

# **Numbers in the News**

---

**Robert L. Raymond**

**Emeritus, University of St. Thomas**

**Milo Schield**

**W. M. Keck Statistical Literacy Project  
Augsburg College**

***ASA JSM 2008***

***Denver, Colorado***

***6 August 2008***

***[www.StatLit.org/pdf/2008RaymondSchieldASA1up.pdf](http://www.StatLit.org/pdf/2008RaymondSchieldASA1up.pdf)***

# 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."*

# Statistical Literacy: Lack of Data

---

Statistics course design should model the use of real data.

*“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 real- world use of numbers must be analyzed; the news media are a good place to start.

# Numbers in the News

---

Selected 899 news articles that used numbers to make inferences: generalize, predict or explain.

- 2007: 250 articles researcher analyzed, 93 traits.
- 2008: 160 articles researcher analyzed, 73 traits
- POOLED: weighted average for common traits.
- 2008: Machine-readable content of 899 articles computer-searched for prevalence of 231 terms.

Excluded numerical articles that didn't involve inferences: sports, weather & stock prices.

# **Pooled Content of Articles**

## **SIMPLE DATA**

---

**Percentage of articles that have the following:**

48%: Percents

33%: Numbers (counts or sums)

31%: Rates (c.f. unemployment rate)

6%: Ratios (e.g., miles per gallon)

4%: Mean/Average<sup>07</sup>

2%: Ranks or percentiles

# **Pooled Content of Articles**

## **COMPLEX DATA**

---

**Percentage of articles that have the following:**

3%: Slope

1%: Range

1%: Effect size or elements thereof

---

**For all 899 articles:**

0%: Standard deviation, z-score

0%: Coefficient of Variation

0%: Relative risk

0%: Odds ratio or Gini coefficient

# **Pooled Content of Articles COMPARISONS**

---

## **Percentage of articles that include:**

65%: Quantitative comparisons<sup>08</sup>

21%: Qualitative comparisons

4%: “Attributed[able] to”

4%: Cases attributed to

# **Pooled Content of Articles**

## **RATIO GRAMMAR**

---

### **Percentage of articles by type of grammar:**

45%: Percent of (X% of <whole> are <part>)

37%: Rates (e.g., birth rate, rate of births)

28%: Chance/risk/probability

10%: Ratios (e.g., miles per gallon)

3%: Percentage\*

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



# **Pooled Content of Articles Statistical Inference**

---

## **Percentage of articles that involve inference:**

62%: Assert cause

59%: Use sample

53%: Give sample size

14%: Mention “significant” or “significantly”

2%: Mention random sample

1%: Give Margin of Error

1%: Mention “Statistically significant”

---

0%: Give a Confidence Interval

0%: Give p-value

# **Pooled Content of Articles: DESIGN OF STUDY**

---

## **Percentage of articles that indicate study design:**

- 26%: Controlled study (two or more groups)
- 22%: Observed several times (longitudinal)
- 15%: Cohort
- 11%: Subject manipulation (experimental drugs)
- 11%: Controlled by selection<sup>08</sup>
- 9%: Factor controlled or taken into account
- 5%: Subject blinded (placebo)
- 4%: Plausible confounder indicated<sup>07</sup>
- 2%: Random assignment

# **Content of Articles: Troublesome Inference**

---

Plausible alternate explanations for numerical results based on researcher judgment:

45%: Assembly\*

42%: Confounding<sup>08</sup>

11%: Bias<sup>08</sup>

9%: Chance or random effect<sup>08</sup>

**\* Choice of definition, groups or measures**

# Conclusion

---

**To help people analyze numbers in the news,  
Statistical Literacy must focus on:**

- Sampling, sample size (60%)
- Association versus causation (60%)
- Assembly (45%)
- Rate, ratio (40%)
- Study design: longitudinal & manipulation (40%)
- Significant (14%) vs. statistically significant (1%)