Cut to the QUIC: What is the Essence of Quantitative Reasoning and How Can We Assess It?

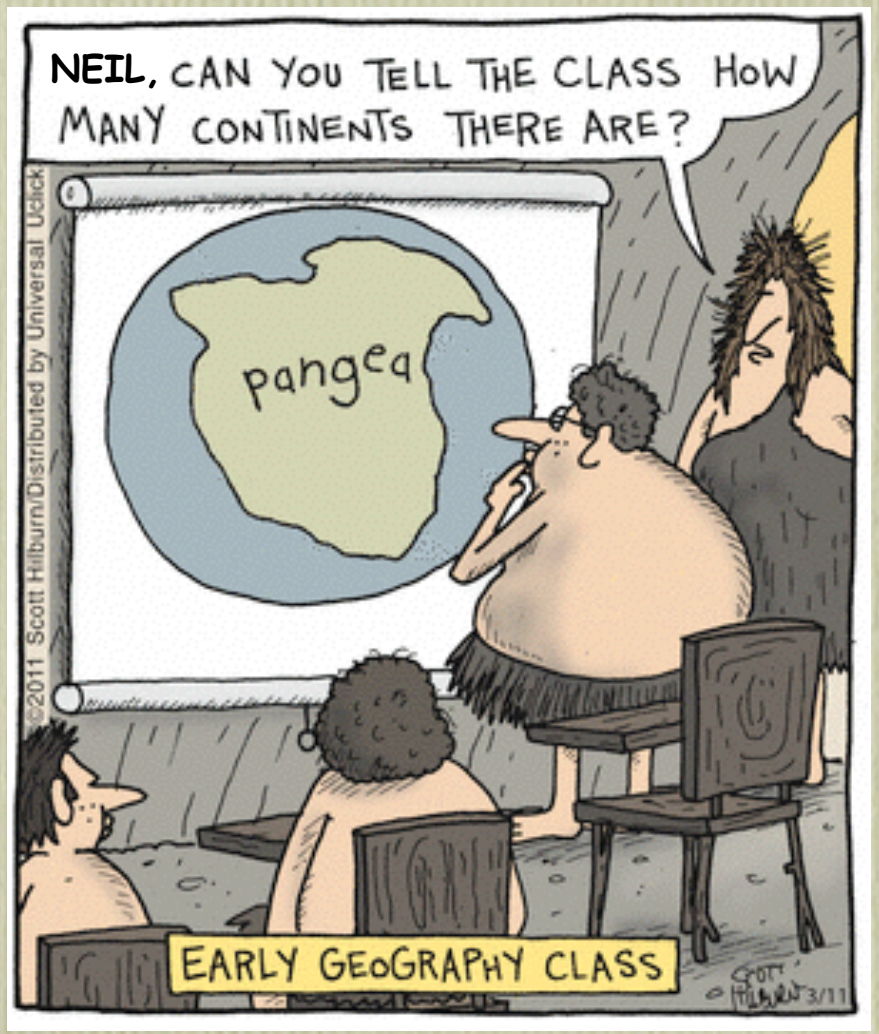
Neil Lutsky, Carleton College

Enhancing QR Across the Curriculum Workshop & NNN Annual Meeting Plenary, October 10, 2014

Presentation Agenda:

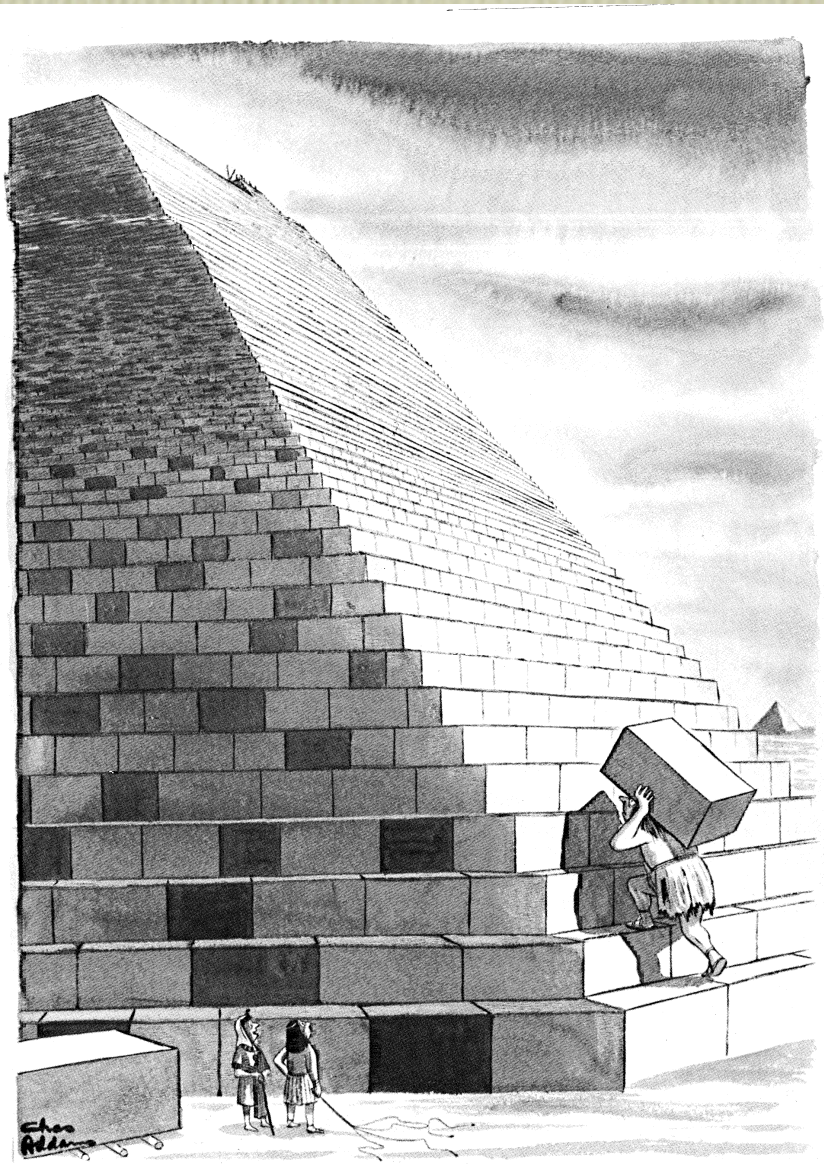
- It *is* possible to enhance QR across a curriculum.
- *One of the features of Carleton's approach continues to challenge regnant conceptions of QR.*
- This has *critical* implications for how QR is assessed.
- *The QUIC, a new measure of QR, may offer direction. I invite you to help refine it.*

QR at Carleton



- *Concern about student attitudes.*
- *Faculty/staff belief in the importance of QR.*
- *But little attention to QR in the educational agenda of the College.*

A *Voluntary* Committee!



"We could never have done it without him."

- *Faculty & Staff.*

— *Read.*

— *Gathered information.*

— *Listened to wise mentors.*

— *Talked.*

— *Articulated a rationale.*

— *Appealed to administrators.*

- *QuIRK: Quantitative Inquiry, Reasoning, and Knowledge Initiative.*

- To help students learn how to evaluate numerical claims thoughtfully.

Critical Thinking

- To help students appreciate the contributions numerical analysis can make to human understanding.

Apprehending the World

- To help students strengthen their abilities to construct principled arguments using quantitative evidence.

Rhetoric & Writing

How did we recruit faculty for our mission?

How did we recruit faculty for our mission?

- *Faculty workshops.*

  Reading @ Medical Research; Writing with Numbers; QRU

- *Campus speakers (notably, at convocations).*

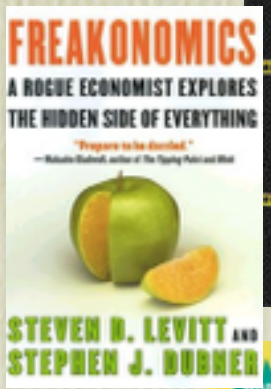
  David Hemenway; Steven Levitt; Joel Best

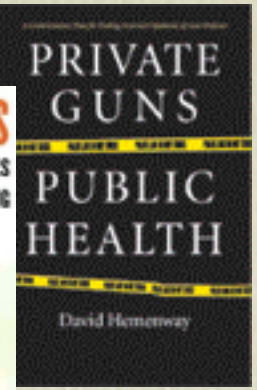
- *Curriculum development grants.*

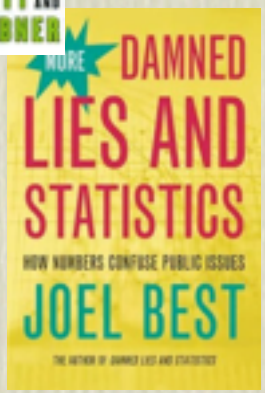
  First Year Courses; Upper Level QR Modules

- *Learning & Teaching Center events.*

  Presentations from colleagues using QR Modules







Were we successful?

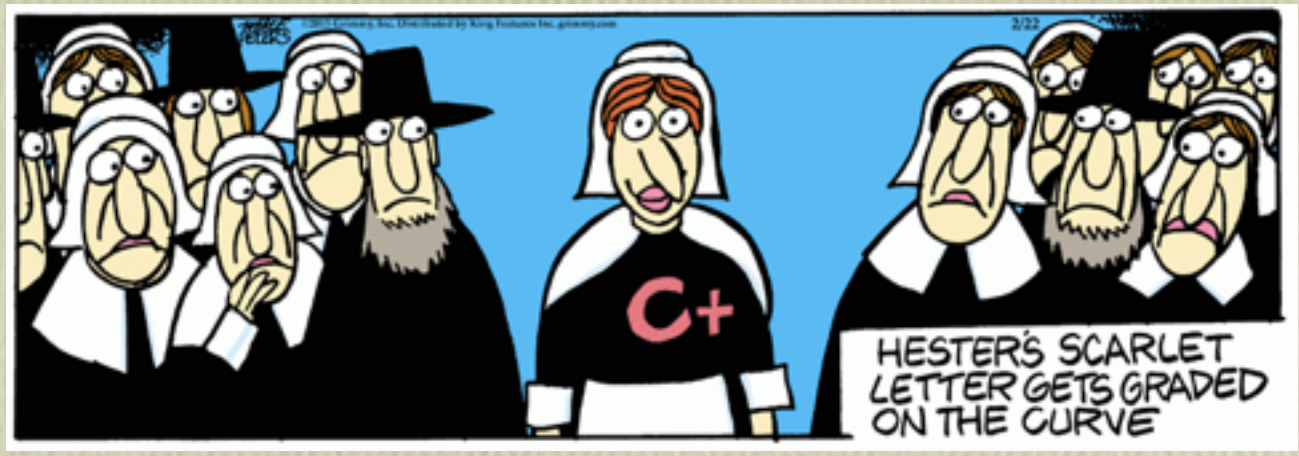
- *Faculty involvement.*
- *Curricular development.*
- *QRE institutionalization.*
- *Student attitudes.*
- *Student demand for statistics courses.*
- *Student use of QR in writing* (Grawe, 2013).
- *Participation in larger conversations.*

HOWEVER…

Were we successful?

