# Letters to the Editor

Dear Editor,

I enjoyed the article "'We All Survived' and Other Failings of Risk Perception" by Stephanie R. Land very much. [*CHANCE*, Volume 21, Issue 4] I vaguely recall seeing the email about all the childhood fun and freedom we oldies (I am from the '50s) had and simply dismissed it as absurd. It was nice to see the graphs in Figure 3 solidly rejecting the myth about how we ALL survived so well without any new-fangled restrictions on our lives.

However, there is a slight problem in the graph in Figure 4 on Page 55. In the graph in the original paper (G. K. Singh and S. M. Yu (1996) U.S. childhood mortality, 1950 through 1993: Trends and socioeconomic differentials. *American Journal of Public Health* 86(4):505–12) both NON-MVA and Homicide were labeled with a solid line, though the NON-MVA should have been dashed. In the copy in *CHANCE*, the label for Homicide got "dashed," leading one to look at the wrong curve.

Looking closely at the graph in Figure 4, the homicide rate shows a rather steady increase between 1968 and 1992, even though the other rates in Figure 4 decreased and the overall child death rates between 1930 and 2005 (Figure 3 a and b) decreased. That disturbed me. All other causes of child deaths improved except for the one of killing kids. There is an accompanying graph in Singh and Yu for years five through 14 to the one for years zero to four reproduced in Figure 4. That graph shows the same trends for the different causes of death. In the graph, there is also a suicide cause, which behaves much like the homicide cause at half the rate. Both the two homicide rates and the suicide rate in the two graphs in Singh and Yu were increasing between 1968 and 1992. So, not only did kids get killed more, they also became so depressed that they committed suicide at a higher and higher rate.


Here's to Your Health

Mark Glickman, Column Editor

'We All Survived' and Other Failings of Risk Perception

Stephanie R. Land

I wondered what happened after 1992. The data on the Maternal and Child Health Bureau web page, *www.mchb.hrsa.gov/mchirc/chusa_04/pages/0436cm.htm*, show that the three rates for homicide and suicide have all decreased from the 1992 levels. Adding the 2002 rates to the 1968–1992 rates shows the P&I (pneumonia and influenza) cause to have bottomed out, while all other rates keep decreasing.

So it continues to be safer and safer to be born now, rather than in the fun, free, good-old days …

Susanne Aref
Aref Consulting Group LLC

---

**Stephanie Land responds:**

I thank the reader for her interesting observations. Regarding the graph, the unfortunate aspect of the original image is that the legend does not distinguish between the dashed line (non-MVA) and solid (homicides).

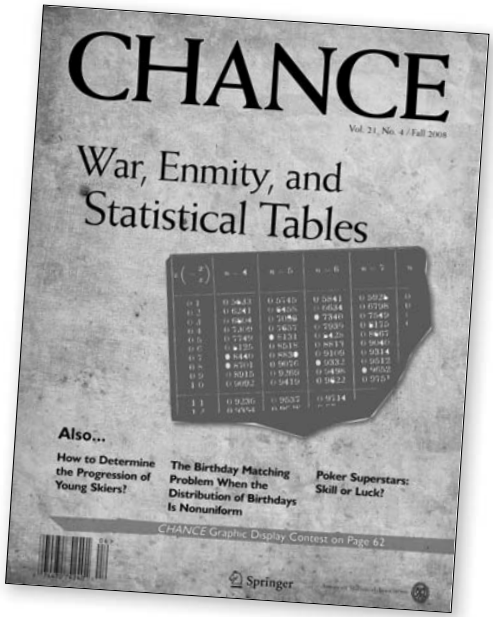
Dear Editor,

In *CHANCE* Volume 21, Issue 4, the articles "War, Enmity, and Statistical Tables" by Brian Clauser and "Fisher and the 5% Level" by Stephen Stigler provided insight into the dysfunctional relationship between R. A. Fisher and Karl Pearson. Hidden within those articles was an equally interesting interaction between William Gosset and Karl Pearson. Gosset had the enviable position of brew master at Guinness Brewery, which objected to him publishing his statistical work under his own name, hence his pseudonym "student." I thought I would use the Excel TDIST command to duplicate the probabilities in Clauser's Figure 1, showing a fragment of Gosset's (Student's) table from "The Probable Error of a Mean," *Biometrica*, 6(1), published March 1908. Gosset's table is parameterized using $z = x/s$, where $x$ is the difference from the mean and $s$ is the standard deviation of $n$ observations. I assumed that Gosset used the unbiased $s^2$, found by dividing by $n$–1 when estimating the variance of $n$ independent observations. To find $t$, as is common practice today, I divided the square root of the unbiased $s^2$ (multiplied $z$) by the $\sqrt{n}$. With $n$–1 degrees of freedom, TDIST did not duplicate Gosset's probabilities. For example, in Table 1 with $z$=.1 and $n$ = 4, then $t$ would be the $\sqrt{4}$ times .1 or .2 with $n$–1 = 3 degrees of freedom. The cumulative probability using TDIST is 0.5729, not 0.5633. Also, with $z$ =.5 and $n$ = 6, then $t$ would be the square root of 6 times .5 or 1.225 with $n$–1 = 5 degrees of freedom. The cumulative probability from TDIST is 0.8624, not 0.8428.

I realized that Gosset must have used the biased $s^2$, found by dividing by $n$; hence, it was necessary to find $t$ by dividing the square root of the biased $s^2$ (multiplying $z$) by the $\sqrt{n-1}$. With $n$–1 degrees of freedom, TDIST duplicated Gosset's probabilities. For $n$ = 4 observations, the values in column one are multiplied by the square root of 3 to get $t$ and using 3 degrees of freedom, we get all the values in column 2. Similarly, for the $n$ = 5 column, the values in column one are multiplied by the square root of 4 to get $t$ and using 4 degrees of freedom, we get all the values in column 3.

I downloaded a copy of Gosset's 1908 paper, and indeed, on page 3, the variance $s^2$ was found by dividing by $n$; but why? The answer is contained in "Student's $z$, $t$, and $s$: What If Gosset Had *R*?" by Hanley, Julien and Moodie in *The American Statistician*, 62(1), February 2008. Here is what they wrote:

"Gosset defined $s^2$ as the sum of squared deviations divided by $n$, rather than $n$–1 (suggested in Airy's textbook) that yields an unbiased estimator of $s^2$–*a* decision influenced by his professor Karl Pearson. Gosset would have preferred to use $n$–1: he wrote to a Dublin colleague in May 1907, 'when you only have quite small numbers I think the formula with the divisor of $n$–1 we used




is better.' Even in 1912 Karl Pearson—still a large sample person—remarked to him that it made little difference whether the sum of squares was divided by n or $n$–1 'because only naughty brewers take n so small that the difference is not the order of the probable error' (Pearson 1939)."

True to his pseudonym, Gosset was the dutiful student to his professor, Karl Pearson. It is noteworthy that "Student" effectively parameterized his own $t$ different from today's practice.

Ray Stefani
California State University, Long Beach

## Correction

In Volume 21, Issue 3, part of the "Children 2–5 year olds" graph for Figure 6 is missing from the article "Healthy for Life: Accounting for Transcription Errors Using Multiple Imputation—Application to a study of childhood obesity."