(b) Concerning the rest of the process, after a manuscript has been accepted, there is no excuse for taking an additional year or two to publish the paper. Journals should use modern technology—Fax, E-mail, electronic transmission of manuscripts—as much as possible to speed up this process, as the genetics journals already do.

I guess what I'm trying to convey is that you folks in statistics have this situation only because you're willing to put up with it! And Wainer, given the visibility of his column, is in a position to lobby for changes. Instead of just advising junior faculty not to submit to the *Annals* because that journal is the worst offender, he could also devote a column to some of the above points; and all of you could speak to your colleagues about it and try to bring about some changes.

OK, I'll stop now: I have to go multiply some proportions by 100.5 before rounding ... .

After writing the above, I contacted the editorial office of the *American Journal of Human Genetics* to see if they had any hard statistics on that journal's performance. Roberta Wilkes, managing editor, told me that for 1996–97, their mean interval from submission to first decision was approximately 6 weeks; mean interval from submission to publication was 6.8 months. So my estimates above were actually conservative, at least for this particular journal.

Susan E. Hodge
Dept. of Psychiatry,
Columbia School of Physicians and Surgeons Division of Biostatistics, Columbia School of Public Health NYSPI, Unit 14
722 West 168th Street
New York, NY 10032, USA
E-mail: seh2@columbia.edu

This letter was received in the Chance *Editorial Office on June 12, 1998, and sent in to ASA for publication on June 19, 1998 –Ed.*

Reply by Wainer

Thanks for your letter. I truly appreciate the kind words.

On publication delays—Obviously I agree with you. I edit a section of an education journal (*The Teachers Corner of the Journal of Educational and Behavioral Statistics*). I try hard to get submitted articles reviewed within two weeks. To do this I am enormously autocratic. I have two very trusted associate editors that I fax articles to and ask for reviews ASAP. Thus we can often get 3 reviews (mine and theirs) in one day. Then, using E-mail, I try to get the letter back to the author immediately. My record for rejections (often an easy decision) is one day (tied many times) but I even have had one acceptance (a tougher and less likely decision) in one day! Many within a week. Authors seem happy to put up with what may be idiosyncratic decision making in return for the blazing speed. But after the editorial decision the accepted papers are put into the standard publication queue (of which I have no control) and it takes a year or more to appear. Alas!

What got Bradlow and me to write the publication delay piece was an experience we had with one of the worst offending journals. After nearly a year's delay we had heard nothing (except replies to our queries that the editor had sent requests to the associate editor for status information). Then, in reply to another of our polite inquiries, the editor said that he now had all reviews in hand and had read the paper himself and we should expect a formal letter from him in six weeks. SIX WEEKS! Why not now? So we started gathering data to find out how typical was our experience.

It is gratifying that our publication delay paper has begun to have an effect, and your input about the practice in another (allied) discipline supports our notion that publishing needn't take this long. We must now watch to see if any statistics journals actually change their policy. If such changes are announced I must prevail on Bradlow to help with a follow up to our survey and measure the effect of the policy change.

Howard Wainer
Measurement-Statistics & Data Research
Educational Testing Service (15-T)
Rosedale Road
Princeton, NJ 08541-0001, USA
E-mail: hwainer@ets.org

Friedrich Pukelsheim: photo courtesy of his coauthor T. Klein.

Rounding Tables on My Bicycle

This morning, as I was riding my bicycle to the office, I was thinking about Howard Wainer's recent Visual Revelations column entitled "Rounding Tables" [*Chance*, Vol. 11, No. 1, Winter 1998 pp. 46–50] and it dawned on me that the handling of ties in Step I of the Webster method is a bit (too) brief. If the fractional part is equal to 1/2, then rounding to the nearest even integer works fine provided the resulting integers sum to 100.

There are tied cases, though, where more is needed. Imagine eight categories of equal size, so each contains 12.5 percent. Remembering Wainer's (1997b, p. 97) 2D-Theorem ("Two digits suffice"), we would like to round to full percentages. The nearest-even-integer advice (Eisenhart 1947) gives eight times 12 percent and sums up to 96 percent, thus falling short of 4 percent.

