Jl. of Computers in Mathematics and Science Teaching (2006) **25**(1), 41-75

Fostering Mathematical Discourse in Online Asynchronous Discussions: An Analysis of Instructor Interventions

LINDA SIMONSEN
JEFF BANFIELD
Montana State University
USA
simonsen@math.montana.edu
banfield@math.montana.edu

The purpose of this article is to describe a typology of instructor discourse interventions to enhance mathematical discourse in online asynchronous discussions in a statistics course for teachers. A completely asynchronous model of distance delivery was used for the course. Emphasis was given to fostering mathematical discourse in that the students were encouraged to state their ideas, elaborate on their thinking, and compare their ideas with previously shared ideas. Using the constant comparative methodology for developing theory, the analysis of asynchronous discussions revealed 5 recurring categories of response with respect to instructor intervention: (a) resolve, (b) validate, (c) redirect, (d) expand, and (e) withhold. With respect to the identified typology, the authors recommend that the default action on the part of the instructor be to withhold, using one of the other interventions only if necessary. Withholding encourages the students to lead and become active participants in the discussion, thus enhancing mathematical discourse. The article also discusses implications of the findings and recommendations for future research.

INTRODUCTION

Distance learning is defined by the National Science Foundation (NSF) (2002) as education where learning occurs all or most of the time in a different place from teaching, and the principal means of communication between learners and teachers is through technology. The type of course delivery, however, can be vastly different from one distance course to another. Types of distance delivery have included one-way (live and pre-recorded) and two-way video, two-way audio, and asynchronous communication via Internet technologies. One of the fastest growing modes of delivery, and the structure of the course studied in this article, utilizes asynchronous communication. Spiceland and Hawkins (2002) define asynchronous communication as communication that is mediated by technology and is not dependent on instructors and students being present at the same location at the same time. The asynchronous mode of delivery has been credited with promoting access and convenience for diverse learners (NSF, 2002), improving cost effectiveness of education (NSF, 2002), increasing active student participation in the learning process (Spiceland & Hawkins, 2002), and providing a system of documenting construction of knowledge (Gunawardena, Lowe, & Anderson, 1997).

Coinciding with the growth of asynchronous communication, fostering mathematical discourse in the face-to-face mathematics classroom has grown considerably over the past 10 years (Elliot & Kenney, 1996; National Council of Teachers of Mathematics (NCTM), 2000). Fostering mathematical discourse means encouraging students to state their ideas, elaborate on their thinking and compare their ideas with previously shared ideas (Sherin, 2003). Although "genuine" mathematical discourse as defined here is scarce in most face-to-face classrooms (Weiss, 1994), mathematical discourse has been identified as an important component in the learning and understanding of mathematical topics (NCTM, 2000; Steele, 1998). In our experience, mathematical discourse is not scarce in mathematics courses taught online using an asynchronous communication teaching strategy. In fact, when participants take mathematics courses using asynchronous communication, they are virtually forced into engaging in some level of mathematical discourse as a requirement of the course. Yet, little or no research exists on the quality, nature, and impact of such asynchronous mathematical discourse. Though not specific to mathematical discourse, much research is being done in Europe regarding successful asynchronous discussion using Web-based tools such as WebCT, Conferencing on the Web (COW), and Moodle. Most relevant to this study, research has shown that asynchronous discussions can enhance higher-level communication (Jarvela & Hakkinen, 2002), as well as help students to process course information at fairly high cognitive levels (Hara, Bonk, & Angeli, 2000).

Sherin (2003) delineates mathematical discourse in a face-to-face mathematics class into two categories: *content* and *process*. The *content* of mathematical discourse corresponds to "the mathematical substance of the comments, questions, and responses that arise." The *process* of mathematical discourse corresponds to the way teachers and students interact in class discussions of mathematical content. In this article, we focused on both the *content* and *process* of mathematical discourse. As online instructors of the course, we were interested in effectively intervening (*process*) in order to promote mathematical discourse (*content*) in the required asynchronous discussion. Relying heavily on the recommendations regarding the teacher's role in discourse set forth by the National Council of Teachers of Mathematics (NCTM, 1991, 2000), we envisioned the asynchronous discussions to be guided by our constant input, asking questions and generally leading the conversation. However, such was not the case. Although students would discuss the questions that we initiated, it was soon obvious that they had their own topics of interest and discussions followed those topics at least as often as our intended topics. Continued observations of the discussions lead us to examine in detail our "instructor discourse interventions" in the asynchronous discussions. The purpose of this article is to describe a research-based typology of instructor discourse interventions in online asynchronous discussions in a statistics course for teachers.

METHOD

Setting

Four years ago, mathematics educators at our university were challenged with meeting the needs of place-bound mathematics teachers who would be teaching advanced placement (AP) statistics the following school year. The professional development need was to provide teachers with a conceptual understanding of statistics as well as provide practical support for the teaching of AP statistics. The impetus for developing an online statistics course for teachers came from this call as well as the realization that this course would fit nicely into the existing repertoire of online courses for the university's Masters of Science in Mathematics Education (MSME) and Masters of Science in Science Education (MSSE) programs. In the four years since the initial course development, we, the authors, have taught the statistics course for teachers every summer via distance delivery. The course has consistently had a maximum enrollment of 25 students. The students in the course have primarily been secondary mathematics teachers preparing to teach AP statistics or secondary science teachers who are interested in the design and analysis of data-driven classroom experiments.

Procedures

A completely asynchronous model of distance delivery was used for the course using the WebCT course management system. The structure of the course was set up as a series of online, multi-element lessons organized in parallel with the AP statistics curriculum and the required textbook. Most recently, we used the popular AP statistics, though collegiate, textbook *Introduction to the Practice of Statistics* by Moore and McCabe (2000). Each overview was approximately 1 week in length and was associated with assigned readings from the text, review of important points and concepts the students should understand, homework, and projects. The course also incorporated weekly timed quizzes using the quiz feature of WebCT, which allows instructors to write multiple-choice questions that are graded by WebCT. We maintained a strict schedule for turning in assignments and taking part in required asynchronous discussions. More specifically, homework and projects were consistently due on the same 2 days each week, and each student was required to make at least three postings to the asynchronous discussions each week.

The main instructional component of the course was the required asynchronous discussions between students and instructors. The discussions were organized into topics that corresponded to the online lessons. For example, while the students were studying hypothesis testing, they reviewed the chapter readings and discussed assigned homework and projects in the "hypothesis-testing" discussion area.

Discussion postings within each topic were organized as threaded discussions. For example, one student could ask a question about the standard error formula from the text while another student could ask a question about a specific example in the text. More specifically, each post (i.e., the standard error formula or the specific example) and all subsequent responses were organized as one group of postings ordered by time of posting. This organization made it very easy to follow a particular discussion thread and ignore those that were not of interest to the individual. Finally, participants could choose to read a single message or read the entire discussion thread.

All students were required to read all threads and make at least three contributions in each of the topic areas. As a result, all students were involved in every discussion topic, although they may not have participated in every discussion thread.

