AP Research: Course Study Guide

******Creator's Note:*** *Hello and welcome to AP Research! This class is very different from all other AP classes, so this will function more as a how-to with tips than as a 'cheatsheet'. For this class, you'll be designing, conducting, and analyzing your own research project; everybody's path will be slightly different.* ***Some tips to keep in mind before you get started****:*

1. **Create and keep a timeframe in mind!** It doesn't go well to get to March and realize that not only have you not started preparing your presentation, but you also haven't finished collecting data.
2. **Rely on a small group of other classmates.** This is *critical* because your teacher isn't allowed to give you direct feedback on your project. Make sure that you trust your group to give you honest feedback, good and bad.
3. **Choose someone that you trust to help edit your paper**, and *don't show them your paper until you are done*. The impact a fresh pair of eyes can have on the editing process is *very valuable*, so be sure to take advantage of that. They'll have a much more neutral and critical viewpoint than someone who has read your paper before.
4. **Build a good relationship with someone who has a strong grasp of statistics** or math (*especially a teacher, if they have time*) so that you can have them look over your data and analysis.

In short, **staying organized** and **relying on the resources you do have access to** will make your year a lot easier and more successful.

Finally, this guide will be following the hypothetical journey of a student called "Joe". He'll be serving as an example so you can see how to put some of these tips into practice.

CHOOSING A RESEARCH TOPIC:

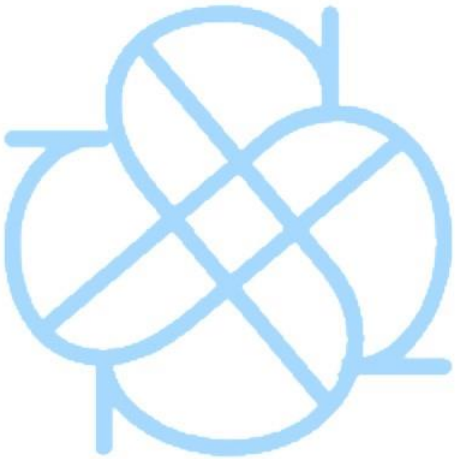
- This is often the most intimidating step of the year. The most important thing to keep in mind is that *you should choose something you're interested in*. You will be working with a very specific part of whatever topic you choose for an entire year. If you don't actually care about your topic, it gets a lot harder to stay motivated and score well.

- **Here's a good way to start choosing a topic:**

1. **Sit down and write down as many things and questions that you're interested in** for 5 minutes. No barriers. Treat this as a stream of consciousness exercise.
    a. Let's say Joe writes down "how is coffee made?", "best tennis swing", "chocolate", "construction", "paint", and "sunrises", as well as many other things in his 5 minutes.
2. **Set your pencil down** after your 5 minutes of throwing random ideas/thoughts/questions onto a page. Do something else for at least 10 minutes to give your brain a break. Then, go back to your list. **Star the top 5** or so things that you think you could or would want to research.
    a. Some of Joe's starred items are "construction", "paint", and "how is coffee made?".
3. **Next, take a closer look at your starred items**. Spend a couple minutes with each of them. For anything that wasn't a question, see what questions you can come up with about that topic. For any questions that you starred, see if you can add any more questions about the topic that you're interested in. At this point, **you should have at least 3-5 questions about each thing you've starred**.
    a. Based off of Joe's thought "paint", he came up with some preliminary questions like "what color of paint is most calming?", "what are the impacts of painting exteriors?", "is lead paint still used?", and so on. Some of these are not very good research questions. "Is lead paint still used?" can probably be answered in one sentence, and "what are the impacts of painting exteriors?" is extremely broad.
4. The fact that these are not refined research questions *does not matter at this stage*. Don't worry about that! Instead, if you, like Joe, still find yourself interested in paint (or whatever your topic is), **spend 30 minutes lightly researching your topic** and/or some of your generated questions. You shouldn't be reading academic papers at this stage, or spending hours checking out all the wikipedia sources linked on the page of your research. This step is just to find out more about your topic so you can **1) see if you're still legitimately interested, 2) come up with more refined research questions, and 3) get a preliminary idea of what's been researched on your topic.**

5. **Decide if you want to keep researching your topic**. If you don't want to research it all year, all you have to do now is go back to your starred items and choose the next most interesting one, and repeat steps 3-4. Eventually, you (and Joe!) will land on a topic that you still care about after your 30 minutes of research.

******Note:** This may be the first topic you think of, or it might be the 6th one you try. Don't worry if it's your 6th! This part of the research process is extremely important. Changing topics later in the year is much more difficult than it is right now, so **take your time.***



WRITING A RESEARCH QUESTION:

- Congrats! You have a research topic you care enough about to get to this step. Before you start trying to create your research question, it's important to know what makes a good and bad research question.

Good questions are:

- **Current, or timeless.** It's always more valuable to have a question about something that is currently important, or that will always matter. Asking yourself if your topic is debatable usually helps find topics with more 'oomph'.
- **Specific**. You should be able to satisfactorily answer your question in one research paper. Your question should have focus.

- **Answerable/feasible**. There's little point in trying to write a research paper about a topic you don't have the equipment to effectively talk about.
- **'Arguably important'**. You need to be able to make a case as to why what you're going to write about matters. What does it matter if you write a paper about *insert your topic here*? Who does your answer affect? What could your answer change? Be sure you have answers to these questions about your topic.
- **Original**. There's no point in asking a question that's been answered many times before. If you want to research a topic or question that's been researched many times before, see if you can think of a new or different lens that might give the field a unique understanding of your topic.
- **FINER**: feasible, interesting, novel, ethical, and relevant. The website Scalelive elaborates.