This is why Balinski and Young's (1982, p. 99) definition of the Webster "method" enumerates multiple solutions, if any. In the little contrived example above, we have 70 choices of four times rounding down to 12 percent and four times rounding up to 13 percent.

This sounds more intimidating than it really is. It has been my experience that looking at the data usually suggests a good way of resolving ties. I do not think that it is a good idea to have ties resolved mechanically by a computer program.

Here is an example from one of my unwritten papers. We are given two apportionments (counts) of 28 seats to four parties, together with their percentages:

15: 53.6   10: 35.7   2: 7.1   1: 3.6
14: 50.1–   10: 35.7+   2: 7.1+   2: 7.1+

In the first line, standardization to the tenth of a percent is unambiguous, as 15 out of 28 rounds to 53.6 percent etc. In the second line, rounding to the tenth of a percent produces a tie, with four equally legitimate solutions. They are the one quoted above, plus the three solutions where a tenth of a percent is transferred from 50.1 to either one of the following three percentages. What to do?

I propose the following. Since 2 is rounded to 7.1 percent in the first line, we are only inviting trouble if we round it to something else in the second line. Same with rounding 10 to 35.7 percent. Therefore, I shall round 14 to 50.1 percent.

Now, if my readers are not only literate but also numerate, as I hope they are, I may be getting E-mails pointing out a typo in my unwritten paper, to the effect that 14 out of 28 evidently equals 50.0, not 50.1 percent. What am I going to respond? I guess I would first try a friendly lie:

"When I rounded those numbers, I noticed a tenth of a percent was missing. So I added it on."

Should people get too inquisitive, I would have to hammer the truth into them:

"I did a minimum Chi-square fit among all distributions whose probabilities are integer multiples of 1/1000."

This stills everybody. The curious true fan has to read up on it in Balinski and Young (1982, pp. 103–104), anyways.

Last, if I may be forgiven an *obiter dictum*, I would like to advertise that more on the subject, like further Tech Reports and the RoundPro program itself, can be retrieved from the Internet at www1.math.uni-augsburg.de/sta/.

Friedrich Pukelsheim
Institut für Mathematik
der Universität Augsburg
D-86135 Augsburg, Germany
pukelsheim@uni-augsburg.de

Follow-up on "A Selection of Selection Anomalies"

A number of sharp-eyed readers of our paper "A Selection of Selection Anomalies" (*Chance*, Vol. 11, No. 2, Spring 1998, pp. 3–7) have pointed out that we erred in attributing the *Literary Digest*'s sampling error associated with using a poll conducted by telephone to the 1948 Truman/Dewey election. It should have been the 1936 election in which this misbegotten poll predicted a victory for Alf Landon over Franklin Delano Roosevelt. We are grateful for their help in straightening things out.

At some point in the process the abstract for our paper was mistakenly appended to the final conclusion instead of a footnote acknowledging the help we received in the preparation of the paper. The concluding paragraph should have ended:

"The multiple imputations may not give you a good answer, but they can provide you with an estimate of how sensitive your inferences are to the unknown. If you do not do this you have not dealt with possible selection biases, you have only ignored them."

And then the missing footnote:

"This research was partially supported by the research allocation of the Educational Testing Service to the Research Statistics Group, and we are grateful to acknowledge this support. At various times several of our colleagues offered important help, principal among these are: Nancy Burton, Nancy Ervin, Fred Marino, Karen McQuillen, Linda Steinberg, and Stephen Stigler. We would like to express our gratitude to you all. Our use of Thurber's fable was suggested by Rosenbaum (1989)."

