TOWARDS MEASURING TECHNOLOGICAL PEDAGOGICAL CONTENT KNOWLEDGE IN STATISTICS: MIDDLE SCHOOL TEACHERS USING GRAPHING CALCULATORS

M. Alejandra Sorto
Texas State University—San Marcos, USA

Lawrence M. Lesser
The University of Texas at El Paso, USA

This paper discusses the process of designing and empirically testing measures of teachers' technological pedagogical statistical knowledge (TPSK), addressing conference theme # 3 -- next steps in statistics education with respect to preparing teachers to implement new (usually, technology-rich) curricula. The authors conceptualize this knowledge in the broader context of technological pedagogical content knowledge and provide some specific examples of items applied to statistics. In creating measures to capture this complex construct, the authors drew upon literature of students' (mis)conceptions, authentic teaching experiences, and alignment with statistical expectations for middle school teachers. Results from a pilot survey of middle school teachers suggest that some items have potential to uncover aspects of teachers' (mis)conceptions associated with learning statistics with common graphing calculator technology.

INTRODUCTION

The *PreK-12 GAISE Report* (American Statistical Association, 2005a), states: "Advances in technology and modern methods of data analysis in the 1980s, coupled with the data richness of society in the information age, led to … introducing statistical concepts into the school curriculum as early as the elementary grades." (p. 3) The *College GAISE Report* (ASA, 2005b) states that technology advances make procedures more accessible, which means teachers 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. Thus, the PreK-12 and College GAISE reports both document the importance to pre-service teachers of technology in teaching and learning statistics.

ASA (2005b) not only lists many types of technologies, but also insists that teachers be able to select the most appropriate tool. To do this, teachers need not only competence in each of these separate technologies, but also technological pedagogical content knowledge (TPCK). TPCK is an area recently gaining much attention from education scholars (e.g., AACTE, 2008). This paper aims to report results of initial work towards developing an empirical instrument to assess the TPCK of pre-service secondary teachers with respect to using graphing calculators to do (or teach) statistics.

APPLYING TPCK TO STATISTICS

Niess (2005) articulates four main components of TPCK: "an overarching conception of what it means to teach a particular subject integrating technology in the learning; knowledge of instructional strategies and representations for teaching particular topics with technology; knowledge of students' understandings, thinking, and learning with technology in a particular subject; knowledge of curriculum and curriculum materials that integrate technology with learning in the subject area" (p. 511).

Mishra & Koehler (2006) and Niess (2005) articulate TPCK as the intersection and interaction of its components. In applying this to statistics (Figure 1), Lesser & Groth (2008) introduced the phrase "technological pedagogical statistical knowledge (TPSK), rather than saying "TPCK in statistics education".

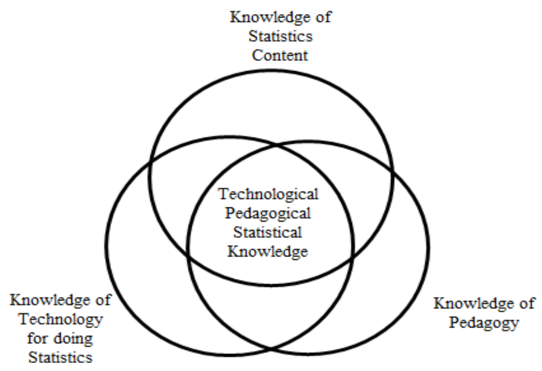


Figure 1. Technological Pedagogical Statistical Knowledge (from Lesser & Groth, 2008)

DEVELOPING MEASURES

Measuring teacher knowledge, let alone knowledge that involves *technology*, is a complex task and there are difficulties in determining what and how to measure (Ball, Lubienski, & Mewborn, 2001; Lee & Hollebrands (2008)). Projects such as Learning Mathematics for Teaching (LMT) at the University of Michigan and the COACTIV project in Germany have developed measures of mathematics for teaching, but these measures do not include technological aspects.

