Today: Chapter 3

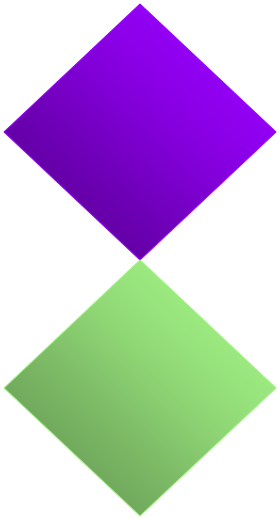
Homework: 3.23b, 3.63a to d

Due Friday, October 16th

Announcements:

No late homework accepted, because solutions are posted. (That's why you get to drop one.)

Please stop talking in class – I've had complaints about people not being able to hear because people around them are talking. Very rude!



Chapter 3

Sampling: Surveys and How to Ask Questions

Some Definitions

Population: Entire group of units about which inference will be made.

(Recall inference = hypothesis tests and confidence intervals)

Example: Political poll, population = all those who will vote in the election

Sample: The units measured or surveyed.

Example: n = 1000 likely voters nationwide

Census: Sample = entire population

More Definitions

Sample Survey: a subgroup of a large population questioned on set of topics. Special type of observational study.

Simple random sample: Every conceivable group of units of the required size from the population has the same chance to be the selected sample.

This is the ideal!

The Fundamental Rule for Using Data for Inference:

Available data can be used to make inferences about a much larger group if the data can be considered *representative* for *the question(s) of interest.*

Example: Our class probably *is* representative of *all college students* for relationship between measurements like hand span and height, but not for something like estimating proportion who have been to Disneyland.

Advantages of Sample Survey over Census

Sometimes a Census Isn't Possible

when measurements destroy units

Speed

especially if population is large

Accuracy

devote resources to getting accurate sample results

US Government conducts a census every 10 years (since 1790). Otherwise, relies on sample surveys to get unemployment rates, etc, etc.

The Beauty of Sampling When Done Right

With proper sampling methods, based on a sample of about 1000 adults we can almost certainly estimate, to within 3%, the percentage of the entire population who have a certain trait or opinion.

This result does *not* depend on how large the (large) population is. It could be tens of thousands, millions, billions….

(1000 and 3% is just an example; % depends on the size of the sample)

Estimating a Population Percent from a Sample Survey: Margin of Error

For a properly conducted sample survey:

The sample percent and the population percent rarely differ by more than the **margin of error**. They do so in fewer than 5% of surveys (about 1 in 20).

$$\textbf{(Conservative) Margin of error} \cong \frac{1}{\sqrt{n}} \times 100\%$$

where ***n*** is the number of people in the sample.

95% Confidence Interval for a population percent

In about 95% of all surveys, the interval

sample percent – margin of error

to

sample percent + margin of error

will cover the *population percent* (a fixed but unknown *parameter*).

Add and subtract the margin of error to the sample percent to create a **95% confidence interval for the population percent**.

Example: December 2008 poll on gays in the military (http://www.pollingreport.com/civil.htm)

CNN Poll of $n = 1013$ adult Americans asked:

"Do you think people who are openly gay should or should not be allowed to serve in the U.S. military?"

Should: 81% Should not: 17% Unsure:2%

Estimate percent of *population* who think "should":

Conservative margin of error is 3%: $\frac{1}{\sqrt{1013}} = 0.03$

Constructing and Interpreting the Confidence Interval

95% confidence interval for the percent of ***all*** adult Americans who would say they should be allowed: **81% ± 3% or 78% to 84%**

Interpretation:

Based on the sample of 1013 people interviewed, we are 95% confident that between 78% and 84% of *all* adult Americans think gays should be allowed in the military.

Interpreting the *Confidence Level*

The interval 78% to 84% ***may or may not*** capture the percent of adult Americans who think gays should be allowed in the military.

But, ***in the long run*** this ***procedure*** will produce interval***s*** that capture the unknown population values about 95% of the time => **95%** is called the **confidence level**.

(In Chapters 10 and 11 you will learn to use other confidence levels, like 90% and 99%.)

Technical Note: 95% Confidence Interval for a population *proportion*

In about 95% of all surveys, the interval

sample proportion – margin of error

to

sample proportion + margin of error

will cover the *population proportion* (a fixed but unknown *parameter*).

