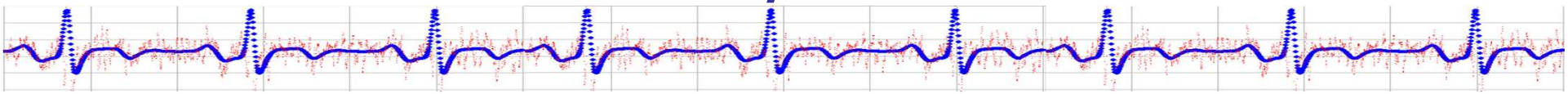
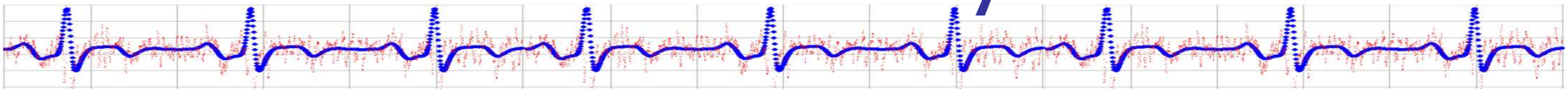
Empirical Research Methods in Information Science

IS4800 / CS6350



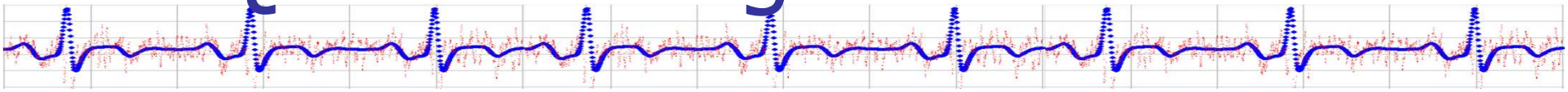
Lecture 3

Overview for today



- Quick reading assessment
- Doing background research
- Sample research plan
- Study recruitment/samples
- Study ethics

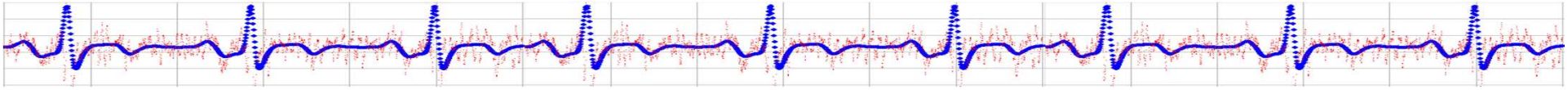
Quick reading assessment



Closed book, closed computer

10 minutes

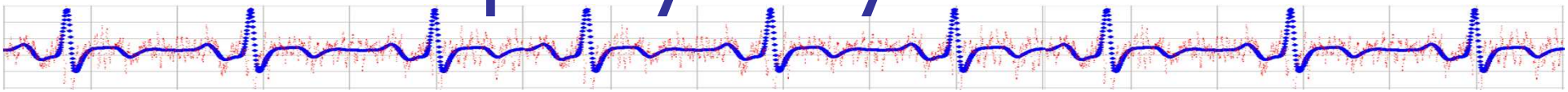
Homework for Wed



- Reading:
  Ethnography (B&A pp 242-247, Fetterman, Klasnja, Chen)
  (links on syllabus)

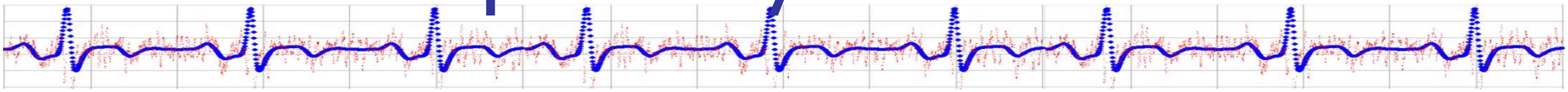
- Postponing until later: Python data graphing introduction

TA help! Sydney Hubbell



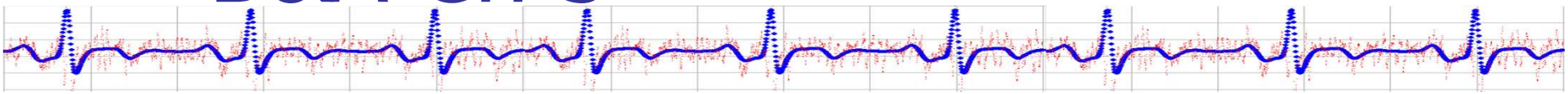
- 5th year CS/IS combined major, with a business minor
- Took IS4800 last spring
- Completed IS capstone project in Information Privacy Awareness
- Taking classes part-time and working part-time at her previous co-op; working in HCI post-graduation
- Will be scheduling office hours

TA help! Aditya Ponnada



- 4th year Personal Health Informatics PhD student
- Very interested in research methods, especially measuring behavior
- May help out a bit as class goes on

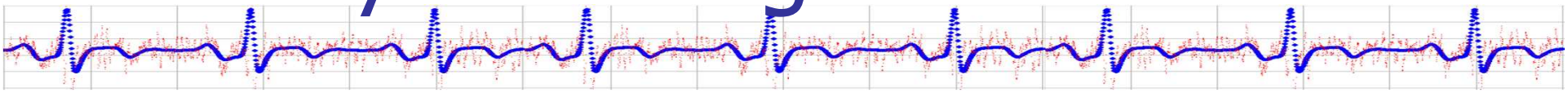
B&A Ch 3



Doing Background Research

(Lit review part … developing research questions later)

Why do background research?

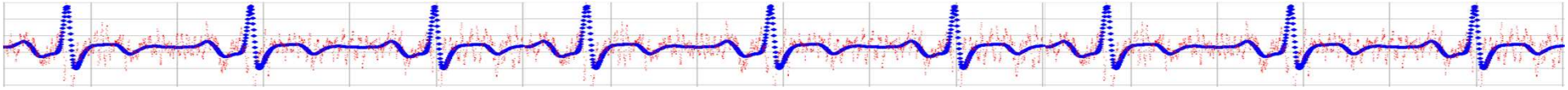


Most important take-away from this chapter…

Has anyone?

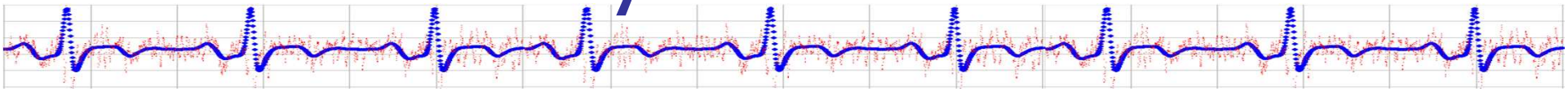
- Answered the question before (may not be definitive, or exact match)
- Done a similar study before (lessons learned)
- Used methods you could reuse?