Data and Analysis

The ongoing and archived asynchronous discussions between the students (practicing teachers) and instructors formed the database for this research. The constant comparative methodology for developing theory served to guide the development of a typology of instructor intervention. The "category construction" discussion in Merriam (1998) provided further assistance in giving structure to the category development. More specifically, during the teaching of the course we continuously looked for key issues, recurrent events, and activities in the asynchronous discussion data that became categories of focus (Bogdan & Biklen, 1992). Categories continuously emerged throughout the data collection period. Emergent categories, based primarily on the data we were gathering, were resolved into a more formal typology at the conclusion of the data collection.

As mentioned earlier, the students were secondary mathematics and science teachers. Course enrollment was limited to a maximum of 25 students; during the 4 years we collected the data, enrollments ranged between 18 and 25. The course ran for 6 weeks each summer. The total number of messages ranged from 1,151 to 1,958 each summer.

Although the categories we will define apply only to a particular response (or non-response) and we may use several types of response in a given discussion thread, the unit of analysis must be the thread. It is only over the course of the entire thread that the need for and consequences of a particular intervention become manifest. In particular, the effect of the non-response that we will advocate must be viewed over the course of the thread.

RESULTS

The analysis of archived asynchronous discussions revealed five recurring categories of response with respect to instructor intervention: (a) resolve, (b) validate, (c) redirect, (d) expand, and (e) withhold. Any discussion thread may have incorporated several different types of instructor interventions. The overriding goal of the instructor interventions was to enhance mathematical discourse in the online discussions. As such, the typology listed is a qualitative tool that can be further developed and eventually be used to evaluate movements towards that goal.

> ***Resolve*** refers to answering a specific and current question from a student that does not warrant class discussion.
>
> ***Validate*** refers to an instructor intervention in which the student has implicitly or explicitly requested clarification or encouragement from the instructor.
>
> ***Redirect*** refers to an instructor intervention that is necessary to refocus the group on the original question or recognize a glaring misconception in the group discussion.
>
> ***Expand*** refers to an instructor intervention that moves the discussion forward, or takes the discussion to a new level.
>
> ***Withhold*** refers to an instructor intervention that is either *non-responsive* in order to let the discussion continue naturally or *minimally responsive* by stating that information is being purposefully withheld in order to see how the discussion develops.

Based on the 2001 course, with 1,151 messages, we responded 134 times. Out of those 134 responses, approximately 28% were *resolve*, 28% were *validate*, 11% were *redirect*, and 33% were *expand*. The *withhold* responses do not, of course, often appear for us to count. However, out of the 1,151 messages, 1,017 were posted by students. If we assume that, perhaps, half of them required no response from us (because approximately half of them may have been responses from fellows students) then that leaves about 500 messages where we might have needed to respond. Therefore, our 134 responses make up about 27% of the total possible response opportunities, which means we *withheld* approximately 73% of the time. The aforementioned percentage is based on speculations only given the numerous unknown factors that exist. We believe, however, that it is a reasonable approximation since discussions are structured in such a way that almost all of the threads are initiated by the students.

DISCUSSION

In this section, we provide an example of each of the categories of the intervention typology. The examples not only provide clarification of the typology categories but also offer a window into the world of asynchronous mathematical discourse.

The first example demonstrates the *resolve* intervention. In this short discussion thread, Mary simply wanted to be told how to read the standard normal probabilities table found in her textbook. In this case, the *resolve* intervention was deemed appropriate as there appeared to be little reason to have a class discussion on this topic. It should be noted, however, that many other students benefited from the intervention that occurred (based on the many reply messages thanking us for the help.)

Message no. 799 posted by Mary: Monday, June 25, 2001 2:13 p.m.

> Could and would someone please remind me what the different columns (.00, .01, .02, ...) stand for in Table A (just inside your book's front cover)? Thanks.

Message no. 801 posted by Simonsen & Banfield: Monday, June 25, 2001 2:26 p.m.

> It is the second decimal place for the z value. For example, in the last row, second column the table value is .4960, the row value is -0.0 and the value at the top of the column is .01 ... all of this means
> $P[Z < -.01] = .4960$
> From the 7th column (headed by .06) of the same row we find
> $P[Z < -.06] = .4761$

The following discussion thread demonstrates the *validate* intervention. In this discussion thread the students were trying to make sense of the general fact, "any linear combination of independent normal random variables is also a normal distribution," found in the textbook on page 401 (Moore & McCabe, 2000). Notice that the instructor intervention occurred only after there was an explicit call for validation in the form of "…somebody out there that really KNOWS for sure if this is right, say something quick before we all get hopelessly confused." It must also be noted that the discussion thread ended once the validation was provided.

Message no. 830 posted by Sandra: Tuesday, June 26, 2001 7:43 a.m.

> Could some one explain the addition of variables on page 401? When they added the means, they subtracted but when adding the variable they added. Any help here would be appreciated.
> *(Two messages of the thread were deleted here. See Appendix A for full discussion thread.)*

Message no. 834 posted by Pam: Tuesday, June 26, 2001 11:31 a.m.

Sandra, I was also having trouble understanding the example. Message 809 from Linda and Jeff explain that the formula was from the Rule for Variance on page 337.

Message no. 837 posted by Mary: Tuesday, June 26, 2001 12:07 p.m.

The box on page 337 gives the formula, and I know it can be proved using some nasty algebra if I had the time (and face it, motivation) to do so--but I think I might even have an mental image grip on why it works this way...look out, this could get wild!

I wish I could draw pictures...

(Mary elaborates on her idea. See Appendix A for full discussion thread.)

Now, somebody out there that really KNOWS for sure if this is right, say something quick before we all get hopelessly confused.

When all else fails you can just use the formula with no understanding. I've had kids that did it for 6 years and somehow managed to get out of high school! Mary

Message no. 850 posted by Simonsen and Banfield: Tuesday, June 26, 2001 4:20 p.m.

Mary, Your explanation is as good as anything I've ever come up (without going through the math ... an unpleasant experience). Sometimes I take a small set of numbers for X (say 3, 4, 5) and another for Y (say 6, 7, 8) and show that the range for X-Y is the same as the range for X+Y and they are both larger than the range of X or Y. This doesn't prove the formula is correct but it often convinces students that when you subtract random variables the variance does actually increase.

In the following example of the *redirect* category, the instructor intervention occurred to help identify a misconception and refocus the group discussion to the actual question being asked. Mary started the discussion by asking if other students in the class would help her understand the topic of degrees of freedom a little more clearly. She did not specifically ask the instructors for validation, but she asked her fellow classmates for discussion. Deanne pushed the discussion forward with her call for clarification of a possible misconception. Although many students tried to help Deanne with her quandary, none of them answered the question to her satisfaction. Again, there was not a specific call for validation, although the instructor felt an intervention was necessary to refocus the group. The *redirect* intervention may have occurred by default, rather than design, as the discussion thread was not being monitored over the weekend when the discussion was taking place. Although it is difficult to predict, had the discussion thread been monitored more closely, it is possible that the instructor intervention would have been to *withhold*. (Note: given the length of this discussion, the bold sentences highlight the key ideas in the abbreviated example.)

