

TECHNOLOGICAL PEDAGOGICAL CONTENT KNOWLEDGE IN STATISTICS

Lawrence M. Lesser
The University of Texas at El Paso (Mathematical Sciences Dept.)
500 W. University Avenue; El Paso, TX; 79968-0514
Lesser@utep.edu

Randall Groth (Salisbury University)

The GAISE College Report (ASA 2005) contains the results of a survey by Joan Garfield suggesting that the biggest recent change in statistics instruction has been the use of technology. GAISE goes on to say that technology advances make procedures more accessible, which means instructors spend less time teaching mechanics, and more time teaching interpretation. The Report frequently mentions technology and the fifth of its six recommendations is to use technology for developing concepts and analyzing data.

Technologies listed in GAISE span a wide range: graphing calculators, statistical packages, educational software, applets, spreadsheets, web-based resources (including data sources, on-line texts, and data analysis routine), and classroom response systems. Having competence with these forms of technology does not itself ensure a teacher is aware of and able to implement all of the ways the GAISE Report suggests teachers can use technology for instruction, including being able to: access large real data sets, automate calculations, generate and modify appropriate statistical graphics, perform simulations to illustrate abstract concepts, explore “what happens if...”- type questions, and create reports. Teachers should also be able to select the most appropriate tool for a situation by using these criteria: ease of data entry, ability to import data in multiple formats, interactive capabilities, dynamic linking between data, graphical, and numerical analyses, ease of use for particular audiences, availability to students, and portability.

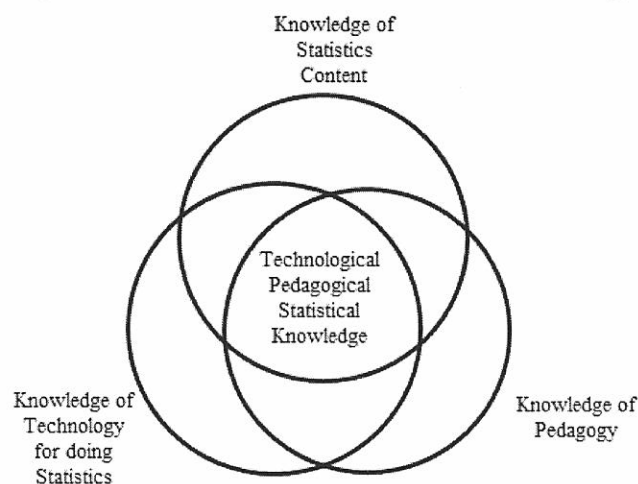
Literature discussing necessary components of the knowledge base for teaching provides further context. Lee Shulman (1986, p. 9) defined *pedagogical content knowledge* (PCK) as the intersection and integration of pedagogy and content knowledge, including “the most useful forms of representation of these ideas, the most powerful analogies, illustrations, examples, explanations, and demonstrations.” Selecting useful and powerful pedagogical devices for teaching statistics is a non-trivial matter, since teachers have to attend to issues like avoiding the portrayal of statistics as fundamentally deterministic rather than stochastic (Moore, 1988), while also understanding students’ potential difficulties with specific statistical ideas (Groth, 2007). Clearly, knowing one’s content (e.g., as a statistician knows statistics) and knowing general principles of sound pedagogy (e.g., as they might be taught in a generic course in a College of Education, not specific to only one domain) do not together guarantee this kind of knowledge. Pedagogical content knowledge is more than the sum of the parts and may be acquired through a

combination of years of experience, ongoing reflection, and reading the educational literature (National Research Council, 2000).

TPSK: Technological Pedagogical Statistical Knowledge

Mishra and Koehler (2006) and Niess (2005) articulate technological pedagogical content knowledge (TPCK) as the intersection of technology, pedagogy and content (see Figure 1) This paper aims to explore this for the specific content of statistics, thus yielding technological pedagogical statistical knowledge (TPSK):

Figure 1. TPSK -- the intersection of knowledge of statistics, pedagogy, and technology



An example of the intersection of statistics and pedagogy would be a repertoire of analogies that are effective in giving intuition or insight about statistics concepts (e.g., Martin 2003). An example of the intersection of statistics and technology would be performing or modeling calculations with a software package commonly used by practicing professional statisticians (e.g., Minitab, SAS, or SPSS). Other examples of this intersection include knowing how to import data from the Internet directly into Excel with each data column going into its own spreadsheet column (Christie 2008) and being aware of workarounds for charts not part of Excel such as stemplots, boxplots, and back-to-back charts (Baker 2004). An example of the intersection of pedagogy and technology would be using data collection devices in class (such as the Calculator Based Laboratory). An example of the three-way intersection of statistics, pedagogy and technology would be using a software package such as Fathom, Tinkerplots, or Probability Explorer in a way that helps students directly engage with their conceptual understanding of fundamental content. The journal most specifically focused on TPSK is *Technological Innovations in Statistics Education*, which was launched in 2007.

Consider the specific example of disjunctive probabilities – that is, probabilities of an event happening at least once in a fixed number of opportunities. From a statistics content perspective, this involves recognizing that it is easiest to utilize rules of probability to find the complement – i.e., the probability the event not happening – and then subtract

that result from 1. From the perspective of pedagogy, a teacher understands the importance of finding examples to which students can readily relate or engage. Pedagogical content knowledge tells us that students tend to underestimate such probabilities. And for the technology dimension, we know how to use specific devices such as calculators and spreadsheets. Putting this all together, we examine a particular familiar disjunctive probability scenario referred to as the birthday problem: “how many people must assemble for there to be at least a 50% chance that at least 2 people share the same birthday (ignoring the year and ignoring leap days)?” The traditional approach involves multiplying out a long string of fractions with a calculator, but with TPSK, Lesser (1999a) shows that this string of fractions is more readily generated with a spreadsheet like Excel and can also be used to model an excellent approximate solution yielding far more intuition into the surprisingly low answer (of 23). Another example of TPSK involving Excel is understanding the limitations and pitfalls of pie charts (Hunt and Mashhoudy 2008).

In addition to allowing teachers to plan effective instruction, TPSK can help teachers anticipate pitfalls related to using technology to teach statistics, so we will explore some of them here. The TI-83/84+ family graphing calculator is commonly used in the secondary and post-secondary setting. For an example from Lee (2005), suppose a teacher pulls a class set of new calculators out of the box (or has just cleared the RAM of existing calculators) and has students simulate 10 rolls of a die using the command `MATH → PRB → randInt(1,6,10)`. Because the pseudorandom number generator depends on a seed value, a situation where all calculators have (or go back to) the factory preset seed value will cause all calculators to generate the same sequence!

For another example (adapted from Lee 2005 and Lesser 2007), consider students being asked to use the `STAT → CALC → LinReg` command to find the correlation value r and equation for the line of best fit to predict L1 from L2 for this set of 5 ordered pairs (L1, L2): $\{(0, 0), (1, 1), (3, 2), (4, 3), (12, 4)\}$. First, students will not obtain a correlation value unless they were previously specifically told to hit the Catalog key, scroll down to DiagnosticOn and do a one-time activation. As for the equation of the line of fit, chances are some students will obtain $y = 0.3x + 0.8$ because they were not aware of one or more of these: (1) the equation is affected by which variable is independent; (2) using the command LinReg without any arguments following it has the same effect as LinReg L1, L2; (3) LinReg L1, L2 assumes that L1 (not L2) is the independent variable, which is different from other forms of technology (e.g., Excel); (4) there is an application (CtlgHelp) that could help recall (3). In the 8th edition of *For All Practical Purposes*, Lesser avoids the “Diagnostic On” and “swapped variables” pitfalls by offering the alternative sequence `STAT → TESTS → LinRegTTest`.

Another example from Lesser (2007): suppose a teacher is teaching about cumulative distribution functions (cdf) using a graphing calculator and asks students to use `2nd → DISTR → normalcdf` to find the probability that a score on a normally-distributed standardized test is below the mean by more than 1 standard deviation. A pitfall here is

that the typical textbook definition of cdf is an area with an upper or right bound only, but the calculator is expecting a lower bound and upper bound.

Beyond awareness of pitfalls and how to address them with students proactively or retroactively, TPSK allows teachers to discern which technology best illustrates a concept. For example, is the concept of simulation (of die rolls) best done using the TI-84 command `PRB→randInt`, the TI-84 application ProbSim, an Internet applet, the Probability Explorer software, etc.? Is the best way to have students explore the effect of outliers on regression slope and correlation coefficients by using an applet such as illuminations.nctm.org/ActivityDetail.aspx?ID=82? (Are teachers prepared to explain to offer ready explanations to confused students who can so readily generate examples where adding a single outlier completely changes the value and sign?) Is an applet such as standards.nctm.org/document/eexamples/chap7/7.4/index.htm the best way to have students explore the meaning of “least squares”? How do statistics instructors best respond to the White Box/Black Box Principle for using Symbolic Computation Software (Buchberger 1989), which recommends introducing technology only after students study the underlying formulas and concepts? Lesser (1999b, 2004) gives us a good look at these in the context of line of best fit. Is it better to introduce the steps for calculating a standard deviation as a spreadsheet-like display of sequential labeled columns, as Vogt (2007, p. 21) does, or verbally and then a single formula as Utts (2005) does?