Sources of research information



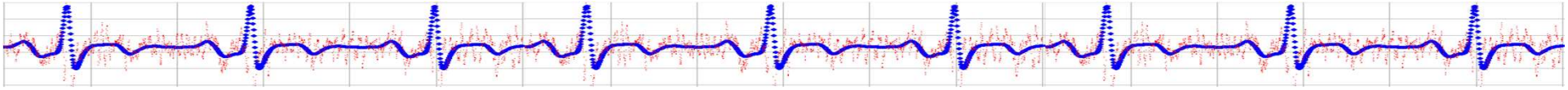
- Primary vs. secondary sources
  - A *primary source* includes a full report of a research study, including methodological details
    - Primary sources are preferred
  - A *secondary source* summarizes information from a primary source
    - These should be used sparingly, because they may be
      - incomplete
      - biased
      - inaccurate

Which is better?
How can you tell?



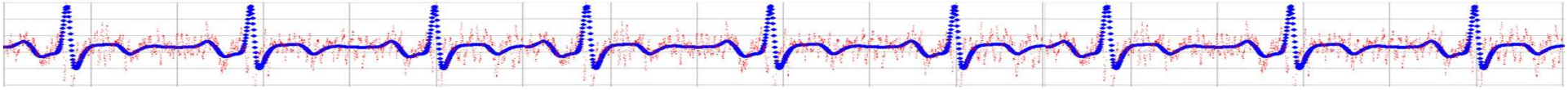
- Magazine – substantive (SciAm)
- Journal – refereed (JAMA)
- Sensationalist Tabloid (Enquirer)
- Book – single author
- Newspaper (Globe)
- Magazine – popular (Wired)
- Conference paper
- Book – edited collection
- Open source / Online Journal

Sources of research information



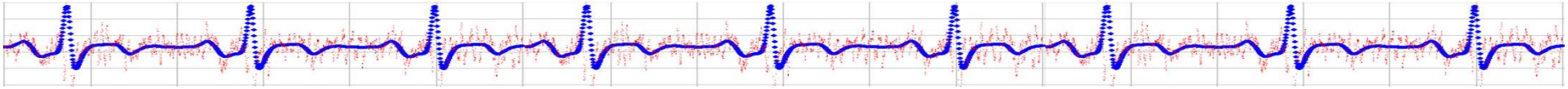
- Books (generalized textbooks, specialized)
  - Anthologies assemble papers that an editor feels are important in a field
    - May not represent the entire field; editor's bias
  - Books are most useful in early literature search
  - Books that provide summaries are secondary sources
  - Books should be used with caution because they may not undergo rigorous review, and information may not be up to date

Sources of research information



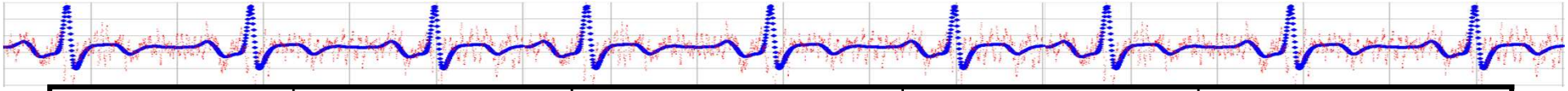
- Scholarly Journals
  - Provide current research and theoretical thinking
  - Papers submitted to a *refereed journal* undergo peer review
  - Papers submitted to a *nonrefereed journal* do not undergo peer review
    - You should prefer referred over nonrefereed sources
  - You can evaluate the quality of a journal by
    - Consulting *Journals in Psychology*
    - Consulting the *Social Science Citations Index*
    - Using the method of authority (asking a trusted authority about the quality of a journal)

Sources of research information



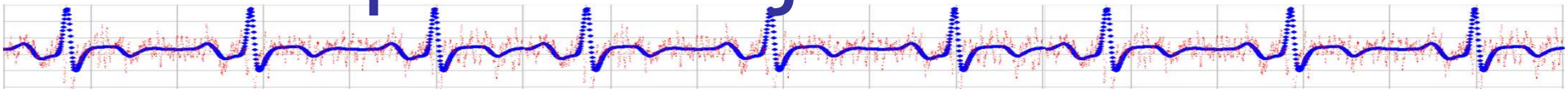
- Conventions and professional meetings
  - Provide the most up-to-date information
  - Researchers from many areas present research findings at conventions and meetings
  - Research may be presented in a paper session (oral session) or a poster session
  - Papers presented at an oral session have a time limit, and methodological details may not be reported
  - Attending a paper or poster session has two advantages
    - Information is at the frontiers of science
    - You can meet others in your field and exchange ideas
  - A drawback is that conventions can be expensive to attend

Differences Between Published Sources



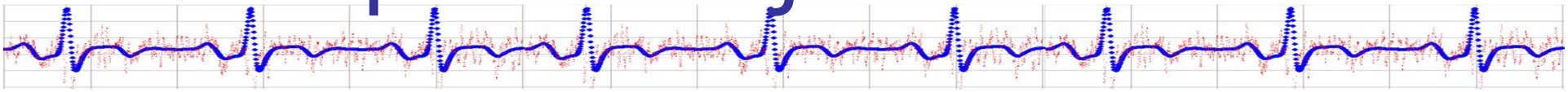
Example?	Scholarly	Substantive	Popular	Sensational
Appearance	Sober and serious	Attractive, with photographs	Attractive with many photos	Newspaper format
Reference Citations	Always provided	Sometimes cited	Rarely provided	Obscure references
Author	Scholar in the field	Scholar, editorial staff, freelance	Wide range of authors	Wide range
Language	Geared to scholars	For educated, no specialty	Simple, for less educated	Elementary for gullible audience
Content	Original research	No original research	Sources mentioned, may be obscure	Pseudoscientific sources
Publisher	Many by professional organizations	Commercial or professional organization	Commercial to entertain	Commercial to arouse curiosity

Respect or reject?



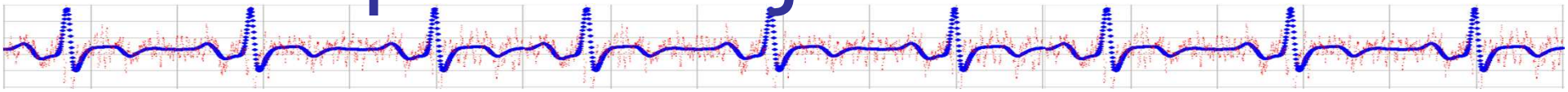
- **King, A.**, **Haskell, W.**, **Taylor, C.**, **Kraemer, H.**, and **DeBusk, R.**: Group vs. home-based exercise training in healthy older men and women. JAMA 266, 1991) 1535-1542.
  - King: Stanford Prof.

Respect or reject?



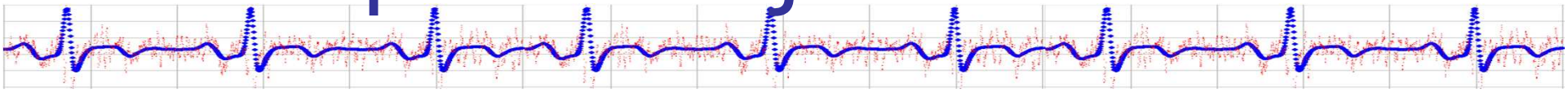
- McClusky, M. *Review: Canon Powershot A640 Digital Camera,* Wired Blog, 1/12/07.

Respect or reject?



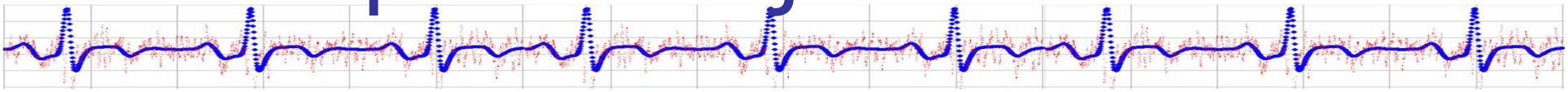
- Nass, C. and Gong, L.: Is Maximization or Consistency the More Social? The Case of Synthesized Voices and Faces. Proceedings of CHI '00 (2000)
  - Prof. Stanford

Respect or reject?



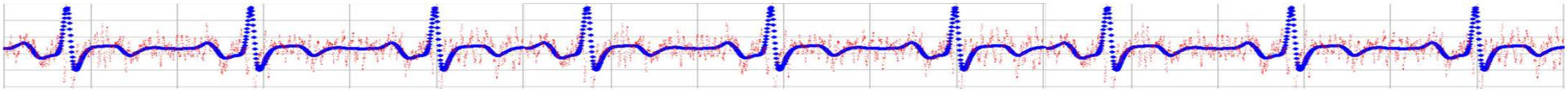
- Dorfman, P. *Measurement and Meaning of Recreation Satisfaction: A Case Study in Camping*, Environment and Behavior, Vol. 11, No. 4, 483-510 (1979)
  - Prof at U. NM

Respect or reject?

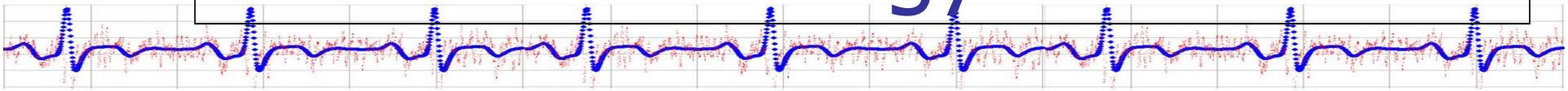


- Grundel, C. Schneider-Hufschmidt, M. *A direct manipulation user interface for the control of communication processes - making call handling manageable*, Proceedings HCI International '99.

How do you do a literature review?

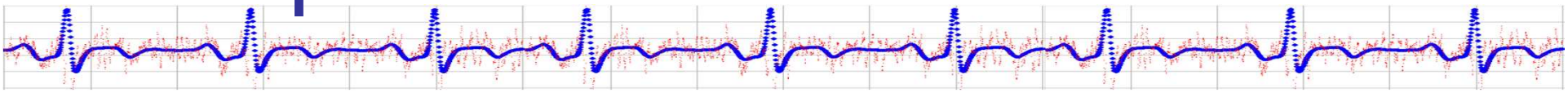


Library research: The basic strategy



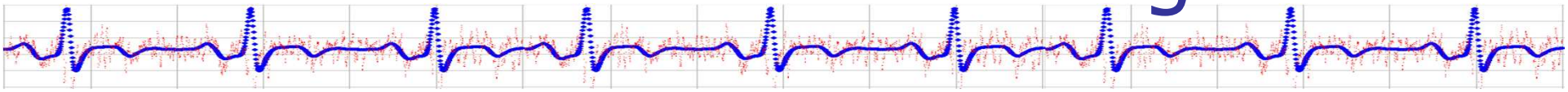
- Find a relevant research article, using the reference section from a textbook or computer database
- Use the reference section from the article to find other articles
- Repeat the first two steps for each article identified until you can find no more
- Use one of the many indexes (e.g., **portal.acm.org, IEEE Digital Library**) to locate more recent articles
- Repeat the entire process as you find more recent articles

Tip



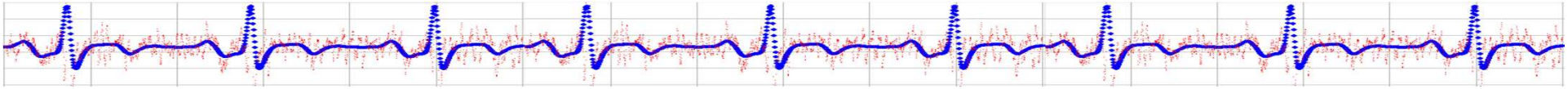
- To find the "seminal" articles in a field, look for number of citations
  - e.g., Web Of Science, Google Scholar

Some literature searching tools



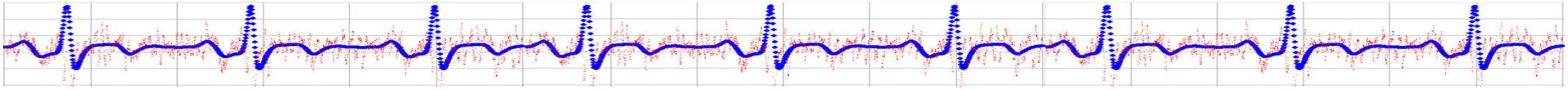
- *ACM Digital Library*
  - ACM affiliated journals, conference proceedings, etc.
- *IEEE/Explore*
  - IEEE affiliated journals, conference proceedings, etc.
- *PsycINFO*
  - Indexes over 1,800 journals in 25 languages (1872-)
- *Google Scholar (scholar.google.com)*
  - Most CS conference & journal articles are on the web
- *Citeseer (citeseer.ist.psu.edu*)
- And for the PHI crowd: PubMed

Example from a *PsycINFO* search



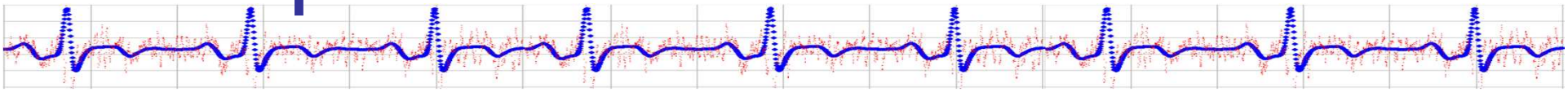
TI: Title
Students' perceptions of sexual harassment: Is it harassment only if the offender is a man and the victim is a woman?
AU: Author
Runtz, Marsha G; O'Donnell, Cedar W
AF: Author Affiliation
U Victoria, Dept of Psychology, Victoria, BC, Canada [Runtz]; U Victoria, Victoria, BC, Canada [O'Donnell]
CI: Contact Individual
Runtz, Marsha G, U Victoria, Dept of Psychology, P.O. Box 3050, Victoria, BC, Canada, V8W 3P5
SO: Source
Journal of Applied Social Psychology. Vol 33(5), May 2003, pp. 963-982
IS: ISSN
0021-9029
PB: Publisher
US: VH Winston & Son, [URL:http://www.bellpub.com/]
AB: Abstract
The present study investigated 3 potential sources of variability in university students' perceptions of sexual harassment in hypothetical professor-student scenarios: raters' gender, gender of the professor and student, and rater's own sexual harassment. Participants were most likely to identify the interactions as harassment when they involved a male offender and a female victim. They were less likely to label the behaviors as harassment when they occurred between members of the same gender or between a female professor and a male student. Women were more open to viewing the scenarios as harassment and men were unlikely to view the interactions between a female professor and a male student as harassment. Personal history of sexual harassment did not influence participants' perceptions. (PsycINFO Database Record (c) 2003 APA, all rights reserved) (journal abstract)

Example from a *PsycINFO* search (continued)



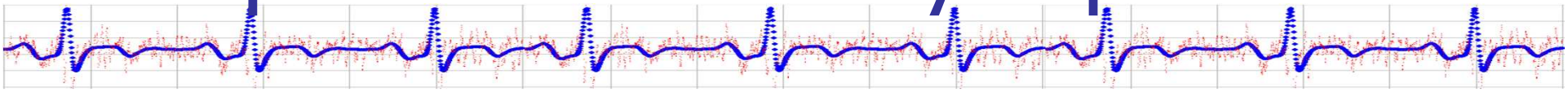
PT: Publication Type
Peer Reviewed Journal; Empirical Study; Journal Article
FV: Format Availability
Print
GN: Genre
article
PO: Population
Human; Male; Female; Adolescence (13-17 yrs); Adulthood (18 yrs & older); Young Adulthood (18-29 yrs); Thirties (30-39 yrs); Middle Age (40-64 yrs)
LO: Location
Canada
DE: Descriptors
*Human Sex Differences; *Sexual Harassment; *Social Perception; *Student Attitudes; *Teacher Student Interaction; College Students; College Teachers; Individual Differences
ID: Identifiers
sexual harassment; student perceptions; gender; professor-student interactions; personal harassment history; social perception; college professors; college students
CL: Classification
3230 Behavior Disorders & Antisocial Behavior; 3560 Classroom Dynamics & Student Adjustment & Attitudes
NR: Number of References
36 reference(s) present, 36 reference(s) displayed
UD: Update
20030818
AN: Accession Number
2003-99700-005

Tip



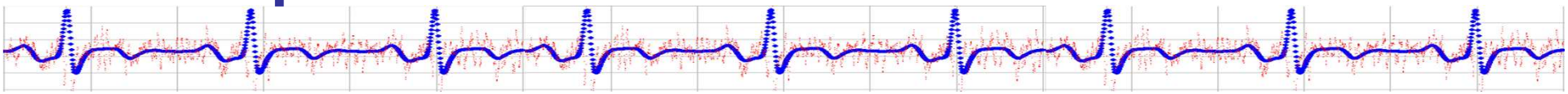
- Stay organized with Endnote!
- "A guide to getting started with the EndNote citation management program"
  http://subjectguides.lib.neu.edu/c.php?g=709685&p=5042991

Standard format for quantitative study report



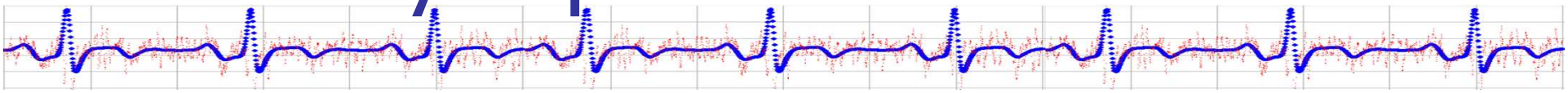
- Introduction
- Methods
- Results
- Discussion

How do you tell if a study report is credible?



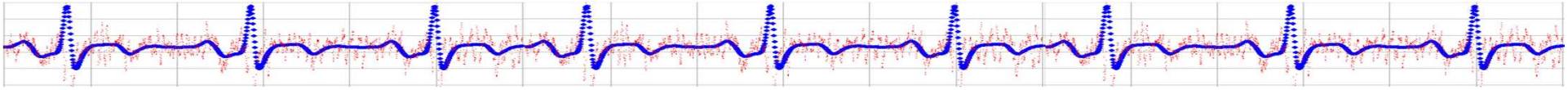
What happened with cold fusion?

Study report evaluation



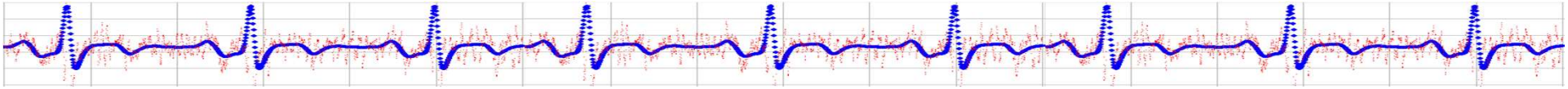
- Publication venue
- Author & affiliation
- Introduction
- Methods
- Results
- Discussion

Preliminary analysis of an article



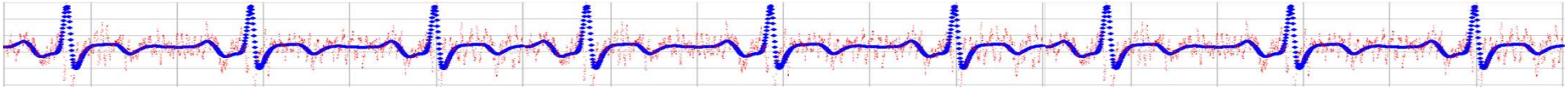
- What are the author's credentials and institutional affiliation?
- Has the author's work been cited by others?
- Is the publication current or dated?
- Is the work published by a publisher of professional books or journals?
- Is the work peer reviewed?
- Is the source a scholarly journal?

Evaluating a research article: The introduction



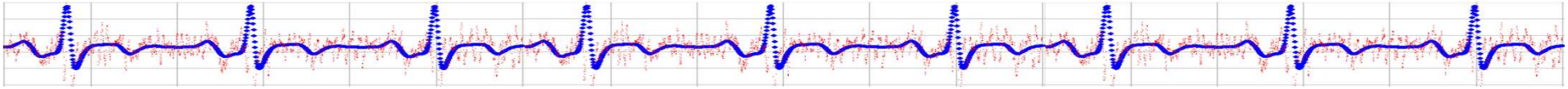
- Has relevant research been adequately reviewed?
- Are assertions supported with the appropriate citations?
- Are the purposes of the study clearly stated?
- Are the hypotheses clearly stated, and do they flow logically from the information in the introduction?

