SIMPSON'S PARADOX

Dr. Stan Hartzler Archer City High School

The examples here are courtesy of Dr. Peter John of the Mathematics Department at The University of Texas at Austin, 1980.

Mathematics survey textbooks, and chapters and sections thereof, are often titled "The Use and Misuse of Statistics." This paper is directed at the misuse category. It is usually omitted in such chapters.

Example 1: There are two treatments given to a disease (dandruff?), with both treatments given to men and women. The percent cured is figured by *number cured* divided by *number treated* and the quotient is written as a per cent.

	TREATMENT A	**TREATMENT B**
MEN	$\frac{20}{100} = 20\%$	$\frac{50}{210} = 24\%$
WOMEN	$\frac{40}{60} = 67\%$	$\frac{15}{20} = 75\%$

At this point, it appears that Treatment B is better for both. But: Treatment A, all totaled, cured 60 out of 160 people, or 38%, while Treatment B cured only 65 out of 230 people, for 28%. Now Treatment A appears to be the best.

Now for an example which (I hope) demonstrates how statistics can be misused to get people to think and act and vote and contribute strongly and wrongly. This example is deliberately close to home and personal bias.

It may be well and good for me to inform you here that I, Hartzler, was delivered at birth by a woman doctor, and that my mother and three sisters were career-minded before husband-minded, thus making me astonished and sad when I began dating girls whose primary source of self-image was what the men thought of them. That aspect at least of equality-of-sexes has my support.

Example 2: A company employs 50 men and 50 women, with the male salary average $15,920 and the female salary average $14,000.

Discrimination? No.

YEARS OF SERVICE	**MEN**		**WOMEN**	
	number	**salary**	**number**	**salary**
LESS THAN 5	10	$10,000	40	$12,500
MORE THAN 5	40	$17,400	10	$20,000

The issue is that women are newer to the job market, and have less overall experience, which is one thing companies pay for.

It might appear by the above breakdown that the men are being discriminated against. Not so: perhaps the women are better trained, which is ANOTHER thing companies pay for.

The attention of future teachers is needed. Simpson's Paradox can, and should, be taught to grade-school students as soon as they can work with per cents.

But most of all: when people throw statistics at us to spur us on to action and agitation, we should be careful.

Which way is best? In the first example, Cure B is best. It is always desirable to look at statistical breakdowns by category or categories. When a pattern exists (or does not exist) across different categories, one may be generally more confident when making inferences about truth and then decisions.

Why do students need to learn mathematics? What can we communicate to them that may motivate them?

Students help themselves by being good students in mathematics, by preparing themselves for a wide variety of good-paying jobs with short working hours, good security, and better vacation options. Students help their communities and nation by being productive workers and managers. Mathematics helps here also.

Students help themselves and communities and nation by being intelligent citizens. Those aware of Simpson's Paradox may be able to address community issues where statistics appear to show a problem or solution, but where a somewhat more detailed breakdown of data would show an opposite problem or a different solution.

The mathematically-literate citizen will use mathematics for the common good, and will be able to help when others might attempt to misuse mathematics for private fraud. For these reasons, mathematics teachers and students should be motivated to succeed.