ISS 305:002
Evaluating Evidence:
Becoming a Smart Research Consumer

4. Science as a Method of Evaluation

Reminder: Turn on your I<CLICKER

Science as a way of evaluating empirical statements

I. Alternative ways of testing empirical statements
II. What *science* is
   A. Empirical
   B. Public
   C. Controlled/systematic observation
III. What science isn't
IV. Does science "reject the idea that people need to be shielded from the truth" (Stanovich)?

Alternative ways of testing empirical statements

Consider the following statement:

- "Russians interfered with the 2016 U.S. presidential election."
  - Is it an empirical question/statement? Why?
- There are many ways to determine whether this is a true or a false empirical statement.

1. Ignore the question/statement.
2. Rely on chance to decide
   - flip a coin
   - rely on mood
3. Rely on dogma – "an official system of principles or tenets concerning faith, morals, behavior, etc.; especially one considered to be absolutely true"

Alternative ways of testing empirical statements

4. Rely on authority
   -
   -
   -
5. Rely on "common sense"
   - Problems:
     -
     -
     -

- Social Psychology and Common Sense Example

Alternative ways of testing empirical statements

6. Rely on heuristics/rules of thumb/cognitive shortcuts (fallacies)
   - Give me some examples:

7. Rely on observation.
   - Q: But what observations, made how?
   - One Answer: Make one's observations "scientifically"
     - But what does that mean?

Real World Example!

- **New Alien Life Claim Far from Convincing, Scientists Say**
  - "In the absence of a mechanism by which large particles like these can be transported to the stratosphere, we can only conclude that the biological entities originated from space," Wainwright added. "Our conclusion then is that life is continually arriving to Earth from space, life is not restricted to this planet and it almost certainly did not originate here."
  - http://news.yahoo.com/alien-life-claim-far-convincing-scientists-163757104.html

Features of science

- Empiricism
- Publicness
- Controlled/systematic observation

Empiricism

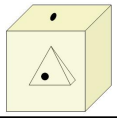
- We're already familiar with some of what this means
  - To test the validity of *empirical* statements (or questions), we must...
    - rely on sensory experience or observation as the final basis for evaluation
    - be able to verify the statement through observation
    - be able to falsify the statement through observation
      - both of which prohibit vague/imprecise language about what it is we need to observe

Empiricism: Uncertainty

- The result of a scientific evaluation of an empirical statement is never certain
- Scientific knowledge is always probabilistic knowledge
  - Goodbye Big Bang, hello black hole? A new theory of the universe's creation
    - http://phys.org/news/2013-09-goodbye-big-black-hole-theory.html

Why?

- Our observations may be incorrect
  - because of limitations of our sensory abilities
    - e.g., perception of sound direction
    - perception of color (under low illumination)
    - perceptual illusions
  - because of our perspective
    - Pyramid in a box



Empiricism: Uncertainty

Why?

- Our observations may be biased
  - we tend to see what we expect to see (the Assimilation Bias; the Confirmation Bias)
  - we tend to see what we want to see
    - Prof example
- Our observations may be distorted
  - By what?

Empiricism: Uncertainty

Why?

- Our observations may be incomplete
  - Maybe other observations (by others; in a different context; of a different type; in the future) will falsify
    - Magellan circumnavigates the world and shows that the world is not flat, as ordinary observation suggests
    - Will gravity bend light? Right observations will show it does.
- **At best**, the probability of an empirical statement being true can only **approach** what?
- Likewise, the probability of an empirical statement being false can only **approach** what?
- To say science is empirical is to say that scientific answers to empirical questions are always uncertain

Empiricism: Uncertainty

- This means that science requires that one always has an **open mind**
  - **However**, this does not mean a willingness to accept statements uncritically
  - Or, that one can't have and express profound doubts
- It **does** mean that one never concludes that an empirical statement is **absolutely** true or false (although we can sometimes get very close to such certainty)
  - examples where P(true)≈ 1.0?
  - When scientists speak of theories being **proven** or of a scientific **law**, they only mean that the probability of the theory/empirical statement is very near (but never at) 100%
  - examples where P(true)≈ 0.0?
  - Contrast with certain value and metaphysical statements, or bogus empirical claims

Example #1: Guaranteed results:

- Many (unregulated) herbal products predict that their products are "100% effective"
  - baldness prevention and cure
  - results are guaranteed, and guarantee reads as such
  - but elsewhere, the advertisement/site/bottle/whatever notes a difference

Example #2. Dr. Death & Expert Testimony of Future Dangerousness in Capital Cases

- In Texas (and most other states') capital cases, 2 phases
  - Jury decides on guilt and then on whether death penalty should be imposed
  - only for those who would commit violent acts in the future
- For determining the latter, Prosecutors rely on testimony from psychiatric/psychology experts
- How accurate is such testimony?