Message no. 338 posted by Mary: Friday, June 15, 2001 5:53 a.m.

> Just a quick question before I get on the road for MT... **Would you all please discuss degrees of freedom for me?** I understand about how if you know n-1 pieces of data, by default you know the nth one. What I don't get is how dividing by n-1 can get you an accurate measure if you have n items being considered. I remember talking about sample sizes in college, "sufficiently large...."--**but I don't remember the WHY**.

Message no. 385 posted by Kim: Friday, June 15, 2001 12:26 p.m.

> **Wow, Mary. This is the exact question I had. Where DOES the n-1 come from? I don't think I totally understand standard deviation** - the formula that is.

Message no. 392 posted by Lisa: Friday, June 15, 2001 2:39 p.m.

> **Mary, I sure wish I could help you with this one, but that whole paragraph has me confused. Thanks for bringing it up. Maybe someone will be able to clarify this.**
>
> *(Twenty-one messages of the thread were deleted here. See Appendix B for full discussion thread.)*

Message no. 504 posted by Simonsen & Banfield: Sunday, June 17, 2001 9:25 p.m.

> **Boy, I take off a weekend for my wife's family reunion and come back to almost 200 discussion messages! Do I have some catching up to do. Let's deal with this n-1 or n for calculating the sample variance (and note that**

I emphasize sample). I'm sorry to hear some of you have been confused by this for years, maybe we can straighten it out.
First, let me say that it is not that big a thing. Divide by n or n-1; it just does not matter in any practical way. It is important in a theoretical sense, dividing by n-1 leads to an **unbiased** estimator, which is highly valued by statisticians but relatively unimportant to the rest of the world. You should all go back and look at the last half of William's message (# 502) because he nails the reason we divide by n-1 ... it is just to adjust the size of the **sample** variance. Before you go back to William's message, look over Ed's (# 488) where he mentions estimation and then Deanne's (# 437) where she has the sentence that may come closest to the true reason we use n-1 ... "Using n-1 gives you a better estimate than using n would."
(A long explanation with and without technical jargon follows.)

Message no. 507 posted by Deanne: Monday, June 18, 2001 5:17 a.m.
Thanks for your explanation. I notice you did not use the nasty, very confusing term degrees of freedom in it. **Does degrees of freedom really have anything to do with why n-1 is the best correction for the estimate? Is n-1 the best correction because it gives the degrees of freedom or is it just a coincidence that the best correction is equal to the degrees of freedom? I don't need to know why, just which is the case. Actual explanation or just lucky (or perhaps unlucky) coincidence?** Thanks again, Deanne

Message no. 513 posted by Simonsen & Banfield: Monday, June 18, 2001 12:07 p.m.
Degrees of freedom is a term that has been handed down from the depths of statistical time…

Message no. 517 posted by Alan: Monday, June 18, 2001 1:19 p.m.
Jeff, This explanation helps me. I believe I am understanding standard deviation somewhat better at this time.

In an effort to move the discussion to a new level, hence, the *expand* intervention, the instructor asked the students to synthesize what they had previously learned about regression and ANOVA and discuss the defining characteristics of when to use ANOVA and when to use regression. Although the initial instructor intervention in this discussion thread was to *expand*, in the remainder of the thread the instructor chose to *withhold* by not replying to the discussion. (Note: the bold sentences highlight the key ideas in the example.)

Message no. 1570 posted by Simonsen & Banfield: Friday, July 13, 2001 7:58 a.m.

> **Just a quick note to relate what you see in ANOVA to what has come before... In the two-sample section (that we just finished) we had two populations, took a sample from each, and asked the question "based on these samples does it look like these two populations have the same mean or is there a difference?" Now, in ANOVA, we extend that to 3 or more populations** (you can use ANOVA for 2 populations if you want ... it gives you the same result as the two sample t-test) and the questions become a little more complicated. We still want to know if the means of the populations are all the same or are there some differences. It is just that there are more options for the differences part ... A and B could have the same mean and it is different from C and D (which might have the same mean); all four means might be different; A, B, and D might have the same mean and it might be different from C; and so forth. Get the picture?
> The idea is the same as in the two-sample setting it is just more complicated and the t-statistic can't deal with it ... enter the F-statistic. The F-statistic is designed to test the null hypothesis that all of the means are the same against the alternative that at least one is different. If you reject the null then you have the difficult question of "If the means are not all the same, which ones are different?" That is what the multiple comparison confidence intervals deal with.
> This section on ANOVA assumes all of the population variances are the same (remember the discussion on pooling the variances in the previous section). The reason

for this is because the formulas are really nasty if you do not assume equal variances (maybe not really nasty but they are nicer if you assume equal variances and the whole problem becomes less complicated). You can use ANOVA if the populations do not have equal variances, but wait until you understand the concepts behind ANOVA at this level before you start adding in additional complications.

A question you should probably discuss: What are the defining characteristics of when to use ANOVA and when to use regression? They are kind of related it seems?

Message no. 1574 posted by Peter: Friday, July 13, 2001 11:28 a.m.

I don't have the whole answer to Jeff's question, but I will try and get us started. Let me start with what I see as the role of Analysis of Variance. **I see ANOVA as testing the definite hypotheses as to whether differences exist under assumptions about the sampling situation.**

(For further discussion by Peter, see Appendix C for full discussion thread.)

Message no. 1577 posted by Charles: Friday, July 13, 2001 2:26 p.m.

As I understand it ANOVA is used when we want to compare several population means.

(Additional arguments by Charles led to the following statement: See Appendix C for full discussion thread.)

I think one of the major differences what sets the ANOVA and regression apart is that ANOVA is comparing variation among group means and variation within groups.

(Ten messages of the thread were deleted here. See Appendix C for full discussion thread.)

Message no. 1682 posted by Lisa: Tuesday, July 17, 2001 10:17 a.m.

Steven, I did take the time to read this "lengthy post." Thanks, I appreciate the help. And, it does make sense. Lisa

Message no. 1708 posted by Ken: Thursday, July 19, 2001 1:53 p.m.
Steven - I also read your lengthy post and it helped me a lot. Thanks.

Message no. 1710 posted by Kevin: Thursday, July 19, 2001 6:34 p.m.
This makes a great deal of sense now that I have finished the homework. I wish I had been here to be a part of this conversation. I would have been done in half the time.