Bad questions are:

- **Yes or no questions.** You shouldn't be able to answer your question in one word or sentence.
- **'Google-able'**. You shouldn't be able to search your question and get a quick answer. These kinds of questions often start with "who", "when", "where", "how much", "how many", etc.
- **Asking for opinions.** Asking "What's the best type of pasta?", for example, is impossible to answer. Instead, "How can pasta support a healthy life?" opens up opportunities to talk about how pasta might be good for you.
- **Contain a nested statement, assumption, or question(s).** These kinds of questions might look like "How did the celebrity survive and how does that impact her life now?" (2 questions in one) or "What did using the disgusting American healthcare system do to the cat?" (nested opinion/assumption).

Use the above lists to check and workshop your question. It will often take many 'rounds' of refinement among your peers and yourself to create a question that satisfies all of these markers and is still interesting to you.

Example: Let's take Joe's question of "What are the impacts of painting exteriors?" and run it through these lists to refine it. We'll start by running it through the **'bad questions' checklist** to find any obvious questions.

- Joe's answer won't be a simple yes or no.
- Although you could google his question and find information about painting exteriors, it's doubtful that he'll find a simple answer from a google pop up.
- His question doesn't ask for an opinion.
- It doesn't contain any nested questions or assumptions.

So far, Joe's question is looking pretty good! Let's take it through the **'good questions list'**.

- His topic isn't necessarily current, but it is 'timeless' in the sense that the answer should be usable now and for the foreseeable future.
- Unfortunately, Joe's question isn't specific at all. It needs to be far more targeted to be effective. Asking "what are the impacts" opens up way too many possibilities -- is Joe trying to write a paper on the fiscal impacts of painting exteriors? The environmental impacts? The psychological impacts? The economic impacts? What exteriors is he talking about? Where are they?

Before moving on with the rest of his list, Joe refines his question to ask the much more specific, **"What are the environmental impacts of painting the exteriors of ocean-front properties?"**.

- Joe's question is now **more feasible** because it is specific enough to be answered in one research paper. Fortunately, he also lives in an ocean-side town and thus has good access to raw data, and also has a good family friend who's an ecologist and willing to lend Joe whatever equipment he might need for his research.
    - From that example, you can pretty easily see that feasibility depends not just on the scope of your topic, but also your *personal situation*. Be sure to take that into account!
- Joe's question is certainly **arguably important**. His answer could affect what kinds of paints are used, who uses them, or even environmental regulations.
- Finally, after a bit of research, Joe finds that few if any researchers have looked into how exterior paint in sea-side structures impacts the environment. This "research gap" means that his question is **original** and will be **providing new information** to his field, which is excellent and important for Joe to keep in mind.

***Note:** Keep in mind that *your research question will probably change!* That's okay. The important thing is to make sure that whatever your question evolves into, it still passes the above checklists, and you're still interested in it.

TOPIC RESEARCH & ANNOTATED BIBLIOGRAPHY:

Note:** It isn't super uncommon to change topics/questions part way through this stage, but it does mean you'll have to repeat all this work for a different question. Try and find a topic/question that you think you can stick with from here on out, and *certainly* by the time you finish your annotated bibliography. This stage is where you stop **brainstorming and refining questions, and start in-depth researching your question.

- A helpful tip when beginning your topic research is to **keep track of what you've read**.
    - It is incredibly frustrating to read 20 different sources and then not be able to remember which ones you might want to use or follow up on. Make sure that you're tracking what you've read. Some students find that keeping a quick spreadsheet full of the sources you've looked at helpful. Joe's spreadsheet might look like this:

Link to access source	Brief summary	Evidence	Citation
www.com	Argued that paint flecks off of buildings and enters waterways.	- "Abdef;ladfj;" (pg 50, para 1) - Paraphrased evidence (pg 36, para 4)	Brown, Smith. "Appropriate citation for your academic style". *Example*, example citation.

- For your citations, be sure you're using the appropriate academic style for your topic.
    - The three main academic styles are MLA, APA, and Chicago. Many students find https://guides.lib.uw.edu/research/citations/citationwhich helpful in determining what style to use for their paper.
- This kind of organization of sources makes your research more effective and easier to use in the long run so you don't end up wasting a lot of time hunting down old sources.

Annotated Bibliography

- An annotated bibliography is a pretty simple idea. It's essentially a **more in-depth version of the above spreadsheet, formatted differently**. While you're more in depth researching your question, cite sources that help you understand your topic or question better, and write some about the source.
- An annotated bibliography **isn't required by Collegeboard itself**, but is often required by AP Research teachers. Whether it is or isn't required for you doesn't change that it is an *incredibly valuable step you should do anyways*.
- Usually, an annotated bib entry looks something like this:
    - Citation of paper/work in whatever manner is appropriate for your academic style.
    - Link if it isn't included in your citation.
    - Paragraph including a *brief* summary of the author's main arguments/insights. Include a sentence establishing the credibility of your source/author (credentials, institute, how many times it's been cited, can all be used).
        - Try your best to include at least a sentence explaining how this source connects to others you've seen in your research (agree, disagree, cited, etc). Explain *how the source is relevant and how you will use the source*.
    - "Directly cite evidence under your paragraph for quick and easy reference/retrieval later when you're writing your paper"
- **Example:** The following is what an actual entry from an annotated bib looked like. This source was used in an AP Research Paper on emojis:

Cifolelli, K. (2018). Harassment by Emojis - The Newest Headache for HR. Retrieved from https://www.aseonline.org/News/Articles/ArtMID/628/ArticleID/1435/Harassment-by-Emojis-%E2%80%93-The-Newest-Headache-for-HR

Cifolelli's 2018 article provides modern context for the way that HR people within companies now view and manage issues with emoticons/emojis within work communications. It also provides the valuable insight that emojis make workers be viewed as less competent. This article is credible because it cites a credible source (National Law Review 1/12/18) and because Cifolelli has previously reported on many employee-employer relationship issues, which establishes her as a knowledgeable person in the field (although not an authority, because she only reports on work others have done). The article is relevant because it provides modern context from the view of someone who understands what HR people are looking at. It will be used to provide context and a modern point of view from someone based in the HR POV.

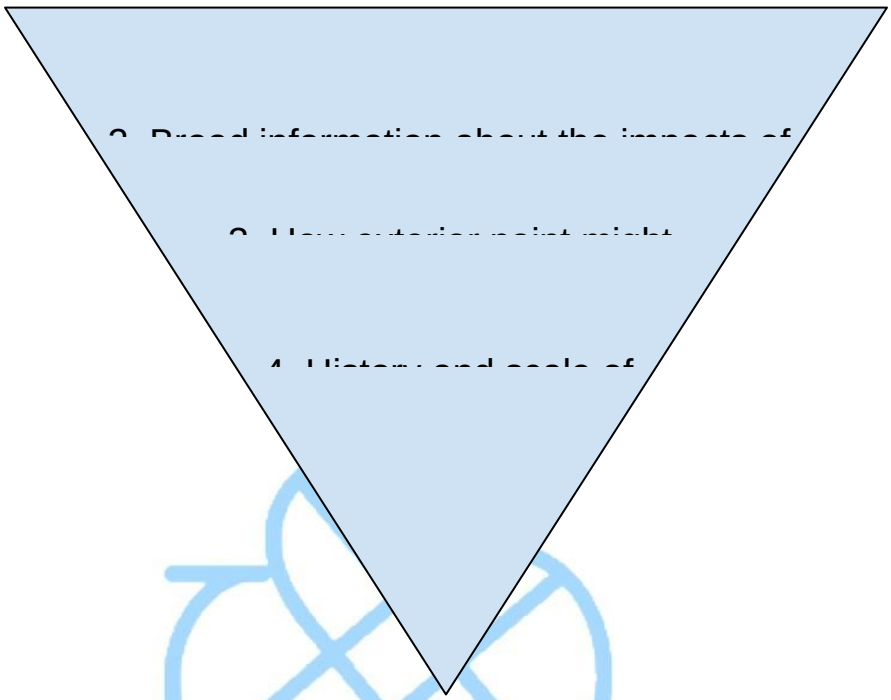
- "when an employee uses emojis in their work communications it can actually make them look less competent. Aside from the impact to an employee's reputation, there is a more sinister issue HR needs to be aware of. Emoji use has now become another potential avenue for sexual harassment." (page ..)

- Remember that **the point of an annotated bib is to make it easier for *you* to write your paper later, and organize it accordingly.** Setting up a good annotated bib *means far less work later on*, and will let you focus more on the finer points of writing instead of trying to remember what your source was saying.
- **Strong annotated bibs draw on multiple lenses**. You should have sources in it you plan to model your research method off of, sources to use to establish historical context for your question, sources that say the answer to your question is A, and sources that disagree to varying degrees, etc. Try and have that kind of diversity while still primarily drawing on reputable academic sources.

WRITING THE INTRO/CONTEXT SECTION:

- The point of your this section is to **'set the scene'.** This will be the leading part of your paper, so you want it to be *impactful*. You want to **explain why your answer- whatever it may be- *will* matter.**
    - For example, let's consider Joe's research question of "What are the environmental impacts of painting the exteriors of ocean-front properties?". Here, it's pretty easy to come up with good ideas for your context section. After all, the question is literally asking about the impacts of something, which is often a major part of your context section. In other cases, it might be more difficult to think of impacts. Trying to connect your research to a broader audience by drawing on **scale** (how many people will be impacted by your issue), **timeframe** (if your issue will worsen or reach a 'breaking point' in time), and **magnitude** (how much it will impact the people that it does affect) is usually a good fix. Joe's section would probably talk about how many exterior ocean-front properties are painted (scale), he might include a projection of how many more properties will be painted (timeframe), and he'd probably want to include some basic citations explaining how exterior paint might hurt the environment (magnitude).
- The contextual section is also often where you can explain the **historical 'record'** of your question or problem. This can help to establish how your issue came to be and/or why it hasn't been solved/answered yet.
    - Joe might want to write a little about the different paints that have been used throughout the years, or why people ever started painting ocean front exteriors in the first place.
- But beyond the broad impacts that could be associated with your study, your context section should also be answering this question: **What's the problem/disagreement/issue?**
    - Once you've set up *why* you're researching your topic, very briefly take the time to explain what exactly the issue is. Start **broad, and narrow down**. Think of your writing as an upside down pyramid. The beginning of your research paper

should be quite broad, but by the end of your context section, you should have narrowed down to your specific issue or question so that you can stay focused on that from thereon. Here's what Joe's context section might look like:



- The important thing to remember is that if somebody reads your context section and doesn't understand what your research is about, or why the answer might matter, you haven't done your job.