Prediction of future dangerousness

- Two methods
  - clinical judgment (by individual clinicians for individual defendants)
  - statistical judgment (odds that a defendant with certain characteristics will offend again)
- Latter more reliable than the former but too few cases (repeat offenders) to make reliable predictions
  -
  -
  -

Dr. James Grigson

- Has made predictions for nearly 400 capital defendants and testified in 150+ cases
- For 80+%, he has predicted that they would be dangerous
- Occasionally (when testifying as a paid witness for the defense) he predicts they will not be dangerous
- Thus, Dr. Grigson received "Dr. Death" as a nickname

Examples of Grigson's Testimony

- Stanley Faulder
  - Grigson never interviewed Faulder
  - At trial, Grigson testified that
    - Faulder was an extreme sociopath,
    - "there is no cure for his condition"
  - Faulder had suffered a massive head injury as a child, and had brain damage (not sociopathy)
  - Showed no violence in 22 years on death row before being executed by lethal injection

Examples of Grigson's Testimony

- Randall Dale Adams
  - Grigson interviewed Adams for 15 min.
    - asked about family and Adams' interpretation of "a rolling stone gathers no moss"
    - no tests, no delving into Adam's background
  - At trial, Grigson testified that
    - Adams was a sociopath
    - "there is nothing in this world today that is going to change this man"
    - Adams would continue to be a threat to society, and have no regard for the lives of others
  - Adams' conviction later reversed (State knowingly used perjured testimony) and lives in Texas. He is married, employed, and has had no arrests or violence

Examples of Grigson's Testimony

- Thomas Barefoot
  - Grigson never examined Barefoot, but testified that
    - Barefoot was in "the most severe category" of sociopaths (on a scale of one to ten, Barefoot was "above ten")
    - there was no known cure for his condition
    - there was a "**one hundred percent and absolute**" chance that Barefoot would commit future acts of criminal violence that would constitute a continuing threat to society
  - based in part on Grigson's testimony, Barefoot sentenced to death

Barefoot v. Estelle (1983)

- Barefoot appealed his conviction based on the unreliability of psychiatric predictions
- American Psychiatric Association filed *amicus* brief stating that psychiatrists could not make reliable predictions
- By 6-3 vote, Supreme Court denied Barefoot's appeal. Their reasoning...
  - APA only showed large error rate, not that psychiatrists were <u>always</u> wrong, so
  - leave it to juries to decide whether or not the psychiatrist is correct in any instance
    - that is, the jury should be capable of separating valid from mistaken predictions (when the experts themselves are wholly unable to do so)

Coda

- Grigson was expelled from the APA on July 9, 1995 for
  - arriving at diagnoses without examining patients
  - using inadequate diagnostic methods
  - maintaining that he could predict future dangerousness **<u>with certainty</u>**
    - you cannot claim to be reaching your conclusions <u>scientifically</u> **and** allow <u>no</u> room for doubt.
- Grigson's reaction?
  - APA is "a bunch of liberals who think queers are normal"

Empiricism: <u>What</u> evidence is considered?

- Empiricism also <u>requires</u> that
  -
- We can't pick and choose which "facts" (i.e., carefully and well established observations) we will pay attention to and which we will ignore
  - some "creation science" advocates ignore evidence like
    - converging evidence shows age of the earth is > 4 billion years
    - the fossil evidence documents the appearance of new species over geological time

Empiricism: <u>What</u> evidence is considered?

- A statement/idea/theory (e.g., creationism) which is contradicted by <u>many</u> observations is less plausible than a competing theory which is contradicted by <u>few</u> observations
- On the other hand, sometimes a few new observations can overcome many old observations, contributing to the **self correcting** nature of science
  - demonstrations that time is relative (Twins' Paradox)
    - led to acceptance of relativity theory over classical physics (at least when high speeds are involved)
  - by contrast, pseudoscience [e.g., astrology, tarot, ESP] typically shows little or no changes as new observations are made

Publicness

- Science doesn't rely on observations by any one person at any one time
- Science assumes that if something occurs once, it will occur again, if the conditions are the same our confidence in any reported observation increases with the number of times that others have made the same observation
- Science only trusts and relies on observations which are
  -
- Thus, science disregards observations which
  -
  -

Publicness

Bargh & Shalev (2012) – Shower Study

•Hypothesis – People use warm showers and baths to compensate for a lack of social warmth.