Define margin of error as *proportion:* $\frac{1}{\sqrt{n}}$

Choosing a Sample Size

Most polling agencies use samples of about 1000, because margin of error ≈ .03 or 3%.

In general:

Desired margin of error = $e = \frac{1}{\sqrt{n}}$

Then $n = (1/e)^2$

Examples:

e = .02 (2%) or 1/50, then n = 2500

e = .05 (5%) or 1/20, then n = 400

Ex: You want interval to be ± 2%, need n = 2500.

Methods of Choosing a Sample

Probability Sampling Plan: everyone in population has specified chance of making it into the sample. Many special cases, such as:

Simple Random Sample: every conceivable group of units of the required size has the same chance of being the selected sample.

Choosing a Simple Random Sample

You Need:

1. List of the units in the population.
2. Source of **random numbers** (usually a computer).

Portion of a Table of Random Digits:

ROW								
0	00157	37071	79553	31062	42411	79371	25506	69135
1	38354	03533	95514	03091	75324	40182	17302	64224
2	59785	46030	63753	53067	79710	52555	72307	10223
3	27475	10484	24616	13466	41618	08551	18314	57700
4	28966	35427	09495	11567	56534	60365	02736	32700
5	98879	34072	04189	31672	33357	53191	09807	85796
6	50735	87442	16057	02883	22656	44133	90599	91793
7	16332	40139	64701	46355	62340	22011	47257	74877
8	83845	41159	67120	56273	67519	93389	83590	12944
9	12522	20743	28607	63013	60346	71005	90348	86615

Simple Random Sample of Students

Class of **200** students.

Want a **simple random sample of 5** students.

1. **Number the units**: Students numbered 001 to 200.
2. **Choose a starting point**: Row 3, 2nd column (10484…)
3. **Read off consecutive numbers**: (3-digit labels here)
   104, 842, 461, 613, 466, 416, **180**, 855, **118**, 314, 577, **002**, 896, …
4. If number corresponds to a label, **select** that unit.
   If not, **skip** it. Continue until desired sample size obtained.

ROW								
0	00157	37071	79553	31062	42411	79371	25506	69135
1	38354	03533	95514	03091	75324	40182	17302	64224
2	59785	46030	63753	53067	79710	52555	72307	10223
3	27475	10484	24616	13466	41618	08551	18314	57700
4	28966	35427	09495	11567	56534	60365	02736	32700

Simple Random Sample of Students

5. Step 4 very inefficient.
   Can give each unit in population **multiple labels**.
   e.g. use 001 to 200 then 201 to 400, 401 to 600, etc.
   so the second 3-digit number of 842 would correspond to unit with label 842 – 800 = 042.

Using method in Step 4 selected units would be:
104, 18, 002, etc.
Using method in Step 5 selected units would be found more efficiently as:
104, 042, 061, 013, 066

Using a Table of Random Digits in a Randomized Experiment (Ch 4)

Randomization plays a key role in designing experiments to compare treatments.

Completely randomized design = all units are randomly assigned to treatment conditions.

Example: Nicotine or placebo patch. Number all 240 people, then chose 120 numbers for nicotine patch.

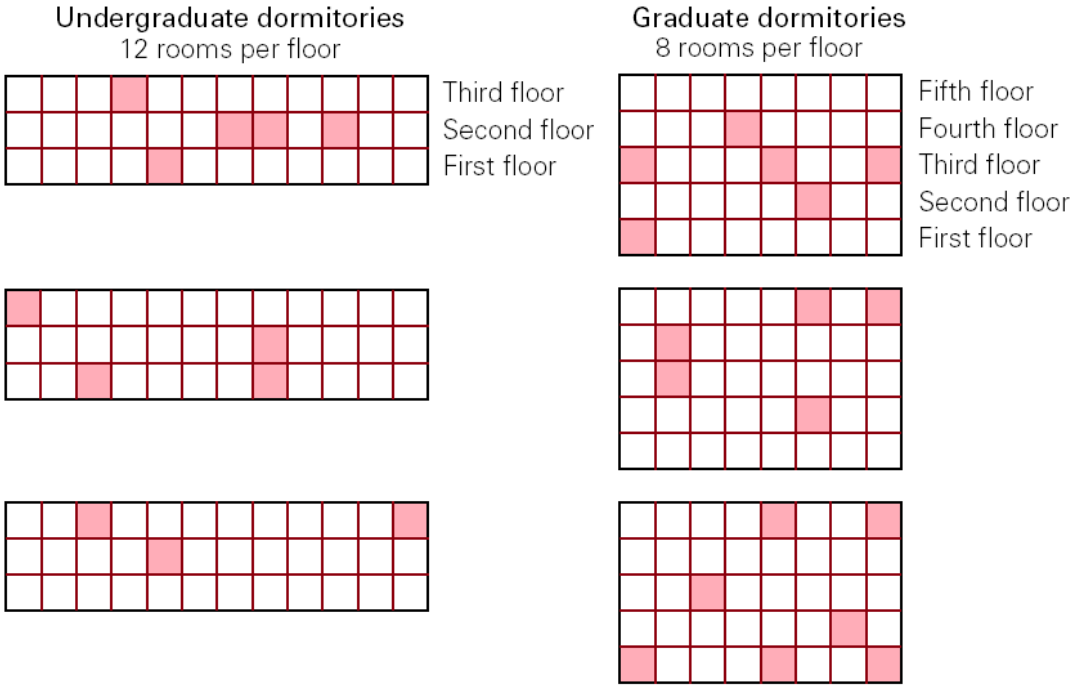
Caution: Do not confuse ***random sampling*** with ***randomization*** = ***random assignment*****.**

3.4 Other Sampling Methods

Not always practical to take a simple random sample, can be difficult to get a numbered list of all units.

Example: College administration would like to survey a sample of students living in dormitories.

Shaded squares show a simple random sample of 30 rooms.



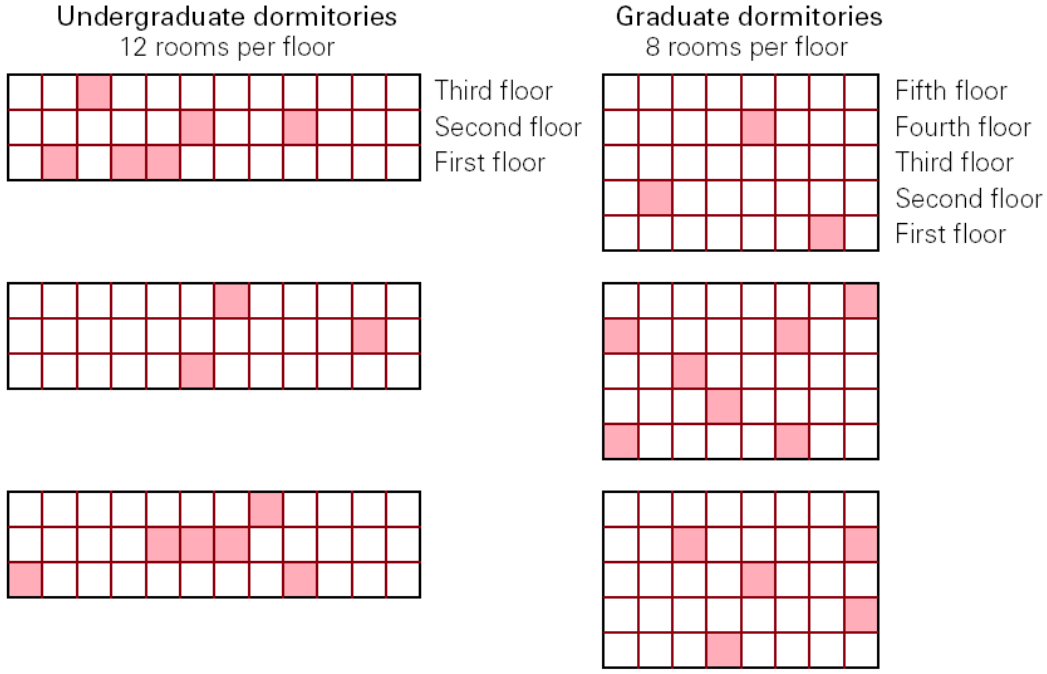
Stratified Random Sampling

Divide population of units into groups (called **strata**) and take a simple random sample from each of the strata.

