

# A Brief Look at Statistical Literacy

Iddo Gal

This paper appeared in the Fall 2004 issue of Math Practitioner, vol 10(2), pages 4 - 8. The Math Practitioner is published by the Adult Numeracy Network: <http://www.literacynet.org/ann>.

*Excerpted from: Gal, Iddo (2002) Adults' Statistical Literacy: Meanings, Components, Responsibilities, International Statistical Review, 70(1), 1 – 25. Dr. Gal is a Senior Lecturer at the Department of Human Services, University of Haifa, Israel.*

Many curriculum frameworks and national and international educational initiatives stress the need for all people to function effectively in an information-laden society. Yet, statistical literacy, critical to becoming more informed citizens and employees, is often neglected. “Statistical literacy”, as posited by Gal, refers to two interrelated components: (1) people’s ability to *interpret and critically evaluate* statistical information, data-related arguments, or probabilistic phenomena, which they may encounter in diverse contexts, and when relevant, and (2) their ability to *discuss or communicate* their reactions to such statistical information, such as their understanding of the meaning of the information, their opinions about the implications of this information, or their concerns regarding the acceptability of given conclusions.

Gal suggests that statistical literacy involves both knowledge elements and dispositional elements described in the enclosed model. Applied together, they form the basis for statistical literacy. A look at this model shows that statistical literacy is based upon the availability of literacy knowledge and skills, as much as about content knowledge in relevant areas in statistics and mathematics. However, other knowledge elements are essential as well, as explained below.

Knowledge Elements	Dispositional Elements
Literacy skills Statistical knowledge Mathematical knowledge Context/World knowledge Critical questions	Beliefs and attitudes  Critical stance
<b>Statistical Literacy</b>	

## Knowledge Elements

### Literacy Skills

Literacy skills are critical for statistical literacy because virtually all statistical messages are conveyed through written text (e.g., in newspapers) or oral text (e.g., on TV). Some messages also require that individuals navigate through and comprehend displays that employ a combination of non-prose text and symbolic or numerical information, such as in tables, graphs, or charts. Many statistical terms, such as “sample” or “correlation” often cannot be represented in numbers when discussed in an everyday context, and are hence communicated only through words. [See statistical example on page 7.]

### Statistical Knowledge

Clearly, knowledge of statistical and probabilistic concepts and procedures is required for statistical literacy. Gal outlines five aspects of the statistical knowledge base:

1. Knowing why data are needed and how data can be produced
2. Familiarity with basic terms and ideas related to descriptive statistics
3. Familiarity with basic terms and ideas related to graphical and tabular displays
4. Understanding basic notions of probability
5. Knowing how statistical conclusions or inferences are reached

### Mathematical Knowledge

Some mathematical knowledge that is not normally associated with the study of statistics is required for statistical literacy. The most basic knowledge needed may be that adults realize that any attempt to summarize a large number of observations (data) by

a concise quantitative statement (such as percent, mean, median, probability, etc.) requires the use of some mathematical tools and procedures. “Number sense” is critical for interpreting mathematical representations of statistical ideas. Percentages are often used to summarize and communicate results of surveys or differences between subgroups, and probabilities are often communicated via proportional representations such as ratios or odds, as well as via percentages.

### **Context/World Knowledge**

Correct interpretation of statistical messages requires that adults place the messages in the appropriate real-life context. For example, if an individual is not familiar with the context in which data were collected, it is difficult to consider reasons for variability in the data, to understand how data collection could go wrong, to come up with ideas for why the validity of the interpretations offered may be questionable, or to make sense of conclusions presented on the basis of a statistical investigation.

### **Critical Questions**

Statistical messages are produced by different sources, such as by the general media, advertisers, politicians, workplaces, or governmental bodies. Often, these sources have their own need or goal for presenting the messages, not necessarily to present an unbiased report of findings or interpretations. For example, those who generate messages regarding diseases, accidents, and the like can make them seem more significant and frightening by choosing to report absolute numbers (e.g. 2500 people nationwide suffer from X) rather than to downplay them by using incidence rates (e.g. 1 in every 100,000 suffer from X).

Because of the possibility for biased or simply incompetent reporting, individuals need to be concerned about the validity of messages. Gal and others have developed a list of “worry questions” that adults should ask themselves regarding statistical information being communicated:

1. Where did the data come from? Is this kind of study (e.g., survey, experiment) reasonable in this context?

2. How many people did actually participate? Did the sample include people/units which are representative of the population or is the sample biased in some way? Overall, was the sample large and representative enough and could it reasonably lead to valid inferences about the target population?