With respect to the identified typology, we believe there are reasons to use each of the categories throughout an online course. Furthermore, we highly encourage strategic use of the *expand* category to advance the discussion. We recommend, however, that the default action on the part of the instructor be to *withhold*, using one of the other interventions only if necessary. *Withholding* encourages the students to lead and become active participants in the discussion. The following abridged version of one asynchronous discussion thread involving the interpretation of the definition of simple random sample (SRS) demonstrates the power of the *withhold* intervention. In this discussion, we meet Harold, who believes a particular problem is an SRS when it is in fact a stratified random sample. We were continually tempted to intervene in a more active manner. However, we believe that by *withholding*, we were successful in engaging students in mathematical discourse by compelling them to help one of their peers understand a problem. Ultimately, the instructors intervened to *resolve* the question by identifying Harold's misconception, but only after several other students in the class benefited from the discussion. Harold's misconception was not obvious in any single posting; it was the ensemble of all postings that revealed the subtle error Harold was making. The example begins when Harold posts a discussion message explaining to the class why he disagrees with an exercise in the book. Specifically, Harold asserts that the method of drawing a sample in this exercise is an SRS when the textbook answer has claimed that it is a stratified random sample. The other students helping Harold try to convince him that it is a stratified random sample because they are sampling groups within the population separately and then combining the samples in the end to form the full sample. (Note: the bold sentences highlight the key ideas in the example.)

Message no. 733 posted by Harold: Saturday June 23, 2001 8:41 p.m.
To everyone: I believe that each faculty member does have an equal probability of being selected. Let me explain.
(A lengthy explanation follows. See Appendix D for full discussion thread.)
I believe this is a SRS. If I am wrong, would someone please convince me?

Message no. 737 posted by Cathy: Saturday, June 23, 2001 9:31 p.m.
Harry, I don't think that this is a SRS. Your explanation and calculation seem to assume that each sex is its own population. However, each sex is a stratum, not a population… the definition of SRS mentions nothing about strata so I think that they are not considered in an SRS.

Message no. 739 posted by Harold: Saturday, June 23, 2001 9:57 p.m.
I looked at the definition on page 262. It doesn't say anything about individuals. It says that each SAMPLE will have the same probability of a given size. **I still think this is a SRS …** does each random stratified sample meet the requirement to be a SRS. My answer is YES, because each random stratified sample meets the requirement of being 250/2500 or 1/10 of the population being studied.

Message no. 761 posted by Deanne: Sunday, June 24, 2001 08:33 a.m.
Here is the definition of simple random sample from the book:
A **simple random sample (SRS)** of size n consists of n individuals from the population chosen in such a way that every set of n individuals has an equal chance to be the sample actually selected.
Especially important in this definition is that every set of n individuals chosen from the population has an equal chance to be the sample selected… **there is not an equal chance for every set to be selected and therefore, this is by definition not a SRS.**

Message no. 763 posted by Harold: Sunday, June 24, 2001 08:55 a.m.
Deanne, Thanks for responding. I agree with all that you say except for the part…

(A long post by Harold explaining why he disagrees with Deanne and laying out his argument again. See Appendix D for full discussion thread.) **Most respondents seem to think it has to be an SRS first. I disagree. That's how I see it. Thanks for trying, Harry**

Message no. 765 posted by Doug: Sunday, June 24, 2001 09:49 a.m.

Harry, here is another definition of SRS… In this problem once you stratify and commit to drawing samples that contain 50 women and 200 men you have gone against the first statement in the above definition. **Stratification is a more complex process but is sometimes used to produce desired samples.**

Message no. 767 posted by Harold: Sunday, June 24, 2001 10:20 a.m.

Doug, I agree with what you said and I believe even more strongly that this is a SRS. Here's why. *(Harold goes on to give a detailed breakdown of the sampling procedure.)* This is a SRS.

Message no. 768 posted by Steven: Sunday, June 24, 2001 11:09 a.m.

By discounting groups 1 and 2, where the number of men and women are not 200 to 50, you move from a SRS to a stratified random sample. **A SRS must include every set if individuals having an equal chance to be the sample actually selected for your study.**

Message no. 769 posted by Simonsen & Banfield: Sunday, June 23, 2001 12:14 p.m.

I think I see what you are thinking, Harry, that for a SRS each sample that can be drawn according to the current sampling plan (such as the one in 3.44) have the same probability. But, that is not quite right. *(A long explanation, referencing Harold's previous posting.)*

Message no. 775 posted by Peter: Sunday, June 24, 2001 6:36 p.m.

Harry, maybe this would be a better definition for SRS?

> ***A SRS of size n consists of n individuals from the population chosen in such a way that every sample of size n selected from the population has an equal chance to be the sample selected.***
> Just a thought, Doug

Message no. 777 posted by Harold: Sunday, June 24, 2001 6:55 p.m.

> **Thanks Jeff. Thanks to all of you who tried to explain it to me. I see that I was only looking at... well as Jeff said. Thanks, all of you, for the discussion.**

Message no. 778 Posted by Harold: Sunday, June 24, 2001 6:59 p.m.

> **Doug, I agree. Your definition of SRS is better than theirs! Thanks, Harry.**

By *withholding* in this discussion, we allowed the conversation to develop naturally, which resulted in deeper understanding for several students. Harold, who apparently knew the definition of a SRS but did not understand how it should be applied, eventually understood his error. In explaining the concept of SRS to Harold, fellow students (Cathy, Deanne, Doug, and Steven) are effectively constructing their own knowledge. This observation is evocative of the Gunawardena, Lowe, and Anderson (1997) model for examining construction of knowledge in computer conferencing. The other students following the discussion are also benefiting by seeing the idea of SRS presented from several different viewpoints. Notice that Doug and Peter even come up with new definitions for an SRS in Message No. 765 and Message No. 775.

After following the discussion, we, as instructors, identified the subtle error that Harold was making in his interpretation of the definition of SRS and finally were able to address his specific problem (Message No. 769). We may never have allowed this much time for student discussion in the classroom; instead, we would probably have asked Harold to stop by after class to discuss it. That may have solved the problem for Harold, but we are now aware how much the rest of the class would have missed.

Finally, there is a luxury in the face-to-face classroom in that the teacher's intentions are more visible to the student. For example, a teacher's intentional *withholding* is evident with, say, a nod of the head or a role of the eyes. We have found that in the online classroom, students value a similar response, such as a head-nodding message stating, "I am here, but I am going to let this discussion continue for awhile before I step in." Although an instructor may choose to *withhold* by not saying anything, we believe this is a disservice to the students because they have no indication whether the instructor is simply "not present" or purposely withholding.

IMPLICATIONS

The mathematics education literature reveals that a student's ability to communicate mathematically is critical to an understanding of mathematical concepts (NCTM, 2000; Sherin, 2003; Steele, 1998). We believe that required asynchronous discussions in online (or possibly face-to-face) mathematics courses have great potential to hone the skills of mathematical discourse. With respect to the *content* of mathematical discourse, we feel there is need for in-depth examination of the nature and quality of such asynchronous mathematical discourse. It was evident in the "Harold" example of asynchronous mathematical discourse that students were constructing mathematical knowledge (Gunawardena, Lowe, & Anderson, 1997). Future research is needed to examine the cognitive development of students involved in asynchronous mathematical discourse. It is also desirable to investigate the potential relationship between the quality of the mathematical discourse and success in the mathematics course.

