Numeracy: Assessing Basic Skills and Knowledge

To do: Add Joel Best. Consider Lutsky, Gal and Bracey.

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

Abstract

An increasing number of colleges have a graduation skills requirement involving numeracy, quantitative reasoning, quantitative literacy or statistical literacy. Colleges need a way to assess the level of numeracy in their students and for their classes. Identifying the associated skills and competencies is a requirement for any grounded attempt at assessment. To be credible and provide a basis for content validity, those skills and competencies must be vetted and validated by an organization involving subject-matter experts. This paper proposes that a national association undertake this process. To ensure that this process is open and transparent, a two stage process is proposed. In the first stage, subject matter experts will review the proposed instrument. In the second stage they will rank related skills and competencies on their importance in assessing numeracy. A draft survey instrument is presented.

1. Background

More colleges are requiring that students complete a course that satisfies a graduation skill in numeracy: in quantitative literacy, statistical literacy or quantitative reasoning. These are envisioned as cognitive skills that are desirable for all college students.

Colleges need a way to assess the level of numeracy in their students and for their classes. Before one can assess this ability, this ability must be defined in ways that are operationally measureable.

This requires that these cognitive skills be clearly defined and that the associated skills and knowledge be clearly identified before any grounded attempt at assessment.

2. Defining Numeracy

Numeracy and related cognitive skills such as quantitative literacy or statistical literacy currently lack clear operational definitions that are generally accepted.

Sommerville (1994) noted this problem for Quantitative Literacy (QL): "QL advocates need to be very clear about what all students need to know and be able to do, starting with where it fits into the mathematics program."

Steen (2004a) noted the same problem: “One clear priority has emerged: the need to develop benchmarks for quantitative literacy that can guide both curriculum and assessment in grades 10-16.”

Madison (2005b) echoed this concern: “determining what quantitative reasoning – quantitative literacy – is and how to measure it is a major national issue.”

Schield (2009) noted this problem for statistical literacy where two different approaches are being used: “the first approach begins by linking statistical literacy with ‘for whom’ (all adults) and ‘for what’ (to be good citizens). The second approach links statistical literacy with cognitive skills that are selected based on expert insight.”

But even if there were agreement that s=numeracy should involve the ability to read and interpret the statistics encountered in everyday life, there still might be major problems for assessment.

3. Assessing Numeracy

According to Bookman (2005), “Assessing QL presents the same challenges of assessing other types of learning but QL presents some particular challenges for assessment:

- Assessment items must be set in a real world context.
- The problems and the contexts must be familiar to all the students but neither so routine that they require very little thought nor so non-routine that students cannot solve them.
- The assessment instruments must not take too much time to administer, must include a multitude of problems and situations, and must be designed to allow for reliable scoring.”

But if these cognitive competencies involve critical thinking in unstructured contexts, then authentic assessment may be difficult if not impossible in a limited amount of time.

Madison (2006) noticed this in commenting about the mathematical reasoning involved in quantitative literacy.

> “By and large, neither students nor faculty act as if they believe that mathematics has much to offer in terms of cross-cutting competencies such as critical thinking or communication. Reasoning is accepted by mathematicians as a critical feature of developing or using mathematics, and students believe that reasoning is an important process for them to master. However, in my experience, students see mathematical reasoning as distinct from reasoning in other domains, another manifestation of the separation of mathematics from the rest of the world of many stu-

dents. To many students, mathematics is a subject all on its own, and faculty are not much different. Although mathematics faculty recognize the incredible array of uses of mathematics in the real world, most of these uses are in contexts well out of reach of beginning college students."

A canonical QL situation involves several steps, some of which are encountered in traditional mathematics or statistics courses, but rarely is the process with all the steps part of these courses. The steps can be described as follows, where I have indicated a critical strand in mathematical proficiency that seems necessary.

- Encountering a challenging contextual circumstance, e.g. reading a newspaper article that contains the use of quantitative information or arguments. (Productive disposition)
- Interpreting the circumstance, making estimates as necessary to decide what investigation or study is merited. (Adaptive reasoning)
- Gleaning out critical information and supplying reasonable data for data not given. (Productive disposition and conceptual understanding)
- Modeling the information in some way and performing mathematical or statistical analyses and operations. (Strategic competence and procedural fluency)
- Reflecting the results back into the original circumstance. (Adaptive reasoning)

These steps often require careful reading of continuous prose and graphical representations or other discontinuous prose, using mathematics or statistics, and then interpreting and critiquing the original prose in light of the mathematical results. Critical reasoning (closely akin to adaptive reasoning) is required throughout. Students are not expecting this complicated process because their previous mathematics experiences have been narrower and better defined. Consequently, one struggles with breaking the process into bits and pieces and teaching these separately. Frequently, the third phase gets the most attention because it is the process of traditional mathematics and statistics courses.