Evaluating a research article: The method section



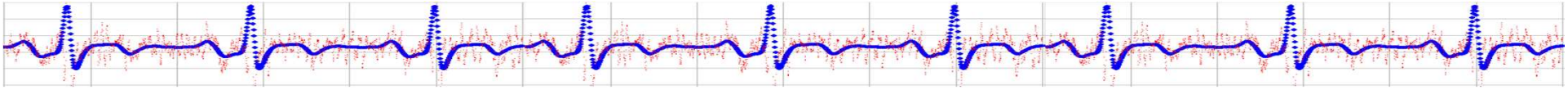
- Was the nature of the subject sample specified?
- Does the design of the study allow an adequate test of the hypotheses?
- Are there any methodological flaws that might affect the validity of the results?
- Is sufficient detail presented to allow one to replicate the study?

Evaluating a research article: The results section



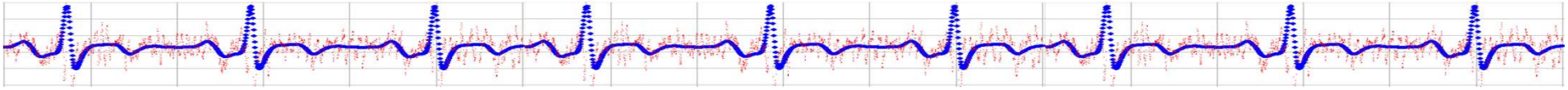
- Did the statistically significant effects support or refute the hypotheses?
- Are the differences reported large or small?
- Were the appropriate statistics used?
- Do the tables, figures, and text match?

Evaluating a research article: The discussion section



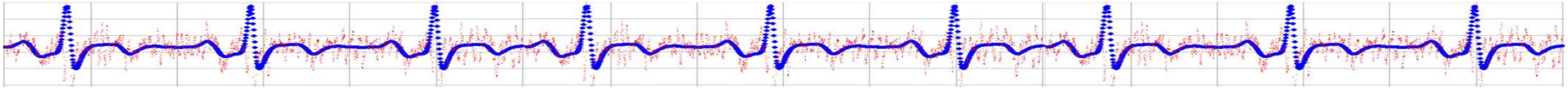
- Do the conclusions presented match the results reported?
- If the author speculates about implications of results, does he or she stray too far from the results reported?
- How well do the results mesh with existing theory and empirical data?
- Does the author point the way to directions for future research?

Factors affecting the quality of research information



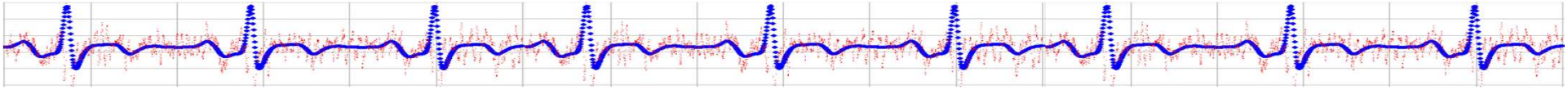
- Statistical Significance
  - Journals typically do not publish findings that do not meet the minimum .05 level of statistical significance
  - *File drawer* phenomenon: Findings that don't reach significance at .05 end up in the file drawer
    - If 100 articles about a phenomena are studied…
    - How to prevent?
  - "Effect Size" also important when interpreting significance – more on this later

Factors affecting the quality of research information



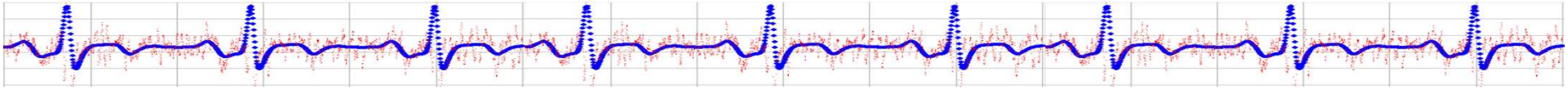
- Consistency with prior knowledge
  - Research results are expected to build on what came before
  - Results that don't fit within current thinking are suspect
  - Anomalous findings are treated with suspicion and may not get published in mainstream journals
  - An anomalous finding may be important, but ignored

Factors affecting the quality of research information



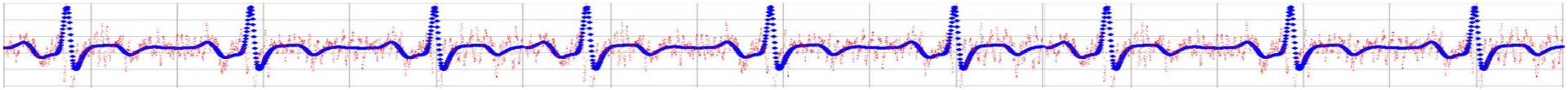
- Significance of the contribution
  - Do findings advance knowledge in science?
  - Currently, many journals will not publish single experiments
    - Require a series of experiments
    - This requirement may delay publication of important results until all studies are completed
- Editorial policy
  - A journal editor may set a policy that favors certain research topics and excludes others
  - An editor's choice of reviewers may be biased

Factors affecting the quality of research information



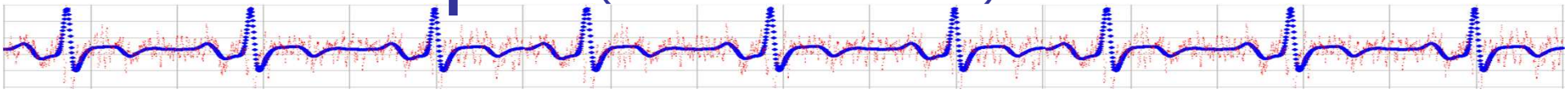
- Peer review
  - Articles submitted for publication are reviewed by experts in the field (usually *blind reviewers*)
  - Intended to ensure quality research is published
  - The process is time-honored, but has problems
    - Personal feelings of reviewers can bias review
    - Agreement with reviewer's view enhances chances of publication
    - Low rates of inter-reviewer agreement

Peer review: Warning! Humans involved!



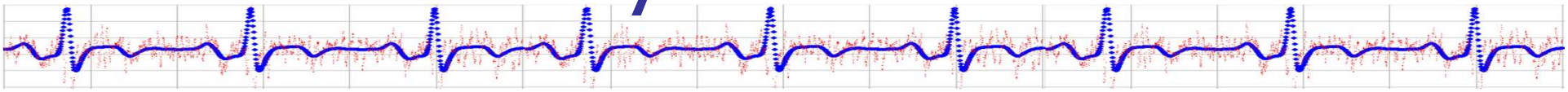
- Discussion at Interactive, Mobile, Wearable and Ubiquitous Technologies (IMWUT) editors meeting
- You have to review these? (single spaced, dense). Your frame of mind?

