Information literacy, statistical literacy and data literacy

By Milo Schield[1]

Introduction

The evaluation of information is a key element in information literacy, statistical literacy and data literacy. As such, all three literacies are inter-related. It is difficult to promote information literacy or data literacy without promoting statistical literacy. While their relative importance varies with one's perspective, these three literacies are united in dealing with similar problems that face students in college. More attention is needed on how these three literacies relate and how they may be taught synergistically. All librarians are interested in information literacy; archivists and data librarians are interested in data literacy. Both should both consider teaching statistical literacy as a service to students who need to critically evaluate information in arguments.

Information Literacy

The need for information literacy has been highlighted in the US by several organizations including the American Library Association (ALA). In 1989 a call for information literacy was issued by the ALA Presidential Committee on Information Literacy.[2] In 1989, the National Forum on Information Literacy[3] was formed. And in 1998, the ALA/ACRL (Association of College and Research Libraries) issued a progress report.[4] Each organization and each report had some differences in their approach to information literacy. But one element was common to all – the need for the critical evaluation of information.

The ALA and ACRL issued a set of information literacy competency standards for higher education.[5]

> Information literacy is a set of abilities requiring individuals to "recognize when information is needed and have the ability to locate, evaluate, and use effectively the needed information." An information literate individual is able to: (1) Determine the extent of information needed, (2) Access the needed information effectively and efficiently, (3) Evaluate information and its sources critically, (4) Incorporate selected information into one's knowledge base, (5) Use information effectively to accomplish a specific purpose, and (6) Understand the economic, legal, and social issues surrounding the use of information, and access and use information ethically and legally.

In their presentation of the standards for information literacy, the American Association of School Librarians Association (AASL) for Educational Communications and Technology presented evaluation as one of three standards[6], and identified several related indicators.

> **Standard 2 The student who is information literate evaluates information critically and competently.**
> The student who is information literate weighs information carefully and wisely to determine its quality. That student understands traditional and emerging principles for assessing the accuracy, validity, relevance, completeness, and impartiality of information. The student applies these principles insightfully across information sources and formats and uses logic and informed judgment to accept, reject, or replace information to meet a particular need.

Indicators: (1) Determines accuracy, relevance, and comprehensiveness. (2) Distinguishes among fact, point of view, and opinion. (3) Identifies inaccurate and misleading information. (4) Selects information appropriate to the problem or question at hand.

Evaluating information can be difficult. Even government sources may have their agenda or ideology. But evaluating information can become much more difficult when the information involves statistics. There are numerous publications on information literacy.[7]

Statistical Literacy

A great deal of information involves statistics. How would we talk about current social issues without using basic statistical ideas such as ratios, percents and rates? And with the advent of high speed computing and the internet, everyone is facing a flood of information in the form of statistics. It seems difficult to be considered information literate in the 21st century without being statistically literate.

Statistical literacy studies the use of statistics as evidence in arguments (Schield, 1998, 1999).

Joel Best (2001, 2004) identified the key to being statistical literate when he noted that all statistics are socially constructed. This isn't some deep philosophical claim. It merely states that people choose what to count or measure, how to assemble those measurements into summary statistics, what comparisons to form from these statistics and how to communicate these statistics.

A key element of statistical literacy is assembly: how the statistics are defined, selected and presented.

- For example, in 1999 the difference in US mean household income before and after taxes was $14,000 whereas the difference in median household income was about $7,000.[8] Just changing the choice of the statistic cut the tax burden in half.
- Similarly, one person might note that in 1999 the mean US household income before taxes was almost $55,000 while another might note that the median household income after taxes was less than $34,000. An unwary reader might presume the entire difference of $21,000 was due entirely to taxes.
- In the US Consumer Expenditure Survey, the 1997-98 table for those under age 25, shows an average expenditure for alcohol of $330 per year. But in many families, this expenditure is zero so the average amount in families that spend money on alcohol can be much higher.

The phrase "lies, damned lies and statistics" is well known (first proclaimed by Benjamin Disraeli and then popularized in the US by Mark Twain). In many ways statistics are just words in a different form. Perhaps the years spent learning arithmetic have led us into thinking that since numbers don't lie then neither do statistics. But statistics are more than numbers. Statistics are numerical summaries about things in reality. The nature of the things being summarized can make a difference. See Schield (20005ab) for examples of statistical prevarication.