WRITING THE LIT REVIEW:

- Writing your Lit Review (aka "Professional Conversation" section) is similar to writing the IRR from AP Seminar. Here, **you'll be explaining what current researchers have found about your topic, their arguments, how their conclusions relate to one another, and the research gap that you plan to fill with your own question**. It should be logically organized (perhaps by different subtopics, or 'chronologically' if that applies) so that it clearly flows.
- You will cite the majority of your sources in this part of the paper. It's also likely that this part of your paper will be the longest section. Having spent enough time and becoming familiar with your sources while researching will make writing this far easier. If you wrote a good annotated bib, you can often use only that document (and those sources for reference if necessary).
- The most important thing to keep in mind throughout this section is to be sure that you're **putting your sources 'in conversation'** with each other. L**ay out their arguments and clearly explain how each relates to the other**. Do *not* simply say that "A and B disagree about X, but B's study is correct because it was better." You must explain *why* B is correct, and/or exactly what A and B really disagree about. A good example of putting sources in conversation can be found in this excerpt from Joe's lit review:

"B wrote that 85% of exterior paint chips off from dirt that blows against exteriors and wears them down (citation). However, A found salt and sand are the most damaging environmental factor to paint (citation). As B's study was conducted in a prairie region, and A's study simply compared the corrosive impacts of different materials on paint in a lab environment, it seems likely that salt and sand are in fact the most corrosive, even if not found in all areas. C's study conducted across prairies and deserts supports A's theory, and quantifies that sandy areas experience 32% quicker decay (citation).

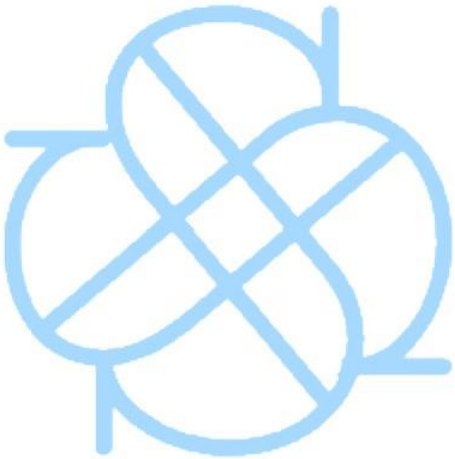
Importantly, both A and B claim that the paint's decay becomes not only a noticeable aesthetic issue once 20% of the paint has worn away, but an environmental concern (citations). D expanded on this and argued that the paint flecks are particularly damaging to the environment once they enter the waterways. That supports the conclusion that paint decay might be a particularly

impactful problem in seaside areas due to quicker decay and the paint flecks' close proximity to the water. However, E argues that.."

    - The highlighted sections in the above example are where Joe himself offered additional commentary or explanation. Note that they **build off** of what the 'researchers' had already claimed, and serve **to clarify the disagreements and conclusions** between the researchers.

- Beyond laying out your sources' arguments and how they relate to each other, it's also important to **explain how the current research relates to your own.** This means establishing the research gap you'll be filling, or the new lens you'll use.
    - Going back to Joe's paper, different researchers have separately established that a) paint erodes off of exteriors quickly in sandy environments and b) that paint flecks in waterways are bad for the environment. In other words, his question "What are the environmental impacts of painting the exteriors of ocean-front properties?" has *almost* been answered. His research gap to establish is relatively narrow:
        - Joe needs to explain that ***although there is certainly evidence available*** about sandy areas, ***it doesn't take into account all of the nuances of his specific question*** about seaside areas, such as the seasalt, winds, and weather patterns common to oceanfront areas, and as such, ***his study is necessary to fill the 'gap' in current literature*** about the environmental impacts of oceanfront paint decay.
- If it's helpful, use the italics in the previous example to form your own section explaining the research gap you're investigating. But, if after working through your lit review, you're finding it difficult or impossible to argue that your original question needs you to answer it, you can always change it. Don't be too drastic, but looking at your topic from a new angle, lens, or scale can often help. For example, if Joe decided his research gap wasn't wide enough to warrant a study, he could ask:

    - "What are the environmental impacts of the fumes released by painting the exteriors of ocean-front properties?" (new angle)
    - "What are the fiscal impacts of exterior paint decay in sandy areas?" (new lens & different scope)
    - "What are the environmental impacts of painting the exteriors of ocean-front properties on a small Northeastern town?" (smaller scale)
- Of course, if you try to rewrite your question, be sure to go back and run it through the checklist in the previous "Writing a Research Question" section.



DESIGNING & CREATING YOUR STUDY:

Validity of Results

- When designing your study, you want to create a study with the 'strongest' and most valid results possible. There are many issues that can hurt your study's ability to make a point, but first, you need to understand two terms.
    - The first is **Internal Validity**, or the extent to which a study establishes a trustworthy cause-and-effect relationship between a treatment and an outcome.
    - **External Validity** (or **Transferability**) is how generalizable your results are to other situations or the broader population.