Example (break into teams of 3)



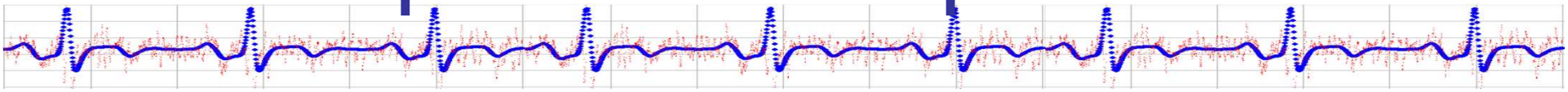
- 15 minutes
  - Identify an article from one of these sources that cites a research result:
    - The New York Times
    - Newsweek
    - Wired
  - Track down the primary source.
    - What kinds of data and statistics are reported?
    - Do the author's conclusions logically follow from the data presented?
    - Is the primary article credible?
- Be prepared for a 2-minute oral report on

Take-aways



- The scientific method is the best way we know to make valid and reliable claims about the physical world
- The first step in doing any research is going to the library (virtually)

Sample research plan

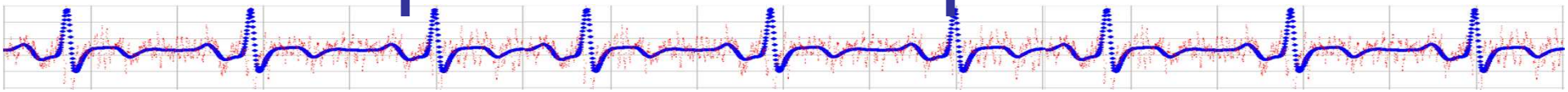


Embodied Conversational Agents to Promote Health Literacy for Older Adults

Sample research plan

- Health literacy
- Older adults
- Diabetes

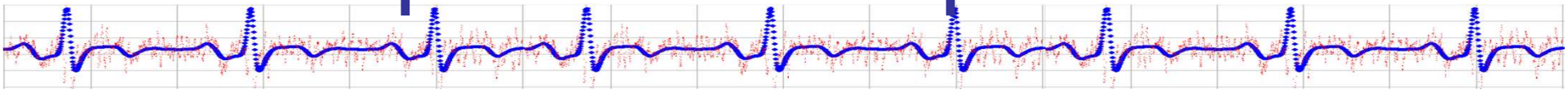
Sample research plan



Hypotheses

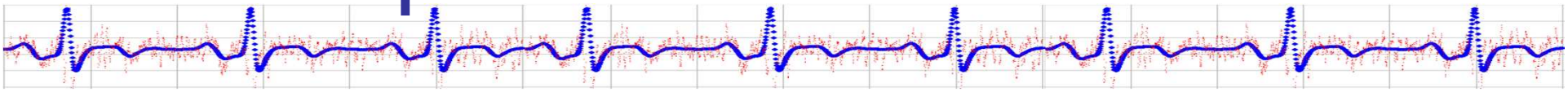
- Agents will lead to greater knowledge gains (regarding diabetes) compared to "standard of care" (brochures)
- Ditto for glycemic control
- Ditto for patient satisfaction

Sample research plan



- What are the units of study?
- What are the variables?
- How measured?
- What are the time points of interest?

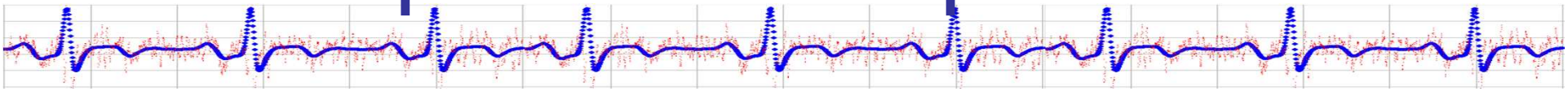
Sample Research Plan



Embodied Conversational Agents to Promote Health Literacy for Older Adults

A. SPECIFIC AIMS
B. BACKGROUND AND SIGNIFICANCE
C. PRELIMINARY STUDIES
D. RESEARCH DESIGN AND METHODS
E. HUMAN SUBJECTS

Sample research plan



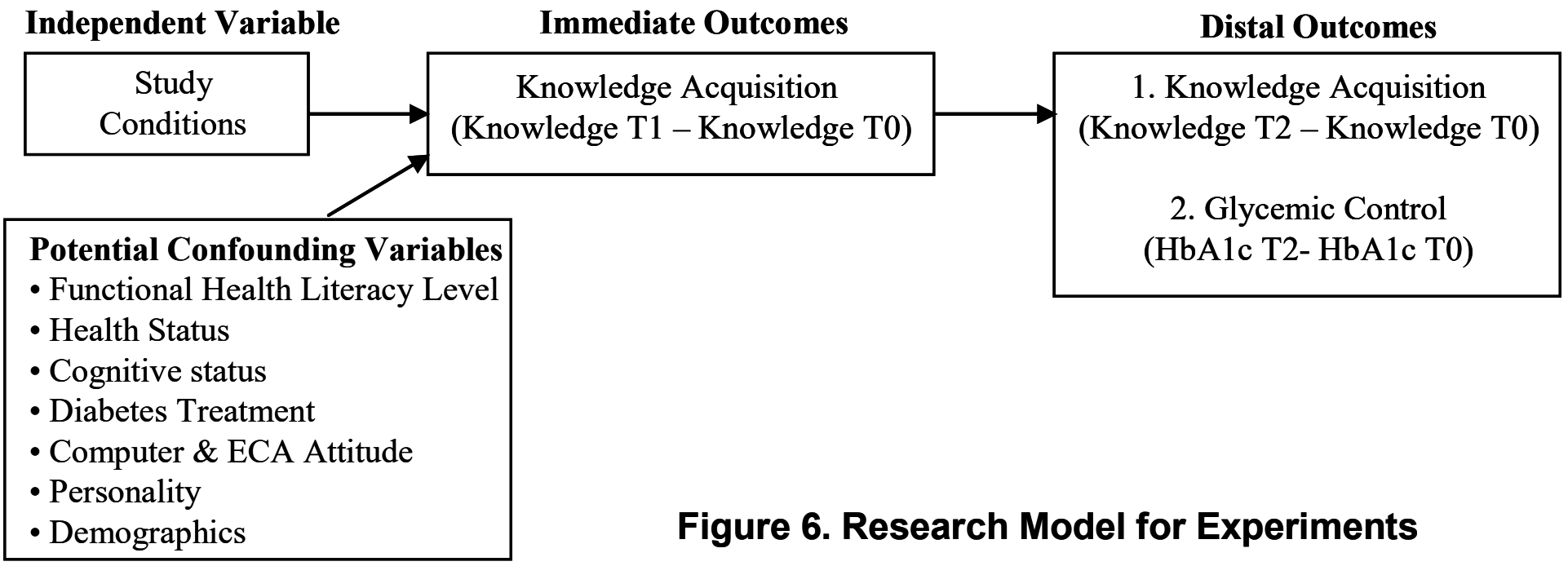
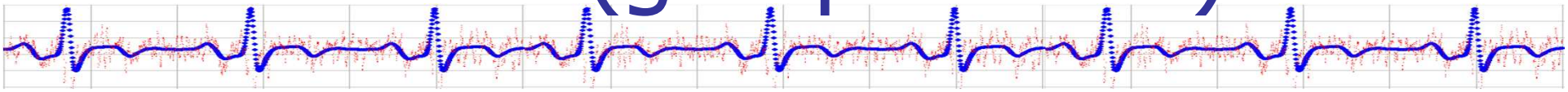


Figure 6. Research Model for Experiments

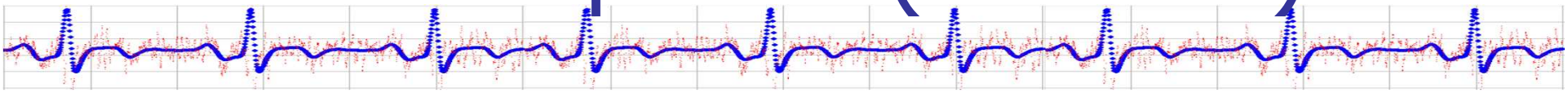
Exercise (groups of 3-4)



Describe how knowledge acquired from conducting the study specified in the sample research plan meets (or does not meet) the criteria for "scientific explanations"

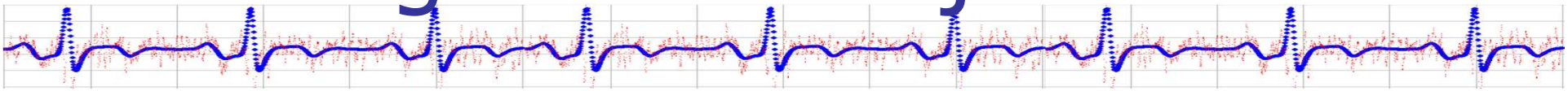
Empirical; Rational; Testable; Parsimonious; General; Tentative; Rigorously evaluated

Exercise part II (at home)



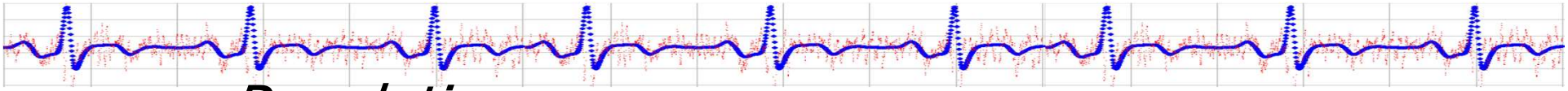
Look for the roles that background research play in the sample research plan.

Using human subjects



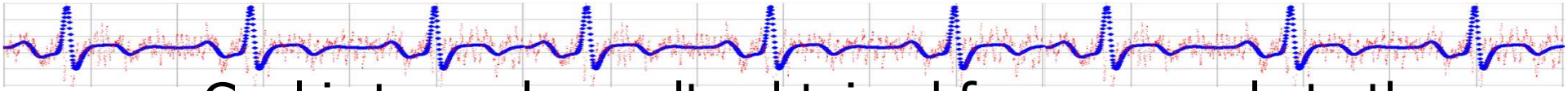
- Sampling
- Ethical issues
- Confidentiality & Identification
- Eligibility
- Recruiting
- Compensation
- NU IRB & student projects

Populations and Samples



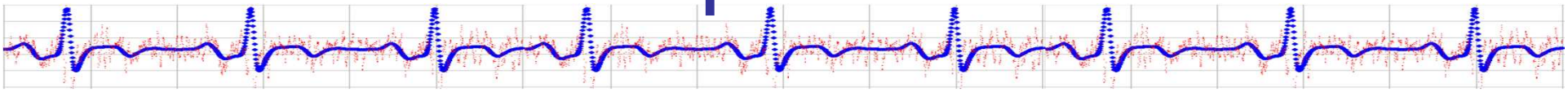
- *Population*
  - Large group including all potential subjects
  - May be defined in many ways
    - All children in day care
    - Children in day care in a particular city
- *Sample*
  - Small subgroup of subjects chosen from the population

Sampling and Generalization



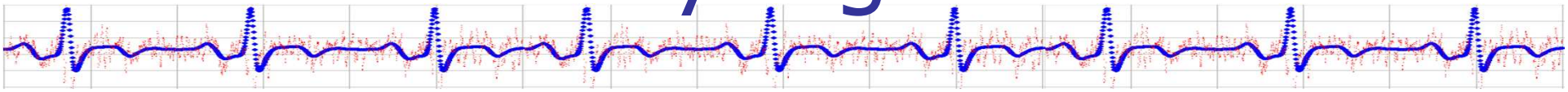
- Goal is to apply results obtained from a sample to the population
- *Generalization* is the ability to apply findings from a sample to the population
  - Aka "External Validity" of a study
- *Random sample:* A sample in which every member of the population has an equal chance of being chosen
  - Ideal that is not often met
- *Nonrandom sample:* A sample from a specialized population (e.g., college students)

Biased Samples



- Northeastern students?
- Other examples?

Biased Sample – Example
The Literary Digest Poll

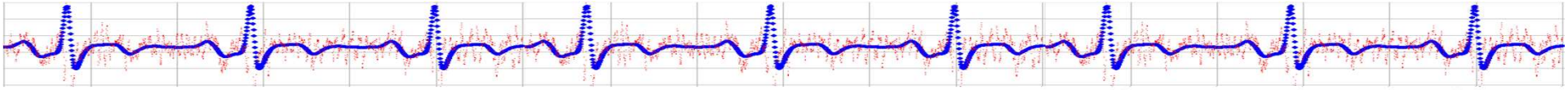


- 1936, depression
- FDR, Democrat, running for re-election against Kansas governor Alfred Landon, Republican.
- Literary Digest did a poll, mailing 10M questionnaires using addresses from the phone book and club membership lists.
- 2.4M people responded (largest number ever sampled)
- The Digest had correctly called the winner in every presidential election since 1916.

	Prediction
Roosevelt	44%
Landon	56%

- Selection Bias, magnified due to depression
- They tended to miss the poor, who did not have phones and did not join clubs. (Only one household in four had a phone

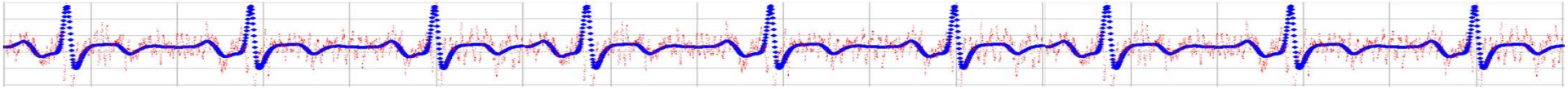
Biased Sample – Example
Are kids worth the trouble?