Howard Wainer
Measurement-Statistics & Data Research
Educational Testing Service (15-T)
Rosedale Rd, Princeton,
NJ 08541-0001, USA
E-mail hwainer@ets.org

Samuel Palmer
63 Governors Lane
Princeton, NJ 08540, USA

Eric T. Bradlow
Marketing Department
The Wharton School
University of Pennsylvania
1400 Steinberg Hall-Dietrich Hall
Philadelphia, PA 19104-6371, USA
E-mail ebradlow@wharton

VISUAL REVELATIONS

Howard Wainer,
Column Editor

Rounding Tables

The first step in the display of data is often compiling the data into a table. The data table has long been the object of derision when it is used to provide an evocative exploratory display. But over the last 20 years several suggestions have been made that, if followed, can make substantial improvements in the ability of a table to communicate information about the latent structure of the data that it contains (Ehrenberg 1977; Wainer 1993, 1997a, b). In this essay I would like to make one small subtraction from the lore of effective table preparation. In particular, I would like to provide the machinery that would allow the subtraction of footnotes like this one:

"**This column adds up to more than 100 due to rounding.*"

> footnote to Table C in
> *Profile of SAT Takers in the Class of 1997.*

This footnote is commonly found whenever tables of proportions are presented. Sums of rounded proportion often fail to add to 1. Consider Table 1 (from Diaconis and Freedman 1979). This table has three categories (the number of categories $c = 3$). Note that the sum of the proportions of families in this table does not equal 1 due to the rounding. This is not unusual. Mosteller, Youtz, and Zahn (1967) showed that the probability that rounded proportions add up to 1 depends strongly on the number of categories; with a large number of categories, it is about $1.38 \times c^{-1/2}$.

In Fig. 1 is a plot of this function for values of c between 2 and 50. In it we can see that with as few as eight categories we would expect about half the columns not to add up to 1. As the number of categories increases the proportion that do not add up increases as well, but at an exponentially decreasing rate. Because we know that predictably often a column of rounded proportions will not add up exactly, we can use the multinomial model that generated the results in Fig. 1 to tell us if there is a fair likelihood that someone has artificially modified the proportions.

The problem of a table of proportions not adding up properly is a minor one in most applications. It is, however, critical in political science, where it has been studied extensively and has been termed the *problem of apportionment*. Balinski and Young (1982) recounted the historical back-

Table 1—Distribution of White Families by Type, United States, 1970

Type	Number (1,000)	Proportion
Husband-wife	40,802	.887
Other male head	1,036	.023
Female head	4,185	.091
Total	46,023	1.001

Column Editor: Howard Wainer, principal research scientist, Measurement-Statistics & Data Research, Educational Testing Service (15-T), Rosedale Road, Princeton, New Jersey, USA 08541-0001; hwainer@ets.org.

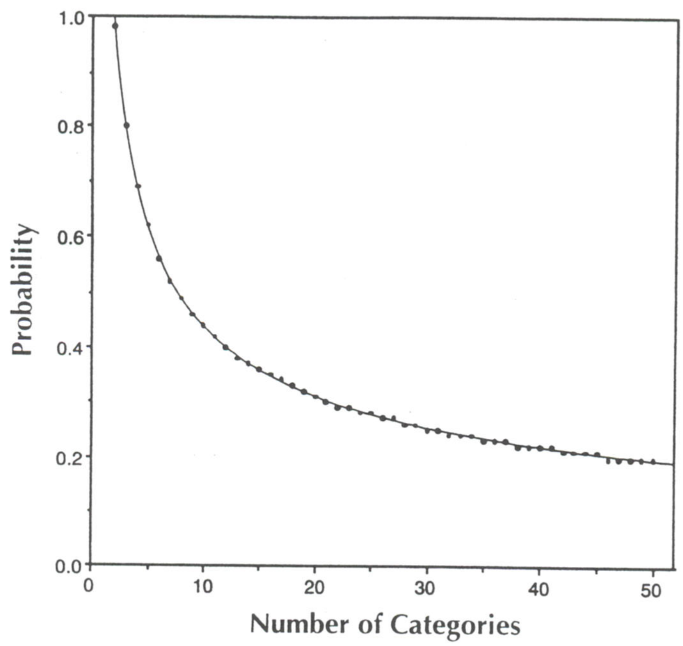

Figure 1. Model-based estimate that the probability of a rounded sum of proportions will equal 1.

ground of the problem and the mathematical issues it generates. These issues have been investigated, and several procedures have been advanced that cleverly force tables to add up properly even when rounded.

There are two classes of methods available that were described by Happacher and Pukelsheim (1994, 1997). We shall not repeat the descriptions of all of the various adjustment schemes used but instead refer interested readers to the appropriate sources. We shall, however, describe the Webster method, which is the simplest and most commonly used. It suffices in most situations and is the one recommended by Balinski and Young (1982).