Building on these efforts, the researchers began to write in fall 2007, and later pilot test, six (mostly multiple-choice) items (see Table 1) intended to reflect TPSK for teaching middle school. The process began by specifying desired domains of teachers' statistical and technological knowledge. Informed by the literature, the authors chose to develop items in the domain of knowledge of students' misconceptions or errors when teaching statistics with technology. This domain includes errors that students make when using technology that obscure the underlying statistical concept and lead to wrong conclusions. For example, item 1 appears in Lesser (2007) and item 3 connects to ASA (2005a). Next, the researchers decided which topics or concept to measure after careful analysis of guidelines from NCTM (2000, 2006) and ASA (2005a). At each stage of the item development process, an external expert assessed the content validity of items, taking into account clarity, correctness, classification within the framework subdomain, and relevance to teacher education.

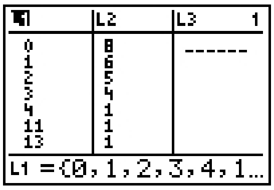
Table 1. Description of items in pilot survey of middle school teachers

Item #	TPSK the item was intended to assess
1	"Switched variables" pitfall of regression
2	Awareness of when data will be within viewing window
3	Connecting mean/median appropriately to numerical and graphical summaries
4	Distinguishing sample and population variance (from 1-Var Stats output)
5	Interpreting frequencies (when going from a given frequency table to the boxplot option from the STATPLOT menu)
6	Knowg whether or not sampling was (and should be) done and without replacement (using a TI-84 command that generates random numbers) in the specific context of a typical 6-ball lottery drawing

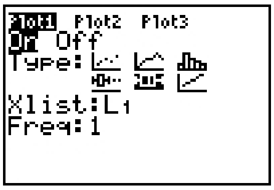
Once specifications were identified, the authors wrote items whose face validity was confirmed by an external statistician's review. Open-ended items were written with the aim of converting them later into multiple-choice, with choices informed by respondents' answers. Items were embedded in authentic teaching situations based on mainstream modern curricula (e.g., *Mathematics in Context: Dealing with Data*) as well as the authors' experiences teaching statistics for pre-service teachers, giving professional development, and making classroom observations. A sample TPSK item appears in Figure 2.

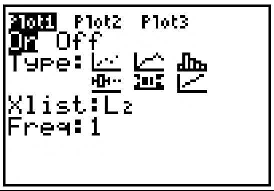
Suppose the students take their TI-84 and enter the data shown on the left into L1 and L2, respectively, as shown in the screenshot:

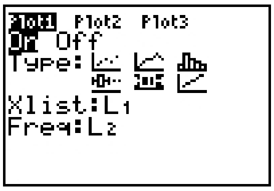
Number of Cans of Soda 8th-Graders Drink per Day	Number of 8th-Graders
0	8
1	6
2	5
3	4
4	1
11	1
13	1

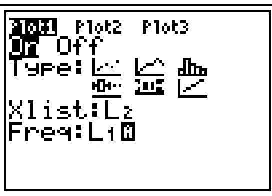


Which of the following screenshots from the STATPLOT menu will produce the appropriate boxplot for number of cans of soda eighth-graders drink per day? (For each item below, circle YES, NO, or I'M NOT SURE.)

a)


b)


c)


d)


	Yes	No	I'm Not Sure
a)	1	2	3
b)	1	2	3
c)	1	2	3
d)	1	2	3

Figure 2. Sample item (item #5 from the pilot study) measuring TPSK.

SAMPLE

Some universities in the US have only recently created new programs for middle school teacher, and so annual enrollment for this population in a large university may be only about 30 students. This study's sample comprised all 20 teachers enrolled in a Master of Education in Middle School Mathematics Teaching degree program in which a large number of courses are delivered online. The degree resides in the Mathematics Department at a Doctorate-granting public university in the south-central United States has more than 25,000 students (30% ethnic minority) and produces the most certified teachers in the state.