Assessment items must be authentic, and according to Grant Wiggins (2003) that requires that they be complex, realistic, meaningful, and creative, and have value beyond school. One can use assessment items that are narrower, say focusing on the basic mathematics or statistics skills and knowledge needed for QL. If one knows what these skills are, then assessing them is only a piece of the bigger assessment task. As Grant Wiggins has pointed out, assessing QL is analogous to assessing whether a person is a good soccer player. One can assess individual skills required in soccer, but the proof comes with actually playing the game.

Even after deciding on authentic assessment items or processes, two challenges remain. What will be valued in scoring? Are reading, interpreting, computing, reflecting, and writing all parts of what will be evaluated? They are all parts of QL, and the challenge of scoring all is substantial. The second challenge is determining levels, or standards, for proficiency in QL. Since QL is society dependent and certainly changes over time and place, the proficiency standards of the past or of other societies are not necessarily appropriate. Few people will be able to successfully handle quantitative issues across all of the possible domains in US society. Consequently, one has to decide on what domains are common enough to be included in setting standards. Clearly, the challenges are quite daunting."

Schield (2008) reviewed these difficulties, argued that students needed a structured approach to achieve any level of competency, and presented a general template for assessing the role of numbers in news stories.

Other approaches to assessing numeracy include:

- The National Assessment of Adult Literacy.[1]
- Dartmouth College Mathematics Across the Curriculum Survey. See Korey (2000).
- the College Proficiency Exam (CPE) Part 2. See Crendall et al (2005).
- The Simpson Quantitative Literacy Competency exam.[2]
- The Q/R instrument at James Madison University. See Sundre (2008).

If assessing the high-level skills involved in numeracy is difficult and time consuming, a quick assessment of the lower level skills and knowledge may still be possible if there is some way to ensure validity.

4. Validity

A valid measure is one that accurately reflects what it is intended to measure. In this context two types of validity are relevant: criterion validity and content validity.

Criterion validity indicates how accurately a measure predicts an observable outcome. Criterion validity is more objective than face or construct validity. At this point there is no general agreement on a suitable outcome, so there is no good criterion on which to base an assessment of numeracy.

[1] http://nces.ed.gov/Pubs2007/2007480.pdf
[2] www.simpson.edu/math/QLC.html

Secondary forms of criterion validity are correlations with those factors that should contribute to greater quantitative literacy. For example, Sundre (2008) reports on the QR test at James Madison University saying that the scores:

- correlate positively with grades in relevant courses.
- increase with greater relevant course exposure.
- discriminate between students who have completed their general education requirements and those who have not.

Content validity is the extent to which experts believe a measure represents a social concept such as depression, honesty or numeracy.

Content validity is similar to – but slightly different from – face validity – the extent to which non-experts think a measure represents a social concept. A question on a survey instrument may have high content validity and low face validity – or vice versa.

Until there is some measure of content validity among subject matter experts for a given instrument, the results of any assessment would lack an objective grounding.

A common method of obtaining content validity is to ask subject-matter-experts (SMEs) to rank job related skills and knowledge in their importance. Using an appropriate cutoff, the rankings can be converted into yes-no results which can be adjusted for random variation to give statistically significant conclusions.

5. Proposal

The goal of this paper is to present a procedure to obtain input from subject matter experts on a list of related skills and knowledge and on their importance in assessing numeracy.

This paper proposes that a national association undertake the process of identifying what is meant by phrases such as numeracy, statistical literacy or quantitative literacy. To insure that the phrase be viewed from a broad general education perspective rather than a narrow discipline-specific perspective, subject matter experts should be recruited from a wide variety of disciplines.

Having a national organization in charge does not mean that the organization is going to define what is meant by the term. Instead, the organization will manage the process whereby subject matter experts give their opinions on which skills and knowledge are considered characteristics of a person who is numerate, quantitatively literate or statistically literate.

To insure that the process is open and transparent, it should allow the subject-matter-experts to review the assessment document prior to giving their opinions, and to make suggestions in the skills and knowledge presented and in the options available for their opinions.

This requires a two-stage process.

- Stage 1: Distribute a draft/knowledge survey for review and comment.
- Stage 2: Distribute an updated survey to the subject-matter experts and collect their rankings.

Perhaps the most challenging step is the stage 1 draft. It is easy to bias the results by the choice of what is included and excluded in the survey. One solution is to use all the proposals of skills and knowledge that might be relevant and let the subject-matter-experts decide which ones to include or exclude.

Once this information has been collected by the association, the results would be summarized and disseminated. This summary could be used by those generating assessment instruments.

6. Hesitation

Before recommending this action, it is useful to review plausible reasons for not trying to identify what is meant by these terms. First, assessing just the basic goals may undercut the existence of higher goals. Second, definitions can divide adherents based on different overall goals or on secondary matters – but these secondary differences may become more essential than any primary sources of agreement. Third, definitions can limit growth and learning when members are at different levels and as an organization grows and changes.

In each case, these divisions can effectively incapacitate a young movement when the group involved is very small.

7. Draft Assessment

Appendices A through G presents some of the key claims involving the skills and knowledge that are relevant. These include claims by Sons (Appendix A), Utts (Appendix B), McKenzie (Appendix C), Madison (Appendix D), Gillman (Appendix E) and Schield (Appendix F).