With respect to the *process* of mathematical discourse, the typology identified in this article provides a baseline for analyzing instructional interventions in fostering mathematical discourse and serves as a springboard to look at the impact of instructor interventions in asynchronous communication. These findings demonstrate the need to carefully examine the literature on discourse in the face-to-face mathematics classroom and develop a framework for analyzing teacher discourse interventions in asynchronous communications.

This research involved analyzing the asynchronous communication of practicing teachers of mathematics in a statistics content course. Sherin (2003) recommended that teachers should be provided with opportunities to engage in mathematical discourse in their own learning in hopes that they will recognize it as an effective tool for their students' learning of mathematical concepts. Further research needs to examine whether teachers transfer their participation in learning via mathematical discourse to their own classrooms.

Finally, the line of inquiry was facilitated by the fact that we could readily examine both the *content* and *process* of mathematical discourse in asynchronous communications because all discourse was documented and archived. We believe that instructors may find it easier to foster mathematical discourse in asynchronous communications than in face-to-face discussions. We are not attempting to extend the findings beyond the scope of this study; however, we recognize that one of the benefits of asynchronous discussions is that the *content* and *process* of mathematical discourse is consciously constructed by both teachers and students. Conscious construction refers to the ability to reflect on the mathematics that is being discussed and respond in a consciously constructed manner. Further investigation is desired on the *process* and *content* of mathematical discourse of consciously constructed discussions (asynchronous) verses immediate response discussions (synchronous or face-to-face).

We initially taught the statistics for teachers course via asynchronous communication because of necessity. As we continue to teach this course in this manner, we will focus on the opportunity such a venue provides for fundamental insights into other areas of mathematical learning.

References

Bogdan, R. C., & Biklen, S. K. (1992). *Qualitative research for education: An introduction to theory and methods*. Boston: Allyn and Bacon.

Elliot, P. C., & Kenny, M. J. (1996). *1996 Yearbook: Communication in mathematics, K-12 and beyond*. Reston VA: NCTM

Gunawardena, C. N., Lowe, C. A., & Anderson, T. (1997). Analysis of a global online debate and the development of an interaction analysis model for examining social construction of knowledge in computer conferencing. *Journal of Educational Computing Research 17*(4), 397-431.

Hara, N., Bonk, C. J., & Angeli, C. (2000). Content analysis of online discussion in an applied educational psychology course. *Instructional Science, 28*(2), 115-152.

Jarvela, S., & Hakkinen, P. (2002). Web-based cases in teaching and learning – the quality of discussions and a stage of perspective taking in asynchronous communication. *Interactive Learning Environments, 10*(1), 1-22.

Merriam, S. B. (1997). *Qualitative research and case study applications in education. Revised and expanded from case study research in education*. San Francisco: Jossey-Bass, Inc.

Moore, D. S., & McCabe, G. P. (2000). *Introduction to the practice of statistics*, Third Edition. New York: Freeman Publications.

National Council of Teachers of Mathematics (NCTM) (1991). *Professional standards for teaching mathematics*. Reston, VA: NCTM.

National Council of Teachers of Mathematics (NCTM) (2000). *Principles and standards for school mathematics*, Reston, VA: NCTM

National Science Foundation (NSF) (2002). *The application and implications of information technologies in postsecondary education: An initial bibliography*, NSF 02-305, Arlington, VA: NSF.

Sherin, M. G. (2003). A balancing act: Developing a discourse community in a mathematics classroom. *Journal of Mathematics Teacher Education, 5,* 205-233.

Spiceland, J., & Hawkins, C. (2002). The impact on learning of an asynchronous active learning course format. *Journal of Asynchronous Learning Networks, 6,* 68-75.

Steele, D. F. (1998) Look who's talking: Discourse in a fourth-grade class. *Teaching Children Mathematics, 4,* 286-292.

Weiss, I. R. (1994). *A profile of science and mathematics education in the United States: 1993.* Chapel Hill, NC: Horizon Research, Inc.

APPENDIX A

Message no. 830 posted by Sandra: Tuesday, June 26, 2001 7:43 a.m.

Could someone explain the addition of variables on page 401? When they added the means, they subtracted but when adding the variable they added. Any help here would be appreciated.

Message no. 831 posted by Tom: Tuesday, June 26, 2001 9:44 a.m.

On page 337, they explain why they do this. I have read it once and still do not understand, but I will read it again. If I can understand it in my terms, I will try and explain it. If anyone else can put it in other terms, it would greatly benefit us all. Tom

Message no. 832 posted by Sandra: Tuesday, June 26, 2001 10:25 a.m.

On page 401, the second sentence above the example says, "The mean and standard deviation aX + bY are found as usual from the addition rules for means and variances." In the example, the mean U has the mean for X - Y and the variance squared has the two variances squared but subtracted. I am confused as to why this happens. Thanks all.

Message no. 834 posted by Pam: Tuesday, June 26, 2001 11:31 a.m.

Sandra, I was also having trouble understanding the example. Message 809 from Linda and Jeff explain that the formula was from the Rule for Variance on page 337.

Message no. 837 posted by Mary: Tuesday, June 26, 2001 12:07 p.m.

The box on page 337 gives the formula, and I know it can be proved using some nasty algebra if I had the time (and face it, motivation) to do so--but I think I might even have an mental image grip on why it works this way...look out, this could get wild!
I wish I could draw pictures...

Suppose this is a picture of the variance of X,
x----x
and this is a picture of the variance of Y,
y--y.

If you subtract the largest spread of y from the smallest value of x, you get a spread like this:
y--yx----x
making the variance larger, not smaller--making it Y+X. Thus the formula has you add variances. Or if pictures don't do it for you, try thinking of subtraction as "adding the opposite"--or in this case as "adding to the left". Either way you're adding on to an amount of variance--spread--and it makes it wider instead of smaller.
Now, somebody out there that really KNOWS for sure if this is right, say something quick before we all get hopelessly confused.
When all else fails you can just use the formula with no understanding. I've had kids that did it for 6 years and somehow managed to get out of high school! Mary

Message no. 850 posted by Simonsen and Banfield: Tuesday, June 26, 2001 4:20 p.m.

Mary, Your explanation is as good as anything I've ever come up (without going through the math ... an unpleasant experience). Sometimes I take a small set of numbers for X (say 3, 4, 5) and another for Y (say 6, 7, 8) and show that the range for X-Y is the same as the range for X+Y and they are both larger than the range of X or Y. This doesn't prove the formula is correct but it often convinces students that when you subtract random variables the variance does actually increase.

APPENDIX B

Message no. 338 posted by Mary: Friday, June 15, 2001 5:53 a.m.

Just a quick question before I get on the road for MT... **Would you all please discuss degrees of freedom for me?** I understand about how if you know n-1 pieces of data, by default you know the nth one. What I don't get is how dividing by n-1 can get you an accurate measure if you have n items being considered. I remember talking about sample sizes in college, "sufficiently large...."--**but I don't remember the WHY**. Off to the interstate.... Thanks for your help--I hope this launches a long discussion that will solve 15 years of confusion and give me a satisfying answer to give to my students!

