TRACKING THE DEVELOPMENT OF STATISTICS KNOWLEDGE IN INTRODUCTORY STATISTICS ®

Mark A. Earley
Bowling Green State University
USA

Discussions of students' understandings of key course concepts typically investigate those understandings at one point in time. This paper reports results from a case study in which eighteen graduate students were interviewed throughout a fifteen-week introductory statistics course. Knowledge structures were assessed once every three weeks, and changes in these structures were discussed. One key finding was that students' understandings of certain course concepts change as the semester progresses, indicating that it may not be enough to assess these understandings at only one point in time. Two concepts in particular, mean and variance, are central to many ideas in the introductory course. Assessing how students "know" these concepts throughout the course may be beneficial to researchers and educators alike. Implications for both statistics education research and teaching introductory statistics are offered.

INTRODUCTION

One of the major outcomes statistics educators should be emphasizing is *understanding* course concepts as opposed to memorizing them or blindly applying formulas with no meaning attached to the results. As a result, assessment practices should be changing to reflect this emphasis: we should be assessing *how* students know statistics, not just *what* they know. Research in statistics education should also be emphasizing understanding, and we should be focused on assisting statistics educators in the development of assessment tools that will tap into this type of knowledge. One way to examine *how* students know statistics is to look at how students relate various course concepts together, what Schau and Mattern (1997) call "connected understanding" (p. 91).

Cognitive psychologists have presented three types of knowledge students may have or acquire during any course: declarative knowledge (knowledge of facts or definitions), procedural knowledge (knowledge of how to complete a process), and structural knowledge (knowledge of the interrelationships among pieces of declarative and procedural knowledge) (Anderson, 1990, 1996; Byrnes, 1992; Jonassen, Beissner, & Yacci, 1993). Understanding can be defined and measured by this third type of knowledge, structural knowledge. Under this guise, a student's understanding of a set of course concepts can be described by how he or she relates these concepts to each other and to other concepts or examples they may generate. Any diagram or pictorial representation of these relationships (e.g., a concept map, hierarchical tree, or knowledge network) would be one way to capture a student's understanding at that particular point in time (Olson & Biolsi, 1991).

Conceptual change is another area of interest to cognitive psychologists. In very general terms, conceptual change is marked by shifts in student understandings, or more specifically by shifts in the relationships students describe among the various concepts they are studying. Accretion has been described as the simple addition of new concepts to existing knowledge structures, with no real attempt to modify the structure in any other way. This is similar to Piaget's notion of assimilation (Chapman, 1988). Revolution is the term most consistently applied to a change in the knowledge structure besides simple addition of concepts – the relationships previously found in the structure change (similar to Piaget's notion of accommodation, Chapman, 1988).

Introductory statistics is a course ripe for conceptual change for a number of reasons. First, many students taking the course are new to the formal learning of statistics. They most certainly bring prior experiences to the classroom, but many researchers have shown these experiences are not always "statistically correct." Second, even if students have had a prior course in statistics, they frequently report forgetting most of that material. Third, introductory statistics is often taught in a way that demonstrates how concepts build on each other: knowledge of earlier concepts is essential to learning later concepts. It is this third reason in particular that emphasizes the need for statistics educators to focus on how students' understandings of these earlier concepts (e.g., the mean and variance) change as the course progresses.

METHODS

Eighteen graduate students enrolled in an introductory level educational statistics course (nine each in two consecutive semesters) volunteered to meet with the researcher five times during the semester, once every three weeks. (Note the researcher was *not* the course instructor). At these sessions, the student completed a variety of structural knowledge activities. Each of these activities was then discussed in an audiotaped clinical interview. Interviews were transcribed for analysis. This paper focuses on one of the tools used during these sessions: the card sorting task.

For this task, students are presented with a stack of 3" x 5" index cards. Each card has a different course concept printed on it. Concepts used in this study were selected by the researcher in consultation with the course professor. At each session, new concepts that had been discussed since the previous session were added to the pile. The first session started with ten concepts covered in the first three weeks of the semester. By the end of the semester, a total of forty-five concepts were in the stack. Students are asked to sort the concepts into piles based on relatedness. Cards in the same pile should be more related to each other than they are to cards not in the pile. The student is left to generate his or her own definition of "relatedness." Students are asked to create at least two piles (i.e., do not just leave all the concepts in one pile and say they are all related), but they may have as many piles as there are cards if the student feels each concept belongs on its own. No time limit is given, and the student is left alone to complete the task.

Once the student is finished, the researcher begins drawing a diagram indicating the final set of piles and which concepts are sorted into each. The student and researcher then discuss the sets of cards, one pile at a time. The researcher generally begins with a question such as, "What do you have here?" or "Tell me what you've got." An attempt was made to discuss each pile, and each concept in each pile, but this was not always accomplished. The resulting diagram is meant to be a summary of that student's knowledge structure at that point in time.

DATA ANALYSIS

Two sets of data are generated from this task: five sort diagrams and the transcript of the discussion. Both sets were used together to summarize each student's knowledge structure. Three different analyses were completed: (a) a summary of each student's structure at that point in the course, (b) a summary of changes in that student's structure from the last session to the current session, and (c) a comparison of all eighteen students' structures. The focus of this paper is on a comparison of the changes made by each of the eighteen students across the entire semester (Jones & Vesilind, 1996). It is impossible to present results from such an analysis for all forty-five concepts included, so this discussion will focus on two specific concepts: the mean and variance.

To track the placement of each concept, I created a grid with each row representing one participant and each column representing the chronological order of the sort. Out of each participant's overall sort diagram, I pulled out the pile in which he or she had sorted the mean (for one grid) and the variance (in a second grid). Next to each pile, I placed the quote or quotes from the transcript in which the participant and I discussed that pile and what it meant.

RESULTS

The Mean

There are two distinct trends in the movement of the concept of the mean from week to week. One group of participants kept the concepts of mean, median, and mode isolated throughout the semester. They always described this pile as being "measures of central tendency" or "averages." This was common in the first sort session (Week 3), as these were really the only main concepts students had learned. One of the assumptions for this study was that as new concepts were introduced, such as the sampling distribution of the mean and hypothesis testing for the mean, that students would see more concepts as related to the mean. This was not obviously so for this first group. They continued to sort the mean just with median and mode, never with any other concepts. About half of the students fell into this category. This does not mean they did not see the relationships, however. What is does indicate is that the students' first choice (potentially the strongest linkages for these students) was to associate the mean as a measure of central tendency, a descriptive statistic similar to the median or the mode.

A second group of participants did allow for the mean to move to other piles during the remainder of the semester. One student, Paul, wanted to place the mean above everything else in his sort, indicating that "it all comes down to this" in statistics. Others indicated that "we're not seeing the median and mode as much any more, just the mean," and so felt the original ties to these concepts were weaker now than they were during the first three weeks of the course. These students related the mean to a variety of concepts, but the most common placement was in a pile with some hypothesis testing concepts.

Variance

Participants seemed much more consistent in their placement of the concept of variance; in part, this was because they really did not have much to say about it. Most frequently, variance was associated with standard deviation, as a descriptive statistic. As the course progressed, few participants included any consideration of variance or standard deviation in piles they created for hypothesis testing. A third concept, standard error of the mean, was included instead. However, no participant related this third concept to variance. One participant did relate variance to population, stating that "there has to be variance in the population before you can do hypothesis testing" – she was the exception rather than the rule. Whereas about half of the participants were able to see the mean as related to future concepts, only five students ever sorted variance in a pile different from their earlier sorts. There are many ways to speculate on why this is so, and future research should explore this more with students through more in depth interviewing.

CONCLUSIONS

The first clear data to emerge from this study are the variety of ways in which students can "know" statistics. Even though this is probably evident to many who teach the course, this is one of the few studies to empirically show this variety and discuss it in any detail. Discussing what to do with this variety of understandings, both in research and the classroom, is less clear. Some of my own observations will be presented, in the hopes that this discussion will continue in further research.

For statistics education researchers, I believe there are at least two important implications of this study. First, it is important for us to continue examining *qualitative* ways in which students understand our course material. Others have done so with specific concepts such as the mean (Pollatsek, Lima, & Well, 1981) or correlation (Ross & Cousins, 1993), but no attempt has been made to integrate these concepts with others in the course. If we as educators expect students to put the material together and see how each set of concepts builds on previous concepts, then we as researchers should be investigating (a) how this happens, and (b) when this happens (Friel, Bright, Frierson, & Kader, 1997).