- The **7 main threats to internal validity** are:
    - "**History**". This is when outside events might change participant's responses to a study. A historical threat to internal validity in Joe's study might be something like a hurricane.
    - "**Maturation**". Maturation is the long-term changes that might change results. In the context of living subjects, a standard example is aging. However, maturation can also refer to changing external factors over the study, such as different economic or political environments.
    - "**Attrition**" refers to the loss of participants or sample size as the study progresses. This is a major factor in some medical studies as participants die.
    - "**Instrumentation**". The tools or methods chosen to measure your variables need to be as accurate as possible. A sleep study could turn out very different if participants self-reported their sleep habits versus wearing sleep monitors.
    - "**Learned Testing/Testing**". Participants which take the same test or complete the same task more than once often simply become better at it with experience. A study about intelligence would need to find a way to show that improved scores were because of the independent variable, not familiarity with the test.
    - "**Regression**" occurs when participants are chosen from the extremes of a population. This principle argues that extremes will always gravitate and move towards the mean of the greater population. If a sports study was trying to find out if players got better when wearing Nikes, but choose only the bottom 10% performers, Schnell argues that the players *would* get better, but because of regression to the mean, not necessarily the shoes.
    - "**Experimenter Bias**". Experimenter bias occurs when an experimenter expects an outcome. It often presents itself when an experimenter treats different participants differently (other than the independent variable), or errs in selecting, recording, or analyzing data.
- **External validity is often threatened by:**
    - "**Ecological Validity**" refers to whether the situations used in your study are representative of the 'real world'. A computer simulation of making a pizza would have very low ecological validity.

    - The "**Situation Effect**" including time of day, location, or setting might lower ecological validity.
- Both internal and external validity are threatened by **Sampling Effects**.
    - Sampling effects not only make it more difficult to argue that results are causatory, they can also make your results less applicable to the broader population. There are **3 main types** of sampling effects:
    - "**Inclusion/Exclusion**", or the decision of who is/isn't included in a study. A study about the impacts of a heart medication might, for example, exclude anybody who doesn't have a specific heart condition. This is often seen in the acronym WEIRD, or Western, Educated, Industrialized, Rich, and Democratic. (The WEIRD group is actually seen so often that Bhandari writes 96% of psychology studies are WEIRD despite the WEIRD group only making up about 12% of the world's population.)
    - "**Probability**". The likelihood of a person being chosen for a study. Studies with many existing parameters to limit the chance of participant selection make that percent smaller. A study lacking equal probability of selection is likely to yield narrow results that are only proven in the case of a small group.
    - "**Sample Size**". The brute size of your study. Studies with low probability or a high rate of exclusion benefit from having much larger sample sizes. That's because the larger sample size is more likely to be representative of your target/general population. Small studies, even those with little to no parameters, often aren't representative of the population and thus have low validity.
- In the end, **an ideal study has no threats to either internal or external validity**. Unfortunately, that is **impossible**. Try to get your study to avoid as many of the threats listed above as possible, but **recognize that no study is perfect**, and that is why the Limitations section is so vital. It helps to explain why your study wasn't perfect, so that future researchers can try to compensate for those factors, or at the very least be aware of them.
- The absolute best way to mitigate these threats is **replication** of your study. This is especially true when the replication is done across different contexts.

Choosing A Study Type

- Choosing a study type should take into account one, obvious, main thing:
    - **How can you gather the data you're looking for?**
- For example, if you're looking for qualitative data, interviews or a survey are often the right format. But once you've chosen what kind of data you're looking for, it's helpful to use the same or similar format that researchers before you have used. At the end of the day, **it's easier to argue that your study type was the best choice to answer your research question if you've based it off of what experts have already done.**
- Research falls into 4 main categories, ranging from most observational to most experimental.
    1. **Descriptive.** Seeks to describe the current state of an identified variable. Ex: Description of teenage vape use.
    2. **Correlational.** This type of research recognizes trends and patterns in data, but doesn't prove causes for these observed patterns. Ex: Relationship between ACT scores and freshman year performance.
    3. **Quasi-experimental.** An independent variable is identified but not manipulated by the experimenter, and effects of the independent variable on the dependent variable are measured. The groups within this are not randomly selected and are instead based off of pre existing conditions or variables. Different from correlational because groups with the variable are compared against those without it and because researchers seek to know the *impact* of a variable, not just their relationship. Ex: the impact of gender on algebra performance.
    4. **Experimental.** Any study where all variables are identified and control is attempted except for the independent effort. Uses the scientific method to attempt to prove a cause-effect relationship. Groups are made up of randomly selected participants. Ex: the effect of positive encouragement on academic performance.
- That said, here are **4 of the main study types** and a brief description of how each works. Any of the following could be used in a very observational or experimental context; it simply depends on how the independent variable is treated.
    - **Cross-sectional:** often use surveys. A representative population is surveyed or examined to find information about them at one point in time.

    - **Case studies:** often use interviews or surveys. A single case (person or small group/incident) is examined to determine information about the situation.
    - **Correlational:** can use surveys or analysis of data. Tries to determine if there is a relationship between factors.
    - **Longitudinal:** look to see how factors might develop/change over time.
    - **Mixed-method:** refers to research which gathers both quantitative and qualitative data.
- **A study could be more than one of the descriptions above.** It is perfectly plausible to conduct a longitudinal mixed-method correlational study about sleep deprivation, for example. The researcher could check in over time (longitudinal) and ask about how participants feel about their sleep deprivation (qualitative) and some health data points (quantitative) without ever manipulating the independent variable of sleep deprivation and instead simply focusing on the relationship between sleep deprivation and wellbeing (correlational).
- Once you've chosen your study type, be sure to **read the analyzation section(s)** that pertain to your study so you can choose what measurement level(s) to use.
- **Example:** In Joe's study, he wants to gather data about paint flecks in the waters (quantitative). Joe's study will fall into the correlational category, because it isn't feasible for him to control the independent variable of how much paint fleck pollution there is, and he isn't trying to prove a cause-effect relationship (which would be a quasi-experimental study). Instead, with the data he collects, Joe plans to investigate the relationship of painted seaside exteriors and pollution levels of paint flecks in the water. From this, Joe decides he is going to conduct a quantitative correlational study.

Ethics

- All studies have to be ethically sound, and many of your studies will involve human participants, which only heightens the stakes. You should check that your study is ethically sound *prior* to actually conducting it.
- Ensuring your study is ethically sound is often best done by running it through an ethics review board. Your school district might have a board or a designated review person to

look over your study and give/deny permission to carry out your study. They'll be able to impartially offer feedback about the ethicality of your study.

- Research ethics principles all stem from the two all important principles of **"Do no harm"** and **"Do good".** Following are the **5 principles of ethical research** with participants.
    1. Obtain informed consent from research participants
    2. Minimize risk of harm
    3. Protect participants' anonymity and/or confidentiality
    4. Avoid deceptive practices
    5. Participants have the right to withdraw from research
        - Laerd (http://dissertation.laerd.com/principles-of-research-ethics.php) offers *excellent* elaboration on how to ethically conduct research with participants. Read it for more information on the above 5 points and how to execute them. The NIH (https://www.niehs.nih.gov/research/resources/bioethics/whatis/index.cfm) elaborates on more ethical principles that apply to all research and provides examples of research where you can try and find ethical/unethical practices. An important principle noted there is the extremely important "Openness" which refers to the need to be clear both with other researchers (and with your participants, if applicable).
- There are different research codes in existence for different kinds of research. The NIH site has a list of many different codes, but if you can't find one that applies to your field/study, look it up.
- Other than research with participants, it of course isn't ethical to misrepresent your data, plagiarize, or infringe on copyrights/patents/intellectual property.

Gaining Informed Consent

- Gaining informed consent from all participants is a vital part to any study with humans. It's important to do so *prior* to actually conducting your study, and to make sure there are **records** verifying that you have gained consent. Your **consent forms** should have at least **6 sections**, explaining:

1. **Purpose of study**- What is your study trying to find out?
2. **Procedures of collecting and storing data-** The method of your study. Survey? Interview? Observation? Explain what you'll be doing to get data and where/how you'll be recording it. What will participants be asked to do?
3. **Right to refuse to participate-** Make an *explicit and obvious* note of the fact that their choice to participate is completely voluntary and can be withdrawn at any time.
4. **Potential benefits/harms of participating-** If you have a very low risk study, you could write something close to "This research is considered to be minimal risk. That means that the risks associated with this study are the same as what you face "every day" here, or you should list any possible risks. As for benefits, you could explain here if there will be any compensation, or how the research field/the participant would benefit from their participation. (Some researchers split this into two sections, benefits and risks.)
5. **Privacy/confidentiality-** Write here about how you plan to protect their privacy, how you will/won't be sharing their information, and what you might do to anonymize any data you gather from them.
6. **Contact information & update opportunity during study-** You should leave your and your advisor's contact information so that participants can reach out to you with any questions/complaints as time goes on. Many researchers also make a note of the fact here that if participants want to see the results of the research once the study is over, the participants should reach out to them.

- **Your consent forms must be in laymens' terms.** Another thing of note is that children cannot consent to a study, though they can 'assent'. That means that if you plan to conduct a study with minors, you need to create parental consent forms as well as assent forms for the minor. The parent form will be very similar to the adult form layout above, but of course should be written in the terms of 'the risks to your child include...', 'your child's information will be stored…', etc. The child assent form will also be very similar, although you should take particular care that it is written at an age-appropriate level.
- At the end of any kind of consent form, include a **place for the participant to print and sign their name with the date.** Directly above that, include a final section that says

something close to: “I understand what I’m being asked to participate in and I give my consent to be a part of the research study”.

- Make sure that you store your signed consent forms in a safe location for future reference if necessary.

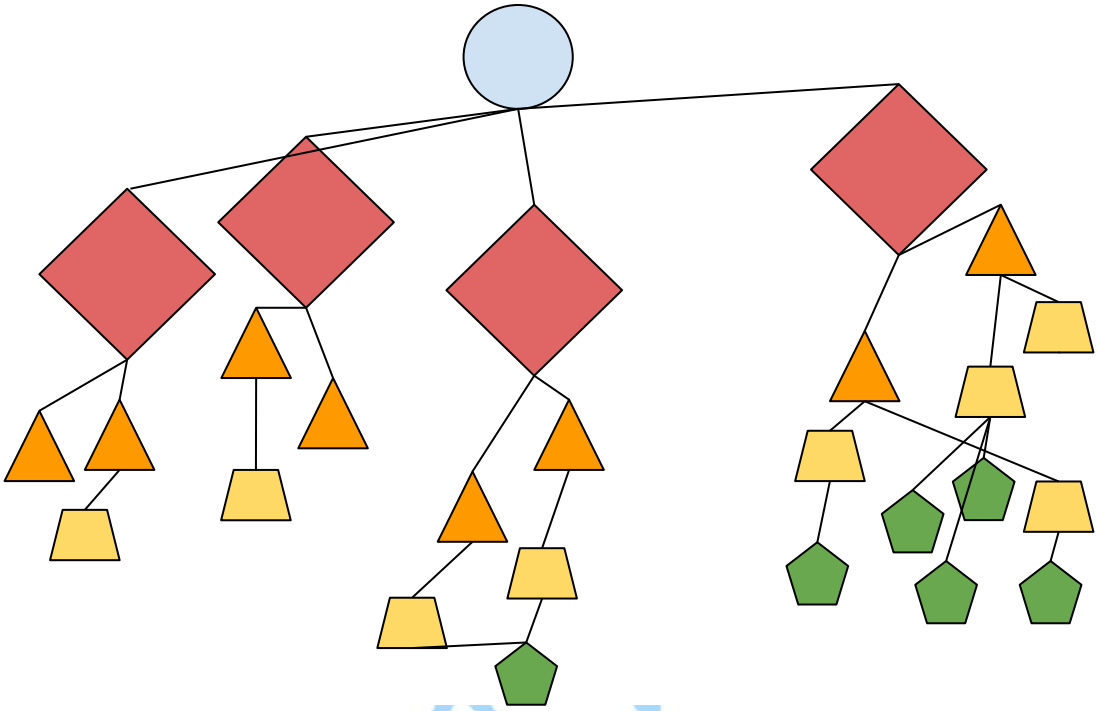
Writing A Survey

- Writing a good survey, unfortunately, sounds easier than it is. The key is to **stay focused** on information that’s actually relevant to your question.
- When you’re writing a survey, there are several different kinds of questions you can use. Most researchers include a **demographics section** so you can explain what kind of participants took your survey. This will probably include questions about race, gender, or age. These questions are closed-ended.
- Finding **numerical** data about people’s opinions is often done using the **Likert Scale** (https://www.simplypsychology.org/likert-scale.html#:~:text=The%20most%20widely%20used%20is,disagree%20with%20a%20particular%20statement). You can also ask more **qualitative** questions that leave room for people to give an open-ended response. Some researchers will also ask agree/disagree questions, multiple choice, matrix questions, or even ranking questions. There are of course advantages and disadvantages to each style, so many researchers will ask multiple types.
- **The most important thing to keep in mind about any question you ask is to not ask it in a biased way**. That includes not asking leading questions (ex: “Hasn’t not sleeping at your normal hours made you more tired?”), double questions (ex: “Does our president make you disappointed in our political system and angry at all nonvoters?”), and questions with an assumed premise (ex: “Why are you angry at the system?” when it hasn’t been previously established that they are in fact angry).
- Here are some other **good tips** to keep in mind:
    1. Ask at least one question that checks if the people who are answering your survey are actually legitimate in their efforts. Ex: Ask participants to “Type the 3rd letter in the word ‘order’”.

2. For age, income, etc, it's easier to have them input a specific number so that you can use google sheets tools to find the average.
3. Do *not* have overlapping categories people can choose from. Ex: If you ask about age, 18-25 and 25-35 should NOT be options because a 25 year old could choose either.
4. Before actually spreading your survey, get a trusted critic to 'take it' so they can give you feedback. Delete their results, especially if you make any changes to it!!

Spreading A Survey

- Spreading a survey is very important because it determines who actually participates in your study. There are several different methods you can use.
- One of the most common is the **snowball method**. That's when you get all of the people who've taken your survey to share it with all of their contacts, and so on.
    - The issue with this type of survey is that in your different 'generations' of research, you're falling for each friend and their group. If you consider the generations laid out below, and consider them to be laid out on a left-right bias map, then you'll notice that the right most red piece from generation 2's resulted in a chain of right-bias throughout all of their acquaintances. And since, of course, it's impossible to control if people actually choose to spread/take the survey, it's easy to get results that don't actually represent the population at large. That's seen in the leftmost chain.



- However, the snowball method is valuable because it's easy and free. If you have a long contact list of people from different areas of your life (family, family friends, your friends, sports, etc), that should help initially diversify your generation 2 and all of the responses after that.

- UK(https://www.healthknowledge.org.uk/public-health-textbook/research-methods/1a-epidemiology/methods-of-sampling-population) provides helpful insight into different sample methods. In short, **the survey method you choose will influence what parts of the population are actually involved in your survey, and thus your biases**.
- A brief overview of **7 different sampling methods**:
    1. **Simple random sample (SRS)**- each and every member of your sample frame of your defined target population has an equal chance of being chosen to participate in your study.
    2. **Systematic**- every nth member of your sample frame or population is chosen to take your study.
    3. **Stratified**- Population is divided into strata (subgroups) based on a characteristic (such as gender), and a certain number of *participants* from each strata is sampled (either a defined number or a certain percent of each strata).

4. **Clustered**- Populations are divided into subgroups, and random *subgroups* are chosen to participate in the study.
5. **Convenience**- Participants are selected based on willingness to take part. (Standing in a public space and asking for participants is a form of convenience sampling). Prone to volunteer bias.
6. **Quota**- Researcher tries to find a certain number of participants with certain characteristics.
7. **Judgement**- Researcher decides who should and shouldn't be part of the sample based on certain characteristics and judgements to try and find a representative sample.

- Practical ways to spread your survey include:
    - SurveySwap (https://surveyswap.io/) is AMAZING!! You take other people's surveys to earn credits so that others will be able to see and take yours. If it would take too long to take a bunch of other people's surveys, you can buy credits from SurveySwap, and whenever you have enough responses you can ask to return any extra credits you have for a full refund. (Unlike many other sites, you do *not* have to create your survey with SurveySwap in order to use it, and you do *not* have to pay them at all).
    - Start a **snowball survey** method through email.
    - There is a reddit thread called **r/SampleSize.** You can repost a request for others to take your survey every 24 hours. Generally friendly and helpful community.
    - There is a facebook group called **Student Survey Swap** which I was invited to join. It's made up of other researchers seeking responses to their surveys. You'll have to request permission to join, but it's worth a shot: it works on a you-do-my-survey, I'll-do-yours basis.
    - Stand in a **public space** and try to get anybody who's interested to take your survey. Or, ask every nth person. (Convenience, systematic).