Message no. 385 posted by Kim: Friday, June 15, 2001 12:26 p.m.

Wow, Mary. This is the exact question I had. Where DOES the n-1 come from? I don't think I totally understand standard deviation - the formula that is.

Message no. 392 posted by Lisa: Friday, June 15, 2001 2:39 p.m.

Mary, I sure wish I could help you with this one, but that whole paragraph has me confused. Thanks for bringing it up. Maybe someone will be able to clarify this.

Message no. 399 posted by Mara: Friday, June 15, 2001 4:34 p.m.

Someone please correct me if I'm wrong, but I just had a brainstorm - I think we can compare this to counting fence posts needed to build a fence...**If this is right, I've cleared up 20 years of muddy water for myself - this makes sense to me. If I'm wrong, let me down easy!**

Message no. 406 posted by Steven: Friday, June 15, 2001 9:03 p.m.

I thought of something similar to that originally and racked my brain trying to answer this question.

Message no. 412 posted by Deanne: Friday, June 15, 2001 10:10 p.m.

I went up to the attic and dug out my old statistics book and the notes I took from the course ... **I still don't**

understand exactly why using degrees of freedom made the estimates better. Why didn't n-2 work for example? What is so special about n-1 that made it the magic formula for correction? Oh well, I think I'm just going to have to accept that it does work until someone can show me otherwise. (For what it is worth, my old text gave this explanation of the idea of degrees of freedom...*Deanne launches into a lengthy explanation from the textbook*):

Message no. 431 posted by Harold: Saturday, June 16, 2001 4:27 a.m.

The n values contain the mean. The reason that we divide the total by n-1 is because we are concerned with the number of the other values which are n-1 values. ...

That is why we divide by n-1 instead of n because the difference, mean minus mean will always equal 0, generating one less value to consider. I hope this helps, and isn't too wordy.

Message no. 433 posted by Mara: Saturday, June 16, 2001 5:07 a.m.

Yes, I see your point. My bubble has burst! So we're back to why n-1. Anyone have a better explanation??

Message no. 434 posted by Deanne: Saturday, June 16, 2001 7:51 a.m.

This sounds great, except, the mean is not always one of the values! ...

It is easy to accept using n-1 when one of the data observations actually equals the mean, but unfortunately, more often than not, this is not the case. Sorry, Deanne

Message no. 436 posted by Brad: Saturday, June 16, 2001 8:04 a.m.

Maybe this will help with the "n" versus (n-1) issue... The standard deviation where you divide by just "n" is for populations only...

Message no. 437 posted by Deanne: Saturday, June 16, 2001 8:15 a.m.

One thought I just had. When you are sampling data, you are trying to estimate what the data set would look like for an entire population. When you calculate the standard deviation for the sample, you are in fact estimating the standard deviation for the whole population. **Using n-1 gives you a better estimate than using n would.?**

Message no. 445 posted by Harold: Saturday, June 16, 2001 9:39 a.m.
Deanne, Thanks for your response. I think you missed the point of what I was saying...
What do you think?

Message no. 457 posted by Doug: Saturday, June 16, 2001 6:43 p.m.
If we divide the square of the differences from the mean by n we would have the mean of the squared deviations which is not the definition of variance. Then why divide by n-1 instead of n? The sum of the deviations = 0; hence, if we know the values of any (n-1) of these values, the last one must have that value that causes the sum of all deviations to be zero. Thus there are only (n-1) "free" deviations. I tell students that the term degrees of freedom is part of a formula used to determine the variance of a data set.

Message no. 462 posted by Deanne: Sunday, June 17, 2001 6:58 a.m.
Yes, this is another example where the mean could be contained in the dataset and where using n-1 is easy to accept. However, …
I would still like to hear an explanation from someone of why dividing by the degrees of freedom (n-1) makes the estimates of the standard deviation for a dataset better for all datasets, not just those where the mean is a member of the dataset.

Message no. 463 posted by Deanne: Sunday, June 17, 2001 7:25 a.m.
But, if there are only n-1 free deviations, why do we include all of the deviations in the summation of the squares of the deviations? ...
I understand degrees of freedom, I just don't understand why using the degrees of freedom creates a better estimate of the variance. Thanks for trying to help, Deanne

Message no. 481 posted by Harold: Sunday, June 17, 2001 12:58 p.m.
Deanne, This would be my question to you at this point: Is your conjecture then that we divide by n-1 when the mean is contained within the data set and we divide by n when the mean is not contained within the data set? Could this be the reason that the calculator asks for the input "to n or to n-1"?

Message no. 483 posted by Deanne: Sunday, June 17, 2001 1:15 p.m.
No, my conjecture is that we always divide by n-1, and that there is some other reason to explain why we do it.

Message no. 484 posted by Harold: Sunday, June 17, 2001 1:33 p.m.
I went back to page 51 where they define what variance is. Here is my conjecture: ...

Message no. 488 posted by Ed: Sunday, June 17, 2001 2:21 p.m.
Deanne, I like what you said in message 412. **When I teach standard deviation to my students, we first do standard deviations for large populations. In that case, we would divide by n and not n-1**

Message no. 489 posted by Deanne: Sunday, June 17, 2001 2:53 p.m.
When you calculate the variance, you sum the squares of the deviations of the observations from their mean not the observations themselves. Even if you divided by n you would not get the mean. **I believe the reason the calculator gives you a choice of using n or n-1 is that you are allowed to use n when working with a very large dataset (a population) but must use n-1 when working with a sample (Which is what we do most of the time).**

Message no. 491 posted by Harold: Sunday, June 17, 2001 3:21 p.m.
Deanne, I agree with you. When dividing by n you would not get the mean of the observations but you would get the mean of the squares of the deviations. You would not get the variance because you need to divide by n-1 to get the variance. **I don't understand the reasoning behind dividing by n or n-1 to get the variance just because**

you have a large data set or a small data set. If your conjecture is that you would always divide by n-1, why are you dividing by n just because the data set is large?

Message no. 502 posted by Doug: Sunday, June 17, 2001 7:12 p.m.

Deanne, Here's an explanation I received several years ago from a stats professor...

Message no. 504 posted by Simonsen & Banfield: Sunday, June 17, 2001 9:25 p.m.

Boy, I take off a weekend for my wife's family reunion and come back to almost 200 discussion messages! Do I have some catching up to do. Let's deal with this n-1 or n for calculating the sample variance (and note that I emphasize sample). I'm sorry to hear some of you have been confused by this for years, maybe we can straighten it out.

First, let me say that it is not that big a thing. Divide by n or n-1, it just does not matter in any practical way. It is important in a theoretical sense, dividing by n-1 leads to an **unbiased** estimator, which is highly valued by statisticians but relatively unimportant to the rest of the world. You should all go back and look at the last half of William's message (# 502) because he nails the reason we divide by n-1 ... it is just to adjust the size of the **sample** variance. Before you go back to William's message, look over Ed's (# 488) where he mentions estimation and then Deanne's (# 437) where she has the sentence that may come closest to the true reason we use n-1 ... "Using n-1 gives you a better estimate than using n would."