Consider this example. We all know that 6 plus 7 is 13 and that 60% plus 70% is 130%. So if a company has a 60% market share in the eastern US and has a 70% market share in the western US, what is their market share in the entire US? The math says 130%, but we all know that is wrong. Market share has a particular meaning or nature. So for statistics, small changes in syntax can create large changes in semantics.

As Joel Best put it, statistics are more like diamonds than like rocks. Diamonds are carefully cut and displayed by people to produce a desired effect. Statistics, like diamonds, are not 100% natural like sand or rocks. They are socially constructed. Helping students see this is a major challenge.

A second key element of statistical literacy is the importance of context and confounding. Consider these examples involving simple rates or percentages.

- In terms of people, the number of unemployed workers is much higher in the US than in Canada. But does this mean unemployment is more prevalent in the US than in Canada? No. The number who are unemployed is strongly influenced by the size of the population. To untangle the influence of population on the number who are unemployed we need to look at the unemployment rates: the percentage of those in the civilian labor force that are unemployed.

Taking into account the influence of a related factor can reverse this association between country and unemployment. The point is that the context is important. Are we talking about counts or rates? What should we be talking about? But to be statistically literate one must go further than simply forming and comparing rates and percentages.

- Suppose someone claimed that Mexico has a better medical system than the US. You might be skeptical. But consider this statistic: the death rate is lower in Mexico than in the US. While there may be variation in how some statistics are assembled, there is little variation in what constitutes death. So what explains this lower death rate in Mexico than in the US? It can't be a difference in the size of the populations; rates take that into account. It could be the difference in medical care. But another relevant factor is the difference in ages. Mexicans are much younger on average than those in the US. Younger people are less likely to die in the next year than are older people. So perhaps taking into account the difference in the average ages in the two countries may decrease – if not reverse – the association between country and deaths.

Untangling the influence of confounding is a major element in being statistical literate. See Schield (2004) and the Statistical Literacy website.[9] See also Gray (2003) and Lackie (2004).

Data Literacy

While all students in majors that deal with information have a need for information literacy, those students in majors that also require a course in data analysis or statistics typically have a need for data literacy. Such majors are often found in the social sciences and business. In these majors, students need training in how to obtain and manipulate data.

Data literacy is supported by the International Association for Social Science Information Services and Technology (IASSIST),[10] by the Association of Public Data Users (APDU)[11] and by related organizations, such as the Inter-university Consortium for Political and Social Research (ICPSR).[12] For more background, see papers by Rice (2001), Gray (2003), Hunt (2004) and Czarnocki and Khouri (2004).

Data literacy may appear less technical than is either Computer Science or Management Information Systems (MIS). Yet students need to understand a wide variety of tools for accessing, converting and manipulating data. These may need to understand structured query language (SQL), relational databases (e.g. MS Access), data manipulation techniques, statistical software

(e.g., SPSS, STATA, Minitab and MS Excel) and data presentation software (e.g., MS Excel and MS PowerPoint).

Inter-Relation

With the advent of the personal computer and the web, information literacy requires both statistical literacy and data literacy. Students must be information literate: they must be able to think critically about concepts, claims and arguments: to read, interpret and evaluate information. Statistical literacy is an essential component of information literacy. Students must be statistically literate: they must be able to think critically about basic descriptive statistics. Analyzing, interpreting and evaluating statistics as evidence is a special skill. And students must be data literate: they must be able to access, assess, manipulate, summarize, and present data. Data literacy is an essential component of both information literacy and statistical literacy. See Linden (2002).

Figure 1 illustrates one way of viewing the relation between the three literacies from a critical thinking perspective.

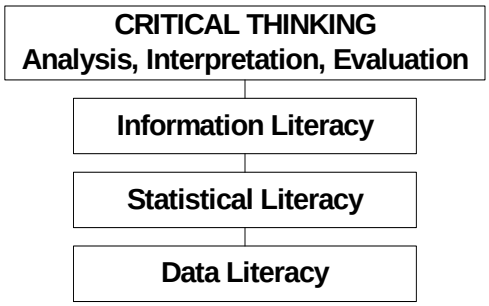

Figure 1: Critical Thinking Perspective

Journalism schools work at integrating these three literacies from this information-literacy perspective. See Ward and Hansen (1997), and Hansen and Paul (2003).

