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

7. Establishing Associations / Relationships

Reminder: Turn on your I<CLICKER

Establishing Relationships.

- Example: **Headache coming On?**
- What variables are mentioned here?
  - amount of coffee consumed
  - amount of caffeine consumed
  - having/not having headaches
  - timing of headaches (weekend vs. weekday)
- What links are suggested between these variables?
  - coffee ⇔ caffeine
  - no coffee ⇔ headaches
  - less coffee on weekends ⇔ more headaches on weekends OR
  - for coffee drinkers, weekday/ weekend ⇔ headaches
    - Alternative explanations for this last relationship?
      - less coffee or more alcohol?

Headache coming on?

☐ Maybe you skipped your coffee this morning. According to new research, people who consume average amounts (one to three cups daily) often suffer debilitating withdrawal symptoms—moderate to severe headaches, depression, fatigue—when they forgo their morning brew. The culprit: caffeine—or, rather, the lack of it. Study findings, reported in the New England Journal of Medicine, explain away the mystery of why some people have headaches only on weekends: They drink less coffee than they do on workdays. Since moderate coffee drinking is considered relatively harmless, the simple antidote to withdrawal symptoms is . . . a cup of java.

Contingency tables:

- Q: If we let the levels or possible values of Variable A be the rows, and the levels of Variable B be the columns, what's the smallest, simplest table possible to establish a relationship?
- A: 2 x 2 table (i.e., each variable has only 2 values)
- Consider the possible example of a relationship between weekday/weekend and headaches.
- **What's "a" here?**
- **What's "d"?**
- **What's "e"?**
- **What's "g"?**
- **What's "N"?**

Variable A Time of the week	Variable B Had a Headache?		Variable A Totals
	Yes	No	
Thursday	a	b	e=a+b
Saturday	c	d	f=c+d
Variable B Totals	g=a+c	h=b+d	N=a+b+c+d =e+f =g+h

Causal Relationships

- When we say that variable A and B are related, **we are not saying that A causes B**
- Many relationships in nature are non-causal:
  - Example: Preventing the Bubonic plague
  - "The pope [Clement VI], in his quarters at Avignon, sat between two large fires. They thought that this would purify the 'bad air' which most blamed for the spread of the plague. Although there was no bad air, the fires actually did prevent the plague..."
  - So it was concluded that bad (cold, damp) air was the cause of the plague
  - The actual cause was a bacteria, carried by fleas that live on rats, but bite humans, too

Errors in judging simple relationships: 1. The "single-cell" error

This comes in many forms. For example, **testimonials**

- The satisfied customer
- Q: what relationship is implied by this testimonial evidence?
  - **A: that this treatment works**
    - **better than no treatment;**
    - **maybe better than any competitor's treatment**

Hair Treatment Testimonials

" I bought the treatment in October 2004; after 4 months of daily usage my hairline restored fully! "

Hans W. Berlin, Germany, age: 44

Errors in judging simple relationships: 1. The "single-cell" error

This comes in many forms. For example, **testimonials**

- The satisfied customer
- What's wrong with such evidence?
  - **it may be made up by the advertisers**
  - **Hans W. may be paid for his endorsement**
  - **BUT, even if these are not problems, there's another *fatal* problem: Observations in only a single cell cannot establish a relationship**
    - **Advertiser's defense? More testimonials.**
    - **but no amount of positive testimonial evidence provides sufficient information**

Another "single-cell" error: Letting an exception disprove the rule

- Several common forms:
- Form 1. **negative testimonials**--dissatisfied customers
  - "I'll never buy another American built car. This Ford of mine is always in the shop."
    - What's the rule/relationship in question?
    - What do we know and what don't we know?
      - » We know that not all American built cars run well, BUT
      - » We don't know whether there is any general relationship
- Moral:
  - **An exception (or even many exceptions) cannot disprove the rule (i.e., show that there is no relationship)**

Variable A: Drive a US built car?	Variable B: Fewer than average rate of repairs?	
	Yes	No
Yes	?	1
No	?	?