Step I of the Webster method is the usual one.

With p_i being the exact proportions, round $100p_i$ to the nearest integer n_i. If the rounded numbers ni sum to 100, they are the final result. [The nearest integer is unique when $100p_i$ has a fractional part distinct from 1/2; if the fractional part is equal to 1/2, then round to the nearest even integer (Eisenhart 1947).]

Step II is called for only if the Step I results fail to sum to 100.

(a) If the sum exceeds 100, then find a small positive quantity ϵ and apply Step I to $(100 - \epsilon)\, p_i$ until the new rounded numbers do sum to 100. (b) If the sum stays below 100, apply Step I to $(100 + \epsilon)\, p_i$ until the rounded numbers sum to 100. There are other ways of carrying out Step II that are computationally more efficient (Balinski and Young 1982; Happacher and Pukelsheim 1997) but in most situations such increased efficiency isn't particularly important.

Table 2—Trends in GRE Subject Test Volumes (1981–1990) (original counts)

	Year									
Subject	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Biology	9,133	8,312	8,094	7,736	7,834	8,451	8,253	7,257	7,555	7,701
Chemistry	3,125	2,999	3,242	3,168	3,039	2,997	2,908	2,595	2,716	2,961
Computer Science	1,325	1,658	1,956	2,372	2,816	3,532	3,393	2,818	3,118	2,999
Economics	1,677	1,560	1,895	1,867	1,755	1,896	1,789	1,379	1,326	1,337
Education	1,337	1,080	1,022	1,077	1,015	1,028	945	769	898	860
Engineering	3,818	3,813	4,296	4,404	4,318	4,509	4,620	4,050	3,883	3,626
Geology	2,139	2,274	2,885	2,475	1,897	1,592	1,085	844	685	642
History	1,405	1,196	1,425	1,422	1,323	1,464	1,612	1,524	1,544	1,809
Literature	2,722	2,531	2,786	2,889	2,996	3,311	3,686	6,732	4,170	4,880
Mathematics	1,573	1,549	1,860	1,960	2,163	2,191	2,124	1,969	2,041	2,181
Music	1,542	1,302	1,112	1,044	970	946	819	772	793	796
Physics	1,953	1,952	2,158	2,213	2,245	2,299	2,459	2,242	2,428	2,799
Political Science	1,446	1,259	1,311	1,392	1,362	1,420	1,458	1,279	1,290	1,316
Psychology	8,722	8,862	8,795	8,810	8,693	8,977	9,567	8,889	10,076	11,175
Sociology	796	646	688	618	543	716	746	642	691	1,089
TOTAL	**42,713**	**40,993**	**43,525**	**43,447**	**42,969**	**45,329**	**45,464**	**43,761**	**43,214**	**46,171**

Table 3—Trends in GRE Subject Test Volumes (1981–1990) (percentages to two decimal places)

	Year									
Subject	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Biology	21.38	20.28	18.60	17.81	18.23	18.64	18.15	16.58	17.48	16.68
Chemistry	7.32	7.32	7.45	7.29	7.07	6.61	6.40	5.93	6.29	6.41
Computer Science	3.10	4.04	4.49	5.46	6.55	7.79	7.46	6.44	7.22	6.50
Economics	3.93	3.81	4.35	4.30	4.08	4.18	3.93	3.15	3.07	2.90
Education	3.13	2.63	2.35	2.48	2.36	2.27	2.08	1.76	2.08	1.86
Engineering	8.94	9.30	9.87	10.14	10.05	9.95	10.16	9.25	8.99	7.85
Geology	5.01	5.55	6.63	5.70	4.41	3.51	2.39	1.93	1.59	1.39
History	3.29	2.92	3.27	3.27	3.08	3.23	3.55	3.48	3.57	3.92
Literature	6.37	6.17	6.40	6.65	6.97	7.30	8.11	15.38	9.65	10.57
Mathematics	3.68	3.78	4.27	4.51	5.03	4.83	4.67	4.50	4.72	4.72
Music	3.61	3.18	2.55	2.40	2.26	2.09	1.80	1.76	1.84	1.72
Physics	4.57	4.76	4.96	5.09	5.22	5.07	5.41	5.12	5.62	6.06
Political Science	3.39	3.07	3.01	3.20	3.17	3.13	3.21	2.92	2.99	2.85
Psychology	20.42	21.62	20.21	20.28	20.23	19.80	21.04	20.31	23.32	24.20
Sociology	1.86	1.58	1.58	1.42	1.26	1.58	1.64	1.47	1.60	2.36
TOTAL	**100.00**	**100.01**	**99.99**	**100.00**	**99.97**	**99.98**	**100.00**	**99.98**	**100.03**	**99.99**