(A long explanation with and without technical jargon follows.)

Message no. 506 posted by Deanne: Monday, June 18, 2001 5:06 a.m.

Wow! This is a great explanation. Thanks! Also, I definitely agree with you on this point. I'm not quite clear on WHY an AVERAGE is:

>...divided by one less than the number of observations, rather than just n. That's a good >question, and one that is not really answered by the common response ("finding the

>sample mean pins down one of the values, so there are only n-1 free values left" What's >that got to do with it?) **It almost seems like a big fat justification for doing the thing that works the best to make your estimate better. I'm sure it is not, but it sure seems that way!** Thanks again for a great explanation, Deanne

Message no. 507 posted by Deanne: Monday, June 18, 2001 5:17 a.m.

Thanks for your explanation. I notice you did not use the nasty, very confusing term degrees of freedom in it. **Does degrees of freedom really have anything to do with why n-1 is the best correction for the estimate? Is n-1 the best correction because it gives the degrees of freedom or is it just a coincidence that the best correction is equal to the degrees of freedom? I don't need to know why, just which is the case. Actual explanation or just lucky (or perhaps unlucky) coincidence?** Thanks again, Deanne

Message no. 513 posted by Simonsen & Banfield: Monday, June 18, 2001 12:07 p.m.

Degrees of freedom is a term that has been handed down from the depths of statistical time…

Message no. 517 posted by Alan: Monday, June 18, 2001 1:19 p.m.

Jeff, This explanation helps me. I believe I am understanding standard deviation somewhat better at this time.

Message no. 534 posted by Alison: Monday, June 18, 2001 11:09 p.m.

Hi Ed, I am a little confused. I was understanding the reason for dividing by n-1 …

APPENDIX C

Message no. 1570 posted by Simonsen & Banfield: Friday, July 13, 2001 7:58 a.m.

Just a quick note to relate what you see in ANOVA to what has come before... In the two-sample section (that we just finished) we had two populations, took a sample from each, and asked the question "based on these samples does it look like these two populations have the same mean or is there a difference?" Now, in ANOVA, we extend that to 3 or more populations (you can use ANOVA for 2 populations if you want ... it gives you the same result as the two-sample t-test) and the questions become a little more complicated. We still want to know if the means of the populations are all the same or are there some differences. It is just that there are more options for the differences part ... A and B could have the same mean and it is different from C and D (which might have the same mean); all four means might be different; A, B, and D might have the same mean and it might be different from C; and so forth. Get the picture?

The idea is the same as in the two-sample setting it is just more complicated and the t-statistic can't deal with it ... enter the F-statistic. The F-statistic is designed to test the null hypothesis that all of the means are the same against the alternative that at least one is different. If you reject the null then you have the difficult question of "If the means are not all the same, which ones are different?" That is what the multiple comparison confidence intervals deal with.

This section on ANOVA assumes all of the population variances are the same (remember the discussion on pooling the variances in the previous section). The reason for this is because the formulas are really nasty if you do not assume equal variances (maybe not really nasty but they are nicer if you assume equal variances and the whole problem becomes less complicated). You can use ANOVA if the populations do not have equal variances, but wait until you understand the concepts behind ANOVA at this level before you start adding in additional complications.

A question you should probably discuss: What are the defining characteristics of when to use ANOVA and when to use regression? They are kind of related it seems?

Message no. 1574 posted by Peter: Friday, July 13, 2001 11:28 a.m.

I don't have the whole answer to Jeff's question, but I will try and get us started. Let me start with what I see as the role of Analysis of Variance. **I see ANOVA as testing the definite hypotheses as to whether differences exist under assumptions about the sampling situation.** The role is to yes or no to the null hypotheses. In summary, the methods of ANOVA set the boundary between sampling variation and real differences in complicated experiments or observations. Our deductions are based on the variances (variability) whether it is explained or not explained. We are analyzing variances. Hope one of you can add to this to get a full answer. Peter

Message no. 1577 posted by Charles: Friday, July 13, 2001 2:26 p.m.

As I understand it ANOVA is used when we want to compare several population means. The null hypothesis is that the population means are all equal and the alternative hypothesis is true if there exists any difference in population means. The ANOVA separates the observed total variation into either variation among group means or variation within group means. A large variation among groups compared to variation within groups will negate the null hypothesis. **I think one of the major differences what sets the ANOVA and regression apart is that ANOVA is comparing variation among group means and variation within groups.**

Message no. 1579 posted by Lisa: Friday, July 13, 2001 2:43 p.m.

I'm no expert at this either, but...**We use ANOVA when we have independent SRS's from each population. We use regression when the SRS's from each population are not independent (I think!). ANOVA can be used with more than two populations. Can regression be**

used with more than two populations? I'm sure there are other defining characteristics (and I'm not sure if I'm right), so if anyone knows of any, please, feel free.

Message no. 1581 posted by Harold: Friday, July 13, 2001 3:10 p.m.

Lisa, One of the distinguishing characteristics that determines whether to use ANOVA with multiple samples is to check the standard deviations, see Rule for Examining Standard Deviations in ANOVA, where you see if the largest standard deviation is less that 2 times the smallest standard deviation. If so, then ANOVA can be used. I think this is a very convenient way to decide to ANOVA or not to ANOVA.

Message no. 1584 posted by Ron: Friday, July 13, 2001 4:43 p.m.

I am still in the fog with ANOVA but your last comment made some sense (I hope it's right.) It does make sense that regression compares how one variable affects another. ANOVA compares independent events. I just had to say it again to convince myself. Ron

Message no. 1590 posted by Chris: Friday, July 13, 2001 6:34 p.m.

Lisa - This makes sense - especially knowing that regression measures the correlation between two variables rather than the means of populations.

Message no. 1613 posted by Steven: Saturday, July 14, 2001 8:12 p.m.

Lisa, You asked if regression can be used with more than two populations? I went and looked back at our reading from section 10.1 on Simple Linear Regression to help answer this question and I am not sure that I will. On pg. 663 it states, "In linear regression the explanatory variable x can have different values. Imagine, for example, giving different amount of calcium to different groups of subjects. We can think of the values of x as defining different "subpopulations," one for each possible value of x. Each subpopulation consists of all individuals in the population having the same value of x." Also, at the bottom of pg. 664, it states, "In the statistical model for predicting (which is a function of linear regression)

body density from skinfold thickness, subpopulations are defined by the x variable. All individuals with the same skinfold thickness x are in the same subpopulation."
My thought about this: It seems that regression deals with one population, but if a person wants to compare any of the subpopulations, that also seems possible, but one may have to use a separate regression for that subpopulation. Whereby, that subpopulation would seem to become the population of the new regression you would plan to analyze. Boy, I hope those of you who took the time to read this lengthy post understood what I tried to say. Steven

Message no. 1637 posted By Pam: Sunday, July 15, 2001 2:52 p.m.
Thanks Chuck, you explanation really helped.