•2 studies found an association between trait loneliness and bathing habits. Lonely people use warmth as a substitute for lack of social connection.

•They claim that individuals "self-regulate their feelings of social warmth (connectedness to others) with applications of physical warmth (through taking warm baths or showers)" (p. 155).

•Bargh and Shalev suggest that a connection between physical warmth and psychological attributes might serve as "a boon to the therapeutic treatment of syndromes that are mainly disorders of emotion regulation" (p. 155).

•What have others found?

•Other examples

- http://blogs.discovermagazine.com/notrocketscience/2012/03/10/failed-replication-bargh-psychology-study-doyen/#.UkOcw1PB_-I
- http://www.nature.com/news/replication-studies-bad-copy-1.10634

Publicness

- Science only trusts and relies on observations that are…
  –
- Thus, science disregards observations which
  –
- Scientific knowledge is **shared**, not **privileged** knowledge (the latter are beliefs or attitude statements)
- This is also why, if we want to test an empirical statement scientifically, we must tell others how they can attempt to repeat our observations

-

Publicness: Who's the "public"?

- http://www.npr.org/sections/health-shots/2017/07/04/535412346/scientists-are-not-so-hot-at-predicting-which-cancer-studies-will-succeed?
- http://willgervais.com/blog/2017/3/2/post-publication-peer-review

- Q: Who is the "public" whose observational consensus establishes the validity of observations?
- **A: Not** just the general public, but
  – those who have the ability (training, skill, opportunity) and motivation to make "good", "valid" observations (more on this soon)
  – although sometimes, for some empirical statements, this could include the general public

Publicness: Peer review

- Scientists place high value on having the evidence for any empirical statement carefully evaluated by other qualified, trained scientists—
  – implicit in publication in a professional, **peer reviewed** journal
  – "*Have the findings been published in a recognized scientific journal that employs some type of peer review procedure? The answer to this question **will almost always** separate pseudoscientific claims from the real thing*" (p. 11, Stanovich)
  – there is little or no peer review in the popular press or Internet
  – and some "peers" may be of questionable qualifications
    - An example from lecture?

Publicness: Peer review

- Is the peer review process a good thing? What are some of the negatives?

A scientific attitude/mindset

- Empirical statements can be false and observations may differ from person to person
- This requires a certain way of thinking for science
- The scientist must always be **a skeptic, one who "habitually doubts, questions, or suspends judgment"**
- Nothing is considered "proved" in science until all significant doubts have been laid to rest

A scientific attitude/mindset

- Science does not just tolerate skepticism, it values and requires it
  - contrast with methods of validating attitude or metaphysical beliefs
- The skeptic always believes that the other guy could be dead wrong
- For the scientist, this must extend to him/herself
  - a scientist must also always accept the possibility that **s/he** could be dead wrong.
- Again, such skepticism and willingness to admit error contributes to the **self-correcting** nature of science
  - If I make mistakes today, either I or other doubters/questioners will be around to catch and correct the error in the future

Controlled/systematic observation

- Not all observations are equally relevant to an empirical statement/question
- For example, consider the statement "having sex without a condom increases your chances of contracting HIV"
  - Many observations turn out to be irrelevant
    - The room temperature
    - The gender identity of your partner
  - Some appear to be relevant but may be misleading
    - Self reports of condom use among HIV positive individuals. Why?

Controlled/systematic observation

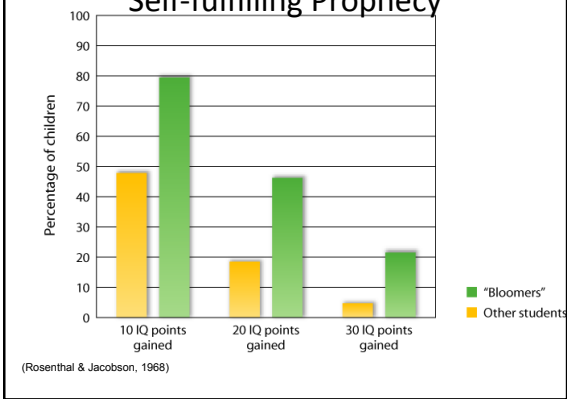
- Only certain, very carefully made observations are directly relevant
  - rate of HIV contraction for 2 groups
    - one ALWAYS used a condom when having sex
    - the other NEVER used a condom when having sex, AND
  - ALL other factors are controlled (held constant)
    - e.g., number and sex of partners
    - use of blood transfusions
    - sharing needles
    - etc.
- Training in the methods of scientific inquiry involves learning how to select/recognize the relevant (and recognize the irrelevant or misleading) observations
  - that's a big part of what this course is about
  - for example....