3. How reliable or accurate were the instruments or measures used to generate the reported data?

4. What is the shape of the underlying distribution of raw data? Can it really be summarized by, say, an average or a median? Could outliers cause a summary statistic to misrepresent the true picture?

5. Is a given graph drawn appropriately, or does it distort trends in the data?

6. How was a given probabilistic statement derived? Are there enough credible data to justify the estimate of likelihood given?

7. Are the claims made here sensible and supported by the data (e.g. is correlation confused with causation, or a small difference made to loom large)? Should additional information or procedures be made available to enable me to evaluate the sensibility of these arguments? (E.g. did the writer “conveniently forget” to specify the base of a reported percent-of-change, or the actual sample size?)

8. Are there alternative interpretations for the meaning of the findings or different explanations for what caused them? (E.g. did an intervening or “moderator” variable affected the results?) Are additional implications not mentioned?

### ***Dispositional Elements***

To be able to act in a statistically literate way, it is not enough to have the cognitive knowledge bases and skills described above. Adults also need to be able to take a critical stance regarding statistical information presented to them, and possess certain attitudes and

beliefs, in order to motivate, sustain, and support their actions. “Actions” in this regard may be an internal process of thinking about the information given or running mentally through the list of worry questions; it may also be an external action such as initiating a discussion or doing some calculations in response to a statistical message encountered on the media.

Certain beliefs and attitudes influence individuals’ critical stance and willingness and ability to take action in response to statistical information. Adults should develop a positive view of themselves as individuals capable of statistical and probabilistic reasoning as well as a willingness and interest to “think statistically”. This suggests that adults have some understanding and appreciation for the power of statistical processes. Metacognitive capacities such as logical reasoning, curiosity, open-mindedness, further enhance statistical literacy. Most importantly, adults need to feel that it is okay to be critical of statistical messages.

## Conclusions

Statistical literacy encompasses a set of knowledge bases and dispositions which somewhat overlap broader notions of adult numeracy or scientific literacy. Yet, it is sufficiently complex and multifaceted to require careful attention from those who want to improve the ability of adults to cope with and understand quantitative information in the world around them. The development of adults’ statistical literacy requires attention to all the five knowledge bases described above, not just to formal knowledge of statistics or percents. Some examples for issues that could be addressed as part of instruction for statistical literacy include the following:

- Understanding results from polls, samples, and experiments as reported in newspapers and other media channels,
- Understanding probabilistic aspects of statements about risk and side effects as reported in newspapers and other media,
- Learning about styles, conventions, and biases in journalistic reporting,
- Becoming familiar with “worry questions” in tandem with experience in applying them to real examples,

- Discussing the usefulness of statistical investigations and the legitimacy of being critical of messages from diverse sources, in order to develop positive beliefs and attitudes about statistics, but also the willingness to be critical of statistical messages one encounters in certain contexts.

---

---

**On page 8** is a graph for you to use with your students to begin to help them to become more critical readers of statistical information. The graph is interesting in that it can give rise to two different discussions:

(a) the presence of “misleading” graphs (the length of bars/pencils is off compared to several of the numbers)

(b) the need that arises in many contexts (have students discuss examples) to compare two or more groups and decide if observed differences are large and significant. Could such a difference be the result of chance fluctuations? The answer depends on sample size.

Also, this graph can be used to discuss the connection between summary numbers and underlying distributions- can we assume they are normal? Could one be bimodal? Etc.

**More information and links to resources for statistical literacy education are offered by the International Statistical Literacy Project**

**(<http://course1.winona.edu/cblumberg/islphome.htm>). Materials of special interest for adult educators appear under “List of ISLP Resource Webpages”. The International Association for Statistics Education ([www.stat.auckland.ac.nz/~iase/index.php](http://www.stat.auckland.ac.nz/~iase/index.php)) has information about conferences, research publications, and links to educational resources.**

# Statistical Example

*Created by Iddo Gal to illustrate points in his Statistical Literacy Article (pp. 4 - 6)*

Consider this news item which appeared on the CNN website on September 23, 2003:

## **Studies: Lou Gehrig's rate higher in Gulf War vets**

WASHINGTON (AP) - Veterans of the 1991 Persian Gulf War were at least twice as likely to be diagnosed with Lou Gehrig's disease as non-Gulf veterans or other people younger than 45, according to two new studies.