- Letter to Dear Abby – "Should we have kids?"
- 10k letters, 70% say "no, don't do it"
- Random sample: 91% say it's worth it

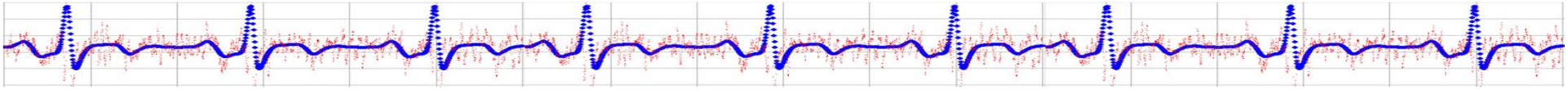
Dear Abby

What happened to the polls in the 2016 election?



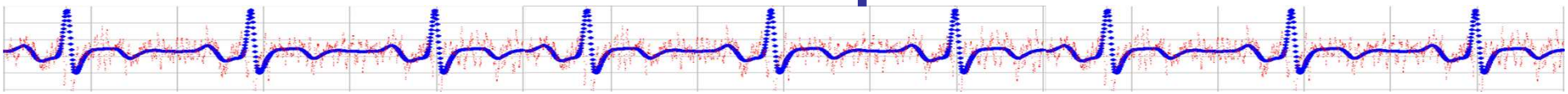
- Proposed reasons:
  - Nonresponse bias. Some groups – including less educated voters – are consistently hard for pollsters to reach
  - Social desirability bias. Support for Trump may have been seen as socially undesirable
  - Errors in identifying likely voters (who actually showed up to polls).
  - Late breaking of undecided voters for Trump
  - Errors in education weighting. Well-educated voters are much likelier to take surveys

Amazon Mechanical Turk



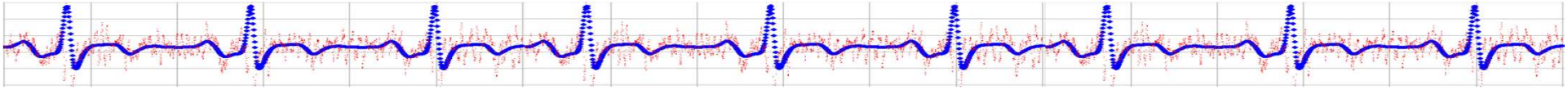
- Younger, more educated, less employed, more liberal
- Does it matter?
- Depends on study!

How do you get a random, non-biased sample?



Wait 'til B&A Ch 9 for more sampling methods

Nonrandom Sampling and Internet Research

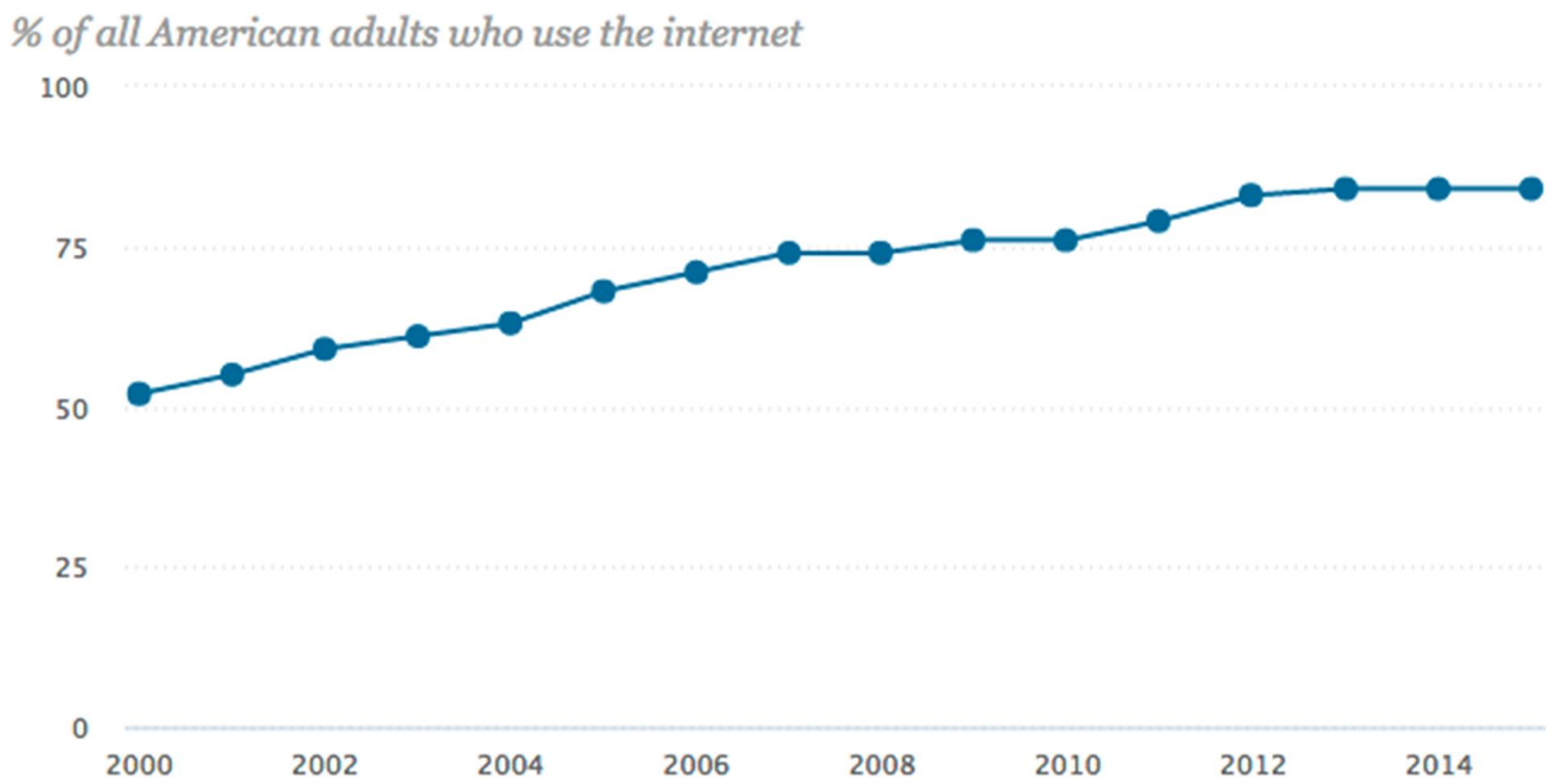


- Internet research uses a nonrandom sample
  - Participants are self-selected volunteers
  - Participants know how to use computers
  - Participants have access to computers
  - Participants are Internet savvy
- Two ways to demonstrate the validity of Internet research
  - Compare Internet with non-Internet results
  - Compare Internet results with theoretical predictions
- Internet and non-Internet samples may not differ significantly

Digital Divide – Pew 2016

87% of American Adults Use the Internet

Chart | Data | Share | Embed



Digital Divide - 2016 Pew
Who does not use the Internet?