Unlike 'critical thinking' (10,116 citations) and 'information literacy' (1,498 citations), 'quantitative literacy', 'statistical literacy' and 'data literacy' are in their academic infancies with less than 65 citations each in the Education Resources Information Center (ERIC) database.[13]

With the internet, statistical summaries are more readily available. So if students are to be information literate, the must be statistically literate. Analyzing, interpreting and evaluating statistics as evidence is a special skill. Statistical literacy must be an essential component of information literacy.

Statistics summarize data. The numerical value of a statistic is heavily influenced by how the underlying data is selected, converted and manipulated. Converting and manipulating data is a special skill that requires in-depth training. Thus, data literacy must be an essential component of both information literacy and statistical literacy.

How one organizes these literacies depends on one's perspective. Consider a perspective of those teaching in the social sciences.

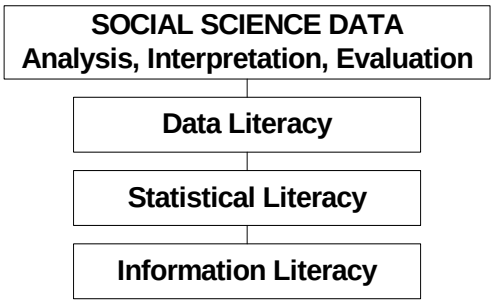

Figure 2: Discipline Perspective

From this disciplinary perspective, students need to be able to analyze, interpret and evaluate social science data. Data literacy is needed to access, manipulate and summarize the data. But statistical literacy is needed to guide in that process while information literacy sets the overall context for evaluating the sources of data and the appropriate manipulations.

Regardless of the arrangement, students need to be able to access, analyze and evaluate information, and they must be able to communicate their findings, conclusions and recommendations accurately and effectively.

The key point is that information literacy, statistical literacy and data literacy are tied together by a common set of problems and a similar level of approach. All three are more general than specific, they each involve interdisciplinary study and they deal with fundamentals. They can be useful to students in any major. As such they should be core elements in a college education.

Organizational Pressures and Needs

Colleges and universities are under increasing financial constraints. At times, staff are sometimes viewed as overhead whereas faculty are viewed as a cost of production. This pressure may tend to reduce funds available for library staff. Yet, librarians have a unique opportunity in view of their training. They are generalists, not specialists. Their focus is not the focus of a particular discipline. As such they are eminently qualified to teach students how to think critically, how to become information literate, how to become statistically literate and how to become data literate.

In a teaching capacity they may obtain a different status in relation to the ultimate mission of higher education than they would as library staff. In helping students think critically about information, statistics and data, their role might be considered mission critical given the importance of critical thinking as a strategic goal of higher education and the difficulties some students have in thinking critically about words – much less about numbers.

Teaching Statistical Literacy

Librarians should consider teaching statistical literacy as a component of information literacy. While they may not have a quantitative background or interest, statistical literacy is typically more about words than numbers, more about evidence than about formulas. Librarians have always tried to be of service to the needs of those seeking to access and understand information.

Data librarians should definitely consider teaching statistical literacy as a component of data literacy. They typically have a quantitative background in the social sciences or they have acquired one on the job. They recognize that students need help in accessing data and in evaluating data. And they recognize that random assignment in the social sciences is often impossible or unethical, so statistical associations are being used as evidence for causal connections. Helping students to think critically about such inferences would be a valuable service to students in the social sciences.

Training teachers in statistical literacy will be a major effort. See Watkins (2004) for a promising approach.

Would the teaching of statistical literacy by trained staff undermine the role of the professional faculty teaching traditional statistics? My answer is "No!" Most faculty teaching statistics are not looking to teach anything 'below' statistical inference and advanced data modeling. I believe the opposite is more likely. Collaboration with such faculty would be welcomed and fruitful to the extent one could show that statistically-literate students are better prepared to understand and appreciate what they are learning in their advanced quantitative courses.

Would resources be available for this teaching? My answer is "Yes!" While such funding is unlikely under existing library or staff budgets, funding is almost guaranteed from the academic budget as soon as the activity involves academic credit. Having statistical literacy, data literacy and information literacy as an academic course that might be required of all students – or at least of those who fall below a certain level of proficiency – will open the door to a new source of funding. From the academic budget perspective, paying instructors on an adjunct or overload basis is much less costly than paying a tenured professor. From the staff perspective, money from the academic budget might go toward the general staff budget, it might go to the staff member teaching the course (provided the staff member is still working full-time on non-teaching activities) or it might involve an allocation between these so as to provide some incentive for staff to teach this kind of course.