The last appendix presents a draft of a stage 1 document. A common five point assessment would be Strongly Disagree, Disagree, Neutral, Agree and Strongly Agree. This paper proposes a skewed five-point scoring rubric: Disagree, Neutral, Agree, Strongly Agree and Absolutely Agree. This skewed scoring seems appropriate since all the claims being reviewed are proposed as being true, important and relevant by a subject-matter expert. Therefore, strongly disagree is not very likely.

8. Analysis

Some might wonder why the comments of noted subject matter experts such as Gal, Lutsky and Bracey are not included. One reason is that their comments focus more on questions or principles than on skills or basic knowledge. See Schield (2008a) for a review.

The proposed survey is lengthy (126 questions) and some of the questions are repetitive or very similar. One suggestion is to eliminate the grouping by author and to combine similar questions. This was not done even though it means the survey is much longer. Combining similar questions may eliminate important differences. For example, confounding is mentioned by both Scheaffer (Q55) and Schield (Q125 and Q139). For Scheaffer, this might refer to just an awareness that confounding can influence an association, while for Schield this might mean that students should be able to work problems.

9. Recommendations

Educators interested in promoting statistical or quantitative literacy should review this proposal and make suggestions for improvement. Organizations interested in promoting assessment should consider managing this process in order to help generate better assessments.

10. Acknowledgments

This work was performed as part of the W. M. Keck Statistical Literacy Project "to support the development of statistical literacy as an interdisciplinary curriculum in the liberal arts." This talk was presented at the 2009 Joint Mathematical Meeting of the MAA.

11. References

Best, Joel (2008). Beyond Calculation: Quantitative Literacy and Critical Thinking about Public Issues. *Calculation vs. Context: Quantitative Literacy and Its Implications for Teacher Education.* Edited by Bernard L. Madison and Lynn Arthur Steen. MAA

Crandall, Gordon, James Giordano, Karlyn Koh and Janet Michello (2005). Across the Curriculum Strategies for Teaching Quantitative Literacy.[3]

Garfield, J. (1995), "How Students Learn Statistics," *International Statistical Review*, 63, 25-34.

Gillman, Richard (2006). *Current Practices in Quantitative Literacy.* MAA Notes #70

Korey, J. (2000), "Dartmouth College of Mathematics Across the Curriculum Evaluation Summary: Mathematics and Humanities Courses," *Mathematics Across the Curriculum at Dartmouth College*[4]

Lutsky, Neil (2006). Quirks of Rhetoric: A Quantitative Analysis of Quantitative Reasoning in Student Writing *Proceedings of the ASA Section on Statistical Education*, pp. 2319-2322.[5]

Madison, Bernard (2005a). Important Mathematical Concepts for Numeracy. Presented at MAA 2005.[6]

Madison, Bernard (2005b). Quantitative Literacy. Presented at MAA QL-SIG 2005.[7]

Madison, Bernard L. (2006). Pedagogical Challenges of Quantitative Literacy. *Proceedings of the ASA Section on Statistical Education*, pp. 2323-2328.[8]

McKenzie, John D., Jr. (2004). *Conveying the Core Concepts.* Proceedings of ASA Statistical Education Section, pp. 2755-2757.[9]

Raymond, Robert and Milo Schield (2008). *Numbers in the News: A Survey.* Proceedings of the ASA Section on Statistical Education.[10]

Schield, Milo (1998). *Statistical Literacy: Thinking Critically about Statistics.* Inaugural issue of the Journal "Of Significance" produced by the Association of Public Data Users (APDU).[11]

Schield, Milo (2004). *Statistical Literacy Curriculum Design.* IASE Curricular Development in Statistics Education Roundtable, pp. 54-74.[12]

Schield, Milo (2007). *Statistical Literacy: Teaching the Social Construction of Statistics.* Presented at the Midwest Sociological Society.[13]

Schield, Milo and Cynthia Schield (2007). *Numbers in the News: A Survey.* Proceedings of the ASA Section on Statistical Education, pp. 2323-2328.[14]

Schield, Milo (2009). *Assessing Statistical Literacy: Take CARE.* Submission for Varieties in Statistical Assessment. International Association of Statistical Educators.[15]

Sundre, Donna (2008). Assessment Resources.[16]

Utts, J. (2003), "What Educated Citizens Should Know About Statistics and Probability," *The American Statistician*, 57(2), 74-79.