This item is an example for articles of interest to many citizens which discuss health-related news (e.g., side effects from using Vioxx, dangers associated with living near power lines, etc), so is worth some attention. Ask your students: *What does this text tell us, in particular about the probability of getting the Lou Gehrig's disease?*

This item was chosen as it illustrates how interpretation of a statistical message requires the joint activation of several knowledge bases and dispositions. To make full sense of this seemingly simple text, readers have to:

1. Notice that this article does make a claim about the likelihood of an event (getting the disease), yet in *relative* terms, that is, "Twice as likely". Hence, the actual probability is *not* known (and for all we know could be very small, such as "1 in 650,000", or "0.003%" or similar).
2. Notice that the statement of probability is made fuzzier because of the use of modifiers. In this case, "at least" is used – but it can have two different meanings: either that the figure provided ("twice") is rounded (i.e., the

actual level is higher than twice), or that the probability is an *average* of two separate ratios, derived from comparisons to nongulf veterans and to the general under-45 population in the two studies mentioned; and

3. Be aware that contextual information *not made explicit* in the text is needed to make sense of the information that is given. The benefit of knowing that the disease is "twice as likely" (without knowing the actual probability) does become apparent, if the reader is aware of the debate as to whether or not the rate of several illnesses among Gulf war veterans was higher than the rate among non-Gulf soldiers, or caused by exposure to toxic chemical agents.

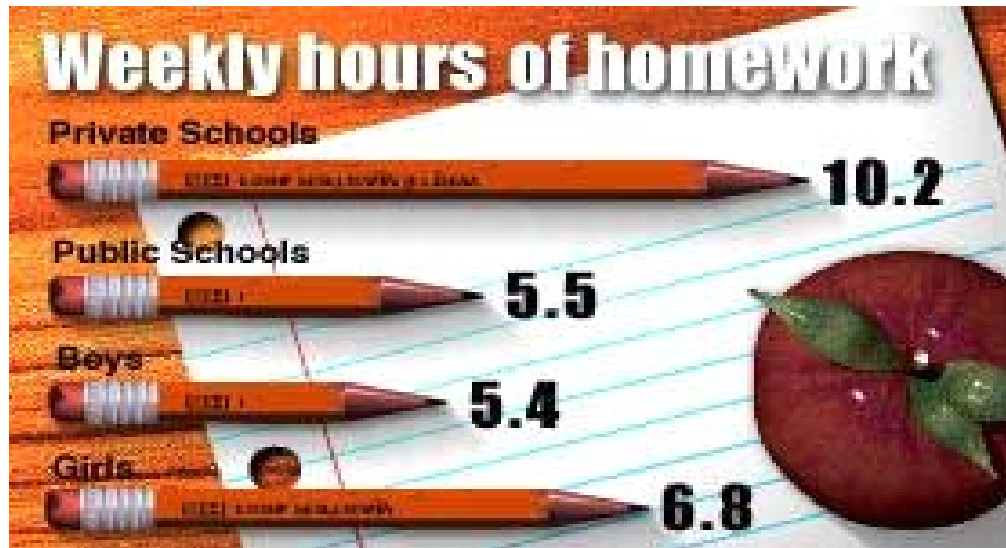
This example illustrates that readers and listeners cannot take probabilistic statements for what they are, but rather have to be able to ask a number of critical questions. It also provides teachers with an opportunity to discuss in an authentic context many issues connected to broader aspects of numeracy.

Sample topics for class discussion are:

- why are verbal phrases describing estimates of probability used in some contexts (point #1) and how should they be interpreted,
- the meaning of very small numbers, such as "0.03%" (point #1),
- the reason for and the interpretation of rounded numbers or averages reported in the media (point #2),
- the importance of understanding the context to help in interpretation (point #3).

More broadly, such an example can help to show students that adults should adopt a critical stance and know what critical questions to ask upon encountering a statement of probability. Hence, the use of such an example, or others which abound in newspapers, also enables teachers to initiate in the classroom an honest discussion of feelings and dispositions about chance, about the relevance of general statistics to an individual person's life.

# What does this graph tell you?



[The graph above is from USA Today (several years ago).]

Without looking at the numbers, what does the graph suggest to you? If you just looked at the length of the pencils, what numbers would you suggest that they might each represent?

Now look at the numbers. What are your thoughts about the graph now?

Compare the data for boys and girls. Based on these data, would you conclude that girls in the USA do more homework than boys? Why or why not? What other information might you want to have in order to help you reach a credible conclusion??