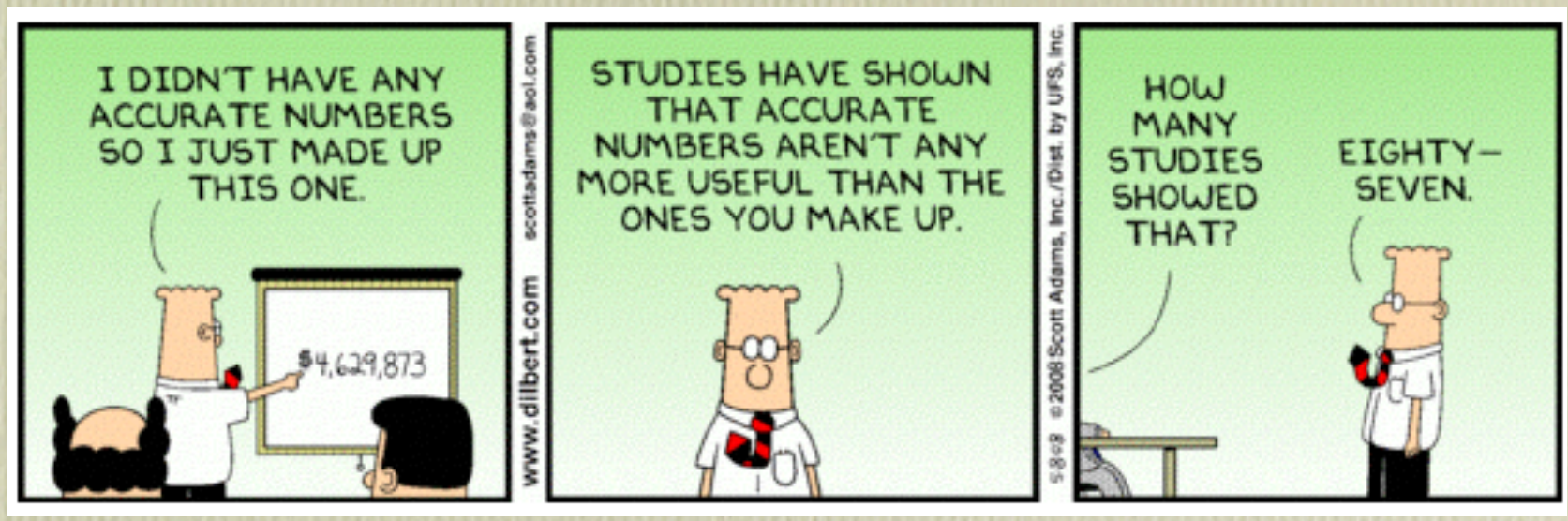
- *Local institutional foreground to background.*
- *Student exposure > systematic training.*
- *Writing assessment approach has had limited appeal at other institutions.*
- We have not overthrown narrow conceptions of Quantitative Reasoning.
- We have not met the need to transform undergraduate QR education.

What accounts for our success?



- ✓ *Colleagues, both in the QR community &*
- ✓ *Collaborative foundation and shared vision.*
- ✓ *$$$$$.*
- ✓ *Positive orientation.*
- ✓ *Interest in learning from assessment.*

Positive Orientation?



The New York Times

December 13, 2011, 8:0[illegible] PM

Millions and Billions

By PHILIP B. CORBETT

Notes from the newsroom on grammar, usage and style. (Some frequently asked questions are here.)

D[illegible]

came [illegible]

...

An ar[illegible] admin[illegible] the Gu[illegible] been [illegible] the fu[illegible]

An ar[illegible] States[illegible] five y[illegible]

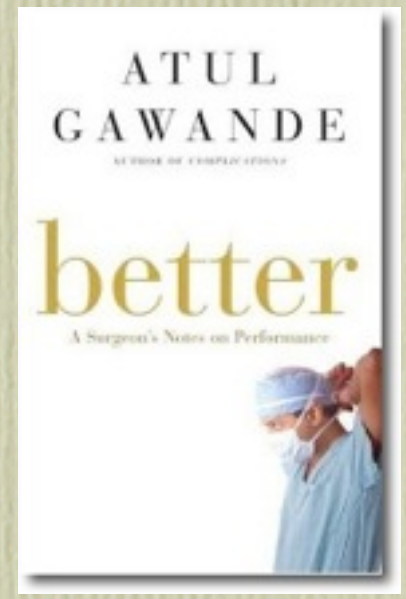
An ar[illegible] Papan[illegible] amid [illegible] installment of foreign aid the country is hoping to sec[illegible] ***not $11 million.***

Positive Orientation?

"Statistics can tell us things about the world that we could not imagine on the basis of our senses alone…The social world is unimaginably more complex than we can see directly from any subjective vantage point within it."

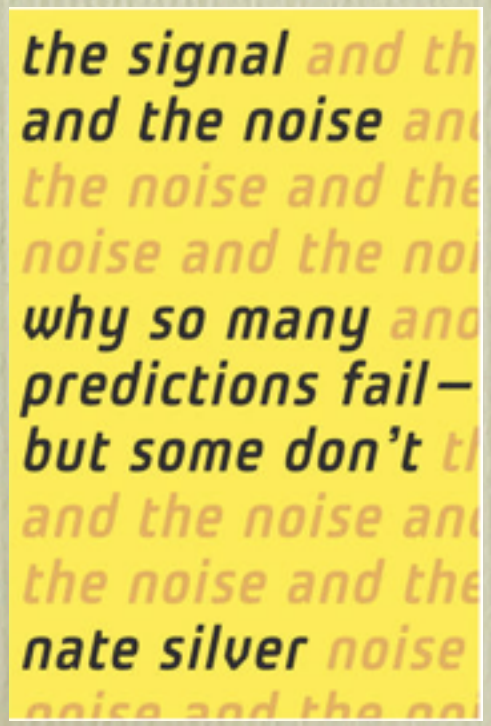
-Paul Seabright, *Times Literary Supplement*

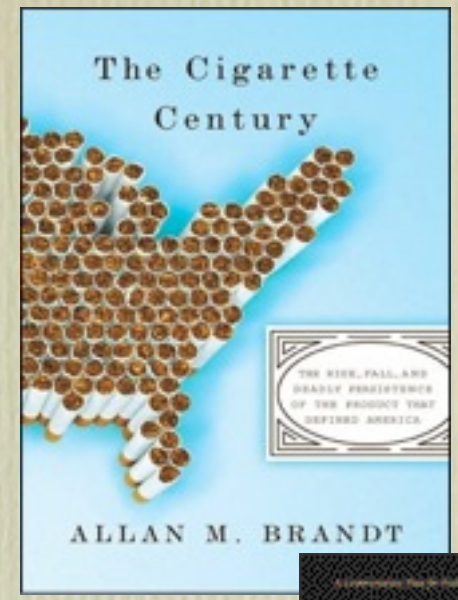
Compelling **Positive** *Examples:*

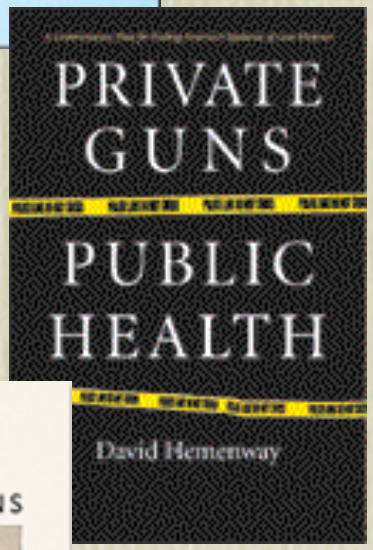


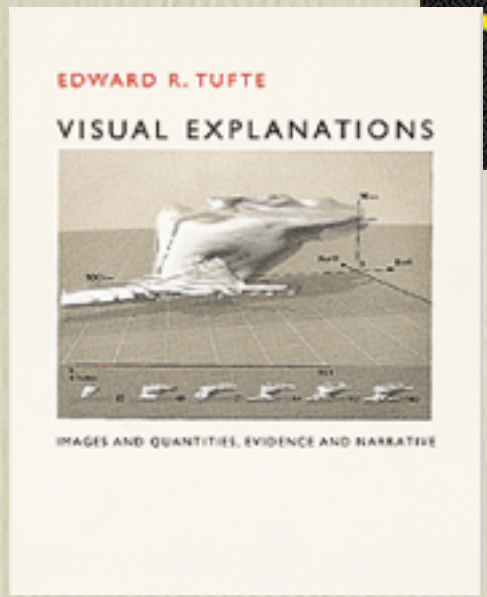
Casualties of War, pp. 51-69.

The Score, pp. 169-200.









A Disability Epidemic Among a Railroad's Retirees

THE RAILROAD A commuter rail line since it began in 1834, retirement policy and disability rate.

The New York Times

10 Arrested in $1 Billion L.I.R.R. Disability Scheme

By WILLIAM K. RASHBAUM
Published: October 27, 2011

Ten people, including a doctor and a former union president, were arrested early Thursday and charged in a major fraud scheme in which hundreds of Long Island Rail Road workers made false disability pension claims costing a federal agency an estimated $1 billion, according to people briefed on the matter. Another doctor charged in the case was being sought, the people said.

The New York Times

Retired, Then Disab

In 2007, 94 perce
employees who r
Long Island Rail
50 then received
benefits, far abov
for all railroads.

The Odds, Continually Updated

By F. D. FLAM SEPT. 29, 2014

In Bayesian statistics, new data is used to shape assumptions, the opposite of the frequentist (classical) approach.

R.R.
4%
ees
er
ALL
HER
ADS
3%

0
'91 '95 '99 '03 '07

Source: New York Times analysis of U.S. Railroad Retirement Board data

THE NEW YORK TIMES

The New York Times

For Today's Graduate, Just One Word: Statistics

By STEVE LOHR
Published: August 5, 2009

MOUNTAIN VIEW, Calif. — At Harvard, Carrie Grimes majored in anthropology and archaeology and ventured to places like Honduras, where she studied Mayan settlement patterns by mapping where artifacts were found. But she was drawn to what she calls "all the computer and math stuff" that was part of the job.