[3] www.lagcc.cuny.edu/OpeningSessions/2005/workshops_I.htm
[4] www.math.dartmouth.edu/~matc/Evaluation/humeval.pdf
[5] www.StatLit.org/pdf/2006LutskyASA.pdf.
[6] www.StatLit.org/pdf/2005MadisonMAA.pdf
[7] www.StatLit.org/pdf/2005MadisonMAAQLSIG.pdf
[8] www.StatLit.org/pdf/2006MadisonASA.pdf.
[9] www.StatLit.org/pdf/2004MckenzieASA.pdf.
[10] www.StatLit.org/pdf/2008RaymondSchieldASA.pdf.
[11] www.StatLit.org/pdf/1998SchieldAPDU1.pdf
[12] www.StatLit.org/pdf/2004SchieldIASE.pdf
[13] www.StatLit.org/pdf/2007SchieldMSS.pdf
[14] www.StatLit.org/pdf/2007SchieldASA.pdf.
[15] www.StatLit.org/pdf/2009SchieldVISA.pdf
[16] www.jmu.edu/assessment/resources/Overview.htm

Appendix A: Sons

Sons (1994): "A quantitatively literate college graduate should be able to

- interpret mathematical models such as formulas, graphs, tables, and schematics, and draw inferences from them.
- represent mathematical information symbolically, visually, numerically and verbally.
- use arithmetical, algebraic, geometric and statistical methods to solve problems.
- estimate and check answers to mathematical problems in order to determine reasonableness, identify alternatives, and select optimal results.
- recognize that mathematical and statistical methods have limits."

Appendix B: Utts

Utts (2003) identified seven topics "commonly misunderstood by citizens, including the journalists who present statistical studies to the public. In fact researchers themselves, who present their results in journals and at the scientific meetings from which the journalists cull their stories, misunderstand many of these topics. If all students of introductory statistics understood them, there would be much less confusion and misinterpretation related to statistics and probability and findings based on them. In fact the public is often cynical about statistical studies, because these misunderstandings lead to the appearance of a stream of studies with conflicting results. This is particularly true of medical studies, where the misunderstandings can have serious consequences when neither physicians nor patients can properly interpret the statistical results."

1. When it can be concluded that a relationship is one of cause and effect, and when it cannot, including the difference between randomized experiments and observational studies.
2. The difference between statistical significance and practical importance, especially when using large sample sizes
3. The difference between finding "no effect" or "no difference" and finding no statistically significant effect or difference, especially when using small sample sizes.
4. Common sources of bias in surveys and experiments, such as poor wording of questions, volunteer response, and socially desirable answers.
5. The idea that coincidences and seemingly very improbable events are not uncommon because there are so many possibilities.
6. "Confusion of the inverse" in which a conditional probability in one direction is confused with the conditional probability in the other direction.
7. Understanding that variability is natural, and that "normal" is not the same as "average."

Appendix C: McKenzie

McKenzie (2004) asked statistical educators in his session at the 2004 JSM to grade the following 30 statistical topics. The numbers shown in Table 1 are percentages: the count per 100 respondents in each category.

Table 1: Statistical Topics Survey Results

	Percentage Of All Reponses[17]	Core Concept	TOP 3 Important	TOP 3 Difficult
1	Variability	**96**	**75**	**12**
2	Association vs. Causation	**82**	**31**	**6**
3	Randomness	**77**	**14**	**8**
4	Significance (Practical/Statistical)	**77**	**14**	**16**
5	Data Collect (Exp, Obs, surveys)	**75**	**24**	**4**
6	Sampling Dist (Law Lg. #, CLT)	**71**	**25**	**66**
7	Hyp. test (crit value, p-value, pwr)	**64**	**22**	**66**
8	Confidence Interval	**63**	**12**	**16**
9	Random Sample	**63**	**10**	**4**
10	Data types	**61**	**8**	**4**
11	Center	**59**	**6**	**0**
12	Assumptions	**55**	**8**	**20**
13	Graphing	**54**	**10**	**0**
14	Uncertainty	**54**	**10**	**2**
15	Distributions	**52**	**10**	**14**
16	Independence	**50**	**4**	**16**
17	Bias	**48**	**2**	**2**
18	Correlation	**48**	**2**	**6**
19	Shape	**45**	**0**	**0**
20	Data Exploration	**43**	**8**	**0**
21	Proportion	**41**	**0**	**0**
22	Least-squares Regression	**39**	**2**	**8**
23	Models	**38**	**4**	**12**
24	Comparisons	**38**	**2**	**2**
25	Prediction	**34**	**2**	**2**
26	Outliers (aspects of robustness)	**32**	**0**	**0**
27	Cross-sectional vs. longitudinal	**11**	**0**	**0**
28	Regression effect	**11**	**0**	**4**
29	Process	**7**	**0**	**0**
30	Transformations	**2**	**0**	**10**

Appendix D: Madison

Madison (2005) described a quantitative literacy course as follows. "For the first semester I put together eleven lessons with the following titles: percent, petty thrift and buying stocks, lower math by Dave Barry, linear and exponential growth, measurement, visual representation of quantitative information, rates of change, weather maps and indices, the odds of that, and risk. As you can probably tell, the course was very loosely organized by mathematical topic, and topics – e.g. rates of change – kept recurring.