An Illustration: GRE Test Volumes

Table 2 (drawn from *Performance at the Top: From Elementary Through Graduate School*, Policy Information Report: Princeton, NJ, 1991) shows the number of examinees who took each of 15 different subject-matter GRE tests during the decade of the 1980s. In Table 3 are the percentages of examinees who took each test by year rounded to two decimal places (1/10,000). Although this is clearly too much precision for communicative purposes, we still see that the summed percentages do not all sum to 100.00%.

Table 4 shows the same percentages rounded to integer percentages. We now see that only one (1990) of the ten years shown still sums to 100%. To get them to sum properly requires only a tiny change in the table entries.

Table 4—Trends in GRE Subject Test Volumes (1981–1990) (integer percents)

	Year									
Subject	1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Biology	21	20	19	18	18	19	18	17	17	17
Chemistry	7	7	7	7	7	7	6	6	6	6
Computer Science	3	4	4	5	7	8	7	6	7	6
Economics	4	4	4	4	4	4	4	3	3	3
Education	3	3	2	2	2	2	2	2	2	2
Engineering	9	9	10	10	10	10	10	9	9	8
Geology	5	6	7	6	4	4	2	2	2	1
History	3	3	3	3	3	3	4	3	4	4
Literature	6	6	6	7	7	7	8	15	10	11
Mathematics	4	4	4	5	5	5	5	4	5	5
Music	4	3	3	2	2	2	2	2	2	2
Physics	5	5	5	5	5	5	5	5	6	6
Political Science	3	3	3	3	3	3	3	3	3	3
Psychology	20	22	20	20	20	20	21	20	23	24
Sociology	2	2	2	1	1	2	2	1	2	2
TOTAL	**99**	**101**	**99**	**98**	**98**	**101**	**99**	**98**	**101**	**100**

Table 5—Rounded to Integer Percents by RoundPro with Webster Method (changed entries boxed and boldfaced)

Subject	Year 1981	1982	1983	1984	1985	1986	1987	1988	1989	1990
Biology	21	20	19	18	**19**	19	18	17	17	17
Chemistry	7	7	7	7	7	7	6	6	6	6
Computer Science	3	4	**5**	**6**	7	8	**8**	6	7	6
Economics	4	4	4	4	4	4	4	3	3	3
Education	3	3	2	**3**	2	2	2	2	2	2
Engineering	9	9	10	10	10	10	10	9	9	8
Geology	5	6	7	6	4	**3**	2	2	2	1
History	3	3	3	3	3	3	4	**4**	4	4
Literature	6	6	6	7	7	7	8	15	**9**	11
Mathematics	4	4	4	5	5	5	5	**5**	5	5
Music	4	3	3	2	2	2	2	2	2	2
Physics	5	5	5	5	5	5	5	5	6	6
Political Science	3	3	3	3	3	3	3	3	3	3
Psychology	**21**	**21**	20	20	**21**	20	21	20	23	24
Sociology	2	2	2	1	1	2	2	1	2	2
TOTAL	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

Table 5 is suitably modified using Webster's method. The entries that are visibly changed are indicated by being boxed in and boldfaced. It is helpful to compare the changed entries with the corresponding entries in Table 3. For example the only change in 1981 and 1982 was for the entry under psychology. In 1981, 20.42 was rounded up to 21 and in 1982 21.62 was rounded down to 21. This amount of rounding (+.08 and −.12) corresponds to a change in the number of examinees taking the test of only 34 and 49, respectively (out of almost 9,000). This amount of rounding does very little damage to the integrity of the data when one remembers that rounding to the

