sense of the conclusion. Critical thinking is required during the evaluation process, and includes weighing the quantitative evidence and contextual knowledge. In real investigations *correct* solutions do not occur, instead statisticians must present their best conclusion fully supported. Abelson (1995) describes statisticians as requiring the narrative and argumentative skills equal to lawyers. Statisticians may require these skills but unlike lawyers their arguments seek to find the truth from the story in the data, not to present one side of a story or a winning argument. Also statisticians' language is tempered by uncertainty whereas lawyers argue with certainty. Hence statisticians' argumentative skills are those of scientific lawyers.

If teachers want to encourage students to engage in argumentation then it is their responsibility to initiate and guide students towards a shared understanding of how the discourse is structured. By presenting and allowing only a single interpretation of the data, which is evaluated only by the teacher, as was the case in this study, the statistical process of inference is not being modelled (Figure 2). Teachers instead need to use questioning and revoicing to support the development and structuring of alternative views and to model critical thinking when evaluating the stances and presenting the argument just as a scientific lawyer would (Figure 2).

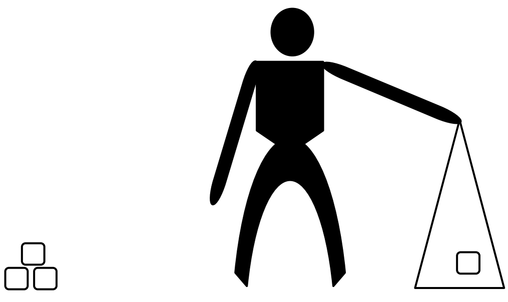
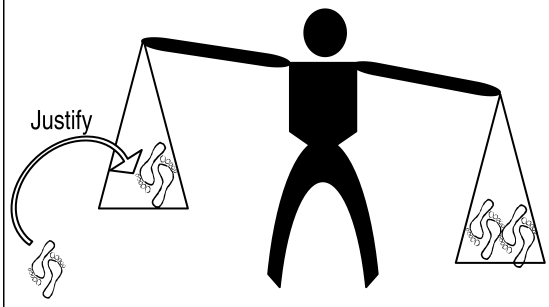
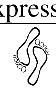
Argumentation of teacher studied	**Proposed argumentation for teacher**
 While the teacher did notice some features of the box plot and did require an explanation for claims made [shown by presence of feature on scale], she only used the median as evidence in the conclusion. No justification was given for using the median. The conclusion was expressed with certainty and not discussed in terms of the problem.	 Features of the box plot are weighed as evidence in the conclusion. Inclusion of features is justified [shown by presence of contexualised features on both sides of scale]. Features are contextualised in order to determine the meaning of the findings. The evidence is discussed and weighed in terms of the problem, its context, and statistical issues such as sampling variability and quality of data. The conclusion is expressed in terms of uncertainty.
Legend:  Represents that feature such as the median is discussed quantitatively.	Legend:  Represents that feature such as the median is discussed within the context.

Figure 2. Summary of argumentation used and proposed.

The teacher was beginning to enculturate her students into a statistical community of practice but her focus on the production of box plots and the formulation of the *correct* conclusion obscured the investigative process and statistical thought, a facet of pedagogic purpose that Mason (2000) has also found. Several researchers (e.g., Biehler, 1997) have found teachers are unsure about how to talk about graphs and so the findings of this study are not unique but contribute to the growing call to discover ways of developing teachers' talk. Statistical thinking is complex and involves searching for the story in the data. Statistics teachers as scientific lawyers need to narrate their thinking, providing an account of how they are reasoning, arguing, and weighing the evidence for the story that they have unlocked, in order to answer the problem posed at the beginning of the investigation.

References

Abelson, R. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Associates.

Ainley, J., Nardi, E., & Pratt, D. (2000). The construction of meanings for trend in active graphing. *International Journal of Computers for Mathematical Learning*, *5*(2), 85-114.

Bakker, A. (2004). *Design research in statistics education: On symbolizing and computer tools*. Utrecht, The Netherlands: CD-ß Press, Center for Science and Mathematics Education.

Bakker, A., Derry, J., & Konold, C. (2006). Using technology to support diagrammatic reasoning about center and variation. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Brazil, July, 2006: Working cooperatively in statistics education.* [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Ben-Yehuda, M., Lavy, L., Linchevski, L., & Sfard, A. (2005). Doing wrong with words. *Journal for Research in Mathematics Education*, *36*(3), 176-247.

Ben-Zvi, D. (2006). Scaffolding students' informal inference and argumentation. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Brazil, July, 2006: Working cooperatively in statistics education*. [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Biehler, R. (1997). Students' difficulties in practicing computer-supported data analysis: Some hypothetical generalizations from results of two exploratory studies. In J. Garfield & G. Burrill (Eds.), *Research on the role of technology in teaching and learning statistics* (pp. 169-190). Voorburg, The Netherlands: International Statistical Institute.

Biehler, R. (2004, July). *Variation, co-variation, and statistical group comparison: Some results from epistemological and empirical research on technology supported statistics education*. Paper presented at the 10th International Congress on Mathematics Education, Copenhagen

Cazden, C. B. (2001). *Classroom discourse: The language of teaching and learning* (2nd ed.). Portsmouth, NH: Heinemann.

Krummheuer, G. (1995). The ethnography of argumentation. In P. Cobb & H. Bauersfeld (Eds.), *The emergence of mathematical meaning: Interaction in classroom cultures* (pp. 229-270). Hillsdale, NJ: Erlbaum.

Lampert, M. (1990). When the problem is not the question and the solution is not the answer: Mathematical knowing and teaching, *American Educational Research Journal, 27*(1), 29-63.

Mason, J. (2000). Asking mathematical questions mathematically. *International Journal of Mathematical Education in Science and Technology*, *31*(1), 97-111.

Perkins, D., Jay, E., & Tshman, S. (1993). Beyond Abilities: A dispositional theory of thinking. *Merrill-Palmer Quarterly*, *39*(1), 1-21.

Pfannkuch, M. (2006). Comparing box plot distributions: A teacher's reasoning. *Statistics Education Research Journal*, 5(2), 27-45. [Online: http://www.stats.auckland.ac.nz/serj]

Pfannkuch, M., & Horring, J. (2005). Developing statistical thinking in a secondary school: A collaborative curriculum development. In G. Burrill & M. Camden (Eds.), *Curricular development in statistics education: International Association for Statistical Education (IASE) Roundtable, Lund, Sweden 28 June-3 July 2004,* (pp. 204-218). Voorburg, The Netherlands: International Statistical Institute.

Rubin, A., Hammerman, K., & Konold, C. (2006). Exploring Informal inference with interactive visualization software. In A. Rossman & B. Chance (Eds.), *Proceedings of the Seventh International Conference on Teaching Statistics, Salvador, Brazil, July, 2006: Working cooperatively in statistics education.* [CD-ROM]. Voorburg, The Netherlands: International Statistical Institute.

Sfard, A. (2000). On reform movement and the limits of mathematical discourse. *Mathematical Thinking and Learning*, *2*(3), 157-189.

Tukey, J. (1977). *Exploratory data analysis*. Reading, MA: Addison-Wesley.

Wild, C.J., & Pfannkuch, M. (1999). Statistical thinking in empirical enquiry (with discussion). *International Statistical Review*, *67*(3), 223-265.