[17] Maximum marks: Core concepts (**54**), Importance (38), Difficulty (38). Sum of marks: Core concepts (833), Importance (**154**), Difficulty (**150**). Respondents were not limited on core topics (the average respondent selected 17 items), but could only vote for three topics for the Top 3. The number of respondents inferred and used above: Core concepts (56 surveys), Top 3 Importance (51), Top 3 Difficulty (50)

During the second semester, the number of lessons declined with the major topics becoming percent and percent change, linear and exponential growth, indices and condensed measures, statistical measures, risk, and graphical interpretation and production."

Madison (2005 MAA): Here are what I think are important mathematical concepts for numeracy.

- Rates and rates of change: Absolutely essential; they occur in many newspaper articles every day.
- A 'percent' – what is it? And percent change.
- Times less. nobody has really told me what that really means.
- Graphs of the first and second derivative. They do come up in the newspaper....
- Linear and exponential rates of growth: Absolutely essential.
- Accumulation. Notice that these – rates of change and accumulation – are the two main ideas in calculus.
- Installment loans, savings and weighted averages: they keep coming up. Students need to learn to handle them.
- Indexes and Condensed Measures: this is something we don't teach at all. I wager that they are very few people in this room who can give me a rigorous definition of an index. I've thought about it and tried to explain it to my class. 'Condensed measures' is the term that I've heard – that's used in the literature – and they are just limited measures of some kind of variation, like 'poverty line.'
- Estimation: In many cases, estimation is the most important lesson of the day. Estimation has become incredibly important, more so because of computers, but very sophisticated
- Plane geometry. My students didn't know any plane geometry by the way. They didn't know the simplest volume and radius formulas. They didn't remember them, because they hadn't used them in real life.
- Graphical production and representation
- Probability: single and conditional, the idea of risk and odds. Students didn't know what odds meant. They didn't know how they could combine odds; they didn't know what risk was. They had never even thought that they could understand it.
- Graphs: One of the things that we spent a lot of time on was looking at graphs out of the newspaper and trying to figure out if they were the same graph that we saw in an algebra book. They didn't look at all like the graphs in an algebra class. In fact to even call them the same name seemed kind of silly. There is so much information in some of these graphs and the presentations are so unorthodox that students see these newspaper graphs as something different from what we deal with in geometry classes or algebra class. We definitely need to make those connections. [Second Derivative Graph, Circle Clock Graph] So my point about this is that our graphs in algebra don't look like those in the news. Now that doesn't make our graphs wrong; it doesn't make graphs in the news wrong either. Sometimes they are wrong; sometimes they are inconsistent. We need to build connections between what we do and what our students meet out there in the real world.

Appendix E: Gillman

According to Gillman (2006), "There is consensus that the mathematical skills necessary to be quantitatively literate include elementary logic, the basic mathematics of financial interest, descriptive statistics, finite probability, an elementary understanding of change, the ability to model problems with linear and exponential models, estimations and approximation, and general problem solving. It is clear that many of our students enter college with minimal mastery of these skills and their application."

Appendix F: Schield

According to Schield (2009), "statistical literacy is defined for adults in a modern society as the ability to understand and interpret the statistics in everyday life – to read and interpret statistics in the everyday media."

"Statistical literacy is the ability of "data consumers" to read and interpret statistics in the everyday media – in graphs, tables, statements and essays found in newspapers, popular magazines and consumer-oriented government reports. Statistical competence is the ability of "data producers" to design surveys and studies and to produce and analyze the statistics that appear in technical reports and professional journals."

"Since statistical literacy is needed by all educated adults as data consumers, it should be needed by students in non-quantitative majors: majors that don't require a math or statistics course."

It is not enough to evaluate the influences on a statistic – where the statistic is the result. To be statistically literate, one must also be able to evaluate the support that statistical evidence provides for a conclusion. To be statistically literate, one must be able to evaluate the statistic as both as an end and as a means to a higher end – as both a conclusion to one argument and as a premise to a second."

"So what would serve as the archetypical activity that could demonstrate excellence in statistical literacy? In the Keck Project, the archetypical activity is the ability to comment intelligently – without prompting – on the nature of, the truth of, the role of, and the support given by any statistic in an argument in everyday life. The assessment of this archetypical activity is not easy ....”

Of the more than 50 suggestions by these subject-matter experts, those that were statistically related were grouped into four categories:

- **C**ontext: (1) Those influences or factors blocked (controlled for) by the study design (cf., controlled vs. uncontrolled; longitudinal vs. cross-sectional; experiment vs. observational study) or by selection (cf., in tables and graphs). (2) Those factors taken into account by averages, comparisons, ratios and comparisons of averages and ratios. (3) Those factors taken into account by models (cf., multivariate regression). (4) Those factors not taken into account. These are plausible confounders: factors that were not controlled for in the study and are not blocked by the study design.

- **A**ssembly: The choices (1) in defining groups or measures, (2) in selecting measures (mean vs. median), comparisons (choice of the basis of comparison and the type: difference vs. ratio), and ratios (e.g., the choice of the denominator and the confusion of the inverse or the prosecutor’s fallacy), and (3) in presenting statistical results and summaries.