Roots of the single cell error

- Q: Why do we commit the single cell error?
- A: **Because we assume that all relationships are perfect ones (rules without exceptions)**
  - Consider the relationship between experience and effectiveness of Presidents
  - Consider the first 40 presidents (may be hard to judge effectiveness of 4 most recent presidents)
    - divide them into the more and less experienced halves
    - and the more or less effective halves (assume that we can make this measurement)
  - If every real rule has no exceptions, then we have only 3 possibilities:
    - Experience and effectiveness always go together
    - Inexperience and effectiveness always go together
    - There's no relationship at all
  - What would these look like in contingency tables?

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	20	0	20
No	0	20	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	0	20	20
No	20	0	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	10	10	20
No	10	10	20
Variable B totals	20	20	40

Roots of the single cell error

- Suppose S. Hess' evaluations of Buchanan and Lincoln are correct. What have we learned?
  - **We could rule out the first possibility.**
- Suppose I can find a single President who is both experienced and effective (e.g. George Washington). What more have we learned?
  - **We could then rule out the 2nd possibility**
  - **More likely we would have ruled out 2nd possibility to begin with as implausible**
  - **Leaving only the third possibility: No Relationship. The exceptions disproves the rule.**

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	20	0	20
No	0	20	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	0	20	20
No	20	0	20
Variable B totals	20	20	40

Variable A: Relatively lots of experience?	Variable B: Good or Bad President?		Variable A totals
	Good	Bad	
Yes	10	10	20
No	10	10	20
Variable B totals	20	20	40

Factors that affect making the "Single cell" error

- Q: (From earlier lectures...) If we view this error as a heuristic, a "short cut", when are we more likely to make the error?
  - **A1: When we <u>lack motivation</u> to think carefully about what evidence we have or need**
    - **we don't really care about AMG or presidential experience**
    - **we don't want to look like we care too much (adopt the "whatever", Fallicist's attitude)**
  - **A2: We lack the <u>ability</u> to think carefully about what evidence we have or need**
    - **we're rushed or stressed or distracted**
    - **we're not smart enough to figure this out**
    - **we're ignorant of what evidence we need**
      - **now all of you should no longer lack that particular ability to recognize that "single cell" evidence is not enough**

Factors that affect making the "Single cell" error: The <u>vividness</u> of the "single cell" evidence

- What would you expect?
  - That Ps in condition 2 would think that prison guards were generally more humane than Ps in condition 1
- But there's more.... Hamill et al. also told some Ps that the nasty guard they saw was very <u>typical</u> of all prison guards
  - Would this likely change how they responded? What do we know?
    - Are most people we know (non-prison guards) nasty, inhumane, and disagreeable? How many out of 100?
    - Now, if we know that the prison guard in the interview <u>was</u> inhumane AND we are told that he is very typical
    - So, out of 100 prison guards, how many should be inhumane?
      - at least more than half. Let's say 2/3
  - So, is there a relationship?
  - Yes, and that's what the Ps said.

Variable A: Prison guard?	Variable B: Inhumane person?	
	Yes	No
Yes	67	33
No	5	95

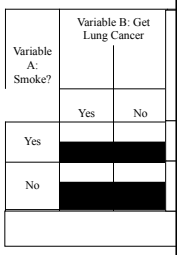
Factors that affect making the "Single cell" error: The <u>vividness</u> of the "single cell" evidence

- But there's still more.... Hamill et al. also told some other Ps the nasty guard in the interview was very <u>atypical</u> of all prison guards
  - Should this change how they responded?
  - What do we know?
    - We know that the prison guard in the interview <u>was</u> inhumane AND that he is very atypical.
    - So, in 100 prison guards, how many <u>should be</u> humane?
      - at least a majority, say 2/3?
      - If he's "<u>Very atypical</u>", then a large majority, say 95%?
  - So, is there a relationship?
    - No, or <u>at least</u>, a much weaker one.
- **But Hamill et al. found that Ps responded exactly the same way, whether they were told the nasty person was typical or atypical**
  - **they always concluded that most guards were nasty.**
- **Likewise, Ps responded the same way to the nice guard, whether they were told that they were typical or atypical**
  - **they always concluded that most guards were nice.**
- **They paid attention only to the vivid information (the interview), not the pallid information (how typical they were)**