College survey: Two strata = undergrad and graduate dorms.

Take a simple random sample of 15 rooms from each of the strata for a total of 30 rooms.

Ideal: stratify so little variability in responses within each of the strata.



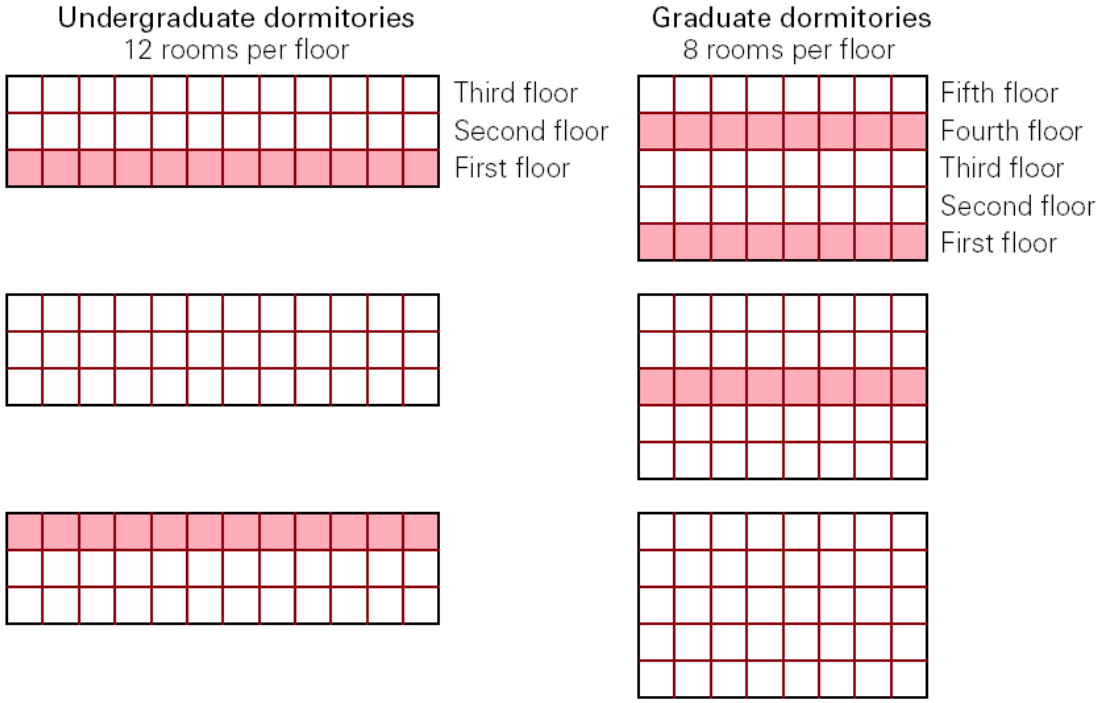
Cluster Sampling

Divide population of units into groups (called **clusters**), take a *random sample of clusters and measure only the items in these clusters*.

College survey: Each floor of each dorm is a cluster.

Take a random sample of 5 floors and all rooms on those floors are surveyed.

Advantage: need only a list of the clusters instead of a list of all individuals.

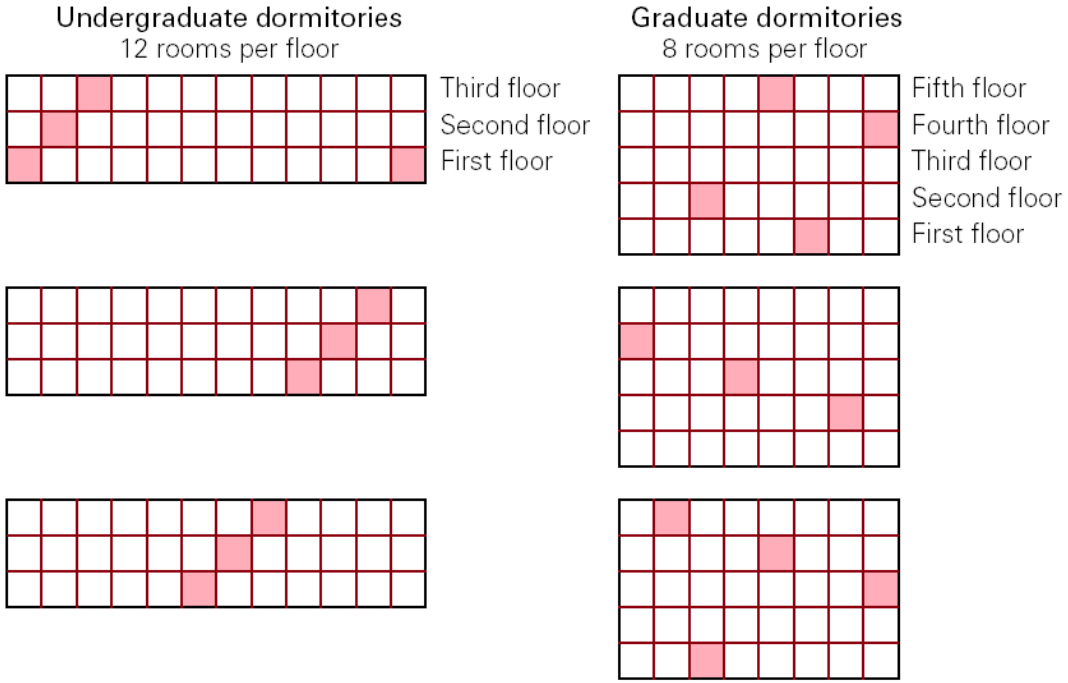


Systematic Sampling

Order the population of units in some way, select one of the first k units at random and then every k^{th} unit thereafter.

College survey: Order list of rooms starting at top floor of 1st undergrad dorm. Pick one of the first 11 rooms at random => room 3, then pick every 11th room after that.

Note: often a good alternative to random sampling but can lead to a biased sample.



Random-Digit Dialing

Method approximates a simple random sample of all households in the United States that have telephones. (Cell phones are now included in most polls.)

1. List all possible *exchanges* (= area code + next 3 digits).
2. Take a sample of *exchanges* (chance of being sampled based on proportion of households with a specific exchange).
3. Take a random sample of *banks* (= next 2 digits) within each sampled exchange.
4. Randomly generate the last two digits from 00 to 99.
5. Once a phone number determined, make multiple attempts to reach someone at that number.

Los Angeles Times Poll FAQ, good source:

http://www.latimes.com/news/custom/timespoll/la-timespollfaq,1,2370162.htmlstory

Times Poll : The Obama Transition and the Economy, December 2008

> "… Although nearly two-thirds of those surveyed believe the country is headed in the wrong direction, the figure represents an improvement from October, when 84% said the country was on the wrong track." *Los Angeles Times*

- Poll 564. December 6-8, 2008. Interviewed 1000 adults including 910 registered voters by telephone. National landline RDD. Margin of sampling error for adults and all registered voters ± 3 percentage points each.

Multistage Sampling

Using a combination of the sampling methods, at various stages.

Example:

- Stratify the population by region of the country.
- For each region, stratify by urban, suburban, and rural and take a random sample of communities within those strata.
- Divide the selected communities into city blocks as clusters, and sample some blocks.
- Everyone on the block or within the fixed area may then be sampled.

Example 3.8* *The Nationwide Personal Transportation Survey*

Nationwide Personal Transportation Survey:
taken every 6 to 7 years by the U.S. Dept of Transportation.

2001 Survey* = 25,000 households. Interviews conducted by telephone using a ***computer-assisted telephone interviewing*** **(CATI)** system, and by mail.

*Example 3.8 is the 1995 survey; 2008 survey now in progress

Multistage Sample:

- U.S. households were ***stratified*** by region of country, size of metropolitan area, and whether there is a subway system.
- Households were then selected by ***random-digit dialing***.
- Everyone in a selected household was included => each household was a ***cluster***.

Bias: How Surveys Can Go Wrong

Results based on a survey are **biased** if the methods used to obtain those results would consistently produce values that are either too high or too low.

Selection bias occurs if the method for selecting participants produces a sample that does not represent the population of interest.

Nonresponse bias occurs when a representative sample is chosen but a subset cannot be contacted or doesn't respond.

Response bias (**biased response**) occurs when participants respond, *but* they provide incorrect information, intentionally or not.

3.5 Difficulties and Disasters in Sampling

Selection bias

- Using wrong sampling frame
- Self-selected (volunteer) sample
- Convenience/haphazard sample

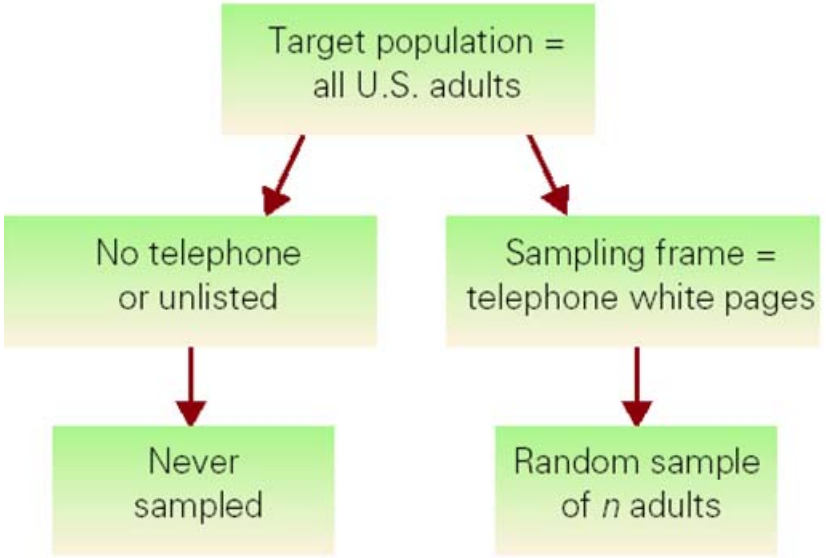
Nonresponse bias

- Not reaching individuals selected
- Non-response or nonparticipation

Using the Wrong Sampling Frame

The **sampling frame** is the list of units from which the sample is selected. This list may or may not be the same as the list of all units in the desired "target" population.

Example: using telephone directory to survey general population excludes those who move often, those with unlisted numbers, cell phones, and those who cannot afford a telephone.
Solution: use random-digit dialing, include cell phones.



Extreme Selection Bias:

Responses from a **self-selected group**, **volunteer sample, convenience sample** or **haphazard sample** often don't represent any larger group.

Example 3.10 *A Meaningless Poll*

"Do you support the President's economic plan?"

Results from TV on-air call-in poll and proper study:

	Television Poll	Survey
Yes (support plan)	42%	75%
No (don't support plan)	58%	18%
Not sure	0%	7%

Those dissatisfied more likely to respond to TV poll. Also, it did not give the "not sure" option.

Not Reaching the Individuals Selected

Failing to contact or measure the individuals who were selected in the sampling plan leads to ***nonresponse bias***.

- Telephone surveys tend to reach more women.
- Some people are rarely home.
- Others screen calls or may refuse to answer.
- **Quickie polls:** almost impossible to get a random sample in one night.

Nonresponse or Volunteer *Response*

"In 1993 the GSS (General Social Survey) achieved its highest response rate ever, 82.4%. This is five percentage points higher than our average over the last four years."
GSS News, Sept 1993

- The lower the response rate, the less the results can be generalized to the population as a whole.
- Response to surveys is voluntary. Those who respond are likely to have stronger opinions than those who don't.
- Surveys often use reminders, follow up calls, small cash award, to decrease nonresponse rate.

Example 3.9 *Which Scientists Trashed the Media? (Answer: Grumpy old men!)*

> "82% (of scientists) trashed the media, agreeing with the statement '*The media do not understand statistics well enough to explain new findings.*' " *Science* (Mervis, 1998)

Science Magazine conducted a poll:

- 1400 professionals (in science and in journalism).
- Only 34% response rate among scientists.
- "Typical" respondent was white, male physical scientist over age of 50 doing basic research.
- Respondents represent a narrow subset of scientists => inappropriate to generalize to all scientists.

Sources of *Response* Bias

1. Deliberate bias
2. Unintentional bias
3. Desire to please
4. Asking the uninformed
5. Unnecessary complexity
6. Ordering of questions
7. Confidentiality and anonymity

Wording is Important and Difficult to Get Right!

Small change of words can lead to big change in answers.

Example 1: How Fast Were They Going?

Students asked questions after shown film of car accident.

- About how fast were the cars going when they **contacted** each other?

  Average response = 31.8 mph
- About how fast were the cars going when they **collided** with each other?

  Average response = 40.8 mph

Example 2: Is Marijuana Easy to Buy But Hard to Get?

2003 Survey of Teens and Drug Use

Two versions of same question.

Half teens were asked about '***buying***' these items and the other half about '***obtaining***' them.

- Which is easiest for someone your age to **buy**: cigarettes, beer or marijuana?
- Which is easiest for someone your age to **obtain**: cigarettes, beer or marijuana?

Example 2: Is Marijuana Easy to Buy But Hard to Get?

Results:

Response	"buy" version	"obtain" version
Cigarettes	35%	39%
Beer	18%	27%
Marijuana	34%	19%
The Same	4%	5%
Don't know/no response	9%	10%

Note:

Beer is easier for teens to '**obtain**' than marijuana, but marijuana is easier for teens to '**buy**' than beer.

Deliberate Bias

Questions can be deliberately worded to support a certain cause.

Example: *Estimating what % think abortion should be legal*

- Anti-abortion group's question: "***Do you agree that*** abortion, the murder of innocent beings, should be outlawed?"
- Pro-choice group's question: "***Do you agree that*** there are circumstances under which abortion should be legal, to protect the rights of the mother?

Appropriate wording should not indicate a desired answer.

Wording of Questions about Cheating (Davis Honors Program Survey)

Version 1:

If you saw a student cheating on an exam, would you betray them and go and tell the professor?

Yes No

Version 2:

If you saw a student cheating on an exam, would you do the honest thing and tell the professor?

Yes No

Results for turning in cheater

Version 1 (Betray):

13 out of 19 said ***no*** they would not turn in the cheater

- **68% no, 32% yes**

Version 2 (Do the honest thing):

14 out of 29 said ***yes*** they would turn in the cheater

- **52% no, 48% yes**

Key Point: Wording indicating a "right answer" is wrong!

Unintentional Bias

Questions are worded such that the meaning is misinterpreted by many.

Example:

- **Do you use drugs? ---** need to specify if you mean prescription drugs, illegal drugs, etc.
- **What is the most important date in your life? ---** need to specify if you mean calendar date or social engagement.

The same word can have multiple meanings.

Desire to Please

Most respondents have a desire to please the person who is asking the question.

People tend to **understate responses about undesirable social habits**, and vice versa.

Example:

Pollsters know that asking people if they plan to vote is a very inaccurate method of identifying "likely voters". Most people say they plan to vote.

Asking the Uninformed

People do not like to admit they don't know what you are talking about.

Example:

"When the American Jewish Committee studied Americans' attitudes toward various ethnic groups, almost 30% of the respondents had an opinion about the **fictional Wisians**, rating them in social standing above a half-dozen other real groups, including Mexicans, Vietnamese and African blacks."

Source: Crossen (1994, p. 24)

Example (Case Study 3.2)

Original Source: Morin, 10-16, April 1995, p. 36.

1995 Washington Post poll #1:
1000 randomly selected respondents asked this question about the *non-existent* 1975 Public Affairs Act:
"Some people say the 1975 Public Affairs Act should be repealed. Do you agree or disagree that it should be repealed?"

- 43% of sample expressed an opinion – with 24% agreeing and 19% disagreeing.

Example, continued…

1995 Washington Post poll #2:

Two groups of 500 randomly selected respondents.

Group 1: "President Clinton (a **Democrat**) said that the 1975 Public Affairs Act should be repealed. Do you agree or disagree?"

Group 2: "The **Republicans** in Congress said that the 1975 Public Affairs Act should be repealed. Do you agree or disagree?"

- **Group 1: 36% of Democrat** respondents agreed, only **16% of Republican** respondents agreed.
- **Group 2: 36% of Republican** respondents agreed, only **19% of Democrat** respondents agreed

Unnecessary Complexity

If questions are to be understood, they must be kept simple.

Examples:

- **Too confusing:** "Shouldn't former drug dealers not be allowed to work in hospitals after they are released from prison?"
- **Asking more than one question at once:** "Do you support the president's health care plan because it would ensure that all Americans receive health coverage?"

Ordering of Questions

The order in which questions are presented can change the results.

Example:

1. How happy are you with life in general?
2. How often do you normally go out on a date?
   about ___ times a month.

Almost no correlation in answers. When order was *reversed*, there was a strong correlation! Respondents seem to think the happiness question was now, "Given what you just said about going out on dates, how happy are you?"

Confidentiality and Anonymity

People answer differently based on degree to which they are anonymous.

- **Confidentiality**: researcher promises not to release identifying information about respondents.
- **Anonymity**: researcher doesn't know identity of respondents. Important to assure respondents of this if possible.

Surveys on sensitive issues like sexual behavior and income are hard to conduct accurately.

Open or Closed Questions: Should Choices Be Given?

- **Open question**: respondents allowed to answer in own words.
- **Closed question**: respondents given list of alternatives from which to choose answer. Often an 'other' choice is provided.

Problems with Closed Questions

Source: Schuman and Scott (22 May 1987).

"What is the most important problem facing country today?"

Open Question Results

Over half of the 171 respondents gave one of these four answers:

- Unemployment (17%)
- General economic problems (17%)
- Threat of nuclear war (12%)
- Foreign affairs (10%)

Closed Question Results

List of choices and percentage who chose them ("other" was an option):

- The energy shortage (5.6%)
- The quality of public schools (32.0%)
- Legalized abortion (8.4%)
- Pollution (14.0%)

} These four choices selected by only 2.4% of respondents in the open-question survey.

Problems with Open Questions

Source: Schuman and Scott (22 May 1987).

"Name one or two of the most important national or world event(s) or change(s) during the past 50 years."

Open Question Results: most common choices

- World War II (14.1%)
- Exploration of space (6.9%)
- Assassination of John F. Kennedy (4.6%)
- The Vietnam War (10.1%)
- Don't know (10.6%); All other responses (53.7%)

Closed Question Results: given top 4 choices above + invention of computer

- World War II (22.9%)
- Exploration of space (15.8%)
- Assassination of JFK (11.6%)
- The Vietnam War (14.1%)
- Invention of Computer **(29.9%)**
- Don't know (0.3%)
- All other responses (5.4%)

Invention of computer only mentioned by **1.4%** in open question survey. Wording of question led to focus on '**events**' rather than '**changes**'.

Open or Closed Form Questions

- Open – hard to summarize results *and* important choice may not come to mind
- Closed – make sure you have the right choices, including "don't know or no opinion"
- To get choices for closed form, do a "pilot survey"

Example – false advertising?

Levi's 501 Report, a fall fashion survey conducted annually on 100 U.S. campuses concluded …

"90% of college students chose Levi's 501 jeans as being 'in' on campus."

List of choices:

- Levi's 501 jeans
- 1960s-inspired clothing
- Overalls
- Decorated denim
- Long-sleeved, hooded T-shirts
- T-shirts with graphics
- Lycra/spandex clothing
- Patriotic-themed clothing
- Printed, pull-on beach pants
- Neon-colored clothing

Levi's 501 jeans were ONLY blue jeans on the list!

Measuring Attitudes and Emotions

How to measure self esteem or happiness?

Common Method: respondents read statements and determine extent to which they agree with statement.

Example for happiness:

"I generally feel optimistic when I get up in the morning."

Indicate level of agreement from:

'*strongly disagree*' to '*strongly agree*'.

Some Concepts Are Hard to Define Precisely

Example: Measuring Stress in Kids

Drug study: ***"How much stress is there in your life?***
Think of a scale between 0 and 10, where 0 means you usually have no stress at all and 10 means you usually have a very great deal of stress, which number would you pick to indicate how much stress there is in your life?"

Results: ***Low stress (0 to 3) = 29%***
Moderate stress (4 to 6) = 45%
High stress (7 to 10) = 26%

Example continued: Stress in Kids

Another study also measured stress:
"To gauge their stress, the children were given a standard questionnaire that included questions like: ***'How often have you felt that you couldn't control the important things in your life?'***"

- **No fixed definition** of stress.
- Important = **reader is informed about how** the researchers measured stress in any given study.

Summary

When you read the results of a poll, ask:

- Who was asked – how were they chosen?
- Who responded (what percent)?
- Exactly what was asked?
- How were people contacted?
- What was the margin of error?
- What might be possible sources of bias?