- **R**andomness: the influence of chance on averages (sampling theory) and on exceptions (hot hand, too unlikely to be due to chance, and regression to the mean). The difference between statistical significance and practical significance in large samples. The difference between “no statistical effect” and “no effect” in small samples. The influence on statistical significance of taking an associated factor into account.

- **E**rror or Bias: The influence of any factor that generates a systematic difference between what is observed and the underlying reality: subject bias (people can lie), measurement bias (instruments fail, questions lead and researchers manipulate) and sampling bias (the sampled population differs systematically from the target population).

Given the extensive influence of human choice on numbers, the Keck project grouped these four sources of influence under the age-old admonition, “Take CARE” where each of the four letters in ‘CARE’ signified a distinct source of influence on any statistic: Context, Assembly, Randomness and Error/Bias. Hopefully, this mnemonic will help students remember these influences on every statistic. If students were to remember just “Take CARE” in analyzing statistics, that would be a considerable achievement.

Basic Skills Survey

Demographics

Instructions: Circle the answer you believe is best or correct. Select only one answer.

1. **What best describes your background or perspective?**

   a. mathematics b. statistics c. psychology d. physical science e. Other social science f. Other.

2. **Have you taught quantitative reasoning, quantitative literacy or statistical literacy? If so, how long?**

   a. none b. one quarter/semester c. 2-5 quarters/semesters d. 6-10 quarters/semesters e. > 10.

3. **Have you taught pre-calculus GenEd math courses such as Math for Liberal Arts? If so, how long?**

   a. none b. one-two years c. 3-5 years d. 6-10 years e. > 10 years.

4. **Have you taught teaching traditional introductory statistics? If so, how long?**

   a. none b. one-two years c. 3-5 years d. 6-10 years e. > 10 years.

5. No Yes **Have you belonged to the National Numeracy Network or attended their meetings/sessions?**

6. No Yes **Have you belonged to the MAA QL-SIG or attended any of their sessions?**

7. No Yes **Have you belonged to the ASA Statistical Education Section or attended their sessions?**

8. No Yes **Have you been a member of the National Numeracy Network?**

9. **Have you taken college statistics? If so, how many?**

   a. None b. one course c. Two or three courses d. four or five courses e. more than five

10. **What best describes your major (or likely major) in college?** *(Select only one.)*

    a. Not applicable
    b. Education: Primary, Special Ed, or Secondary with a non-quantitative emphasis (English, History).
    c. Education: Secondary with a quantitative emphasis (Math or science).
    d. Mathematics or statistics.
    e. Other quantitative major with researcher manipulation (E.g., lab sciences or psychology, clinical trials)
    f. Other quantitative major with just researcher observation (E.g., observational or social sciences)
    g. Other non-quantitative major (E.g., Visual and performing arts, Communications, Journalism, Physical or Sports Education, English, History, Political Science, Religion or Philosophy)

Basic Skills Survey: Topics

D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; AA = Absolutely Agree; NO = No Opinion

For each claim about numeracy, circle the acronym indicating your opinion.

D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; AA = Absolutely Agree; NO = No Opinion

11. D N A SA AA – NO: **To be a real habit of mind, numeracy must deal with unstructured combinations of prose, numbers and numerical representations.**

12. D N A SA AA – NO: **To be teachable, numeracy must focus on teachable skills and knowledge that allow students continuing practice.**

13. D N A SA AA – NO: **Generating any kind of assessment for numeracy WILL TRANSFORM IT FROM an open-ended critical thinking habit of mind activity TO a simple problem/skill-based course**

14. D N A SA AA – NO: **Numeracy should not be operationally defined UNTIL there is widespread agreement by representatives from all groups that could make it a viable activity or discipline.**

15. D N A SA AA – NO: **High-level proficiency CANNOT be assessed using a standardized instrument.**

16. D N A SA AA – NO: **IF high-level proficiency CANNOT be assessed using a standardized instrument that would be so BECAUSE a standardized instrument only allows one right answer.**

17. D N A SA AA – NO: **Basic-level proficiency CAN be adequately assessed using a standardized multiple choice instrument with right-wrong answers.**

18. D N A SA AA – NO: **If basic-level proficiency can be assessed using a standardized instrument, then it SHOULD BE assessed.**

19. D N A SA AA – NO: **Even if basic-level proficiency can be assessed it should not be assessed using a standardized instrument because that will shift the focus from higher-level critical thinking skills to lower level skills involving calculation and memorization.**

Here are statements about statistical or quantitative literacy. Circle your agreement with the claim.

Sons (1994) identified the activities a quantitatively literate person should be able to do. See Appendix A.

Numeracy should include or feature …..