This introduces the second implication for research, which is the need for more longitudinal studies in statistics education. How our students "know" statistics at one point in time is not enough – we must also learn how they use this knowledge later in the course and how they modify this knowledge as the course progresses. Tracking the development of introductory statistics students' knowledge is a profoundly interesting topic that I hope to explore in future research, and I urge other statistics educators to do the same. Longitudinal studies will also help us as researchers help educators with another implication from this study, which is course assessment.

In some college courses, such as writing, students are graded as much on improvement as on what they demonstrate at any particular point in time. An essay written in the tenth week of the semester, for example, will be graded not only on its individual merits but also on how much growth the student demonstrates. Introductory statistics is another course where this type of assessment should be considered. Very often, we ask students to learn concepts and then immediately apply them to new situations, and then know them well enough to build on them in subsequent chapters. Would it be possible, then, to grade students not only on the knowledge they demonstrate at any particular point in the semester, but also on whether or not that knowledge shows improvement over prior assessments? For example, if a student does not immediately learn or understand what a sampling distribution is, would it be possible for him or her to learn that later in the course? I believe so, in which case we should be continuing to offer these students opportunities to show when they have learned concepts.

As seen in this study, it is possible for students to know a statistics concept in many different ways. Handling and organizing these different "ways of knowing" statistics may be the challenge our students face in the classroom. Just when they feel they are comfortable with a concept, we put a whole new spin on it and present them with some cognitive challenges to overcome. In other research, and in my own classroom, students frequently report that their day-to-day anxiety comes from these new situations in which they feel they have to start all over with a particular concept. Statistics education researchers have a powerful role here, both in finding ways to capture these challenges and finding ways to help statistics educators meet and help their students overcome these challenges.

ACKNOWLEDGEMENTS

The research described here was completed in 2001 as part of the author's dissertation research at the University of Toledo, Toledo, Ohio, USA. The results presented here were not discussed in the final manuscript.

REFERENCES

Anderson, J.R. (1990). *The adaptive character of thought.* Hillsdale, NJ: Lawrence Erlbaum Associates.

Anderson, J.R. (1996). ACT: A simple theory of complex cognition. *American Psychologist, 51,* 355-365.

Byrnes, J.P. (1992). The conceptual basis of procedural learning. *Cognitive Development, 7,* 235-257.

Chapman, M. (1988). *Constructive evolution: The origins and development of Piaget's thought.* Cambridge: Cambridge University Press.

Friel, S.N., Bright, G.W., Frierson, D., & Kader, G.D. (1997). A framework for assessing knowledge and learning in statistics. In I. Gal and J.B. Garfield (Eds.), *The Assessment Challenge in Statistics Education* (pp. 55-63). Amsterdam: IOS Press.

Jonassen, D.H., Beissner, K., & Yacci, M. (1993). *Structural knowledge: Techniques for representing, conveying, and acquiring structural knowledge.* Hillsdale, NJ: Lawrence Erlbaum Associates

Jones, M.G., & Vesilind, E.M. (1996). Putting practice into theory: Changes in the organization of preservice teacher pedagogical knowledge. *American Educational Research Journal, 33,* 91-117.

Olson, J.R., & Biolsi, K.J. (1991). Techniques for representing expert knowledge. In K.A. Ericsson and J. Smith (Eds.), *Toward a general theory of expertise: Prospects and limits*. Cambridge: Cambridge University Press.

Pollatsek, A., Lima, S., & Well, A.D. (1981). Concept or computation: Students' understanding of the mean. *Educational Studies in Mathematics, 12,* 191-204.

Ross, J.A., & Cousins, J.B. (1993). Patterns of student growth in reasoning about correlational problems. *Journal of Educational Psychology, 85,* 49-65.

Schau, C., & Mattern, N. (1997). Assessing students' connected understandings of statistical relationships. In I. Gal and J.B. Garfield (Eds.), *The assessment challenge in statistics education* (pp. 91-104). Amsterdam: IOS Press.