Enlarge This Image

Thor Swift for The New York Times

Carrie Grimes, senior staff engineer at Google, uses statistical analysis of data to help improve the company's search engine.

"People think of field archaeology as Indiana Jones, but much of what you really do is data analysis," she said.

Now Ms. Grimes does a different kind of digging. She works at Google, where she uses statistical analysis of mounds of data to come up with ways to improve its search engine.

Ms. Grimes is an Internet-age statistician, one of many who are changing the image of the profession as a place for dronish number nerds. They are finding themselves increasingly in demand — and even cool.

"I keep saying that the sexy job in the next 10 years will be statisticians," said Hal Varian, chief economist at Google. "And I'm not kidding."

"I keep saying that the sexy job in the next 10 years will be statisticians."

"It is easy to lie with statistics, but easier to lie without them."

-Frederick Mosteller

What's your *positivity* ratio?

Assessment


WE WILL BE ADOPTING THE BEST PRACTICES IN OUR INDUSTRY, JUST LIKE EVERYONE ELSE.
BEST PRACTICES
www.dilbert.com scottadams@aol.com
IF EVERYONE IS DOING IT, BEST PRACTICES IS THE SAME THING AS MEDIOCRE.
©2006 Scott Adams, Inc./Dist. by UFS, Inc.
STOP MAKING MEDIOCRITY SOUND BAD!
SORRY.

QuIRK emphasized assessment, and *learned from assessment*.

Lady Gaga in Tilmann Grawe

Tomorrow, 11:00-12:15

Nathan Grawe,

Tomorrow, 11:00-12:15

Formative → Summative Assessment

Can we show students are developing QR?

Well, two out of 10 ain't bad

British church officials have come to the defense of priests who cannot rattle off all Ten Commandments, saying it is substance, not words, that count.

A poll by the (London) Sunday Times found only 34 percent of 200 Anglican priests polled could recite all 10 without help. "When people are put on the spot like this of course they can't remember," a Church of England spokesman said. "Given time they would recall them."

The poll found that most clergy know the commandments prohibiting adultery and coveting one's neighbor's wife, but get a little fuzzy on the details of some of the other eight.

— ***Reuters***

What we test conveys a claim @ what QR is!

What do summative QR measures assess?

- Berlin Numeracy Test (Cokely, Galesic, Schulz, Ghazal, & Garcia-Retamero, 2012).
- Quantitative Literacy & Reasoning Assessment (Gaze, Montgomery, Kilic-Bahi, Leoni, Misener,, & Taylor, 2014).
- Abbreviated Numeracy Scale (Weller, Dieckmann, Tusler, Mertz, Burns, & Peters, 2013).
- Others: Subjective Numeracy Scale (Fagerlin et al., 2007), QL Assessment Instrument (Follette, 2014).

Out of 1,000 people in a small town 500 are members of a choir...

In the BIG BUCKS LOTTERY, the chances of winning...

The age dependency ratio is computed by...

Two married alumni are calculating their federal income tax...

Probability

Ratios

Graphs/Tables

Exponential Growth

Algebra

Multiplication

Mean, Median

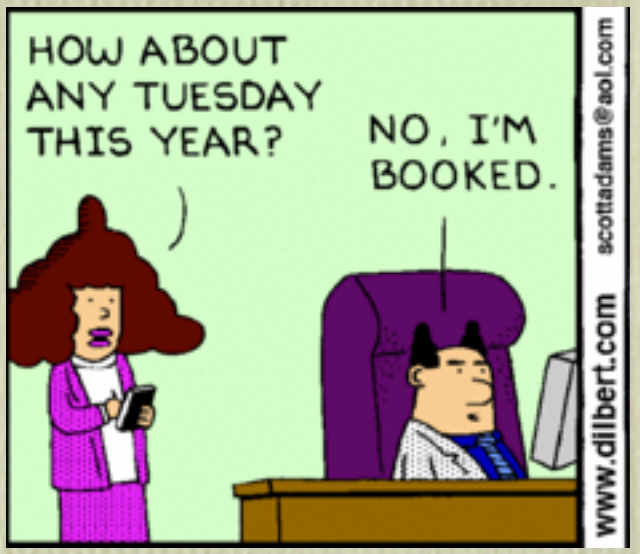
Percentages

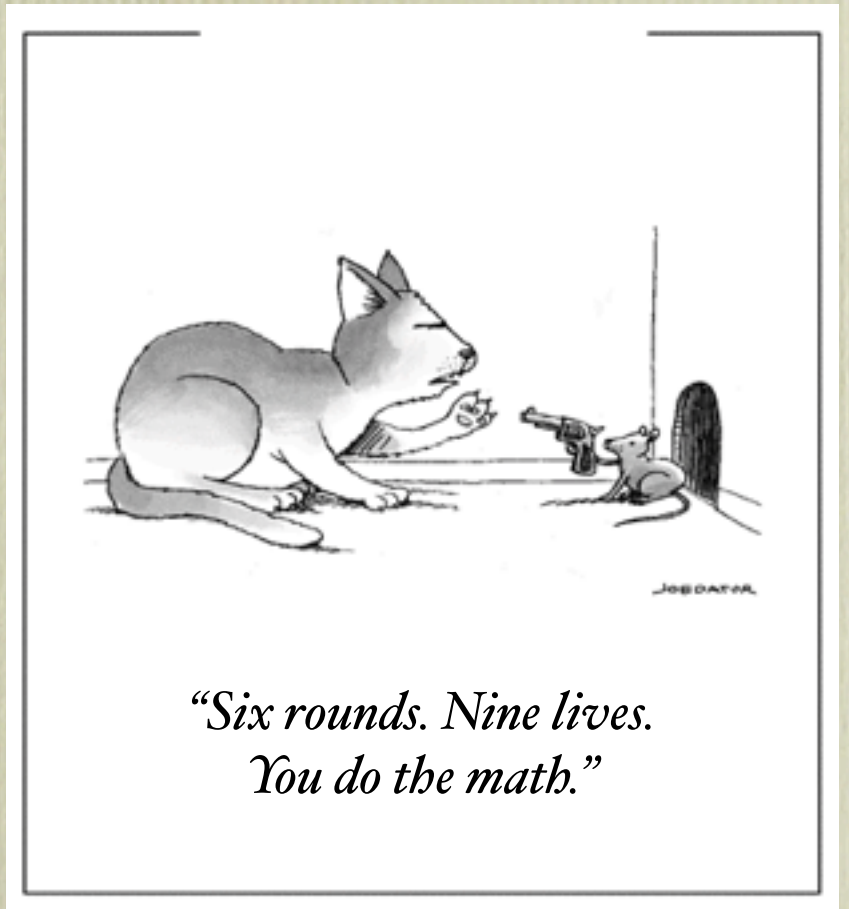
Deborah Hughes Hallett

Quantitative Literacy as:

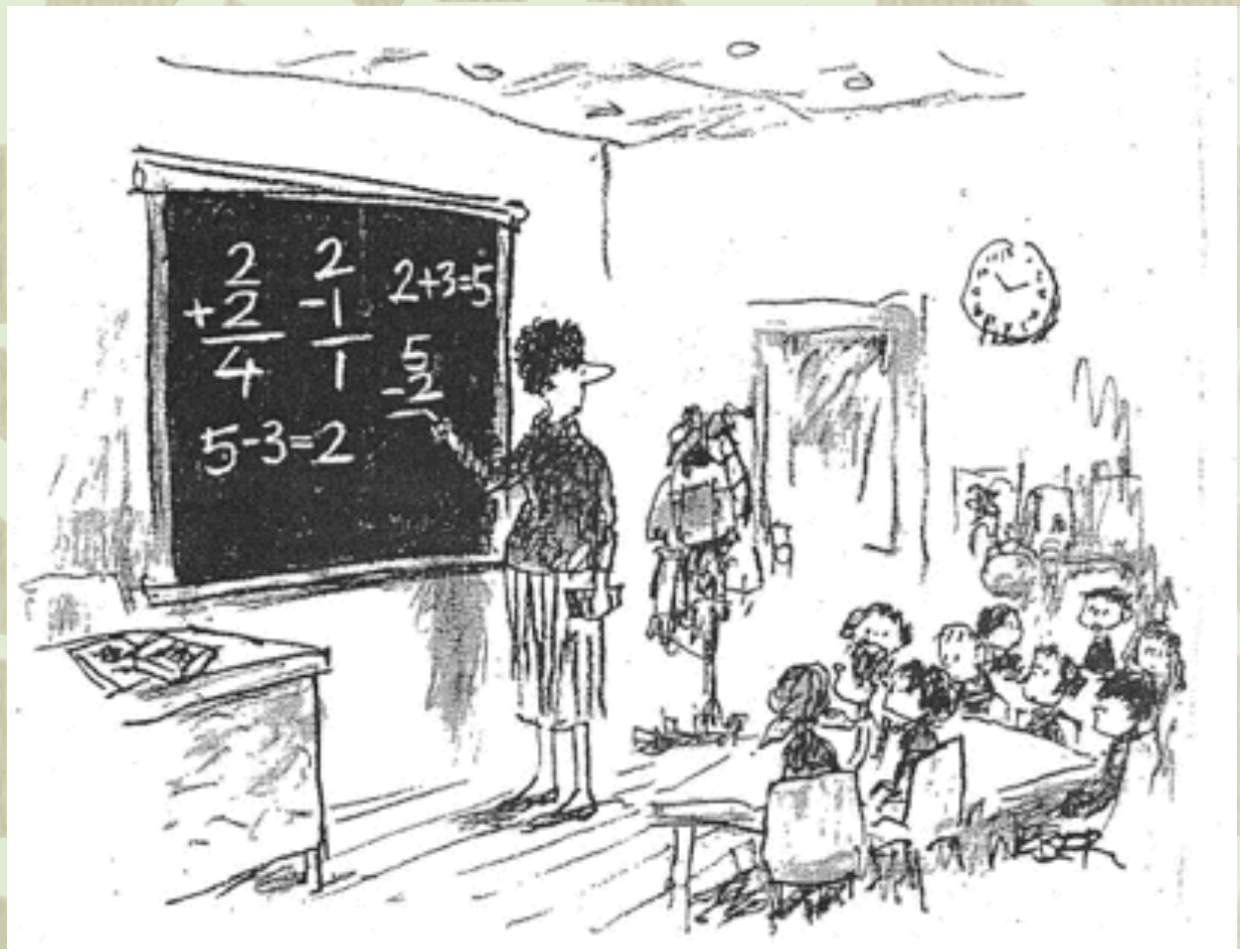
"The ability to identify, understand, and use elementary mathematics in everyday contexts."

- *Arithmetic.*
- *Estimation.*
- *Elementary Probability and Statistics.*
- *Geometry and Measurement.*
- *Elementary Growth Patterns.*


HOW ABOUT ANY TUESDAY THIS YEAR?
NO, I'M BOOKED.
www.dilbert.com
scottadams@aol.com


"Six rounds. Nine lives.
You do the math."

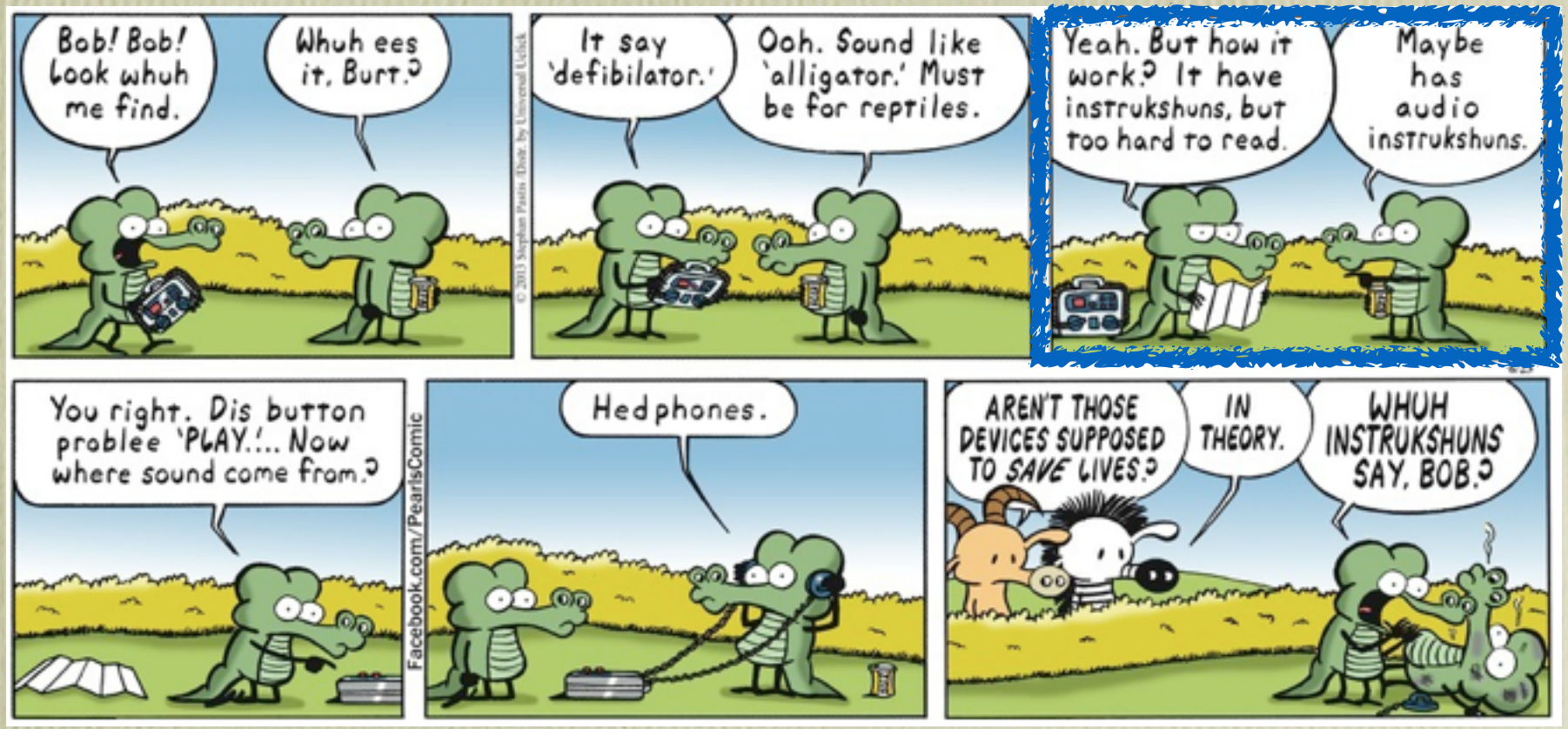
Mathematical concepts and calculation skills *can* contribute to *certain* applications of QR...



"Please, Ms. Sweeney, may I ask where you're going with all this?"

...*but* that doesn't make them the equivalent of QR.

Attentive Reading
contributes to QR.



Attentive Reading

Numeracy or QR?

Questions 1 through 7 refer to the following passage:

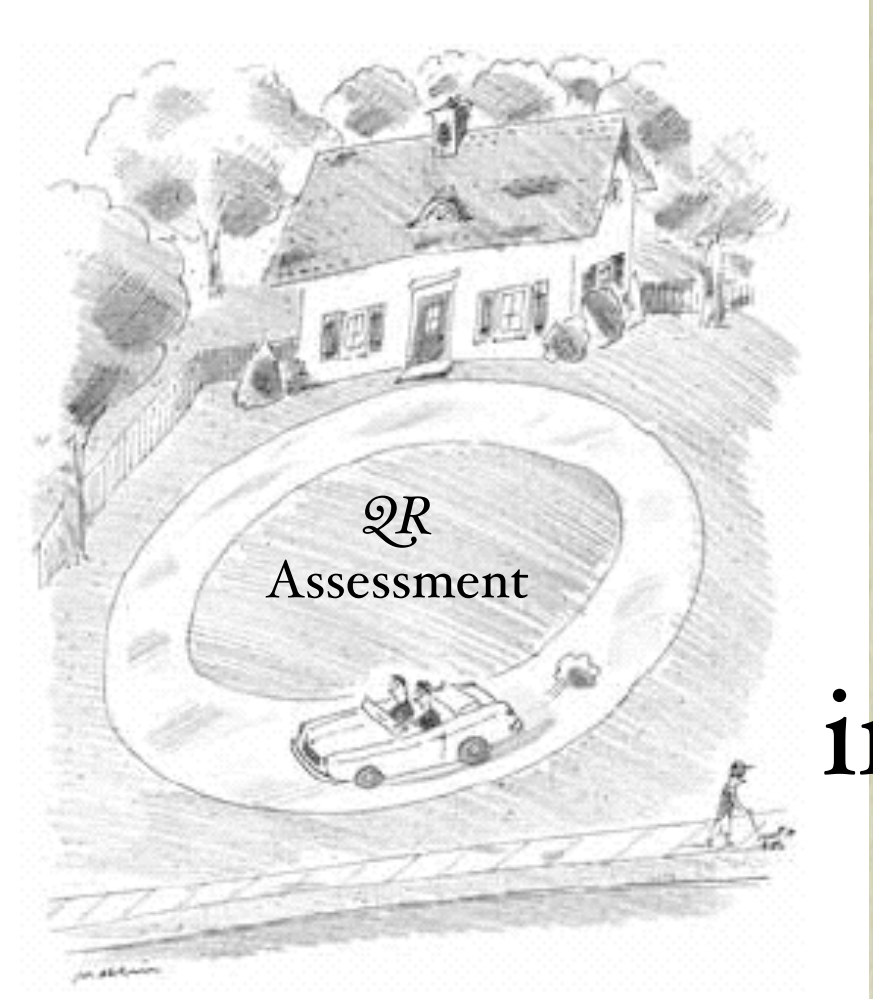
In the 16th century, an age of great marine and terrestrial exploration, Ferdinand Magellan led the first expedition to sail around the world.

1. The 16th century was an age of great _____ exploration.

A. cosmic
B. land
C. mental
D. common man
E. None of the above

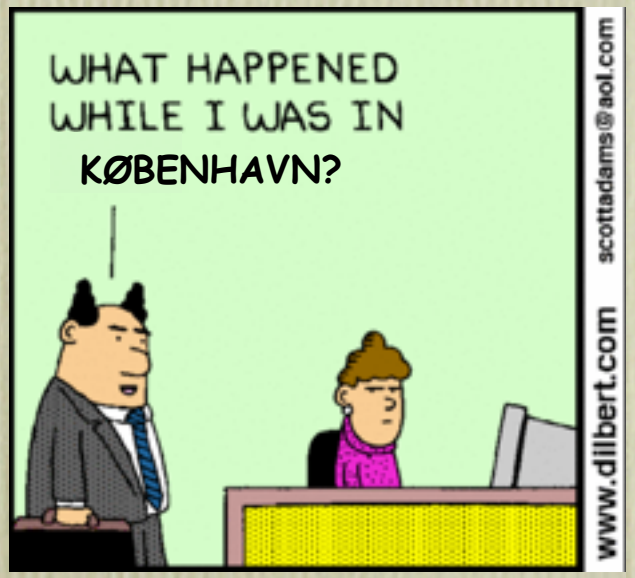
Existing "QR" tests, by and large, assess math concepts and operations.

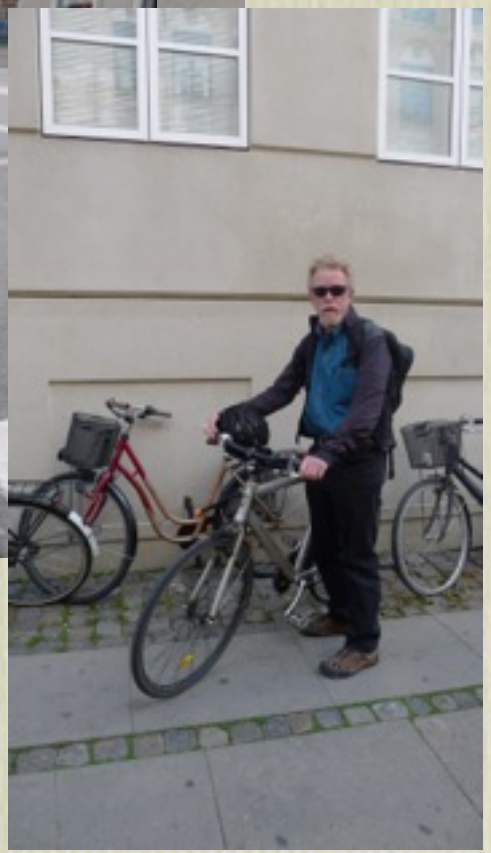
- ✓ Useful for evaluating certain math skills.
- — Problematic and misleading for the measurement of *QR*.



Why are we stuck driving around in mathematical circles in the QR community?


COPENHAGEN


WHAT HAPPENED WHILE I WAS IN KØBENHAVN?
www.dilbert.com
scottadams@aol.com



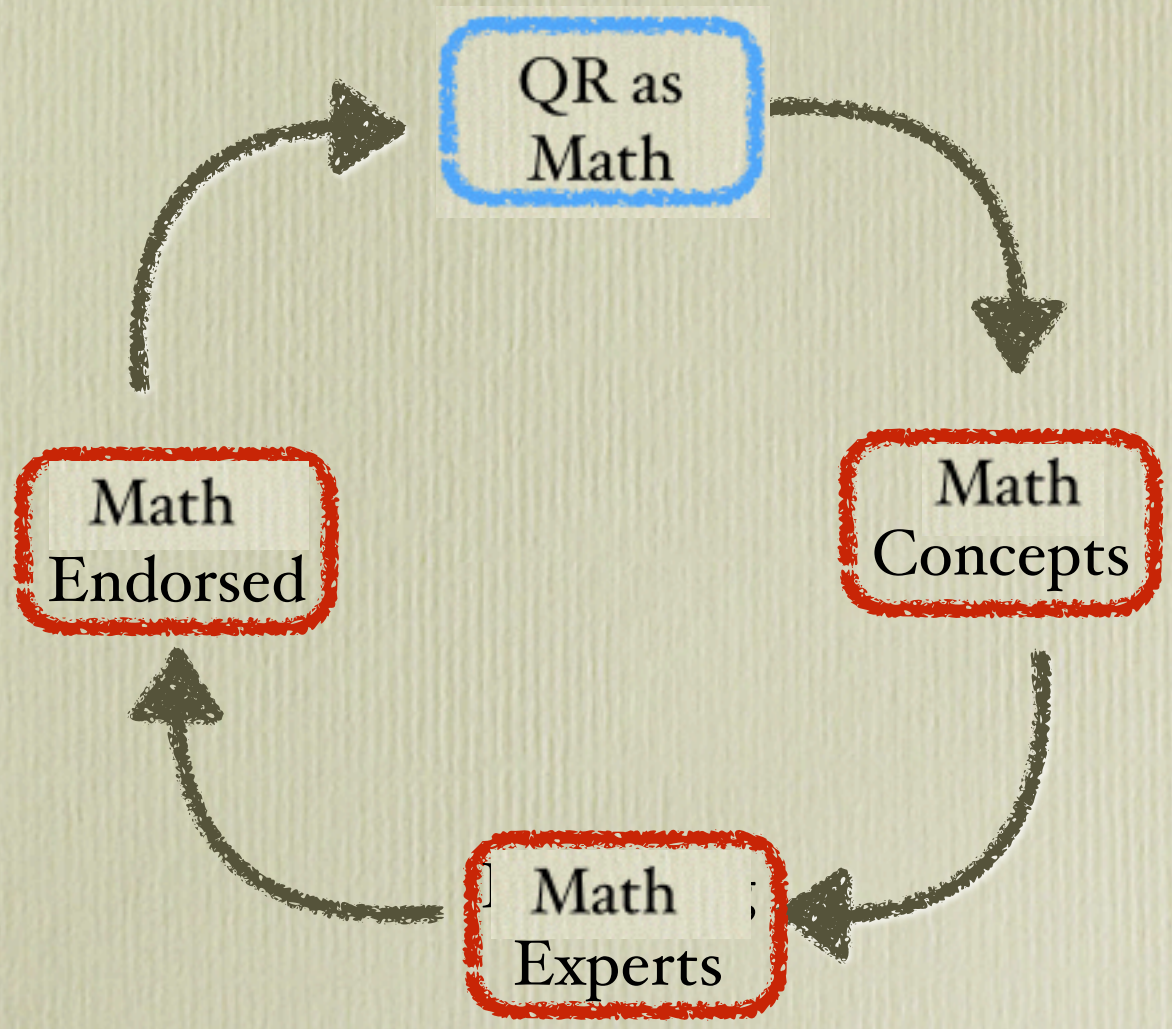
The research procedure makes a huge difference!

40% wore helmets

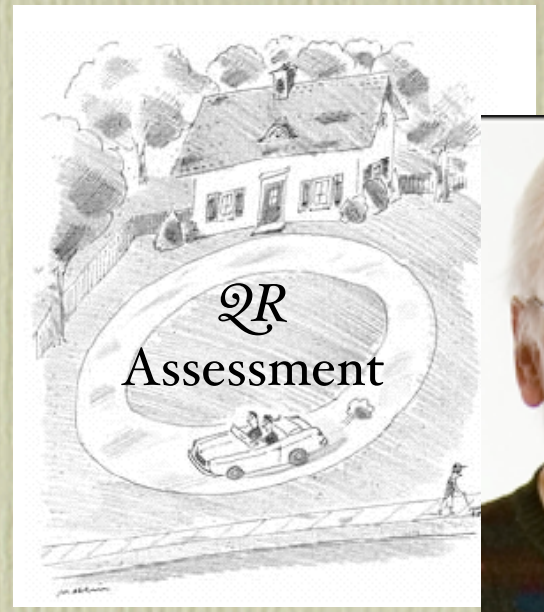
80% serious cyklists wore helmets

20% commuters wore helmets

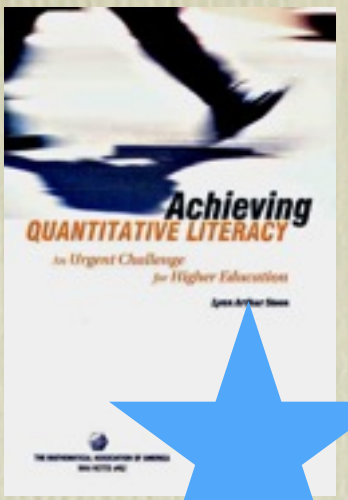
Test Development Procedures, or How to Reinvent the QR Assessment Wheel



The research procedure makes a huge difference!



Lynn Steen



"QL is primarily about reasoning with data, mathematics is about patterns, numbers, and space."

What kinds of data are part and parcel of everyday life (or, at least, should be)?

Court's Gay Marriage Decision Is in Line With Public Opinion

The Supreme Court's decision to deny review of all five pending same-sex marriage cases, thereby clearing the way for such marriages in several ... puts it in line with an American public that supports same-sex marria... says its legality should be left to each individual state to decide.

A majority of Americans say it should be legal for same-sex couples to ... — in a New York Times/CBS News poll last month, 56 percent said so. The issu... wo-thirds of those under 45 sup... st Ame opp... most bac... Repul

Polling Data

Chinese Return to the Waldorf, With $2 Billion

By DAVID BARBOZA OCTOBER 8, 2014 9:28 PM 2 Comments

Business Data

2014 Wins Above Replacement Leaders

1.	Clayton Kershaw (LAD)	8.0
2.	Mike Trout (LAA)	7.9
3.	Josh Donaldson (OAK)	7.4
4.	Corey Kluber (CLE)	7.4
5.	Adrian Beltre (TEX)	7.0
6.	Michael Brantley (CLE)	7.0
7.	Cole Hamels (PHI)	6.9
8.	Felix Hernandez (SEA)	6.7
9.	Jonathan Lucroy (MIL)	6.7
10.	Chris Sale (CHW)	6.6
	Alex Gordon (KCR)	6.6

Sports Data

Drugs Cause Most Fatal Allergic Reactions, Study Finds

By NICHOLAS BAKALAR OCTOBER 6, 2014 11:44 AM 18 Comments

The most common cause of fatal allergic reactions in the United States are medicines, especially antibiotics and radiocontrast agents used in imaging studies, a new analysis found.

Medical and Health Data

SCICURIOUS NEUROSCIENCE

Melatonin and the watery beginnings of sleep

The nightly behaviors of plankton may give insights to how sleep ...

Science Data

Spain Quarantines 3 More Over Ebola

By RAPHAEL MINDER OCT. 9, 2014

Descriptive Counts

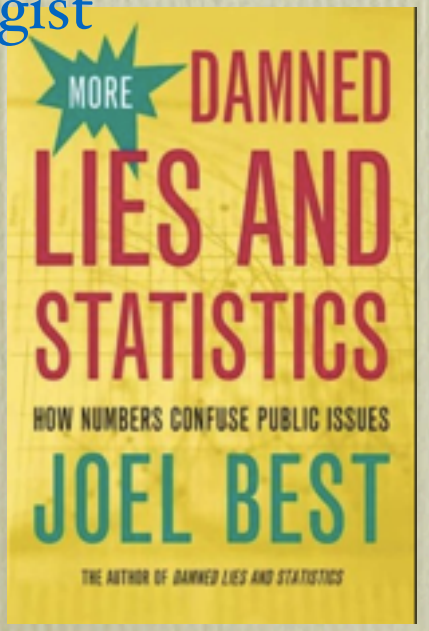
QR involves understanding *quantitative data*, and not, simply, manipulating numbers.

Data, Social Science, and *QR*

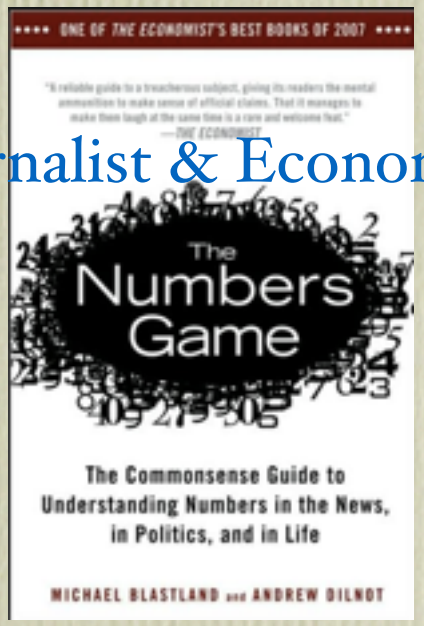
- *QuIRK* is rooted in the so
sciences, not in mathematic

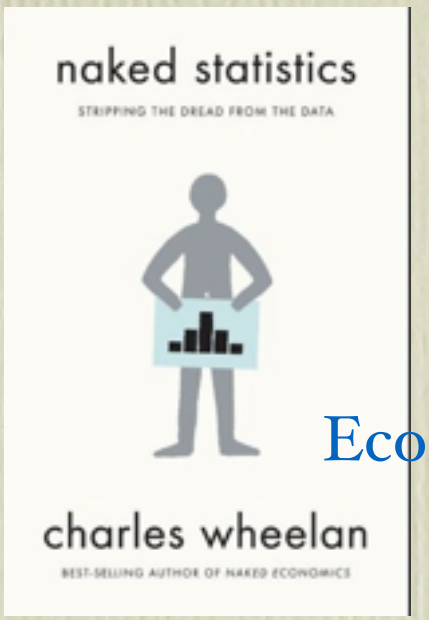
- The public face of *QR* is rooted in the social sciences.

Sociologist



Journalist & Economist

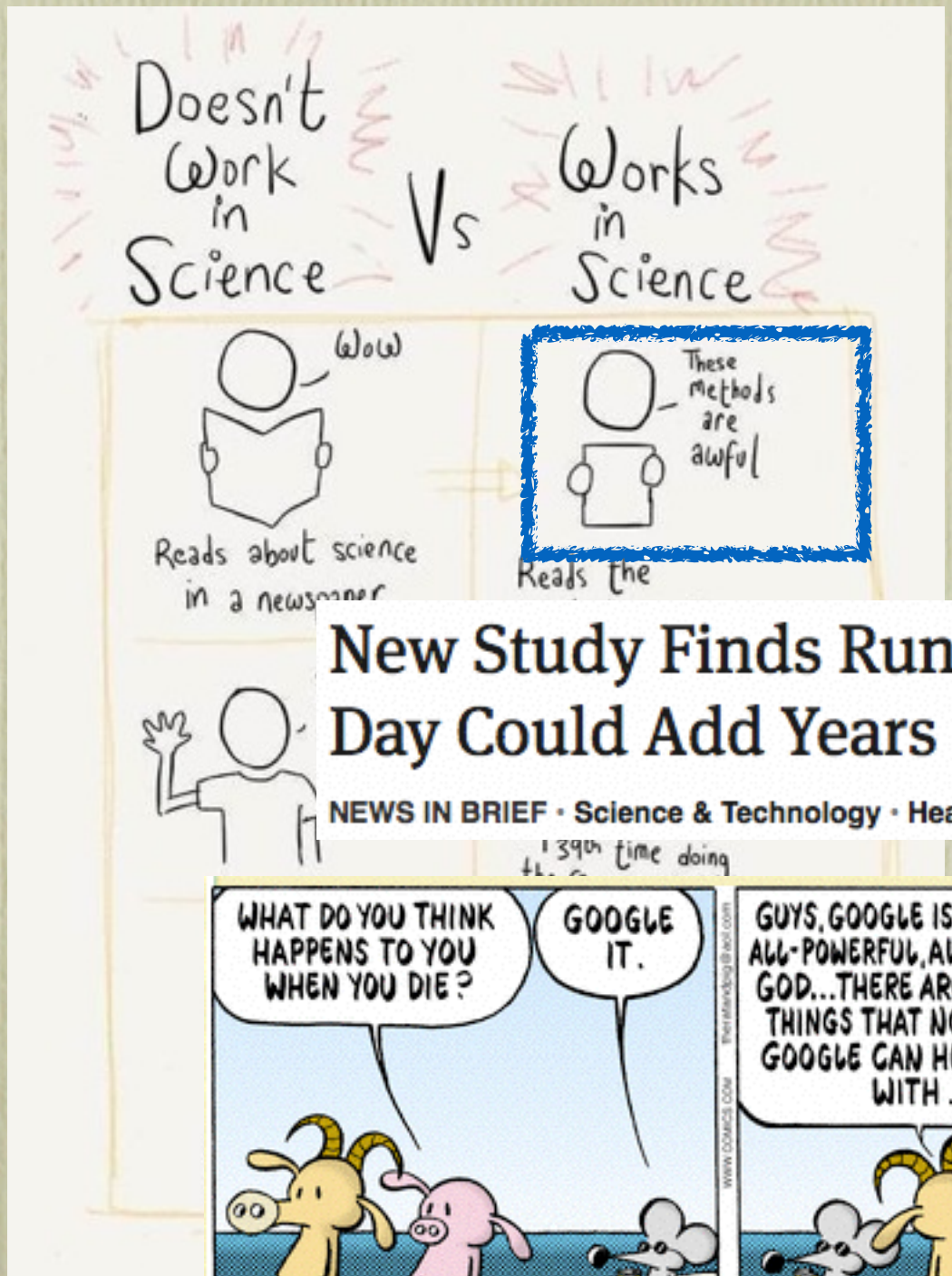




Economist

A primary concern in the social sciences (and in other sciences) is

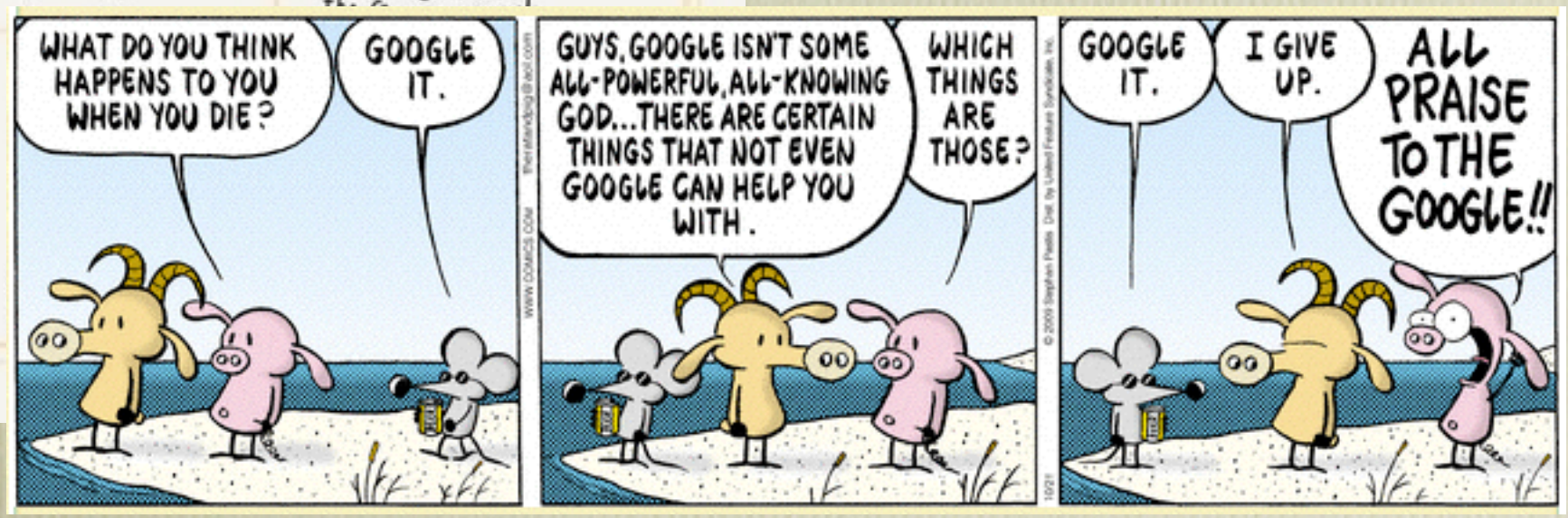
how the data were collected.



How the data were collected

New Study Finds Running For 20 Minutes Each Day Could Add Years Of Soreness To Life

NEWS IN BRIEF · Science & Technology · Health · Fitness · ISSUE 50·30 · Jul 31, 2014



Grounds for a *Data-Based Approach* to Assessing QR

- *10 QR Questions at the Ready* (Lutsky, 2004, *http://serc.carleton.edu/quirk/CarletonResources/10questions.html*)

10 QR Questions at the Ready:

- What do the numbers show?
- How representative is that?
- Compared to what?
- Is the outcome statistically significant?
- What's the effect size?
- Are the results those of a single study or of a literature?
- What's the research design (correlational or experimental)?
- How was the variable operationalized?
- Who's in the measurement sample?
- Controlling for what?

SERC.Carleton.edu

Grounds for a Data-Based Approach to Assessing QR

- *10 QR Questions at the Ready* (Lutsky, 2004, *http://serc.carleton.edu/quirk/CarletonResources/10questions.html*)
- *What Educated Citizens Should Know about Statistics and Probability* (Utts, 2003, The *American Statistician, 57(2)*, 74-79).
- *20 Tips for Interpreting Scientific Claims* (Sutherland, Spiegelhalter, & Burgman, 2013, *Nature*, *503*, 335-337).

Twenty tips for interpreting scientific claims

This list will help non-scientists to interrogate advisers and to grasp the limitations of evidence, say **William J. Sutherland**, **David Spiegelhalter** and **Mark A. Burgman**.

No measurement is exact. Practically all measurements have some error. If the measurement process were repeated, one might record a different result. In some cases, the measurement error might be large compared with real differences. Thus, if you are told that the economy grew by 0.13% last month, there is a moderate chance that it may actually have shrunk. Results should be presented with a precision that is appropriate for the associated error, to avoid implying an unjustified degree of accuracy.

Bias is rife. Experimental design or measuring devices may produce atypical results in a given direction. For example, determining voting behaviour by asking people on the street, at home or through the Internet will sample different proportions of the population, and all may give different results. Because studies that report 'statistically significant' results are more likely to be written up and published, the scientific literature tends to give an exaggerated picture of the magnitude of problems or the effectiveness of solutions. An experiment might be biased by expectations: participants provided with a treatment might assume that they will experience a difference and so might behave differently or report an effect. Researchers collecting the results can be influenced by knowing who received treatment. The ideal experiment is double-blind: neither the participants nor those collecting the data know who received what. This might be straightforward in drug trials, but it is impossible for many social studies. Confirmation bias arises when scientists find evidence for a favoured theory and then become insufficiently critical of their own results, or cease searching for contrary evidence.

Bigger is usually better for sample size. The average taken from a large number of observations will usually be more informative than the average taken from a smaller number of observations. That is, as we accumulate evidence, our knowledge improves. This is especially important when studies are clouded by substantial amounts of natural variation and measurement error. Thus, the effectiveness of a drug treatment will vary naturally between subjects. Its average efficacy can be more reliably and accurately estimated from a trial with tens of thousands of participants than from one with hundreds.

Correlation does not imply causation. It is tempting to assume that one pattern causes another. However, the correlation might be coincidental, or it might be a result of both patterns being caused by a third factor — a 'confounding' or 'lurking' variable. For example, ecologists at one time believed that poisonous algae were killing fish in estuaries; it turned out that the algae grew where fish died. The algae did not cause the deaths[2].

Regression to the mean can mislead. Extreme patterns in data are likely to be, at least in part, anomalies attributable to chance or error. The next count is likely to be less extreme. For example, if speed cameras are placed where there has been a spate of accidents, any reduction in the accident rate cannot be attributed to the camera; a reduction would probably have happened anyway.

Extrapolating beyond the data is risky. Patterns found within a given range do not necessarily apply outside that range. Thus, it is very difficult to predict the response of ecological systems to climate change, when the rate of change is faster than has been experienced in the evolutionary history of existing species, and when the weather extremes may be entirely new.

Beware the base-rate fallacy. The ability of an imperfect test to identify a condition depends upon the likelihood of that condition occurring (the base rate). For example, a person might have a blood test that is '99% accurate' for a rare disease and test positive, yet they might be unlikely to have the disease. If 10,001 people have the test, of whom just one has the disease, that person will almost certainly have a positive test, but so too will a further 100 people (1%) even though they do not have the disease. This type of calculation is valuable when considering any screening procedure, say for terrorists at airports.

Controls are important. A control group is dealt with in exactly the same way as the experimental group, except that the treatment is not applied. Without a control, it is difficult to determine whether a given treatment really had an effect. The control helps researchers to be reasonably sure that there are no confounding variables affecting the results. Sometimes people in trials report positive outcomes because of the context or the person providing the treatment, or even the colour of a tablet[3]. This underlies the importance of comparing outcomes with a control, such as a tablet without the active ingredient (a placebo).

Randomization avoids bias. Experiments should, wherever possible, allocate individuals or groups to interventions randomly. Comparing the educational achievement of children whose parents adopt a health programme with that of children of parents who do not is likely to suffer from bias (for example, better-educated families might be more likely to join the programme). A well-designed experiment would randomly select some parents to receive the programme while others do not.

Separate no effect from non-significance. The lack of a statistically significant result (say a P-value > 0.05) does not mean that there was no underlying effect: it means that no effect was detected. A small study may not have the power to detect a real difference. For example, tests of cotton and potato crops that were genetically modified to produce a toxin to protect them from damaging insects suggested that there were no adverse effects on beneficial insects such as pollinators. Yet none of the experiments had large enough sample sizes to detect impacts on beneficial species had there been any.

Effect size matters. Small responses are less likely to be detected. A study with many replicates might result in a statistically significant result but have a small effect size (and so, perhaps, be unimportant). The importance of an effect size is a biological, physical or social question, and not a statistical one. In the 1990s, the editor of the US journal *Epidemiology* asked authors to stop using statistical significance in submitted manuscripts because authors were routinely misinterpreting the meaning of significance tests, resulting in ineffective or misguided recommendations for public-health policy[4].

> *"The question to ask is: 'What am I not being told?'"*

Study relevance limits generalizations. The relevance of a study depends on how much the conditions under which it is done resemble the conditions of the issue under consideration. For example, there are limits to the generalizations that one can make from animal or laboratory experiments to humans.

Feelings influence risk perception. Broadly, risk can be thought of as the likelihood of an event occurring in some time frame, multiplied by the consequences should the event occur. People's risk perception is influenced disproportionately by many things, including the rarity of the event, how much control they believe they have, the adverseness of the outcomes, and whether the risk is voluntarily or not. For example, people in the United States underestimate the risks associated with having a handgun at home by 100-fold, and overestimate the risks of living close to a nuclear reactor by 10-fold[5].

Dependencies change the risks. It is possible to calculate the consequences of individual events, such as an extreme tide, heavy rainfall and key workers being absent. However, if the events are interrelated, (for example a storm causes a high tide, or heavy rain prevents workers from accessing the site) then the probability of their co-occurrence is much higher than might be expected[6]. The assurance by credit-rating agencies that groups of subprime mortgages had an exceedingly low risk of defaulting together was a major element in the 2008 collapse of the credit markets.

Data can be dredged or cherry picked. Evidence can be arranged to support one point of view. To interpret an apparent association between consumption of a food during pregnancy and subsequent asthma in offspring[7], one would need to know whether the authors set out to test this sole hypothesis, or happened across this finding in a huge data set. By contrast, the evidence for the Higgs boson specifically accounted for how hard researchers had to look for it — the 'look-elsewhere effect'. The question to ask is: 'What am I not being told?'

Extreme measurements may mislead. Any collation of measures (the effectiveness of a given school, say) will show variability owing to differences in innate ability (teacher competence), plus sampling (children might by chance be an atypical sample with complications), plus bias (the school might be in an area where people are unusually unhealthy), plus measurement error (outcomes might be measured in different ways for different schools). However, the resulting variation is typically interpreted only as differences in innate ability, ignoring the other sources. This becomes problematic with statements describing an extreme outcome ('the pass rate doubled') or comparing the magnitude of the extreme with the mean ('the pass rate in school x is three times the national average') or the range ('there is an x-fold difference between the highest- and lowest-performing schools'). League tables, in particular, are rarely reliable summaries of performance. ■

William J. Sutherland *is professor of conservation biology in the Department of Zoology, University of Cambridge, UK.* **David Spiegelhalter** *is at the Centre for Mathematical Sciences, University of Cambridge.* **Mark Burgman** *is at the Centre of Excellence for Biosecurity Risk Analysis, School of Botany, University of Melbourne, Parkville, Australia.*
e-mail: wjs32@cam.ac.uk

1. Doubleday, R. & Wilson, J. *Nature* **485**, 301–302 (2012).
2. Boruck, M. E., Stow, C. A. & Reckhow, K. H. *J. Water Res. Plan. Manage.* **129**, 271–282 (2003).
3. Hróbjartsson, E. C.-B. *Med. J.* **4**, 196–200 (1974).
4. Miller, R. B. & Anderson, M. J. *Fish. Res.* **70**, 357–407 (2004).
5. Marvier, M. *Risk Anal.* **12**, 1119–1124 (2002).
6. Fidler, F., Cumming, G., Burgman, M., Thomason, N. *J. Socio-Economics* **33**, 615–630 (2004).
7. Fischhoff, B., Slovic, P. & Lichtenstein, S. *Am. Stat.* **16**, 240–255 (1982).
8. Billinton, R. & Allan, R. N. *Reliability Evaluation of Power Systems* (Plenum, 1984).
9. Maslova, E., Halldorsson, T. I., Strøm, M., Olsen, S. F. *J. Nutr. Sci.* **1**, e5 (2012).

Grounds for a Data-Based Approach to Assessing QR

- *10 QR Questions at the Ready* (Lutsky, 2004, *http://serc.carleton.edu/quirk/CarletonResources/10questions.html*)
- *What Educated Citizens Should Know about Statistics and Probability* (Utts, 2003, The *American Statistician, 57(2)*, 74-79).
- *20 Tips for Interpreting Scientific Claims* (Sutherland, Spiegelhalter, & Burgman, 2013, *Nature*, *503*, 335-337).
- *Assessing Statistics and Research Methodology in the MCAT Exam* (Zhao, Dowd, & Searcy, 2012, *Chance*, *25(3)*, 11-17).

"The proposed changes to the assessment of *statistical thinking and quantitative reasoning skills* by the MCT2015 exam recognize the significance of these pre-medical knowledge and skills."

-Zhao, Dowd, & Searcy (2012)

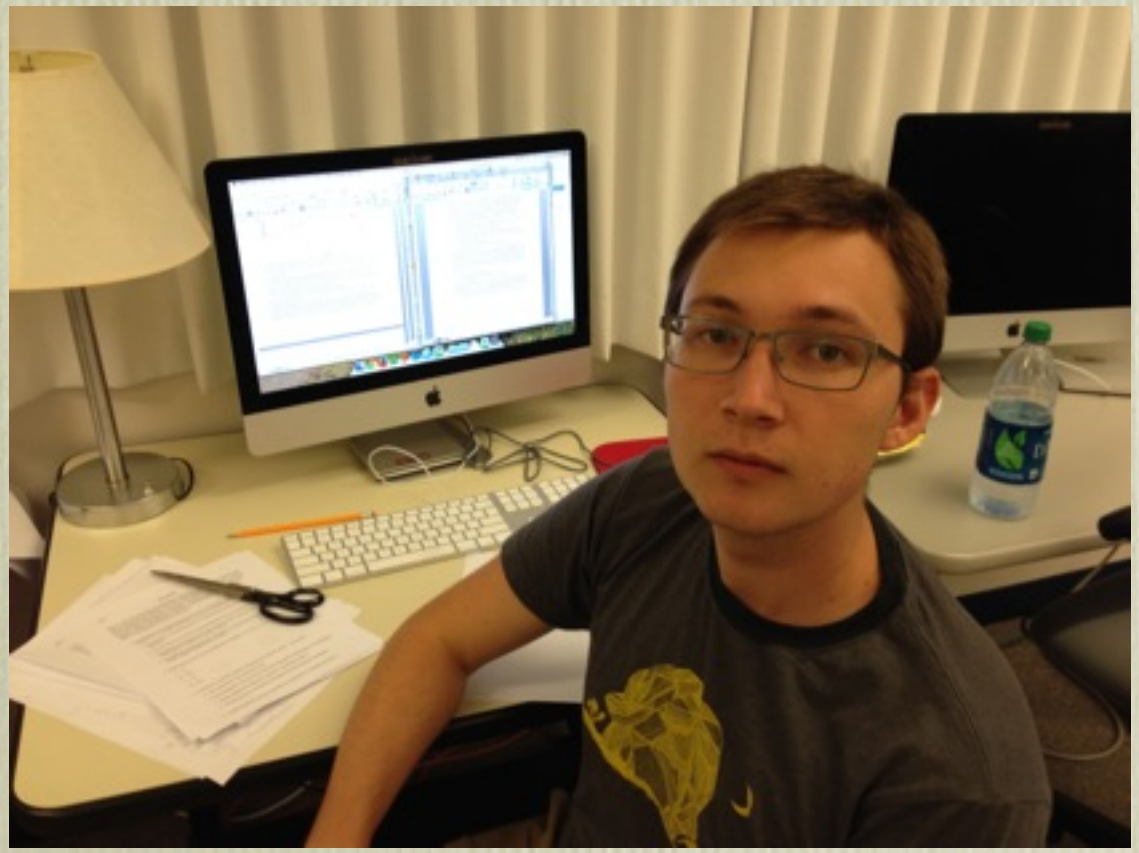
Table 4—Sample Tasks for SIRS 3 and SIRS 4 of the Scientific Inquiry and Reasoning Skills for the MCAT2015 Exam

SIRS 3: Reasoning About the Design and Execution of Research
• Identify foundational aspects of research design (e.g., experimental vs. non-experimental design, independent and dependent variables)
• Critique different aspects of a research design (e.g., identify sources of potential bias, confounds, adequacy of sample)
• Evaluate research designs to determine if conclusions based on the research study are appropriate
• Recognize ethical issues inherent in research investigations
• Make predictions about expected results based on the features of a research design
SIRS 4: Data-Based and Statistical Reasoning
• Use data to describe phenomena in the natural world and/or to describe results of a research study
• Use descriptive statistics to summarize data (e.g., mean, median, and standard deviation) and relate it to the population (e.g., standard error)
• Interpret data or patterns in data to draw conclusions or evaluate the conclusions made at the end of a research study (e.g., notice whether conclusions logically follow based on the data presented)
• Interpret data, or patterns in data, to make predictions
• Use statistics to answer research questions and evaluate the strength of the evidence provided in support of given hypotheses
• Interpret data patterns presented in tables, figures, and graphs (e.g., histograms, scatter plots) to interpret results, make comparisons, and draw conclusions

This is a nonexhaustive, partial list for illustrative purposes. The information is current as of the publication date, but refinements to SIRS could be made during the continued development of the new MCAT exam.

Developing the *QUIC*:
A new measure of QR

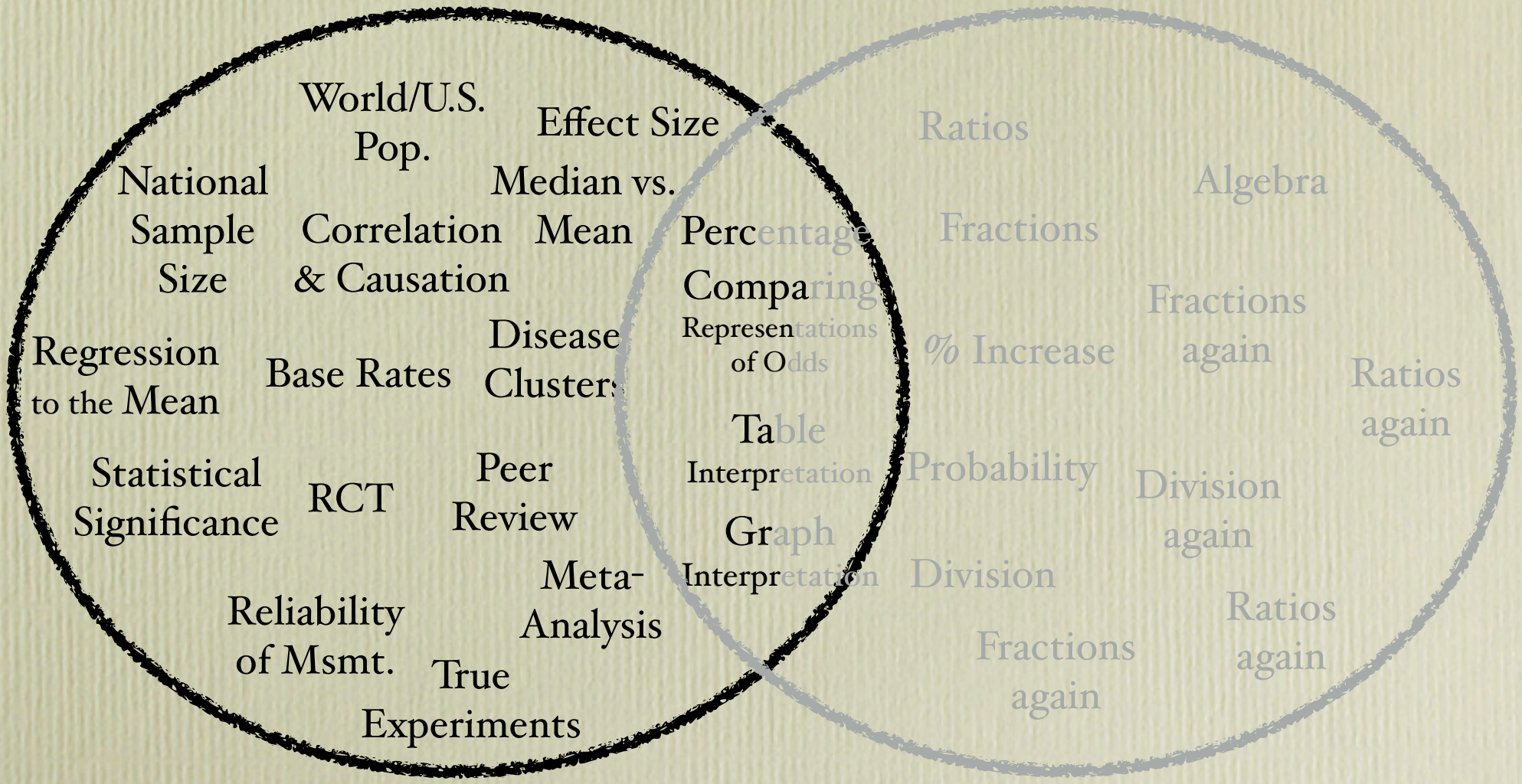
- 25-item multiple choice test.
- *Students* Quinn Batten & Sam Hayward.



Developing the *QUIC*:
A new measure of QR

- Reviewed existing measures.
- Reread QR books and articles, coding for prominent topics.
- *Generated items to address those topics.*
- Tested versions of the *QUIC* on students attending Carleton's summer academic institutes.

A *QUIC* Topic Overview:



- Statistical & Methodological
  - Relevant to Research & Data-Based Findings

QLRA 2012

QUIC Item

- A defining factor of a true experiment, as opposed to a correlational study, is:

a. The presence of a control group.

b. Evidence for the reliability of the measures used.

c. Random assignment to the different conditions of the study.

d. Completion of a study in a laboratory.

QUIC Item

- Which of the following represents the approximate percentage of the world's population living in the United States?

a. 4%

b. 12%

c. 20%

d. 33%

QUIC Item

- A charity reports that it received 1,000 donations during a recent campaign and that the average donation was $500. You wonder whether that average figure was affected by a few quite large contributions. What *other* number would best help you evaluate that?

a. The median contribution.

b. The modal contribution.

c. The mean contribution.

d. The total sum of contributions.

QUIC Item

- A new disease, Vacheritis, has a 1% probability of occurring in the population. A random sample of individuals is given a test for Vacheritis. This test gives a true positive (i.e., detects the disease) 80% of the time in those who actually have Vacheritis and gives a false positive result 10% of the time for people who get tested but don't have the disease. Someone you know was in that sample and tested positive for Vacheritis. How concerned should your acquaintance be that he or she has Vacheritis?

a. The chances are this person does not have the disease.

b. There is a moderate (near 50/50) chance the person has the disease.

c. The person probably has the disease, but there is till a good chance he or she doesn't.

d. The chances are high the person has the disease.

QUIC Item

- A polling firm is hired to conduct a nationally representative survey in the United States to determine Americans' likely preference in a political contest. If the firm uses random sampling procedures, approximately how many individuals are needed to yield reasonably sound findings (e.g., a margin of error of ± 4 percentage points). Again, the answer to this question does not require any calculations. You should have a sense of the approximate answer.

a. 100

b. 1,000

c. 10,000

d. 100,000

QUIC Item

- A study of marathon runners found that there is a relationship between finishing times and how much broccoli the runners' reported consuming the week before the marathon. Runners eating more broccoli tended to have shorter finishing times. Which headline below would best represent this finding?

a. EAT BROCCOLI, RUN FASTER!

b. EATING BROCCOLI NEGATIVELY IMPACTS MARATHON SPEED!

c. EATING BROCCOLI POSITIVELY IMPACTS MARATHON SPEED!

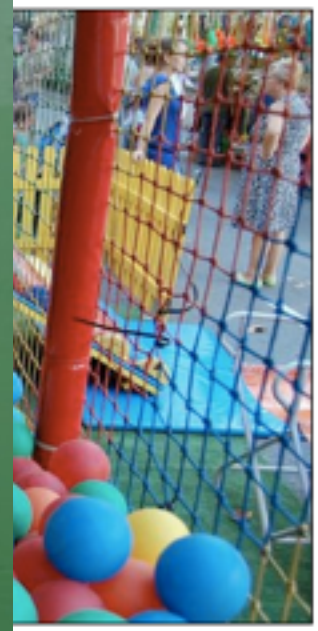
d. EATING BROCCOLI ASSOCIATED WITH MARATHON SPEED!

A few more *QUIC notes:*

- You can get the *QUIC* at: http://apps.carleton.edu/quirk/assets/QUIC.1.1.pdf
- We're in development and would appreciate any suggestions, feedback, and data. My e-mail is *nlutsky@carleton.edu*
- *Teach to the QUIC (or to something like it)!*
- *QUIC = "Quantitative Understanding Instrument from Carleton"*

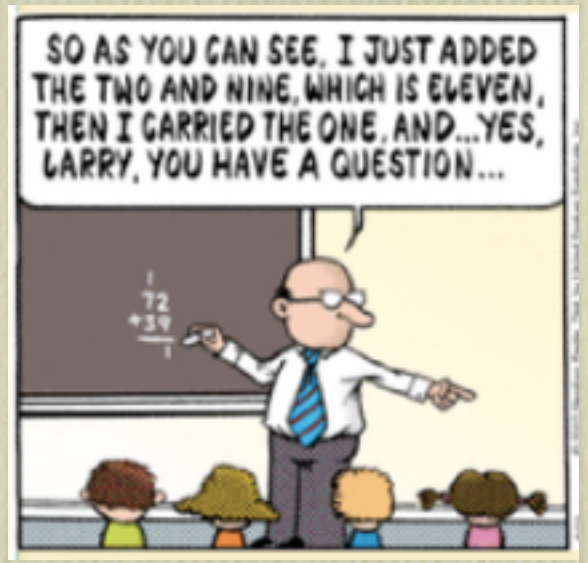
In conclusion...

r Citizen



has multiple exits in case seniors become disoriented or scared. "We have a slide wide enough to accommodate wheelchairs, and on Saturdays, Ronald himself stops by to make balloon animals and just talk to the old folks. They like talking to Ronald." McDonald's

Questions?


SO AS YOU CAN SEE, I JUST ADDED THE TWO AND NINE, WHICH IS ELEVEN, THEN I CARRIED THE ONE, AND...YES, LARRY, YOU HAVE A QUESTION...
1
72
+39
1