20. D N A SA AA – NO: **reading and interpreting graphs.**
21. D N A SA AA – NO: **reading and interpreting tables.**
22. D N A SA AA – NO: **reading and interpreting schematics.**
23. D N A SA AA – NO: **representing information visually.**
24. D N A SA AA – NO: **representing information numerically.**
25. D N A SA AA – NO: **representing information verbally.**
26. D N A SA AA – NO: **solving problems using arithmetical methods.**
27. D N A SA AA – NO: **solving problems using algebraic methods.**
28. D N A SA AA – NO: **solving problems using geometric methods.**
29. D N A SA AA – NO: **solving problems using statistical methods.**
30. D N A SA AA – NO: **estimating to determine reasonableness..**
31. D N A SA AA – NO: **the limits of mathematical methods.**
32. D N A SA AA – NO: **the limits of statistical methods.**

Basic Skills Survey: Topics

D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; AA = Absolutely Agree; NO = No Opinion

Porter (1997): “Rarely do they [students] learn what a stratified sample is, or how an unemployment rate is determined, or what the smog index measures. The sorts of numbers that modern citizens are likely to confront in their lives as citizens and voters have little place in the modern curriculum.”

Numeracy should feature or include …

33. D N A SA AA – NO: **different kinds of random sampling.**
34. D N A SA AA – NO: **the construction of social statistics.**
35. D N A SA AA – NO: **the construction of indexes.**
36. D N A SA AA – NO: **the sensitivity of index values to the methods used.**

Moore (1998) thought that statistical literacy should involve two clusters of “big ideas”: 1) “The omnipresence of variation, conclusions are uncertain, avoid inference from short-run irregularity, [and] avoid inference from coincidence.” 2) “Beware the lurking variable, association is not causation, where did the data come from? [and] observation versus experiment.”

Statistical literacy should feature or include …

37. D N A SA AA – NO: **the omnipresence of variation.**
38. D N A SA AA – NO: **the uncertainty of conclusions.**
39. D N A SA AA – NO: **the error of inferring from short-run irregularity.**
40. D N A SA AA – NO: **the error of inferring from coincidence.**
41. D N A SA AA – NO: **lurking variables.**
42. D N A SA AA – NO: **association is not causation.**
43. D N A SA AA – NO: **data production.**
44. D N A SA AA – NO: **experiments vs. observational studies.**

Utts (2003) identified “seven topics that are commonly misunderstood by citizens, including the journalists who present statistical studies to the public. See Appendix B.

Statistical literacy should feature or include …

45. D N A SA AA – NO: **difference between causation and association.**
46. D N A SA AA – NO: **conditions for validating causation.**
47. D N A SA AA – NO: **difference between randomized and observational studies.**
48. D N A SA AA – NO: **difference between practical and statistical significance.**
49. D N A SA AA – NO: **difference between ‘no difference’ and ‘no statistically significant difference’.**
50. D N A SA AA – NO: **common types of bias in surveys.**
51. D N A SA AA – NO: **coincidences as not being uncommon.**
52. D N A SA AA – NO: **the “confusion of the inverse.” $P(A|B) \neq P(B|A)$**
53. D N A SA AA – NO: **the fact that variation is normal.**
54. D N A SA AA – NO: **difference between ‘normal’ and ‘average.’**

D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; AA = Absolutely Agree; NO = No Opinion

Schaeffer (2004), "Many aspects of statistical thinking are not about numbers as much as about concepts and habits of mind. For example, the idea of a lurking variable upsetting an apparent bivariate relationship with observational data is a conceptual idea, part of statistical thinking, but not particularly about numbers.

Statistical literacy should feature or include …

55. D N A SA AA – NO: **the importance of a lurking variable.**
56. D N A SA AA – NO: **the importance of a lurking variable even if it means less emphasis on statistical inference.**

McKenzie (2004) presented these 30 topics as being important in introductory statistics. See Appendix C.

Statistical literacy should feature or include …

57. D N A SA AA – NO: **the importance of variability.**
58. D N A SA AA – NO: **the difference between association and causation.**
59. D N A SA AA – NO: **randomness.**
60. D N A SA AA – NO: **difference between practical and statistical significance.**
61. D N A SA AA – NO: **how data is collected: experiment, observational study or survey.**
62. D N A SA AA – NO: **sampling distributions: law of large number & Central Limit theorem.**
63. D N A SA AA – NO: **hypothesis testing: p-value, critical value and power.**
64. D N A SA AA – NO: **confidence intervals.**
65. D N A SA AA – NO: **random samples.**
66. D N A SA AA – NO: **the different types of data.**
67. D N A SA AA – NO: **measures of center.**
68. D N A SA AA – NO: **importance of assumptions.**
69. D N A SA AA – NO: **graphing.**
70. D N A SA AA – NO: **importance of uncertainty.**
71. D N A SA AA – NO: **distributions.**
72. D N A SA AA – NO: **independence.**
73. D N A SA AA – NO: **importance of bias.**
74. D N A SA AA – NO: **correlation.**
75. D N A SA AA – NO: **shapes of distributions.**
76. D N A SA AA – NO: **data exploration.**
77. D N A SA AA – NO: **data exploration.**
78. D N A SA AA – NO: **the importance of proportion.**
79. D N A SA AA – NO: **least-squares regression.**
80. D N A SA AA – NO: **importance of models.**
81. D N A SA AA – NO: **importance of comparisons.**
82. D N A SA AA – NO: **importance of prediction.**
83. D N A SA AA – NO: **importance of outliers and robustness.**
84. D N A SA AA – NO: **difference between longitudinal and cross-sectional studies.**
85. D N A SA AA – NO: **regression effect.**
86. D N A SA AA – NO: **importance of a statistical process.**
87. D N A SA AA – NO: **importance of transforming a variable.**

D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; AA = Absolutely Agree; NO = No Opinion

Madison (2005b) identified important mathematical concepts for numeracy. See Appendix D.