Table 6—Table 5 With Rows Sorted and Spaced by Average Volume Over the Decade

Subject	Year 1981	1982	1983	1984	1985	1986	1987	1988	1989	1990	Mean
Psychology	21	21	20	20	21	20	21	20	23	24	**21**
Biology	21	20	19	18	19	19	18	17	17	17	**19**
Engineering	9	9	10	10	10	10	10	9	9	8	**9**
Literature	6	6	6	7	7	7	8	15	9	11	**8**
Chemistry	7	7	7	7	7	7	6	6	6	6	**7**
Computer Science	3	4	5	6	7	8	8	6	7	6	**6**
Physics	5	5	5	5	5	5	5	5	6	6	**5**
Mathematics	4	4	4	5	5	5	5	5	5	5	**5**
Geology	5	6	7	6	4	3	2	2	2	1	**4**
Economics	4	4	4	4	4	4	4	3	3	3	**4**
History	3	3	3	3	3	3	4	4	4	4	**3**
Political Science	3	3	3	3	3	3	3	3	3	3	**3**
Music	4	3	3	2	2	2	2	2	2	2	**2**
Education	3	3	2	3	2	2	2	2	2	2	**2**
Sociology	2	2	2	1	1	2	2	1	2	2	**2**
TOTAL	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**	**100**

nearest integer percent is essentially putting numbers into bins 213 wide anyway.

Last, if we may be forgiven an *obiter dictum*, note that Table 5 can be improved still further by including a measure of the popularity of each test and by reordering and spacing the rows in a way to communicate it. Table 6 is one version of such a table.

This work was supported by the ETS research allocation to the Research Statistics Group, and I am grateful for the freedom that this provides. I would like to thank Friederich Pukelsheim and Xiaohui Wang for their help in this exposition.

References and Further Reading

Balinski, M. L., and Young, H. P. (1982), *Fair Representation. Meeting the Ideal of One Man, One Vote*, New Haven, CT: Yale University Press.

Diaconis, P. and Freedman, D. (1979), "On Rounding Percentages," *Journal of the American Statistical Association*, 74, 359–364.

Dorfleitner, G., Happacher, M., Klein, T., and Pukelsheim, F. (1996), *Roundpro: An Emacs LISP Implementation of Rounding Procedures*, Technical Report 347, Augsburg, Germany:Institut für Mathematik, Universität Augsburg.

Ehrenberg, A. S. C. (1977), "Rudiments of Numeracy," *Journal of the Royal Statistical Society, Ser. A*, 140, 277–297.

Eisenhart, C. (1947), "Effects of Rounding or Grouping Data," in *Selected Techniques of Statistical Analysis for Scientific and Industrial Research and Production and Management Engineering*, eds. C. Eisenhart, M. W. Hastay, and W. A. Wallis, New York: McGraw-Hill.

Happacher, M. (1996), *Die Verteilung der Diskrepanz bei stationären Multiplikatormethoden zur Rundung von Wahrscheinlichkeiten*, Augsburg, Germany: Wissner.

Happacher, M. and Pukelsheim, F. (1994), *Rounding Probabilities: Unbiased Multipliers*, Augsburg, Germany: Report 525, Institut für Mathematik, Universität Augsburg.

——(1997), "And Round the World Away," in *Proceedings of the Conference in Honor of Shayle R. Searle, Aug. 9–10*, 1996, Ithaca, NY: Biometrics Unit, Cornell University, 93–108.

Mosteller, F., Youtz, C., and Zahn, D. (1967), "The Distribution of Sums of Rounded Percentages," *Demography*, 4, 850–858.

Wainer, H. (1993), "Tabular Presentation," *Chance*, 6 (3), 52–56.

——(1997a), "Improving Tabular Displays: With NAEP Tables as Examples and Inspirations," *Journal of Educational and Behavioral Statistics*, 22, 1–30.

——(1997b), *Visual Revelations: Graphical Tales of Fate and Deception From Napoleon Bonaparte to Ross Perot*, New York: Copernicus Books.