To obtain academic credit in higher education, a statistical literacy course must be viewed as non-remedial. For one approach, see Schield (2004b). Developing valid statistical literacy assessments will also be critical. To see how all this might work, review the structure and operation of Quantitative Literacy centers, programs and courses at selected US colleges. See http://www.statlit.org/QL2.htm.

Conclusion

Both information literacy and data literacy should be expanded to include critical thinking and statistical literacy. Expanding information literacy to include statistical literacy will help students deal with information that involves statistics. Expanding data literacy to include statistical literacy will help students in the social sciences deal with inferring causation from associations. As such, including statistical literacy with information literacy and with data literacy will provide more opportunities for librarians to be of service in helping students think critically. See Schield (2005b).

Acknowledgments: To IASSIST for promoting data literacy. To Dianne Harmon, Joliet Public Library, and to Boyd Koehler, Augsburg College Library, for recommendations on information literacy. This work was conducted under a grant to Augsburg College from the W. M. Keck Foundation "to develop statistical literacy as an interdisciplinary curriculum in the liberal arts."[14]

References

American Library Association (1998). A Progress Report on Information Literacy: An Update on the American Library Association Presidential Committee on Information Literacy: Final Report.

Best, Joel (2001). *Damned Lies and Statistics: Untangling Numbers from the Media, Politicians, and Activists*. University of California Press.

Best, Joel (2004). *More Damned Lies and Statistics: How Numbers Confuse Public Issues.* University of California Press

Bruce, Christine (1997). *The Seven Faces of Information Literacy*. Auslib Press, 1997. 203p. Paperback

Corti, Louise (2003) "*Exploiting UK Survey Data Sources for Teaching Political Science: Experiences from the Classroom*", Presented at 2003 IASSIST conference, Ottowa. Available at: http://datalib.library.ualberta.ca/~humphrey/iassist2003/sessions/I03sessions.html#A20

Czarnocki, Susan and Anastassia Khouri (2004). *Filling the Gap: Doing Stats in the Library* Presented at IASSIST 2004 conference, Madison. Available at: http://dpls.dacc.wisc.edu/iassist2004/program.html

Eisenberg, Michael B., Carrie A. Lowe, Kathleen L. Spitzer (2004). *Information Literacy: Essential Skills for the Information Age Second Edition*. Libraries Unlimited; 2nd edition

Goad, Tom W. (2002). *Information Literacy and Workplace Performance*. Green Wood Publishing Group, 2002. 248p.

Gray, Ann (2003). *Data and Statistical Literacy for Librarians*. Presented at 2003 IASSIST conference, Ottowa. Abstract available at: http://datalib.library.ualberta.ca/~humphrey/iassist2003/sessions/I03sessions.html#A20

Hansen, Kathleen A. and Nora Paul (2003). *Behind the Message : Information Strategies for Communicators*. Allyn & Bacon.

Hunt, Karen (2004). *The Challenges of Integrating Data Literacy into the Curriculum in an Undergraduate Institution.* Presented at IASSIST 2004 conference, Madison. Available at: http://scholar.uwinnipeg.ca/khunt/iassist2004/index.cfm

Lackie, Paula (2004). *Understanding and Using Data: A Discussion of the Jargon and Trends in "Quantitative Literacy."* Presented at IASSIST 2004 conference, Madison. Available at: http://dpls.dacc.wisc.edu/iassist2004/program.html

Linden, Julie (2002). *Finding, Evaluating and Using Numeric Data* . Presented at IASSIST 2002 conference, Storrs, Connecticut. Available at: http://ropercenter.uconn.edu/iassist2002/program.html

Loertscher, David V. and Blanche Woolls (2001). *Information Literacy: A Review of the Research: A Guide for Practitioners and Researchers,* 2nd edition.. Hi Willow Research & Publishing, 2001.

Middle States Commission on Higher Education (2003). Developing Research & Communication Skills: Guidelines for Information Literacy in the Curriculum. 112p.

Moore, Penny (2002). *Information Literacy: What's it all about*? New Zealand Council for Educational Research (NZCER).

Nims, Julia, R. Baier, E. Owen, and R. Bullard (2003). *Integrating Information Literacy into the College Experience*. Library Orientation Series, Pierian Press.

