Two Applications of Venn Diagrams

KEYWORDS:
Teaching;
Area proportional Venn diagrams;
Conjunction fallacy.

Glenn Waddell, Jr.
North Valleys High School, Reno, Nevada, USA.
e-mail: gwaddellnvhs@gmail.com

Robert J. Quinn
University of Nevada, Reno, USA.
e-mail: quinn@unr.edu

Summary
The Venn diagram is suggested as a graphical solution to conjunction fallacies and a modification of it is suggested to more fully communicate set relations.

VENN DIAGRAMS AND CONJUNCTION FALLACIES

Consider the two statements:
P: Linda is a bank teller.
Q: Linda is a bank teller and active in the feminist movement.

When people are presented with these two propositions, P and Q, and are asked to choose which proposition is more likely, they often choose Q. Clearly, however, the set of bank tellers is larger than the conjunction of bank tellers and feminists; therefore, it is more likely that Linda is a bank teller than a bank teller and a feminist (Tversky and Kahneman 1983). The mistake of saying that a conjunctive statement is more likely than any of its components is referred to as the "conjunction fallacy".

The literature in cognitive science and psychology is rife with discussion concerning why people succumb to this fallacy. Some authors argue that people do not understand what the conjunction is, while others demonstrate that different questions lead to different percentages of people answering the questions correctly. Published articles discussing this fallacy tend to be published in the fields of cognitive science and psychology (Bonini et al. 2004; Gaynor et al. 2007; Tentori et al. 2004). While this fallacy is purely a mathematical issue, the number of published articles directed towards mathematics teachers is relatively quite small (Tomlinson and Quinn 1997; Shaughnessy 1993). This is an interesting void, given that the problem can be modelled with Venn diagrams (Venn 1881), which, of course, requires little mathematical background.

The Venn diagram, of course, may be used to teach conjunction and disjunction – alternatively: AND and OR, or intersection and union (Kennedy et al. 2001; Schultz et al. 2001). Returning to the "Linda problem", we can use the diagram in figure 1. If this picture is presented when the question is asked, individuals may be much less likely to say the conjunction is more probable, because the region corresponding to bank tellers *and* feminists – namely, the intersection of the two circles, is strictly smaller than the region corresponding to just bank tellers. The use of Venn diagrams makes what seemed like a puzzling problem much easier.

AREA PROPORTIONAL VENN DIAGRAMS

Such Venn diagrams, however, may be made even more useful making problem information even more apparent and easier to see. To this end, we propose that *area proportional* Venn diagrams, discussed below, should be used to a much greater extent than they currently are. Such diagrams – not ordinarily presented in US schools, are certainly not unheard of (see e.g. Frankcom 2008).

To illustrate the use of area proportional Venn diagrams, we present an example in the medical literature given by Marshall (2001). In particular, he discusses the occurrence of carditis (inflammation of the heart) and chorea (an involuntary movement disorder) in a sample of 271 patients with acute rheumatic fever. The data are given in table 1.

	Chorea	*No chorea*	*Total*
Carditis	14	115	129
No carditis	11	131	142
Total	25	246	271

Table 1. Occurrence of carditis and chorea among 271 patients with acute rheumatic fever

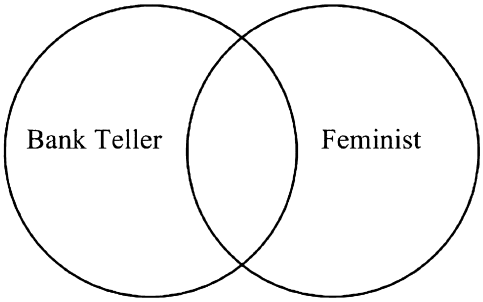

Fig 1. The "Linda problem"

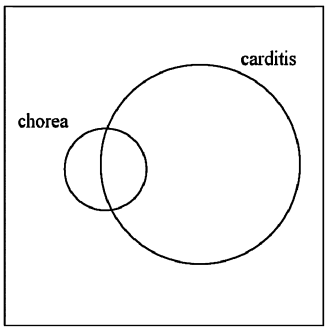

Fig 2. Area proportional Venn diagram for carditis and chorea occurrence among those having rheumatic fever

Letting CH denote chorea and CA denote carditis, note that we have, with respect to the population of rheumatic fever individuals, the estimates:

$$P(\mathrm{CA}) = \frac{129}{271} \cong 0.48, P(\mathrm{CH}) = \frac{25}{271} \cong 0.09$$

This is represented in figure 2 by having the carditis circle contain an area just under half of that of the rectangle. Likewise, the chorea circle should contain an area a bit under one-tenth of that of the rectangle. To correctly determine the amount of overlap between the carditis and chorea circles, any one of several conditional probability estimates may be used. In particular, consider the estimate:

$$P(\mathrm{CA}|\mathrm{CH}) = \frac{14}{25} = \frac{P(\mathrm{CA} \cap \mathrm{CH})}{P(\mathrm{CH})}$$

As $14/25 \cong 0.56$, the intersection of the two circles should have an area that is 56% of that in the chorea circle. The rectangle, of course, represents the population of interest – in our example, it corresponds to individuals having rheumatic fever.

To summarize, the relative *areas* of the circles correspond to the numbers that appear in table 1. With some practice, then, given such a diagram, students can at least roughly estimate the proportions used to construct it. We claim that these area proportional Venn diagrams are easy to read, and clearly show the relations between the various events.