Variable A: Prison guard?	Variable B: Inhumane person?	
	Yes	No
Yes	5	95
No	5	95

Factors that affect making the "Single cell" error: The vividness of the "single cell" evidence

- Consider the following question: "Is smoking related to lung cancer?"
  - -1964 first Surgeon General's report was published.
- However, for at least two decades after that report was published, there was no per capita decrease in cigarette consumption
  - -things have improved in the US more recently
- However, one profession did reduce their smoking after the report came out. Can you guess which one?
  - **•Physicians**
- **Conclusion: The general, pallid information about a relationship linking smoking and cancer was not convincing enough to make most people "kick the habit", but treating smokers for cancer provided more vivid and, therefore, convincing evidence.**



Factors that affect making the "Single cell" error

Some implications of the "vividness" effect:

•We're more likely to pay attention to vivid, testimonial evidence than pallid base rate information

- E.g. 1. If we're trying to decide whether or not to buy a Subaru, which evidence will be more convincing:
  - statistics that show that Subarus are above average in rate of repairs, OR
  - some **person who** tells us that their Sabaru is always breaking down?
    - **In several studies, Nisbett finds that the latter kind of information is paid more attention to**
- E.g. 2. If we're trying to decide whether it's safer to fly or drive to New York, which evidence will be more convincing:
  - statistics that show you're 37 times more likely to be killed in each mile you travel in your car vs. in a commercial airliner, OR
  - pictures of the plane crashes of 9/11?
    - **Which helps us understand why vivid media reports make many afraid of flying, or believe that a terrorist attack is the biggest risk they face**
    - **But the problem is not just in the media, it's in us**

Illusory Correlations: Hamilton & Gifford (1976)

- Ps given 1100 cards describing individuals.
- Each card has 2 bits of information
  - what group the person belongs to
    - a majority group (91%)
    - a minority group (9%)
  - whether the person has committed a crime
    - most people don't (95%)
    - but a few do (5%)
- H&G's data can be summarized in a contingency table
- Is there a relationship between group membership and crime in this table?
  - **No. Only 5% commit crimes in each group**

Group Membership	Commits a crime?		Group Membership Totals
	No	Yes	
Majority	950	50	1000
Minority	95	5	100
"Behavior" Totals	1045	55	1100

Illusory Correlations: Hamilton & Gifford (1976)

- BUT, Hamilton & Gifford's Ps generally concluded that there was a relationship
  - **that members of minority groups were MORE likely to commit crime**
- **Why?**
  - **the combination of minority group membership and rare event**
  - **these folks are distinctive in two ways**
  - **We overestimate the association between variables that are already expected to go together.**
- **Such paired distinctiveness can underlie other such stereotypes**

Group Membership	Commits a crime?		Group Membership Totals
	No	Yes	
Majority	950	50	1000
Minority	95	5	100
"Behavior" Totals	1045	55	1100

Remember the Confirmation Bias?

- What is this bias?
  - **confirmation bias is a tendency to search for or interpret new information in a way that confirms one's expectations and avoid information that contradicts our expectations**
  - Shown dramatically in Wason's rule-discovery task
    - 2-4-6 https://www.youtube.com/watch?v=vKA4w2O61Xo
    - people tend to select only confirming/positive tests of their original rule (e.g., 8-10-12)
    - and never/rarely look at possible disconfirming/ negative tests (e.g., 2-1-5)

Remember the Confirmation Bias?

- When we're looking for a relationship,
  - **the expected information may attract our special attention, and be especially vivid and noticeable**
  - **and unexpected information may be ignored**
- So, what if we're trying to determine if there were a relationship between smoking and lung cancer
- So, if we are trying to decide whether A & B are positively related, do we pay special attention to the "present-present" (Yes/ Yes; Hi A/Hi B) cell?

