

Quantitative Competency Made Simple: 10 Questions to Ask about Numbers

Neil Lutsky, Kenan Professor of Psychology and Director of the Carleton
Quantitative Inquiry, Reasoning, and Knowledge (Quirk) Initiative*

[Handouts in bold]

- Seelye, K. Q. (2006, February 10). Lurid numbers on glossy pages! (Magazines exploit what sells). *The New York Times*, p. A1.
- Lohr, S. (2006, April 9). This boring headline is written for Google. *The New York Times*, p. 14.**
- More or Less (BBC News Programme). http://news.bbc.co.uk/2/hi/programmes/more_or_less/default.stm
- An Inconvenient Truth*. <http://www.climatecrisis.net/>
- Story, L. (2005, September 20). Many women at elite colleges set career path to motherhood. *The New York Times*, p. 1..
<http://www.nytimes.com/2005/09/20/national/20women.html?ex=1284868800&en=6a8e0c413c09c249&ei=5090&partner=rssuserland&emc=rss>
- Shafer, J. (2005, September 20). Weasel-words rip my flesh! *Slate*. <http://slate.msn.com/id/2126636/>**
- Stein, R., & Kaufman, M. (2005, October 21). New diabetes drug poses major risks, panel says. *The Washington Post*, p. A02. <http://www.washingtonpost.com/wp-dyn/content/article/2005/10/20/AR2005102001145.html>**
- Goldacre, B. (2005, September 8). Don't dumb me down. *The Guardian*. <http://www.guardian.co.uk/life/badscience/story/0,12980,1564369,00.html>**
- Levitt, S. D., & Dubner, S. J. (2005). *Freakonomics: A rogue economist explores the hidden side of everything*. NY: William Morrow.
- Johnson, S. (2006). *The ghost map*. New York, NY: Penguin.
- Abelson, R. P. (1995). *Statistics as principled argument*. Hillsdale, NJ: Lawrence Erlbaum Press.
- Gross, D. (2006, June 4). When sweet statistics clash with a sour mood. *The New York Times*, Business, p. 1.**
- Norris, F. (2006, October 7). Dow set a record in height, but weight had a lot to do with that. *The New York Times*, Business, p. 1.**
- Tufte, E. R. (1997). *Visual explanations*. Cheshire, CT: Graphics Press.
http://www.edwardtufte.com/tufte/books_textb
- Hemenway, D. (2004). *Private guns public health*. Ann Arbor, MI: University of Michigan Press.
- Miller, J. D., Scott, E. C., & Okamoto, S. (2006). Public acceptance of evolution. *Science*, 313 (5788), pp. 765-766.
- Rosato, D. (2002, September 15). Investing, worried about corporate numbers? How about the charts? *The New York Times*, Section 3, p. 71. [Deanna Oxender Burgess, Florida Gulf Coast University]
- Taleb, N. N. (2004). *Foiled by randomness: The hidden role of chance in markets and life, 2nd Edition*. NY: Random House.
- Levy, S. (2006, January 31). Does your iPod play favorites? *Newsweek*.
<http://www.msnbc.msn.com/id/6854309/site/newsweek/>

- Reifman, A. (2006). The hot hand in sports. <http://thehothand.blogspot.com/>
- Saul, S. (2005, October 21). Article says diabetes pill would increase coronary risks. *The New York Times*.
- Tavernise, S., & McNeil, Jr., D. G. (2006, October 11). Iraqi dead may total 600,000, study says. *The New York Times*, p. 1.
- Pilcher, J. J., & Huffcutt, A. I. (1996). Effects of sleep deprivation on performance: A meta-analysis. *Sleep*, 19(4), pp. 318-326.
- Ioannidis, J. P. A. (2005). Contradicted and initially stronger effects in highly cited clinical research. *Journal of the American Medical Association*, 294(2), pp. 218-228.
- Burnham, G., et al. (2006). The human cost of the war in Iraq. *The Lancet* (in press). Available at <http://web.mit.edu/cis/human-cost-war-101106.pdf>
- Conrad, P. (1996). *The rooster's gift*. NY: HarperCollins.
- Rivlin, G. (2006, August 13). In vino veritas? Why wine ratings might not pass the sobriety test. *The New York Times*, Section 3, p. 1.
- Newport, F., Saad, L., & Moore, D. (1997). How are polls conducted? <http://media.gallup.com/PDF/FAQ/HowArePolls.pdf>
- Porter, T. (1996). *Trust in numbers: The pursuit of objectivity in science and public life*. Princeton: Princeton University Press.

*Carleton's Quantitative Inquiry, Reasoning, and Knowledge (Quirk) Initiative is supported by grant P116B040816 from the Fund for the Improvement of Postsecondary Education (www.ed.gov/FIPSE), U. S. Department of Education.

The 10 Questions (with some elaborations and expansions):

**I. *What do the numbers show?* Where's the numerical evidence?
What were the exact figures?**

II. *How representative is that?* What's the central tendency or base rate?

"For instance is no proof."

Mean, Mode, and Median.

Interrogating averages:

Are there extreme scores?

Are there meaningful subgroups?

Who's in the denominator?

What's the variability (**standard deviation**)? $\sigma = \sqrt{\sum(x_i - \mu)^2 / N}$

III. *Compared to what?*

Interrogating a graph:

What's the Y-axis? Is it zero-based?

Does it K.I.S.S., or is it filled with ChartJunk?

IV. *Is the outcome statistically significant?* Is it chance or is the outcome unlikely to have come about by chance?

"Chance is lumpy."

Criterion of sufficient rarity due to chance: $p < .05$

V. *What's the effect size?* How can we take the measure of how substantial an outcome is?

How large is the mean difference? How large is the association?

Standardized mean difference (d): $d = (\mu_1 - \mu_2) / \sigma$

VI. *Are the results those of a single study or of a literature (e.g., meta-analysis)?*

VII. *What's the research design (correlational or experimental)?*

What's the source of the numbers: PFA, peer-reviewed, or what?

Design matters: Experimental vs. correlational design.

Experimental Design: Random assignment to the conditions of the **Independent Variable** with possible effects on the **Dependent Variable** evaluated.

Randomized Controlled Trials (RCT): Research trials in which participants are

randomly assigned to the conditions of the study.

Double blind trials are those in which neither the researcher nor the patient know the treatment condition.

Correlational Design: Measuring existing variation and evaluating co-occurrences, possibly controlling for other variables. Two variables are correlated if the values of the one variable relative to its mean are similar to the values of a second variable relative to that second variable's mean.

*What variables were controlled for or were not considered?
Does the design justify a causal claim?*

Interrogating associations (correlations):

Are there extreme pairs of scores (outliers)?

Are there meaningful subgroups?

Is the range of scores in a variable restricted?

Is the relationship non-linear?

VIII. How was the variable operationalized? What meaning and what degree of precision does the measurement procedure justify?

What elements and procedure result in the assignment of a score to a variable?

What exactly was asked?

What's the scale of measurement?

How might we know if the measurement procedure is a good one?

Reliability: Repeated applications of the procedure result in consistent scores.

Validity: Evidence supports the use to which the measure is being put.

IX. Who's in?

What domain is being evaluated? Who's in? Who's not?

Is the sample from that domain representative, meaningful, and sufficient?

Is the sample random?

X. To whom can I turn?

<http://apps.carleton.edu/collab/quirk/resources/consultants/>