It is also worth pointing, by the way, that area proportional diagrams may be used to help students distinguish between $P(\mathrm{A}|\mathrm{B})$ and $P(\mathrm{B}|\mathrm{A})$. Both of these conditional chances may be expressed as ratios with a common value of $P(\mathrm{A} \cap \mathrm{B})$ for the numerator but, of course, with the denominators $P(\mathrm{B})$ and $P(\mathrm{A})$, respectively, which may be quite different.

To summarize, the use of equal-sized circles in Venn diagrams is perfectly sensible if there is no knowledge of the relative frequency of the corresponding events. However, if there is some knowledge of these relative frequencies, then it is certainly sensible to incorporate such information. Doing so allows others to readily grasp the underlying proportions related to the events displayed.

Acknowledgement

Figure 2 is from Marshall (2001). The use of such was kindly granted by Wiley-Blackwell.

References

Bonini, N., Tentori, K. and Osherson, D. (2004). A different conjunction fallacy. *Mind and Language*, **19**(2), 199–210.

Frankcom, G. (2008). Statistics teaching and learning: The New Zealand experience. http://www.tsg.icme11.org/document/get/489 (accessed 16 February 2009).

Gaynor, S. T., Washio, U. and Anderson, F. (2007). The conjunction fallacy: A derived stimulus relations conceptualization and demonstration experiment. *The Psychological Record*, **57**, 63–85.

Kennedy, P. A., McGowan, D., Schultz, J. E., Hollowell, K. A. and Jovell, I. (2001). *Algebra One Interactions: Course 1*. New York: Holt, Rinehart and Winston.

Marshall, R. J. (2001). Displaying clinical data relationships using scaled rectangle diagrams. *Statistics in Medicine*, **20**, 1077–1088.

Schultz, J. E., Kennedy, P. A., Ellis, Jr., W. and Hollowell, K.A. (2001). *Algebra 1*. New York: Holt, Rinehart and Winston.

Shaughnessy, J. M. (1993). Probability and statistics. *Mathematics Teacher*, **86**(3), 244–248.
Tentori, K., Bonini, N. and Osherson, D. (2004). The conjunction fallacy: A misunderstanding about conjunction? *Cognitive Science*, **28**, 467–477.
Tomlinson, S. and Quinn, R. (1997). Understanding conditional probability. *Teaching Statistics*, **19**(1), 2–7.
Tversky, A. and Kahneman, D. (1983). Extension versus intuitive reasoning: The conjunction fallacy in probability judgment. *Psychological Review*, **90**(4), 293–315. doi: 10.1037/0033-295X.90.4.293. http://content2.apa.org/journals/rev/90/4/293 (accessed 16 February 2009).
Venn, J. (1881). *Symbolic Logic*. London: Macmillan.

NEWS AND NOTES

For those of you who enjoy controversy, I would invite you to search Internet postings concerning a paper of Dr Daryl Bem on extrasensory perception (ESP) to appear in *The Journal of Personality and Social Psychology* this calendar year. Apparently, using hypothesis testing, a claim is made for the existence of ESP. One, among several, postings you may wish to examine concerning this is an article by Benedict Carey published 10 January 2011 in *The New York Times*. The article by Carey briefly discusses some of the limitations of significance tests.

Quite a number of activities were held around the world in connection with World Statistics Day last year (20 October 2010), among them a preview of a BBC film entitled "Joy of Stats" featuring Hans Rosling shown at the Royal Statistical Society in London. To see excerpts of this film – which shows the power and beauty of statistical graphics, visit The Open University site http://www.open.ac.uk/openlearn/whats-on/the-joy-stats. The entire hour-long film may be viewed http://www.gapminder.org/.

Two sites concerned with statistical literacy worth visiting if, indeed, you haven't already visited them, are http://www.statlit.org/ (where you can join the Statistical Literacy News e-mail list) and 'getstats' site, http://www.getstats.org.uk/, of The Royal Statistical Society.

The Census at School program, which has been promoted in this journal before, began in the UK and now includes Australia, Canada, Ireland, Japan, New Zealand, South Africa and (most recently) the US. The goals of the program, from the Census at School International site, include: encouraging children to get involved with data handling and learn statistical skills, provide real data handling activities and increase awareness of what national censuses are and what they are for. See the site http://www.censusatschool.com/ for further details.

The 58th World Statistics Congress of the International Statistical Institute (ISI) will be held in Dublin 21–26 August 2011. According to Gerard O'Hanlon, the Director General of the Central Statistics Office in Ireland, "The biennial meeting of the ISI is probably unique in bringing together statisticians from across the disciplinary spectrum – official statistics; business, financial, environmental and industrial statistics; surveying and sampling computational and mathematical statistics; probability and stochastic modeling and statistical education" (http://www.isi2011.ie). Further details are available at the aforementioned Web site.

The 5th International Conference on Mathematics & Statistics will be held 13–16 June 2011 in Athens, Greece. This conference is sponsored by the Mathematics & Statistics and Education Research Units of the AThens INstitute for Education and Research (ATINER). Further details may be found at http://www.atiner.gr/mathematics.htm.

Finally, because of an ongoing family medical issue, I will be stepping down as Editor of *Teaching Statistics* at the end of the calendar year. Applications for the Editor position are now being taken. Please see the inside back cover of this issue or the Royal Statistical Society Centre for Statistical Education Web site http://www.rsscse.org.uk for details of the application process.

Roger Johnson
Editor