Table 2. Demographics of Participants

AREA		DISTRIBUTION	
Gender	female (16)	male (4)	
Grade Level Band	K–4th grade (2)	4^{th} – 8^{th} grade (12)	6^{th} – 12^{th} grade (6)
Instructional area	all areas (6)	mathematics (17)	science (2)
College-level statistics	1 course (11)	2 courses (8)	3 or more courses (1)
Statistics Pedagogy	0 courses (13)	1 course (6)	2 or more courses (1)
Professional development in statistics in past year	none (15)	1-5 hours (4)	6 – 15 hours (1)

Data was collected in summer 2008 after teachers completed an online introductory probability and statistics course. The class had typical topics coverage and was taught by an instructor not on the research team. The course incorporated activities that required the use of technology such as the TI-84+ graphing calculator and *Fathom* (but technology was not allowed on the pilot survey). Most of the activities had connections with major middle school statistics and probability curricula.

RESULTS

Item 1 was an open-ended item asking what could have caused two students to obtain different equations for line of best fit if they both used a LinReg command after correctly entering the data. Both authors independently coded the participants' responses and identified five different types of responses, which are listed (along with their respective frequencies) in Table 3. The codes were a mixture of "descriptive codes" and "interpretive codes" (Miles & Huberman, 1994), and the very few discrepancies between the authors' codes were all reconciled through discussion. These responses informed the development of a forced-choice item for a subsequent survey. Just under 2/3 of the teachers (65%) identified correctly that the reason the two linear models differed was because variables were swapped in the LinReg command syntax. Surprisingly, the most common incorrect response was to say that students simply mistyped something, though the item states explicitly that data were typed in correctly. These teachers may not be aware what types of errors are "just typos" and what errors reveal conceptual misconceptions (e.g., about regression model variables' asymmetry; see Lesser, 2007).

Table 3. Responses for Item 1

Categories of participants' open-ended responses	Frequency of responses (n = 20)
The variables were swapped.	13
Command keys or data were typed in incorrectly.	3
Difference caused by how data was sorted after it was entered.	1
Difference caused by using different LinReg command options ($a + bx$; $ax + b$)	1
"I don't use this with my students" or "no idea"	2

Item 2 asked why the GRAPH button could produce a window that showed the line of fit but not the dataset. This open-ended item yielded greater variation of responses (see Table 4), which the researchers grouped into two main categories. Category #1 is a description of what the student did wrong and most teachers in this category correctly said that the window settings were not appropriate. Category #2 is a suggestion of how to fix the problem under the assumption it is clear what the student did wrong. In this category, teachers suggested a variety of solutions-- some appropriate and some not. A few teachers fall into both categories -- describing what the hypothetical student in the problem did wrong and how she can fix the problem. These responses informed the development of a forced-choice item for a subsequent survey.

Table 4. Responses for Item 2

Participants' open-ended responses	Category	Frequency
Window settings are not appropriate for the data	1	7
The graph plotted does not allow to see the data because the data lie exactly on the line	1	1
Did not use command that display both graphs	1	1
Check for StatPlot command activated	2	2
Use ZOOM command to fix the problem	2	2
Enter 2nd Calc [sic] command to see the data values	2	2
Move the cursor or use the trace command to see the data	2	3
Graph the Scatterplot first and then graph the line	2	1
No answer	n/a	1

Item 3 asked the teachers to select all of the screenshots (of various numerical and graphical summaries) which would be relevant to a lesson about the mean and the median of a given dataset. The choices of screenshots to illustrate the mean and the median are shown in Table 5 below. The result that stands out is the choice of all teachers for the box plot. The justifications for their choices fall into two categories. Some teachers chose the screenshots based on the information in relation to the mean and median. Others justified in relation to the middle grades curriculum. For example, some teachers said that they would not use the screenshot that shows the numerical mean because it is "too complicated for 7th graders." Teachers did not specify what made it complicated, but the research team conjectures that the $\bar{x}$ and Σ notation made it appear beyond middle school. Though the screenshots were created on a TI-84, it should be noted that middle school students commonly use the TI-73 graphing calculator and its 1-Var Stats option yields output that also uses the *x*-bar and summation notation. We note that teachers underrecognized the potential of estimating mean (or median) from a histogram (e.g., they may not know the mean as a histogram's "balance point"). And, of greater concern, some teachers found value in some of the extraneous screenshot choices that the researchers included (e.g., the arbitrary "time series plot"). To give this item greater focus for a subsequent survey, the question asked about the mean only, and the third and seventh screenshots were removed.

Table 5. Responses for Item 3 (multiple answers allowed per person)

Description of screenshot	Frequency
First screen of 1-Var Stats output (includes $\bar{x}$)	12
Second screen of 1-Var Stats output (includes Med)	16
Histogram of data	9
Boxplot of data	20
Scatterplot of data values versus their (arbitrary) position in the dataset	5
Time series plot (a connected line graph of the preceding scatterplot)	3
Normal probability plot (of data values versus the corresponding quantile of the standard normal distribution)	6

Item 4 was a forced-choice item in which the output from 1-Var Stats was shown and the teacher is asked how to respond to a student asking whether s_x or σ_x is the "right answer." The item explicitly said to choose ONE answer, and the response of the one teacher who exceeded this was discarded. As Table 6 indicates, the teachers' modal response is correct, but almost as many chose that both standard deviations were correct. This suggests that teachers may not be clear that there is never a situation in which both are correct answers at once, and this is a major concern. The two students who chose the population standard deviation, however, may have exhibited an honest, less serious confusion about whether the data were to be viewed as a sample from a (larger) population, and this was made more explicit in the revised instrument for a subsequent survey.

Table 6. Responses for Item 4

Fixed answer choice selected	Frequency
You made a mistake when entering the data; they both should be the same.	1
They are both acceptable in this case.	7
s_x is the right answer	9
σ_x is the right answer	2

Item 5 (see Figure 2) involved choosing which settings and options would produce the appropriate boxplot to represent a set of data given as a frequency table. This item had a format different from the other items because it allowed teachers to choose more than one screenshot even though there is only one correct answer and also gave them the option to say “I’m not sure”. That last option was intended to reduce the frequency of guessing. One interesting result is that out of the 11 that chose the correct screenshot (option c), three also chose the screenshot with frequency equal to 1 (option a), and two were unsure about the screenshot that shows them the list swapped (option d). Another interesting result from Table 6 is that the number of “I’m not sure” selections increased when Freq had a column of values other than 1. In the revision of this survey, item 5 became as a forced-choice question and all the questions (not just item #5) asked the respondent to indicate a “level of confidence.” This should allow better assessment of the extent or effect of guessing.

Table 7. Responses for Item 5

Settings on screenshot (see Figure 2)	Yes	No	I’m not sure
a) Xlist: L1 ; Freq: 1	6	10	2
b) Xlist: L2 ; Freq: 1	2	14	3
c) Xlist: L1 ; Freq: L2	11	4	4
d) Xlist: L2 ; Freq: L1	1	13	5

Item 6 asked teachers to select which set(s) of numbers would be reasonable outputs for drawing a set of six lottery balls. One of the six-number sets had a repeated number. This item seemed to play out the most clear to teachers in terms of the high percentage (80%) who recognized that the set with a repeated number was inconsistent with a drawing without replacement (which a typical state or multi-state lottery drawing is), but without narrative data, it’s hard to know how to interpret those three teachers who said ‘yes’ (and one said ‘unsure’) for that set. The researchers conclude that because this item did not ask for an accompanying explanation, it is hard to tell if teachers’ answers were simply a function of their level of familiarity with the setting of lotteries. Perhaps some teachers were unaware that lotteries involve a “without replacement” feature and that it might have made a difference if the question said “a six-ball drawing” instead of “drawings of six lottery balls.” Perhaps this item is not assessing technology, but rather just whether they understand the real-world context of lotteries. Also, Item Response Theory suggests that a question where almost everyone answers correctly is of little use for discriminating between those with high TPSK and low TPSK. For these reasons, it was decided to delete the question in future survey versions.

In the hope of identifying factors that might relate to TPSK, questions about educational and professional background such as number of statistics courses taken, level of certification, etc.

were included at the end of the instrument, but no statistically significant relationships were found between these variables and level of TPSK.

FUTURE DIRECTIONS

As frequently noted, the results of the pilot survey led to many refinements of individual items and the revised instrument was administered to a new set of (n_2 = 15) middle school teachers who were demographically very similar (except that four were Special Education specialists) and were taking the same class in summer 2009. In general, the pattern of results was quite similar to that from the first pilot study (n_1 = 20), thus suggesting that the convenience of having more items in more streamlined closed form did not result in major loss. On Item 4, fewer students fell for choice 2, but were evenly split between choices 3 and 4, despite the clarification in the statement of the problem to view the data as a sample from a larger population. Further analysis of the second pilot study is continuing. The authors are also in the process of adding several new survey items to span technologies in addition to graphing calculators (e.g., statistical packages and spreadsheets). Testing an instrument with more items on more people will support computing reliability statistics and a factor analysis to identify any distinct components of TPSK. It is conjectured that items will vary in their blend of statistics content, technology knowledge, and pedagogical knowledge. The authors also conjecture that years of teaching experience, statistics coursework, and level of understanding may help predict level of TPSK.

The authors also envision contributing to efforts to measure teacher knowledge by using these items as an independent measure or sub-scale within a larger instrument. Statistics and technology appear to have been absent in existing instruments to date such as Content Knowledge for Teaching Mathematics Measures developed by the LMT Project (Hill, Schilling, & Ball, 2004). Adding this unique construct to preexisting measures will allow teacher educators to assess more comprehensively their pre-service and in-service programs.

REFERENCES

American Association of Colleges for Teacher Education (2008). *Handbook of technological pedagogical content knowledge*. New York: Routledge.

American Statistical Association (2005a). *Guidelines for assessment and instruction in statistics education (GAISE) report: A pre-K-12 curriculum framework.* Alexandria, VA: ASA. [www.amstat.org/education/gaise/]

American Statistical Association (2005b). *Guidelines for assessment and instruction in statistics education (GAISE) college report*. [www.amstat.org/education/gaise/]

Ball, D. L., Lubienski, S. T., & Mewborn, D. S. (2001). Research on teaching mathematics: The unsolved problem of teacher's mathematical knowledge. In V. Richardson (Ed.), *Handbook of research on teaching* (4th ed., pp. 433-453). Washington, DC: American Educational Research Association.

Encyclopaedia Britannica Educational Corporation (1998). *Mathematics in context*: *Dealing with data*. Chicago: Encyclopaedia Britannica Education Corp.

Hill, H. C., Schilling, S. G., & Ball, D. L. (2004). Developing measures of teachers' mathematical knowledge for teaching. *Elementary School Journal, 105*, 11-30.

Lee, H., & Hollebrands, K. (2008). Preparing to teach mathematics with technology: An integrated approach to developing technological pedagogical content knowledge. *Contemporary Issues in Technology and Teacher Education*, *8*(4). http://www.citejournal.org/vol8/iss4/mathematics/article1.cfm

Lesser, L. (2007). Using graphing calculators to do statistics: A pair of problematic pitfalls. *Mathematics Teacher*, *100*(5), 375-378.

Lesser, L., & Groth, R. (2008). Technological pedagogical content knowledge in statistics. Paper presented at the twentieth International Conference on Technology in Collegiate Mathematics, San Antonio. http://archives.math.utk.edu/ICTCM/i/20/S118.html

Miles, M. B., & Huberman, A. M. (1994). *Qualitative data analysis: An expanded sourcebook* (2nd ed.). Thousand Oaks, CA: Sage Publications.

Mishra, P., & Koehler, M. J. (2006). Technological pedagogical content knowledge: A framework for teacher knowledge, *Teachers College Record*, *108*(6), 1017-54.

Mishra, P., & Koehler, M. J. (2008). Introducing TPCK. In *Handbook of technological pedagogical content knowledge (TPCK) for educators* (pp. 1-30). New York: Routledge.

National Council of Teachers of Mathematics (2000). *Principles and Standards for School Mathematics*. Reston, VA: NCTM.

National Council of Teachers of Mathematics(2006). *Curriculum focal points for prekindergarten through grade 8 mathematics: A quest for coherence*. Reston, VA: NCTM.

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.