Controlled/systematic observation

- Illustrative techniques for making systematic observations
  - Recognizing the risks in trusting *anecdotal evidence* (e.g., personal experiences, stories, isolated incidents)
    - "My Uncle Bob smoked 3 packs of cigarettes a day and never got lung cancer."
  - Knowing that observers'/experimenters' observations can be affected more by their expectations than by reality (these reflect so-called *experimenter effects*)
    - Self-Fulfilling Prophecy

Self-fulfilling prophecies
A 3-step process

Self-fulfilling Prophecy



(Rosenthal & Jacobson, 1968)

Controlled/systematic observation

- Illustrative techniques for making systematic observations
  - Knowing that observations are often altered by chance and coincidence, and correcting for such effects (primarily through *statistics*)
    - If I get 5 heads in a row with a coin, can I conclude that it isn't a fair coin?
  - Knowing that when humans know that they are being observed, they may change their natural behavior and act in ways designed to create a favorable impression (we'll talk about *demand characteristics* and *reactive measures*)
    - Hawthorne/Observer Effect

What science is NOT

Science often uses, but does not require...

- ... mathematical and statistical techniques
  - Mathematics adds to the precision of empirical statements, and hence, to their verifiability and falsifiability
  - Statistics help to summarize observations and to eliminate chance as a source of error
  - BUT, one can still do systematic, public, empirical research without these tools
    - Learned helplessness research

What science is NOT

Science often uses, but does not require...

- ...the use of complicated, elaborate, laboratory equipment
  - Sophisticated equipment can make possible certain difficult observations and can assist in making precise observations
  - BUT, one can still often do systematic, public, empirical research without these tools
    - field research on helping (e.g., Pomazal & Clore, 1973)

What science is NOT

Science often uses, but does not require...

- ...large and useful bodies of knowledge,
  - Even very immature disciplines (like psychology), with relatively little certain knowledge, can be wholly scientific
  - And even very mature disciplines (like astrology), with large bodies of (apparent) knowledge, can be wholly unscientific
  - it is the way in which knowledge is pursued, not how successful that pursuit has been, that defines a science

What science is NOT

Science often uses, but does not require...

- ...findings which agree with our common sense, our intuitions, or our strongly held beliefs
  - Such common sense is often contradictory
    - **"The more the merrier" vs. "Two's company, three's a crowd"**
  - Such common sense is often based on private, careless observation, or upon wholly non-empirical bases (e.g., folklore; parental teaching, stereotypes, etc.)
    - lawyer's advice for picking juries
      - **"Women forgive male criminal defendants, but men are better jurors when counsel wants to avoid intuitive and sympathetic thinking" (Belli, 1954).**

What science is NOT

Science often uses, but does not require...

- ... findings which **DIS**agree with our common sense, our intuitions, or our strongly held beliefs
  - some students of psychology only credit unexpected findings as "scientific"; seeing the rest as "just common sense"
  - Our common sense can be and is often valid
  - Just because science tells us something "everybody knows", doesn't diminish the validity or possible utility of that knowledge
  - Often it is only in hindsight that "everybody knows"

Must science always reveal the truth?

- Some (e.g., Stanovich) says yes:
  - "*Scientists try to describe the world as it really is, as opposed to what our prior beliefs dictate it should be like*." (p. 27, Stanovich)
  - *Given the essentially skeptical nature of scientific inquiry, science will "...continually challenge previously held beliefs by subjecting them to empirical tests in such a way that they can be shown to be wrong*." (p. 26, Stanovich)
  - "*Psychology is thus like other sciences in* ***rejecting the idea that people need to be shielded from the truth***." (p. 27)

Must science always reveal the truth?

- Some behavioral scientists (e.g., Shelley Taylor, Jonathan Brown, Martin Seligman) have recently argued that people are sometimes happier and healthier if "shielded from the truth"; that is, "ignorance can be bliss"
  - For example, blaming failure on temporary factors outside your personal control (even when they're not) tends to prevent depression in the face of failure
- Others, like Stanovich (e.g., Robyn Dawes; Albert Ellis; Thomas Jefferson) argue that effective human functioning usually or always is promoted through accurate knowledge of the world; that is "ignorance is **not** bliss"
- What do you think? Are there situations in which people do need to be "shielded from the truth"?