Numeracy should feature or include ….

88. D N A SA AA – NO: **rates and rates of change.**
89. D N A SA AA – NO: **percents and percent change.**
90. D N A SA AA – NO: **comparisons: times as much, times more/less.**
91. D N A SA AA – NO: **graphs.**
92. D N A SA AA – NO: **graphs of the first and second derivative.**
93. D N A SA AA – NO: **linear rates of growth.**
94. D N A SA AA – NO: **exponential rates of growth.**
95. D N A SA AA – NO: **accumulation.**
96. D N A SA AA – NO: **installment loans and savings.**
97. D N A SA AA – NO: **weighted averages.**
98. D N A SA AA – NO: **indexes.**
99. D N A SA AA – NO: **condensed measures (e.g., poverty line).**
100. D N A SA AA – NO: **estimation.**
101. D N A SA AA – NO: **plane geometry.**
102. D N A SA AA – NO: **graphical production.**
103. D N A SA AA – NO: **graphical representation.**
104. D N A SA AA – NO: **probability, risk and odds.**
105. D N A SA AA – NO: **conditional probability, risk and odds.**

Madison (2006) added a few more topics:

Numeracy should feature or include ….

106. D N A SA AA – NO: **counting.**
107. D N A SA AA – NO: **weather maps.**

Gillman (2006) identified the mathematical skills necessary to be quantitatively literate. Appendix E.

Quantitative Literacy should feature or include ….

108. D N A SA AA – NO: **elementary logic.**
109. D N A SA AA – NO: **the basic mathematics of financial interest.**
110. D N A SA AA – NO: **descriptive statistics.**
111. D N A SA AA – NO: **finite probability.**
112. D N A SA AA – NO: **elementary understanding of change.**
113. D N A SA AA – NO: **ability to model problems.**
114. D N A SA AA – NO: **modeling problems with linear models.**
115. D N A SA AA – NO: **modeling problems with exponential models.**

Basic Skills Survey: Topics

D = Disagree; N = Neutral; A = Agree; SA = Strongly Agree; AA = Absolutely Agree; NO = No Opinion

Schield **(2009)** identified important concepts for statistical literacy. See Appendix F.

Statistical literacy should feature or include ….

116. D N A SA AA – NO: **study design: controlled vs. uncontrolled**
117. D N A SA AA – NO: **study design: longitudinal vs. cross-sectional**
118. D N A SA AA – NO: **study design: experiment vs. observational study**
119. D N A SA AA – NO: **data selection: rows/columns in tables; series in graphs.**
120. D N A SA AA – NO: **comparisons of numbers: difference, times ratio and percent change**
121. D N A SA AA – NO: **ratios: percentages and rates**
122. D N A SA AA – NO: **comparisons of ratios: percentages and rates**
123. D N A SA AA – NO: **models: linear (constant difference) and logarithmic (constant ratio)**
124. D N A SA AA – NO: **confounders: plausible (not excluded)**
125. D N A SA AA – NO: **confounders: influence on averages and rates**
126. D N A SA AA – NO: **influence of definitions on counts and sums. [social construction]**
127. D N A SA AA – NO: **influence of definition on rates and ratios. [social construction]**
128. D N A SA AA – NO: **influence of definition on rates and ratios. [social construction]**
129. D N A SA AA – NO: **influence of choice of group on rates and ratios. [social construction]**
130. D N A SA AA – NO: **influence of choice of basis on comparison. [social construction]**
131. D N A SA AA – NO: **influence of choice of type of comparison. [social construction]**
132. D N A SA AA – NO: **influence of choice of conditional in probability. [social construction]**
133. D N A SA AA – NO: **influence of choice in presentation. [social construction]**
134. D N A SA AA – NO: **influence of choice on averages. [randomness]**
135. D N A SA AA – NO: **influence of chance on exceptions and runs. [randomness]**
136. D N A SA AA – NO: **influence of chance: regression to the mean. [randomness]**
137. D N A SA AA – NO: **'no difference' vs. 'no statistically significant difference' [randomness]**
138. D N A SA AA – NO: **'statistical significance' vs. 'practical significance' [randomness]**
139. D N A SA AA – NO: **influence of confounder on statistical significance [randomness]**
140. D N A SA AA – NO: **influence of subject bias [error/bias]**
141. D N A SA AA – NO: **influence of measurement bias [error/bias]**
142. D N A SA AA – NO: **influence of sampling bias [error/bias]**

ENTER ANY COMMENTS OR SUGGESTIONS HERE: