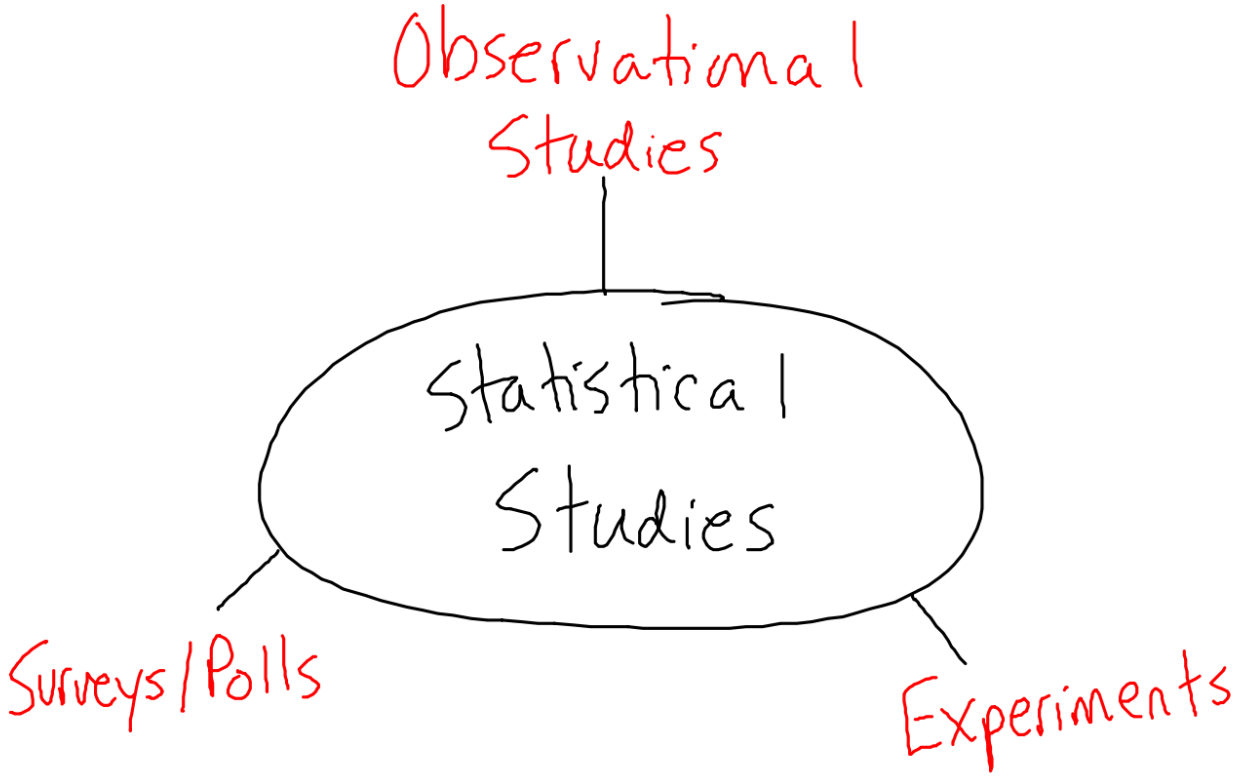
Sec 4.1


Observational Studies
Statistical Studies
Surveys/Polls
Experiments

IN DEPTH | Convention wrap-up

ical clarity remains elusive

THE WORDS THEY USED

The words that speakers used at the two political conventions showed the themes that the parties have highlighted.

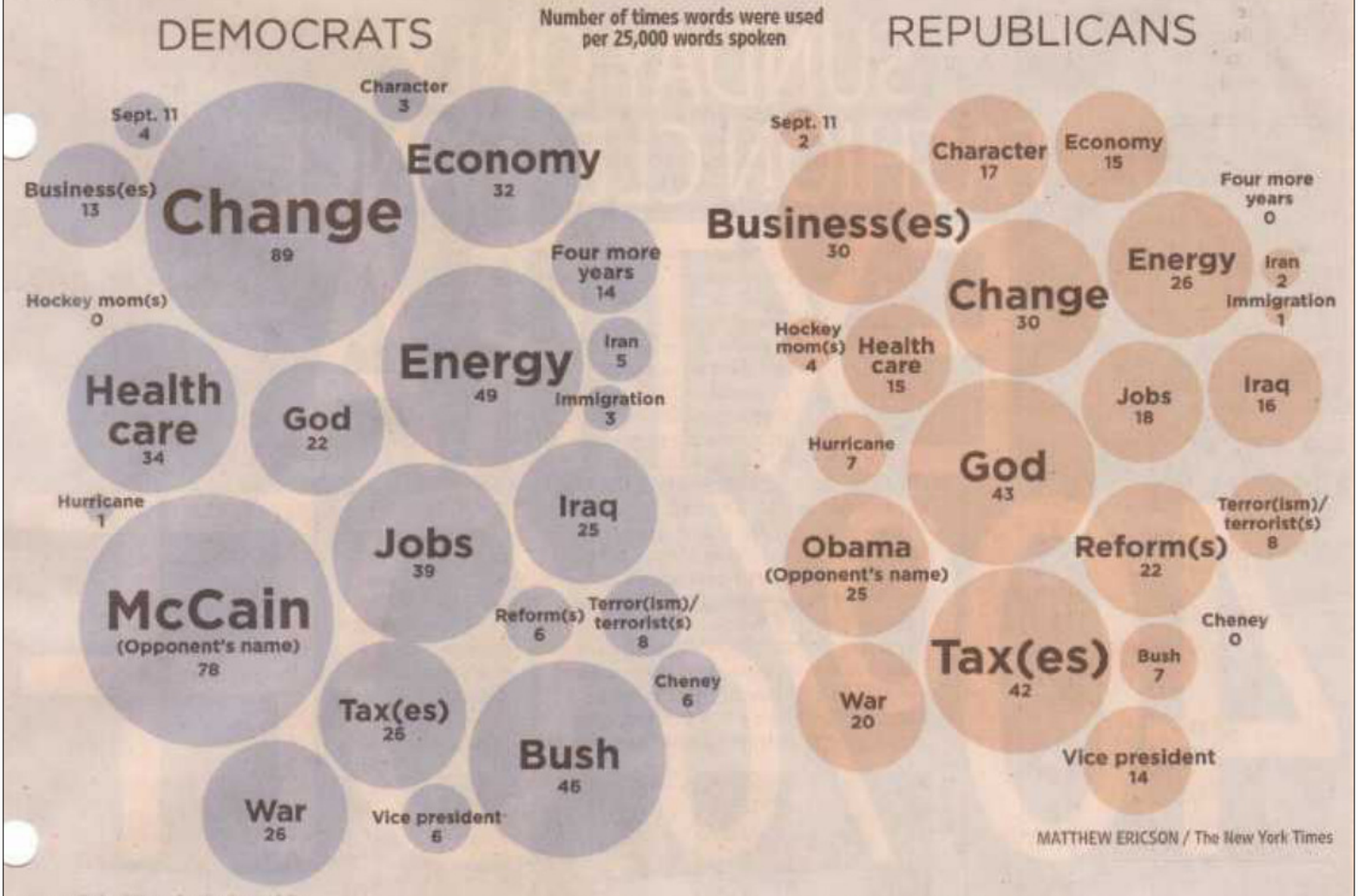
Republican speakers talked about reform and character far more frequently than the Democrats.

And Republicans were more likely to talk about businesses and taxes, while Democrats were more likely to mention jobs or the economy.

Unlike in 2004, this year the Republicans mentioned the opposing candidate less frequently than the Democrats, mentioning Sen. Barack Obama about a third as often as the Democrats mentioned Sen. John McCain.

The Democrats also pushed a couple of catch phrases of the Obama campaign in their speeches: the need for "change" and that the country doesn't need "four more years" of a Republican administration.

And the nomination of Gov. Sarah Palin of Alaska as the Republican vice presidential candidate brought a new term into the convention vocabulary: "hockey mom," which popped up in several speeches.



MATTHEW ERICSON / The New York Times

Source: Federal News Service transcripts

Sample

Part of the population of interest

Bad (Biased) Samples

- Voluntary Response Samples
- Convenience Samples

Good Samples

1) Simple Random Sample (SRS)

Consists of n individuals from the population chosen in a way that every set of n individuals has an equal chance to be part of the sample

Choosing An SRS

1) Give each individual a numerical label

2) Use random digit table or calculator

↑ Same # digits

01 - 10

11 - 20

0 - 9

↓ MATH

↓ PRB

↓ RandInt (min, max)

↓ ENTER

Ex Spring Break (P. 213)

TABLE B Random digits

Line								
101	19223	95034	05756	28713	96409	12531	42544	82853
102	73676	47150	99400	01927	27754	42648	82425	36290
103	45467	71709	77558	00095	32863	29485	82226	90056
104	52711	38889	93074	60227	40011	85848	48767	52573
105	95592	94007	69971	91481	60779	53791	17297	59335
106	68417	35013	15529	72765	85089	57067	50211	47487
107	82739	57890	20807	47511	81676	55300	94383	14893
108	60940	72024	17868	24943	61790	90656	87964	18883
109	36009	19365	15412	39638	85453	46816	83485	41979
110	38448	48789	18338	24697	39364	42006	76688	08708
111	81486	69487	60513	09297	00412	71238	27649	39950
112	59636	88804	04634	71197	19352	73089	84898	45785
113	62568	70206	40325	03699	71080	22553	11486	11776
114	45149	32992	75730	66280	03819	56202	02938	70915
115	61041	77684	94322	24709	73698	14526	31893	32592
116	14459	26056	31424	80371	65103	62253	50490	61181
117	38167	98532	62183	70632	23417	26185	41448	75532
118	73190	32533	04470	29669	84407	90785	65956	86382
119	95857	07118	87664	92099	58806	66979	98624	84826
120	35476	55972	39421	65850	04266	35435	43742	11937
121	71487	09984	29077	14863	61683	47052	62224	51025
122	13873	81598	95052	90908	73592	75186	87136	95761
123	54580	81507	27102	56027	55892	33063	41842	81868
124	71035	09001	43367	49497	72719	96758	27611	91596
125	96746	12149	37823	71868	18442	35119	62103	39244
126	96927	19931	36809	74192	77567	88741	48409	41903
127	43909	99477	25330	64359	40085	16925	85117	36071
128	15689	14227	06565	14374	13352	49367	81982	87209
129	36759	58984	68288	22913	18638	54303	00795	08727
130	69051	64817	87174	09517	84534	06489	87201	97245
131	05007	16632	81194	14873	04197	85576	45195	96565
132	68732	55259	84292	08796	43165	93739	31685	97150
133	45740	41807	65561	33302	07051	93623	18132	09547
134	27816	78416	18329	21337	35213	37741	04312	68508
135	66925	55658	39100	78458	11206	19876	87151	31260
136	08421	44753	77377	28744	75592	08563	79140	92454
137	53645	66812	61421	47836	12609	15373	98481	14592
138	66831	68908	40772	21558	47781	33586	79177	06928
139	55588	99404	70708	41098	43563	56934	48394	51719
140	12975	13258	13048	45144	72321	81940	00360	02428
141	96767	35964	23822	96012	94591	65194	50842	53372
142	72829	50232	97892	63408	77919	44575	24870	04178
143	88565	42628	17797	49376	61762	16953	88604	12724
144	62964	88145	83083	69453	46109	59505	69680	00900
145	19687	12633	57857	95806	09931	02150	43163	58636
146	37609	59057	66967	83401	60705	02384	90597	93600
147	54973	86278	88737	74351	47500	84552	19909	67181
148	00694	05977	19664	65441	20903	62371	22725	53340
149	71546	05233	53946	68743	72460	27601	45403	88692
150	07511	88915	41267	16853	84569	79367	32337	03316

2) Stratified Random Sample

a) Divide population into groups/strata

b) Choose SRS from each strata

c) Combine

Ex LN Population → Divide By Grade (9-10-11-12)

↓ SRS

Combine

3) Cluster Sampling

- Divide population into smaller groups (clusters)
- SRS of clusters
- Include all individuals in clusters

Ex All GI Classes $\xrightarrow{SRS}$ n classes → All students in each chosen class comprise sample

Inference (40% of AP Exam)

- Drawing conclusions about a population based on a sample
- Random samples (SRS) reduce bias
- Larger samples provide more accurate results

Sampling Errors

1) Using a sampling method that is not random

2) Undercoverage - leaving out members of the population from the sampling frame (the list of individuals from which the sample is taken)

Nonsampling Errors

1) Nonresponse

Selected members of sample cannot or refuse to respond

2) Response Bias

Interviewer affects outcome ... due to age, race, gender, position etc, the respondant does not answer truthfully

3) Wording Effect

- Questions confusing or biased

↑ slanted toward specific response

- May or may not be intentional!

How Are Polls Conducted?

Public opinion polls would have less value in a democracy if the public -- the very people whose views the polls represent -- didn't have confidence in the results. This confidence does not come easily. The process of polling is often mysterious, particularly to those who don't see how the views of 1,000 people can represent those of hundreds of millions. Many Americans contact Gallup each year:

1. to ask how Gallup's results can differ so much from their own personal impressions of what people think;
2. to learn how Gallup selects people for inclusion in its polls; and
3. to find out why they have never been interviewed.

The public's questions indicate a sincere desire to find out more about how polling is actually conducted, which this FAQ section will address.

The Sampling Issue

Probability sampling is the fundamental basis for all survey research. The basic principle: If selected correctly, a randomly selected small sample of a population of people *can* represent the attitudes, opinions, or projected behavior of all of the people from which the sample is obtained.

The fundamental goal of a survey is to come up with the same results that would have been obtained had every member of a population been interviewed. For national Gallup polls, in other words, the objective is to represent the opinions of a sample of people who hold the same opinions that would have been obtained if it were possible to interview all adult Americans in the country. There are difficulties involved at every step of the process of attempting to reach this goal, but this overall objective stands as the central organizing principle for Gallup's methodological procedures.

The key to reaching this objective is a fundamental principle called *equal probability of selection*, which states that if every member of a population has an equal -- or in some instances a *known* -- probability of being selected in a sample, then that sample will be representative of the population.

Thus, it is Gallup's goal in selecting samples to allow every adult American an equal or known chance of falling into the sample. How that is done, of course, is the key to the success or failure of the process.

Selecting a Random Sample

To conduct a national opinion poll, Gallup first chooses a method by which all or most Americans have an equal or known likely chance to be selected. Of course, most Americans have a place of residence, making it immediately obvious that if it were possible to sample for a list of all residential addresses in the country, the principle of comprehensive sampling frame could be achieved.

With this in mind, Gallup interviewers conducted the earliest polls in person by selecting places of residences from all possible geographic areas within the country and then fanning out across the country knocking on Americans' doors. This was the standard method of interviewing for nearly 50 years, from about 1935 to the mid-1980s, and it was a demonstrably reliable method.

Gallup polls across the 12 presidential elections held between 1936 and 1984 were highly accurate, with the average error in Gallup's final estimate of the election being less than three percentage points.

It became obvious over time that household sampling with in-home interviewing had problems. For one thing, it was increasingly expensive. Americans became increasingly resistant to allowing interviewers to come into their homes to conduct surveys. And the need for rapidly gathered data made the in-home interview procedure less and less attractive.

Gallup then turned its attention to the telephone. By 1986, a sufficient proportion of American households had at least one telephone to make telephone interviewing a viable and substantially less expensive alternative to the in-person method. By the end of the 1980s, Gallup was conducting the vast majority of its national surveys by telephone.

Another important change was becoming obvious by the mid- to late 2000s: the increasing use of cell phones by Americans. As a result, by 2008, Gallup had shifted its interviewing to include traditional landline and cell phone sampling. Today, with increasing shifts in communication that may eventually move some Americans beyond any type of phone toward texting and written communication, there has been increased attention on the part of survey professionals to the possible benefits of a move back to "old-fashioned" residential, address-based sampling.

For now, however, the vast majority of Gallup surveys intended to represent the national population are based on interviews conducted by landline and cell telephones. This method builds off of the central assumption that most Americans still either live in a residence with a telephone or own a personal cell phone. So, reaching people on their telephones is the starting place for current national surveys.

Procedurally, Gallup includes several steps in putting together its poll with the objective of letting every American household and every American adult have an equal chance of falling into the sample.

Initially, Gallup clearly identifies and describes the population that a given poll is attempting to represent.

If Gallup was polling about baseball fans on behalf of the sports page of a major newspaper, the target population might be all Americans aged 18 and older who say they are fans of the sport of baseball. If the poll were being conducted on behalf of Major League Baseball, however, the target audience the client requires might more specific, such as people aged 12 and older who watch at least five hours of MLB games on television, or in-person, each week. And so on.

In the case of Gallup polls that track elections and major political, social, and economic questions of the day, the target audience is generally referred to as "national adults." Strictly speaking, the target audience is all adults, aged 18 and older, living in United States. In effect, the population represented becomes all Americans aged 18 and older who have a telephone.

Next, Gallup chooses or designs a method to sample the target population randomly.

The findings from Gallup's telephone surveys are based on Gallup's standard national telephone samples, consisting of directory-assisted random-digit telephone samples using a proportionate, stratified sampling design. This complicated process starts with a computerized list of all

telephone exchanges in America, residential and cellular, along with estimates of the number of phones these exchanges have attached to them. The computer, using a procedure called random-digit-dialing (RDD), actually creates phone numbers from those exchanges and then generates telephone samples from those. In essence, this procedure creates a list of all possible household phone numbers and all possible cell phone numbers in America and then selects a subset of numbers from that list for Gallup interviewers to call.

It's important to go through this complicated procedure for two reasons. A significant percentage of residential phones are unlisted. And, almost all cell phone numbers are unlisted in the sense that there is no phone book or other comprehensive listing of them. The random-digit-dial procedure allows telephone exchanges to be the main medium of sampling, with the digits of particular phone numbers added randomly.

Selecting the Individual to Be Interviewed

Within each contacted household reached via landline, an interview is sought with the adult 18 years of age or older living in the household who has had the most recent birthday. (This is a method pollsters commonly use to make a random selection within households without having to ask the respondent to provide a complete roster of all adults living in the household). Gallup does not use the same respondent selection procedure when making calls to cell phones because they are typically associated with one individual rather than shared among several members of a household.

The Number of Interviews or Sample Size Required

One key question faced by Gallup statisticians is how many interviews does it take to provide an adequate cross section of Americans? The answer is not many -- if the respondents to be interviewed are selected entirely at random, giving every adult American an equal or known probability of falling into the sample. The typical sample size for a Gallup poll, either a traditional stand-alone poll or one night's interviewing from Gallup's Daily tracking, is 1,000 national adults.

Broadly speaking, the actual number of people that need to be interviewed for a given sample is to some degree less important than the soundness of the fundamental equal probability of selection principle. This is something many people find hard to believe, if respondents are *not* selected randomly, Gallup could have a poll with a million people and still be significantly less likely to represent the views of all Americans than a much smaller sample of 1,000 people selected randomly.

To be sure, there is some gain in sampling accuracy that comes from increasing sample sizes. Using common sense and sampling theory, a sample of 1,000 people is most likely going to be more accurate than a sample of 20. Surprisingly, however, once the survey sample approaches 500, 600, 700, or more, there are fewer and fewer accuracy gains that come from increasing the sample size. Gallup and other major organizations use sample sizes of between 1,000 and 1,500 for standard surveys because they provide a solid balance of accuracy against the increased economic cost of larger and larger samples. If Gallup were to use a sample of 4,000 randomly selected adults each time it did a poll, the increase in accuracy over a well-done sample of 1,000 would be minimal, and for sample accuracy reasons, would not justify the increase in cost.

There are, however, other reasons why large sample sizes can be important including, in particular, the ability to subset the sample into small population segments for analysis purposes. This is the great advantage provided by aggregates of Gallup's Daily tracking samples.

Statisticians over the years have developed specific ways of measuring the accuracy of samples -- so long as the fundamental principle of equal or known probability of selection is adhered to when the sample is drawn.

For example, with a sample size of 1,000 national adults (derived using careful random selection procedures), the results are highly likely to be accurate within a margin of error of ±4 percentage points. Thus, if a president's approval rating is 50%, the margin of error indicates that the true rating (that is, the rating that would be obtained had Gallup interviewed every adult in American) is likely to be between 54% and 46%. It is unlikely to be higher or lower than that.

To be more specific, the laws of probability say that if Gallup was to conduct the same survey 100 times, asking people in each survey to rate the job President X is doing as president, in 95 out of those 100 polls, his rating would be between 46% and 54%. In five of those surveys, his rating would be higher or lower than that due to chance error.

If Gallup increases a poll sample size to 2,000, the results would then be accurate within ±2% of the underlying population value, a gain of two percentage points in terms of accuracy, but with a 100% increase in the cost of conducting the survey. These are the cost value decisions that Gallup and other survey organizations make when they decide on sample sizes for their surveys.

Weighting the Sample

After Gallup collects and processes survey data, each respondent is assigned a weight so that the demographic characteristics of the total weighted sample of respondents match the latest estimates of the demographic characteristics of the adult population available from the U.S. Census Bureau. Gallup weights data to census estimates for gender, race, age, educational attainment, and region.

Sec 4.2

Experimental Studies

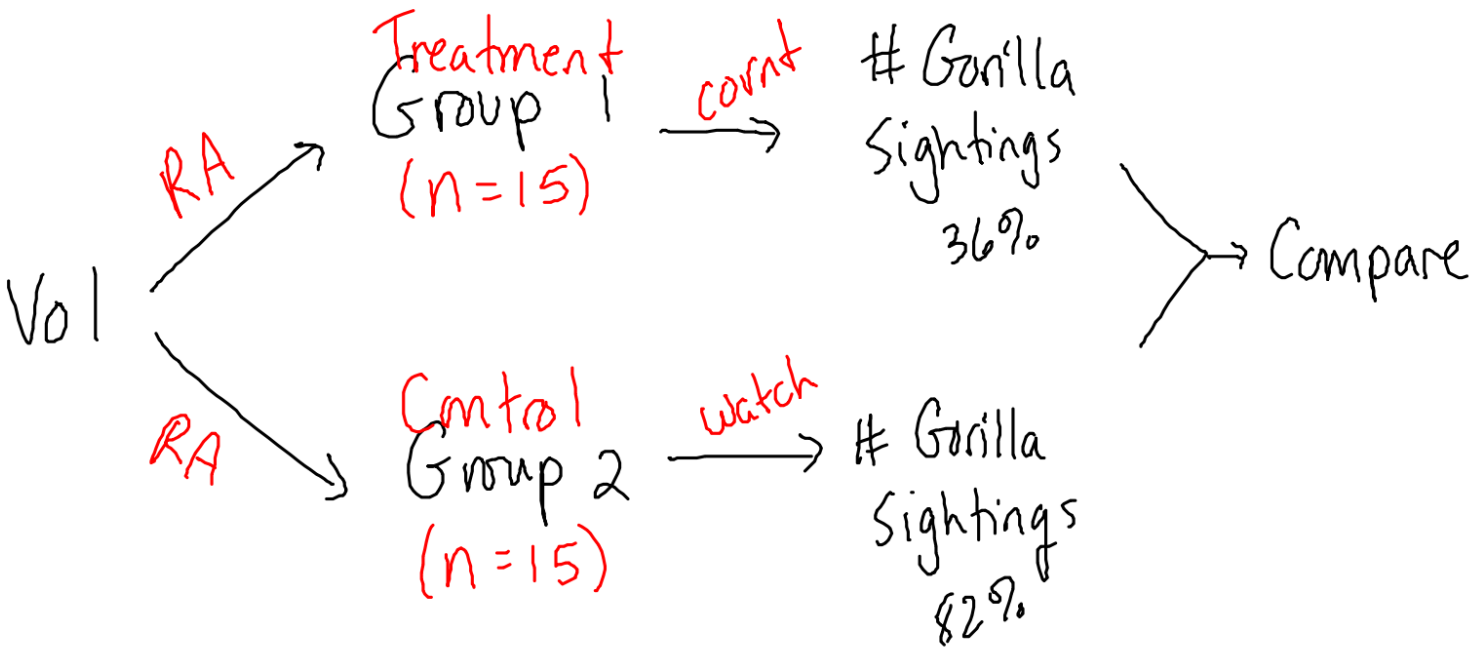
- Deliberately imposes a "treatment" to influence a response

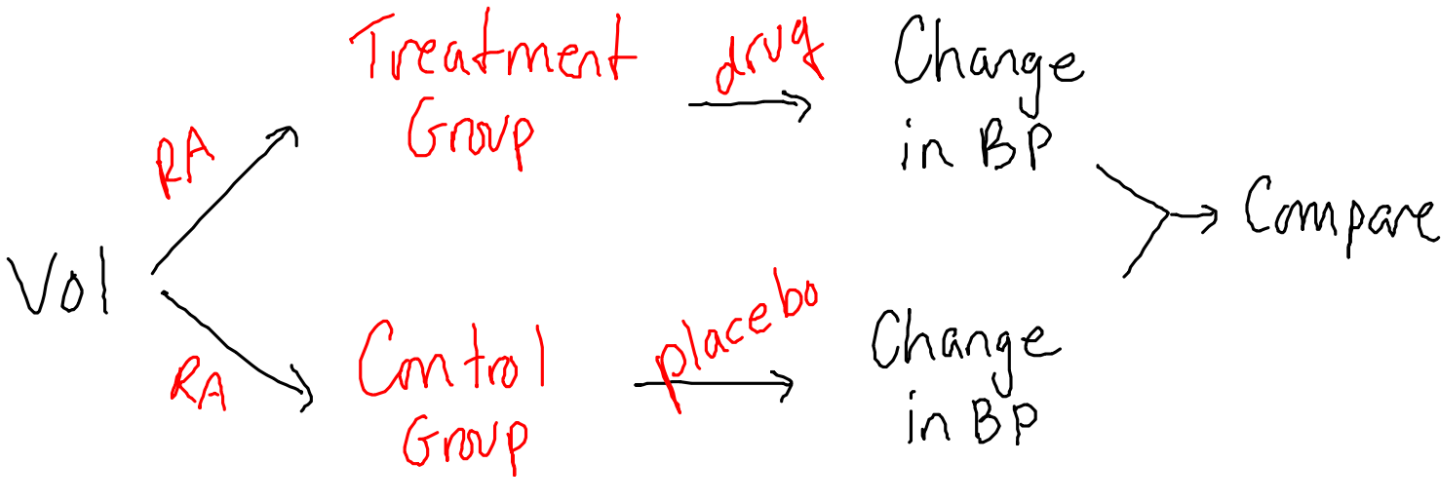
- Common on AP Exam

Common Designs

- Between groups / Independent samples
- Within groups / Repeated Measures
- Matched Pairs

Between Groups

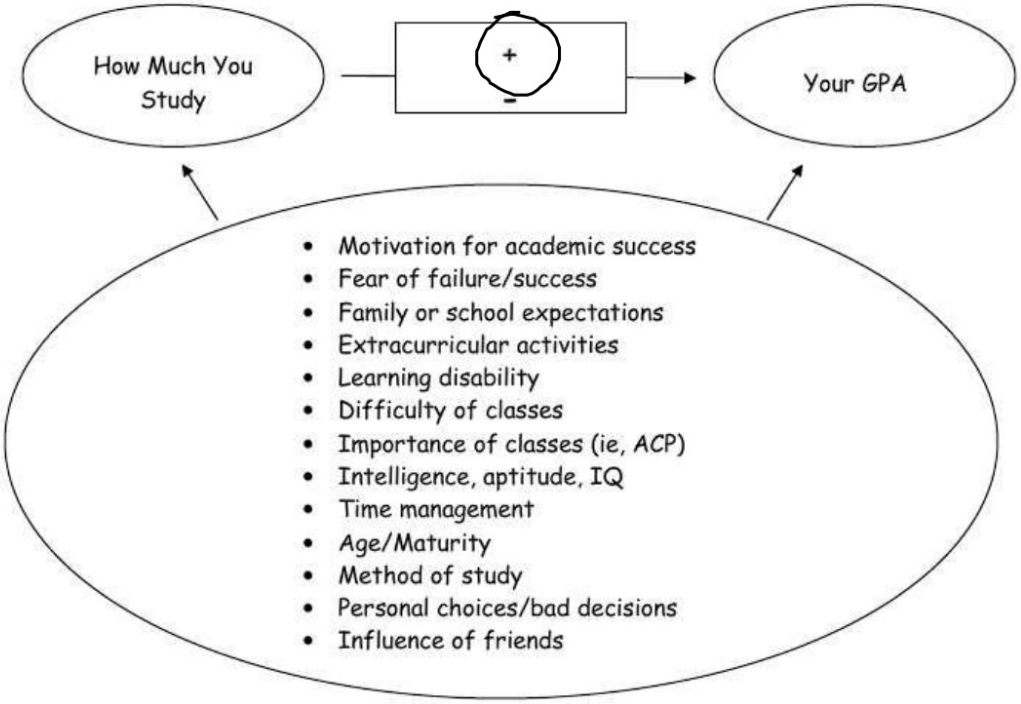




Reducing Placebo Effect

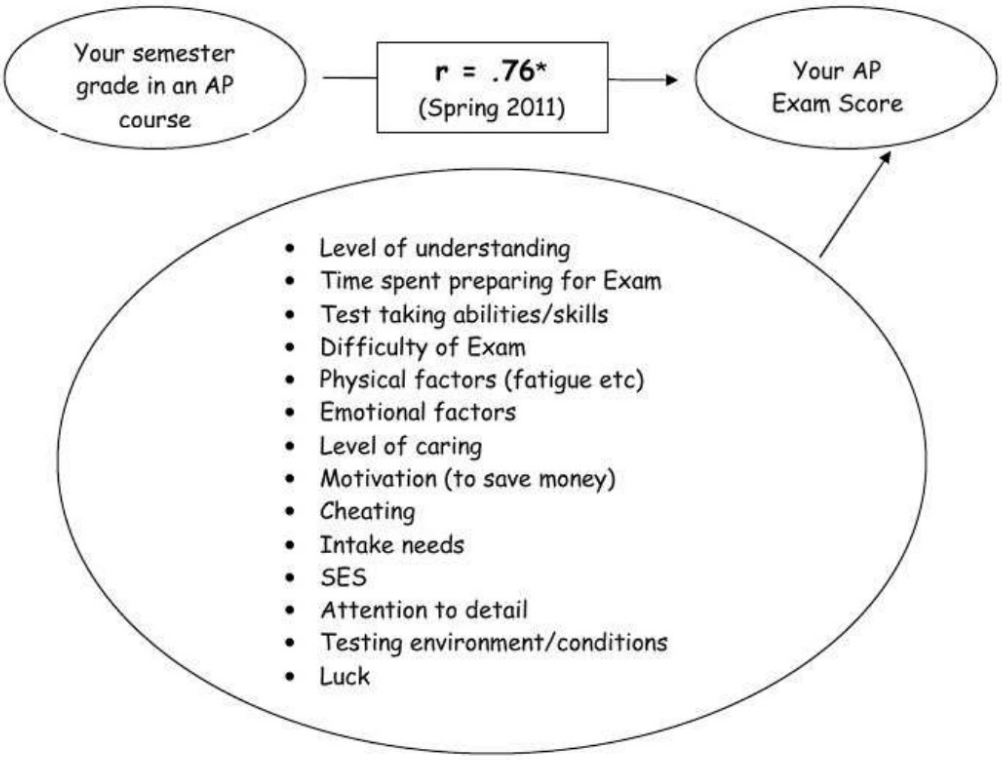
1) Use large # subjects
2) Double Blind Study - neither subject nor experimenter knows placebo

LURKING VARIABLES
(Common Response)

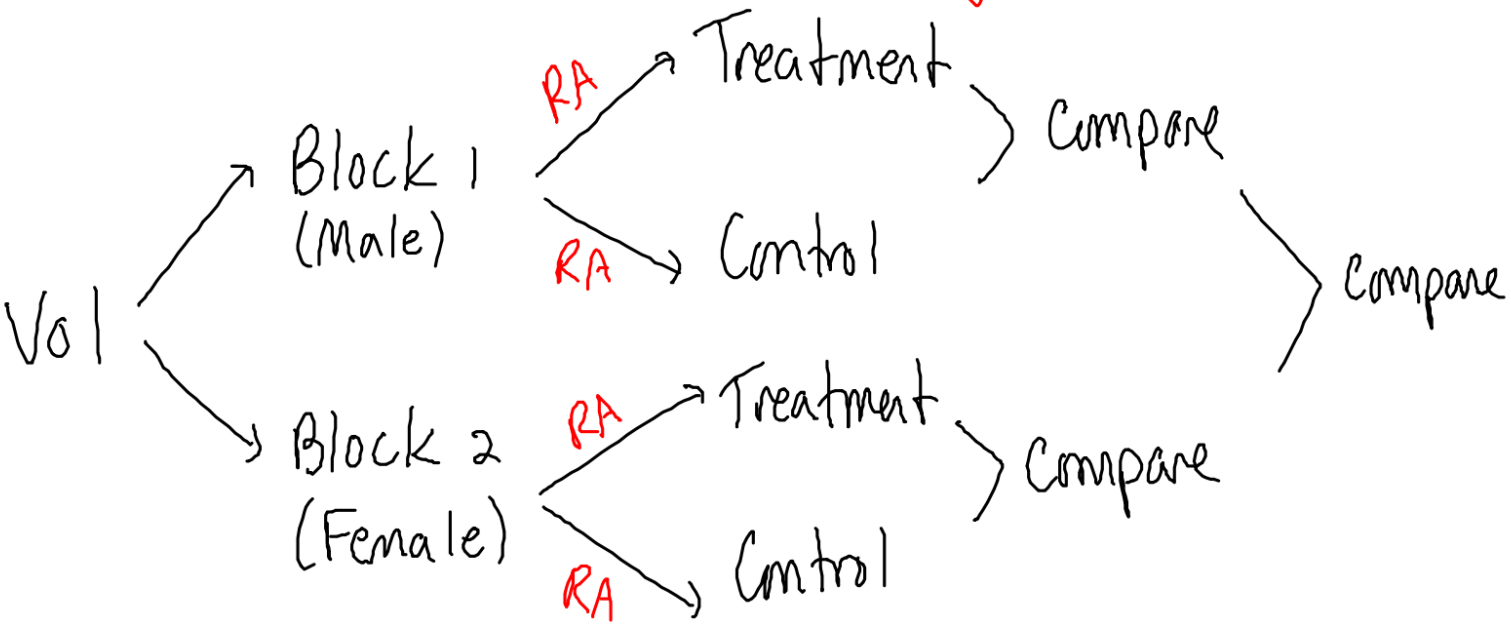


LURKING VARIABLES
(Confounding)

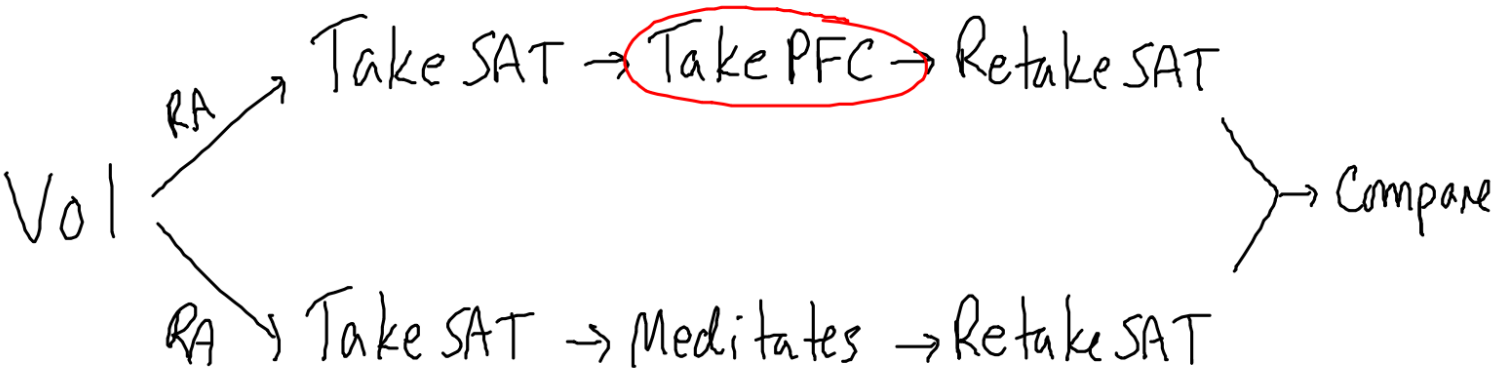
*Predicted AP Exam Score = -9.0157 + .1421(Semester Grade)



Between Groups Block Design



Does PFC improve SAT Scores?



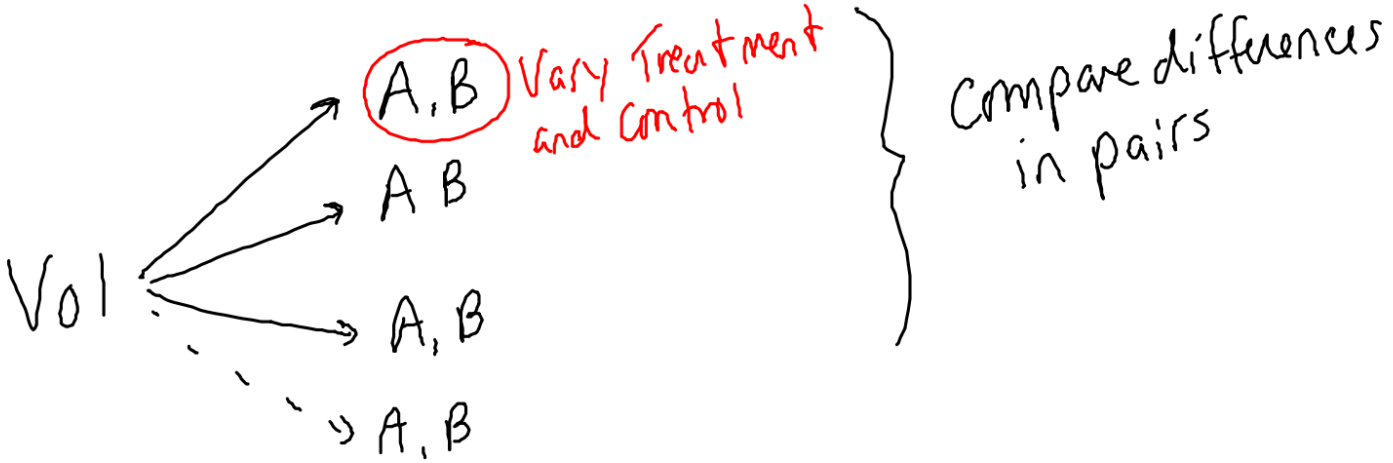
Within Subjects / Repeated Measures Design

- Individuals are own control

	1	2	3	4	n
Take SAT	1000	1200	800	700	X
↓ PFC ↓					
Retake SAT	1100	1200	1000	750	Y
	100	0	200	50	Y-X

Matched Pairs

- Match pairs of "experimental units" with similar characteristics (Gender, IQ, SES...)
- Randomize order of treatment



Inference For Experiments

- If difference(s) between treatment and control group is larger than we expect to find by chance, then we say the results are statistically significant

- Cause/effect can only be determined through replication!

Sec 4.3

Using Studies Wisely

1) Inference about a population requires:
   a) randomly selected individuals (observational studies / surveys and polls)
   b) randomly assigned treatments (experiments)

2) Cause and effect inferred through

a) Replication of well designed experiments

b) Evidence of causation through correlation studies (Ex, P.264)

3) Don't automatically believe a "study" is true!