How do teachers best exploit the interactive nature of certain applets? One example of such an applet is “Plop It!” (at www.shodor.org), which explores how mean, median and mode change as the student chooses additional values to add to the picture. Other applets have sliders which allow the student to continuously vary, say, the bin width of a histogram and see how that impacts its shape and mode(s). Lesser and Melgoza (2007) note an applet with sliders which seems effective for introducing students to the distinction between within-group variation and between-group variation and their interplay in a one-way ANOVA: www.psych.utah.edu/stat/introstats/anovaflash.html.

Ideally, people teaching college statistics have substantial experience with both theory and practice of statistics, along with TPSK. It is hoped that this paper has started to identify ideas, issues and pitfalls to inform college instructors still in the process of acquiring TPSK. In addition to continuing to build their own TPSK, college instructors need to decide upon the type of instructional sequences and curricula that will best help prospective secondary teachers acquire TPSK, in a spirit similar to Niess (2006) for TPCK. Secondary teachers are quite likely to be asked to teach statistics, given the increasing demand for high school statistics courses (it is becoming increasingly common that students are required to take 4 years of high school math), as well as the increased demand for AP Statistics. Helping secondary teachers acquire TPSK to the greatest extent possible will require examining (and possibly revising) existing teacher preparation sequences to meet the challenge of helping teachers acquire this increasingly important aspect of the knowledge base for teaching statistics.

Acknowledgment: We appreciate the feedback of Ball State’s Rebecca Pierce.

References

- American Statistical Association (2005), "Guidelines for Assessment and Instruction in Statistics Education," <http://www.amstat.org/education/gaise>
- Baker, J. (2004), "The Charts that Excel Cannot Do," *Spreadsheets in Education*, 1(3) <http://epublications.bond.edu.au/ejsie/vol1/iss3/6/>.
- Buchberger, B. (1989). "Should Students Learn Integration Rules?" *SIGSAM Bulletin*, 24(1), 10-17.
- Christie, D. (2008), "Extracting Data off the Internet," *Teaching Statistics*, 30(1), 23-5.
- Groth, R. E. (2007), "Research Commentary: Towards a Conceptualization of Statistical Knowledge for Teaching," *Journal for Research in Mathematics Education*, 38(5): 427-437.
- Hunt, N. and Mashhoudy, H. (2008), "The Humble Pie – Half-Baked or Well Done?," *Teaching Statistics* 30(1), 6-12.
- Lee, T. (2005), "Fool Me Twice, Shame on Me," *Mathematics Teacher*, 98(8), 560-3.
- Lesser, L. (1999a), "Exploring the Birthday Problem with Spreadsheets," *Mathematics Teacher*, 92(5), 407-411.
- Lesser, L. (1999b), "The 'Ys' and 'Why Nots' of Line of Best Fit," *Teaching Statistics*, 21(2), 54-55.
- Lesser, L. (2004), "Translated Least Squares," *Mathematics Teacher*, 98(2), 68-69.
- Lesser, L. (2007), "Using Graphing Calculators to Do Statistics: A Pair of Problematic Pitfalls," *Mathematics Teacher*, 100 (5), 375-378.
- Lesser, L. and Melgoza, L. (2007), "Simple Numbers: ANOVA Example of Facilitating Student Learning in Statistics," *Teaching Statistics*, 29(3), 102-105.
- Martin, M. A. (2003), "It's Like... You Know": The Use of Analogies and Heuristics in Teaching Introductory Statistical Methods," *Journal of Statistics Education* 11(2): www.amstat.org/publications/jse/v11n2/martin.html
- Mishra, P. & Koehler, M. J. (2006), "Technological Pedagogical Content Knowledge: A Framework for Teacher Knowledge," *Teachers College Record*, 108(6), 1017-54.
- Moore, D. S. (1988), "Should Mathematicians Teach Statistics?" *College Mathematics Journal*, 19(1), 3-7.
- National Research Council (2000), *Educating Teachers of Science, Mathematics, and Technology: New Practices for a New Millennium*. Washington, D.C.: National Academy Press. Available: <http://books.nap.edu/catalog/9832.html>.
- Niess, M. L. (2005), "Preparing Teachers to Teach Science and Mathematics with Technology: Developing a Technology Pedagogical Content Knowledge," *Teaching and Teacher Education* 21, 509-523.
- Niess, M. L. (2006), "Preparing Teachers to Teach Mathematics with Technology," *Contemporary Issues in Technology and Mathematics Teacher Education*, 6(2). www.citejournal.org/vol6/iss2/mathematics/article1.cfm
- Shulman, L. S. (1986), "Those Who Understand: Knowledge Growth in Teaching," *Educational Researcher* 15, 4-14.
- Utts, J. M. (2005), *Seeing Through Statistics* (3rd ed.), Belmont, CA: Thomson Brooks/Cole.
- Vogt, W. P. (2007). *Quantitative Research Methods for Professionals*. Boston: Pearson.