Rice, Robin (2001). *Data Support for Learning & Teaching: the Final Frontier?* . Presented at IASSIST 2001 conference, Amsterdam. Available at: http://datalib.ed.ac.uk/projects/datateach/iassist2001/tsld001.htm

Rockman, Ilene F. and Associates (2004). *Integrating Information Literacy into the Higher Education Curriculum: Practical Models for Transformation.* The Jossey-Bass Higher and Adult Education Series

Schield, Milo (1998). *Statistical Literacy and Evidential Statistics.* ASA Proceedings of the Section on Statistical Education, p. 137. Available at: www.StatLit.org/pdf/1998SchieldASA.pdf

Schield, Milo (1999). *Statistical Literacy: Thinking Critically About Statistics as Evidence.* Of Significance, Volume 1, Issue 1. Association of Public Data Users (APDU). Available at: www.StatLit.org/pdf/1999SchieldAPDU.pdf.

Schield, Milo (2004a). *Statistical Literacy and Liberal Education at Augsburg College.* Peer Review Summer Issue, Assoc. of American Colleges and Universities. Available at: www.StatLit.org/pdf/2004SchieldAACU.pdf and www.augsburg.edu/statlit.

Schield, Milo (2004b). *Statistical Literacy Curriculum Design.* Curriculum Design Roundtable sponsored by the International Association of Statistical Educators (IASE) in Lund Sweden. Available at: www.StatLit.org/pdf/2004SchieldIASE.pdf and www.augsburg.edu/statlit.

Schield, Milo (2005a). *Statistical Prevarication: Telling Half Truth Using Statistics.* Communicating Statistics conference sponsored by the International Association of Statistical Educators (IASE). Available at: www.StatLit.org/pdf/2005SchieldIASE.pdf and www.augsburg.edu/statlit.

Schield, Milo (2005b). *Statistical Literacy: An Evangelical Calling for Statistical Educators.* Invited paper at conference sponsored by the International Statistical Institute (ISI) in Sydney. Available at: www.StatLit.org/pdf/2004SchieldIASE.pdf and www.augsburg.edu/statlit.

Timms-Ferrara, Lois, (2003). *Public Opinion Matters: A New Roper Center Program Designed to Promote Classroom Use of Public Opinion Data*. Presented at the 2003 IASSIST conference, in Ottawa, Canada. Abstract available at: http://datalib.library.ualberta.ca/~humphrey/iassist2003/sessions/I03sessions.html#A20

Ward, Jean and Kathleen A. Hansen (1997). *Search Strategies in Mass Communications.* Longman.

Watkins, Wendy (2004*). Do It Yourselves: A Peer-to-peer Approach to Professional Training.* Presented at IASSIST 2004 conference, Madison. Abstract available at: http://dpls.dacc.wisc.edu/iassist2004/program.html

Endnotes

[1] Contact: Milo Schield, Professor of Business Administration and MIS, Director of the W. M. Keck Statistical Literacy Project, Augsburg College, Minneapolis, MN. Information available at: www.augsburg.edu/statlit. Email: schield@augsburg.edu

[2] Available at: www.ala.org/ala/acrl/acrlpubs/whitepapers/presidential.htm and at www.infolit.org/documents/89Report.htm

[3] Available at: www.infolit.org/

[4] Available at: www.ala.org/ala/acrl/acrlpubs/whitepapers/progressreport.htm

[5]Available at: www.ala.org/ala/acrl/acrlstandards/informationliteracycompetency.htm

[6] www.ala.org/ala/aasl/aaslproftools/informationpowerInformationLiteracyStandards_final.pdf

[7] See Bruce (1997), Loertscher and Woolls (2001), Goad (2002), Moore (2002), Nims et al (2003), Rockman et al (2004), and Eisenberg et al (2004).

[8] 2001 US Statistical Abstract. Table 664, p. 435.

[9] Statlit website at www.StatLit.org

[10] International Association for Social Science Information Services and Technology (IASSIST). Available at: www.iassistdata.org and http://datalib.library.ualberta.ca/

[11] Association of Public Data Users (APDU). Available at: www.APDU.org

[12] ICPSR. Available at: www.icpsr.umich.edu/org/ and www.icpsr.umich.edu/access.

[13] Education Resources Information Center (ERIC). Available at: http://www.eric.ed.gov/

[14] W.M. Keck Foundation at: www.wmkeck.org/