Variable A: Smoke?	Variable B: Get Lung Cancer	
	Yes	No
Yes		
No		

Framing relationships: Crocker (1982)

- Crocker asked her participants to consider 2 questions:
  - Question 1a: Is there a relationship between working out the day before a tennis match and winning that match?
  - Question 2a: Indicate any information needed to answer Question 1a; what information in this contingency table must you have to make an accurate judgment about Question 1a?
- How do you think her participants answered?
  - this is the % of participants who said each cell was needed
  - **the positive/positive cell is seen as more informative than any others**

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes	a	b
No	c	d

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes	75%	55%
No	35%	10%

Framing relationships: Crocker (1982)

- Crocker then asked another group to consider 2 questions:
  - Question 1b: Is there a relationship between working out the day before a tennis match and losing that match?
  - Question 2b: Indicate any information needed to answer Question 1; what information in this contingency table must you have to make an accurate judgment about Question 1?
- Note: Question 2b is really the same question as Question 2a
- Now how do you think her participants answered?
  - this is the % of participants who said each cell was needed
  - **with this way of framing the question, it was again the positive/ positive combination (working out and losing) that was seen as most informative**
- **Moral: What information we think we need is determined in part by how the question is stated.**

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes	a	b
No	c	d

Variable A: Work out day before match?	Variable B: Win the match?	
	Yes	No
Yes	55%	85%
No	20%	35%

Age & single cell errors: Kuhn et al. (1985)

- Is there a relationship between grade and getting the answer right?
  - **Yes. College students do better than earlier grades,**
  - **but still make the single-cell error 1/3 of the time**
- Why might this be true?

TABLE 2
Judgments Following Initial (Single-cell) Presentation

	Yes	*No*	*Can't Tell*	*Total*
Grade 4	14	1	2	17
Grade 7	13	0	3	16
Grade 10	10	2	3	15
College	5	0	10	15
(total)	42	3	18	63

Factors affecting single row/column errors

- Same as before (for single cell error)
  - low ability or motivation to think carefully
  - "vivid" row/column information
  - individual differences,

Age & single row/column errors: Kuhn et al. (1985)

- Here's how Kuhn et al.'s Ps answered
  - Note a new response: "Sometimes" = not a perfect relationship
- Is there a relationship between grade and getting the answer right?
  - **Yes. older students do better than earlier grades,**
  - **but college students still make the single-row/column error one-third of the time**

Variable A: Use the additive?	Variable B: Car starts? Yes	Variable B: Car starts? No
Yes	6	2
No		

TABLE 3
Judgments Following Second (Two-cell) Presentation

	Yes	*Sometimes*	*No*	*Can't Tell*	*Total*
Grade 4	8	4	3	2	17
Grade 7	7	3	4	2	16
Grade 10	5	3	2	5	15
College	5	1	1	8	15
(total)	25	11	10	17	63

Other biases in detecting relationships: Our expectations about relationships

- Why?
  - **Probably because we tend to notice and remember the confirmations, and ignore or forget the disconfirmations**
- Implications (e.g., for advertisers)?
  - **If we already believe there is a relationship, it takes very little evidence to convince us that this is so**

Other biases in detecting relationships: Our expectations about relationships

- Q: What if we already believe that there IS a relationship?
  - $|r| > 0$.
- A: **Perceived *r* > actual *r* when we expect that *r* is large**
  - **We tend to overestimate the degree of relationship**
  - **And even to see relationships when none actually exists**

- Q: What if we have no reason to think that there is a relationship?
  - i.e., Expected $r = 0$
- A: **Perceived *r* < actual *r* when we expect that *r* is small or zero**
  - **we tend to underestimate the degree of relationship**
  - **or be too "conservative"**

- Implications for the last question?
  - **We seem to need extra strong evidence to convince us two variables are related IF we don't expect them to be.**

Practice in judging relationships from word problems:

- The mean per capita contribution to the United Way from men is $54, the same as the overall mean for both genders.

- What do we know?
- Men mean= $54
- Overall mean= $54
- Women mean=
- **No relationship.**
- Restriction of range or floor/ceiling effect plausible?
- **Not really.**

Practice in judging relationships from word problems:

- The rate of dropping out of the Herbert Hoover High School for girls is 12%, twice as high as the rate for boys
- What do we know and not know?
- **b/h = 12% = 2 * a/g means that**
- **a/g = 6%**
- **b/h is different than a/g,**
- **so there is a relationship**

Variable A: Drops Out?	Variable B: Gender	
	Boys	Girls
Yes	a	b
No	c	d
Variable B totals	g	h