Message no. 1639 posted by Mara: Sunday, July 15, 2001 3:25 p.m.
Steven - The way I see it, if you looked at these subpopulations, there would be only one x, thus no longer "variable." So in that case I don't think you could do a regression line for it. Am I reading this right?
I don't know if there is a fine line or not, but I'm thinking that we use regression lines when we're looking for cause and effect. ANOVA seems to be not so much why do we get different values for y, but rather, *Do we get different values for y*. Does that make any sense?

Message no. 1645 posted by Lisa: Sunday, July 15, 2001 4:50 p.m.
Oh Yeah! Thanks Harry. I remember reading this. I will have to remember to do this. Thanks again, Lisa

Message no. 1649 posted by Steven: Sunday, July 15, 2001 9:01 p.m.
Mara, This seems to make sense to me. Thanks for the response. One question I have then. **Could we then run ANOVA on the subpopulations?** Let's say that a treatment group was given 1500 mg of calcium and another treatment group was given 2000 mg of calcium, and yet another was given 1000 mg of calcium. **It would**

seem that we could run ANOVA with a setup like this where we could check to see if there are any significant difference in (take your pick).

Message no. 1665 posted by Mara: Monday, July 16, 2001 12:47 p.m.
If you mean to compare the means of the subpopulations, I would think that would work. Within a single subpopulation, though, I don't see anything to compare. Sorry so short - gotta run. Mara

Message no. 1682 posted by Lisa: Tuesday, July 17, 2001 10:17 a.m.
Steven, I did take the time to read this "lengthy post." Thanks, I appreciate the help. And, it does make sense. Lisa

Message no. 1708 posted by Ken: Thursday, July 19, 2001 1:53 p.m.
Steven - I also read your lengthy post and it helped me a lot. Thanks.

Message no. 1710 posted by Kevin: Thursday, July 19, 2001 6:34 p.m.
This makes a great deal of sense now that I have finished the homework. I wish I had been here to be a part of this conversation. I would have been done in half the time.

APPENDIX D

Message no. 733 posted by Harold: Saturday, June 23, 2001 8:41 p.m.

To everyone: I believe that each faculty member does have an equal probability of being selected. Let me explain. The strata of there being either male or female gives each group a 1/2 chance of being selected. The probability of being a selected female is the same, which is 50 out of 500, which is a probability of 1/10, the required probability. The probability of being a selected male is the same for each male, which is 200 out of 2000, again the required 1/10 probability. Finding the probability of being a selected female is 1/2 x 1/10 = 1/20. Finding the probability of being a selected male is 1/2 X 1/20 = 1/20. The probability of being a selected male or female is 1/20. To ignore the strata and to say that a male and a female do not have the same chance as 250/2500 is correct even though the probability is 1/10 for an individual of the population. They would not have the same probability in this situation. **I believe this is a SRS. If I am wrong, would someone please convince me?**

Message no. 737 posted by Cathy: Saturday, June 23, 2001 9:31 p.m.

Harry, I don't think that this is a SRS. Your explanation and calculation seem to assume that each sex is its own population. However, each sex is a stratum, not a population… the definition of SRS mentions nothing about strata so I think that they are not considered in an SRS.

Message no. 739 posted by Harold: Saturday, June 23, 2001 9:57 p.m.

I looked at the definition on page 262. It doesn't say anything about individuals. It says that each SAMPLE will have the same probability of a given size. **I still think this is a SRS …** does each random stratified sample meet the requirement to be a SRS. My answer is YES, because each random stratified sample meets the requirement of being 250/2500 or 1/10 of the population being studied.

Message no. 761 posted by Deanne: Sunday, June 24, 2001 08:33 a.m.

Here is the definition of Simple Random Sample from the book:

A **simple random sample (SRS)** of size n consists of n individuals from the population chosen in such a way that every set of n individuals has an equal chance to be the sample actually selected.
Especially important in this definition is that every set of n individuals chosen from the population has an equal chance to be the sample selected… **there is not an equal chance for every set to be selected and therefore, this is by definition not a SRS.**

Message no. 763 posted by Harold: Sunday, June 24, 2001 08:55 a.m.
Deanne, Thanks for responding. I agree with all that you say except for the part concerning the 50 more or less or the 200 more or less. To me the stratified random sample demands that there be exactly 50 women and 200 men. Each one of these members has a 1/20 chance whether men or women of being selected. That is what is generated in each 250 member sample. Each of these 250 member samples has a 250/2500 chance of being chosen. This is the 1/10 chance required to be an SRS. I guess I don't see the connection made to claim this stratified random sample is not an SRS. I see that it is. I am trying to see it from the viewpoint that it is not an SRS, but I don't. I see it this way; Each individual has an equal chance of being selected into the sample (1/20) and each sample has an equal chance of being the sample selected (250/2500). This is what makes it an SRS. **Most respondents seem to think it has to be an SRS first. I disagree. That's how I see it. Thanks for trying, Harry**

Message no. 765 posted by Doug: Sunday June 24, 2001 09:49 a.m.
Harry, here is another definition of SRS… In this problem once you stratify and commit to drawing samples that contain 50 women and 200 men you have gone against the first statement in the above definition. **Stratification is a more complex process but is sometimes used to produce desired samples.**

Message no. 767 posted by Harold: Sunday, June 24, 2001 10:20 a.m.
Doug, I agree with what you said and I believe even

more strongly that this is a SRS. Here's why. *(Harold goes on to give a detailed breakdown of the sampling procedure.)* This is a SRS.

Message no. 768 posted by Steven: Sunday, June 24, 2001 11:09 a.m.

By discounting groups 1 and 2, where the number of men and women are not 200 to 50, you move from an SRS to a stratified random sample. **A SRS must include every set if individuals having an equal chance to be the sample actually selected for your study.**

Message no. 769 posted by Simonsen & Banfield: Sunday, June 23, 2001 12:14 p.m.

I think I see what you are thinking, Harry, that for a SRS each sample that can be drawn according to the current sampling plan (such as the one in 3.44) has the same probability. But, that is not quite right. *(A long explanation, referencing Harold's previous posting.)*

Message no. 775 posted by Peter: Sunday, June 24, 2001 6:36 p.m.

Harry, maybe this would be a better definition for SRS?

A SRS of size n consists of n individuals from the population chosen in such a way that every sample of size n selected from the population has an equal chance to be the sample selected.

Message no. 777 posted by Harold: Sunday, June 24, 2001 6:55 p.m.

Thanks Jeff. Thanks to all of you who tried to explain it to me. I see that I was only looking at… well as Jeff said. Thanks, all of you, for the discussion.

Just a thought, Doug

Message no. 778 posted by Harold: Sunday, June 24, 2001 6:59 p.m.

Doug, I agree. Your definition of SRS is better than theirs! Thanks, Harry.