---

30 July 2007 ASA   1

Numbers in the News

Milo Schield
W. M. Keck Statistical Literacy Project
Augsburg College

Cynthia Schield
Former co-publisher of the *Napa County Record*

ASA JSM 2007

Salt Lake City, Utah   30 July 2007
www.StatLit.org/pdf/2007SchieldASA6up.pdf

---

30 July 2007 ASA   2

GAISE Guidelines

The ASA GAISE College report recommended:

- *"introductory courses in statistics should, …strive to emphasize statistical literacy ....*

- assessing statistical literacy by students *"interpreting or critiquing articles in the news and graphs in media*."

---

30 July 2007 ASA   3

Statistical Literacy: Lack of Data

"*Clarity on the building blocks of statistical literacy are needed*" before addressing "*assessment and instruction…*" Gal (2001)

"*No comparative analysis has ... mapped out the ... statistical ... concepts and topics ... that adults may encounter…*" Gal (2003)

If statistical literacy is to be empirically-based, the use of "numbers in the news" must be analyzed.

---

30 July 2007 ASA   4

Numbers in the News

Selected 100 news articles using numbers if they:
- had the word 'study' in the title,
- involved samples (surveys) or medical tests,
- involved longitudinal data or random assignment,
- involved taking into account confounders, OR
- used statistics as evidence for causation.

Excluded articles: sports, weather & stock prices.

Analyzed 93 characteristics per story.

---

30 July 2007 ASA   5

Content of Articles
SIMPLE DATA (69%)

Percentage of articles that have the following:
- 48%: Percentages
- 29%: Numbers (counts or sums)
- 4%: Mean/average
- 4%: Rates (c.f. unemployment rate)
- 3%: Ratios (e.g., miles per gallon)
- 2%: Ranks or percentiles
- 1%: Percentage points
- 0%: Median

---

30 July 2007 ASA   6

Content of Articles
COMPLEX DATA (5%)

Percentage of articles that have the following:
- 4%: Slope
- 1%: Range
- 1%: Correlation (qualitative or numeric)
- 1%: Effect size or elements thereof

- 0%: Standard deviation, z-score
- 0%: Coefficient of Variation
- 0%: Relative risk or relative ratio
- 0%: Odds ratio or Gini coefficient
- 0%: R-squared

Content of Articles
COMPARISONS (22%)

Percentage of articles that compare these:

- 8%: Numbers (counts or measures)
- 7%: Cases attributed (statistical deaths)
- 5%: Percentages
- 4%: Rates
- 1%: Means/averages

- 0%: Change in numbers
- 0%: Medians
- 0%: Ratios
- 0%: Percentage attributable

Content of Articles
RATIO GRAMMAR (68%)

Percentage of articles by type of grammar:

- 40%: Percent (X% of <whole> are <part>)
- 22%: Likely/Prevalent
- 18%: Ratios   (e.g., miles per gallon)
- 15%: Chance/risk/probability
- 15%: Rates   (e.g., birth rate, rate of births)
- 6%: Percentage*
- 4%: Attributable/attributed

*   The percentage of <whole> who are <part>
    Among <whole>, the percentage of <part>

Content of Articles
Statistical Inference (49%)

Percentage of articles that involve inference:

- 44%: Use sample
- 40%: Give sample size
- 11%: Mention "significant" or "significantly"
- 9%: Give Margin of Error
- 7%: Mention random sample
- 1%: Give a Confidence Interval
- 1%: Mention "Statistically significant"

- 0%: Give p-value

Content of Articles:
DESIGN OF STUDY (45%)

Percentage of articles that indicate study design:

- 25%: Controlled study (two or more groups)
- 20%: Observed several times (longitudinal)
- 10%: Subject manipulation (experimental drugs)
- 9%: Cohort (high school graduating class)
- 8%: Subject blinded (placebo)
- 7%: Factor controlled or taken into account
- 5%: Plausible confounder indicated
- 1%: Researcher blinded
- 1%: Random assignment

Content of Articles:
CAUSATION & ASSEMBLY

Percentage of stories that imply causation (60%)

Action verbs: Change, alter, explain. Increase, improve, save, prevent. Reduce, cut, kill, hurt

Adjectives: Harmful, safe, effective

Percentage of articles that use assembly* (68%)

affordable, excessive, discrepancy, dangerous rich, poor, unhealthy, dysfunctional, poverty…

* Choice of definition, groups or measures

Conclusion

The goal of statistical literacy is to help people analyze numbers in the news.

Statistical Literacy must focus on:

- assembly / construction of categories (68%)
- ratio grammar (68%): percent and likely
- association versus causation (60%)
- study design (45%): longitudinal & manipulation.
- significant (11%) vs. statistically significant (1%)
- margin of error (9%) and cases attributable (7%)