    - Get an **authority** (teacher, principal, boss) to mandate/suggest their underlings take your survey. (Likely to result in clustered- if only classrooms A and C take the survey but B doesn't, for example).
    - Leave the **link/access to your survey on poster boards** and anywhere your target community might see it and want to take it. Convenience.
- The most important thing to remember with whatever method you choose is that **there will be inherent biases** to it. Make sure you recognize and describe these in your limitations section.

Interviewing

- Interviewing is a helpful research method when you're trying to gather qualitative information. Although you're likely to **gain much more in depth information** than you would through a survey, you'll also **have a much smaller sample size**.
- When asking someone to take part in an interview, be sure that you're clear about how **long** you expect it to take, the **purpose** of the interview, and if you expect **any sensitive topics** will come up. Wherever you interview the other person, you should make sure that they feel **safe, comfortable**, and that their **confidentiality** is preserved. If you want to record the interview, be sure to get their explicit permission to do so. Either way, keep careful notes throughout.
- During the interview, **keep careful notes**. At the beginning, it's useful to ask fact-based questions to help your participant get 'warmed up' and quickly engaged. After that, asking **open ended, bias free questions** is usually your best bet to get the most and most relevant information. Refer back to the "Writing a Survey" section above for more help writing unbiased questions.
- After the interview, send an initial thank you! It's common courtesy, and you never know if your participant will become a valuable contact later. It's also a good idea to send a final thank you and your paper once your research paper is completed.
- McNamara's (https://managementhelp.org/businessresearch/interviews.htm) article provides more in-depth information on how to conduct a successful research interview.

Research With Existing Data

- This type of research can be tricky because although you're using information that's already been gathered (and likely analyzed one way or another!), **you need to compile it and/or attack it from a new angle or lens in order to give any benefit to your research community.**
    - However, *these can also be easier than the other methods* in some ways, namely because it's usually much more time and cost efficient to use data that has already been gathered.
- There are three main types of this research, as described by Grady, Cummings, and Hulley's invaluable paper.
    1. **Secondary data analysis**. The use of existing data to answer a research question which wasn't the purpose of the original data collection. Cheong's paper elaborates.
    2. **Systematic Reviews**. Combines the data gathered by different studies investigating the same research question and offers a more specific/thorough analysis of answer(s).
        a. **Fair warning:** systematic reviews may seem easy upfront, but they are *NOT*. They require an *extensive and expert* knowledge of the existing body of knowledge and a *strong* grasp of statistical analysis methods. In the real world, these are *very rarely if ever* conducted by a single researcher. AP Research students should be very cautious of pursuing a systematic review design unless they have previous experience or have at very least read and thoroughly understood Semantic Scholar's PDF.*** That said: they often present some of the most valuable data and conclusions, as their compiled research allows for more thorough investigation and helps to limit the effect of any certain bias found in a given study.
    3. **Ancillary Studies.** Add measurement(s) to part or all of the original population to answer an additional research question.
- To choose between the above methods, you first need to see if there's an existing data set which could answer your question.

1. **If so, but the data wasn't intended to answer your question, you'll be conducting a secondary data analysis.** To conduct a secondary data analysis, it's extremely helpful to have a mentor figure who can help point you to a data set that will help you answer your question. Often, the dataset you'll use for a secondary data analysis will be very large/broad- for example, the death registry of the US, because it could be used to answer many research questions.
2. **If so, and the data comes from multiple data sets all intended to answer your question, then you'll conduct a systematic review**. To begin a systematic review, you must create and follow a *well defined method of choosing what research/datasets to include*. This is what separates a systematic review from a literature report: you *will* follow the method you define to include/exclude all possibilities from your review. Generally, data will be abstracted from the included studies by two different independent abstractors, with disputes settled by a third. However, this isn't really a possibility for AP Research, so you'll have to abstract your own data in the most unbiased manner possible. The guidelines you create for choosing which data to extract must be well defined and followed to the letter. If necessary information (such as if the study was blind or not) that would help you calculate confidence in their results cannot be found, those studies are generally excluded.
    a. After your data has been abstracted, **important characteristics** (such as population, sample size, methods, etc) **of each included study should be presented in tables**. Next, the analytic findings of each study should be presented in tables or figures. These often include relative risk, odds ratio, risk difference, confidence intervals, P values, etc.
    b. Lastly, you should present **summary estimates and confidence intervals along with subgroup analyses and sensitivity**. This summary effect estimate is the main outcome of your meta analysis, but should only be discussed in the context of the individual studies. Importantly: readers should always be able to form their own conclusions off of the data presented without relying on your meta analysis.

    c. For a *far* more in depth look at how to conduct a systematic review, check out Semantic Scholar's section on it.

3. **If not, you'll be conducting an ancillary study,** where you have to add data to fully answer your question. Ancillary studies are extremely helpful because you're almost definitely adding something to the research field by adding question(s), but you also get to rely for the most part on already gathered data. This makes them very efficient. The first step of an ancillary study is to identify another study that you could add your question(s) to. For example, a study on the learning styles of gifted children with a scheduled follow-up in 6 months presents a good chance to add a quick survey to their follow-up asking your own question(s) about their mental health. Once you've identified said study, get in contact with the researchers and obtain their cooperation -- they'll often agree if your ancillary study does not interfere with theirs and seems like it'll help answer another important question. (Many places will have an established procedure for applying to conduct an ancillary study, take care to follow it.)

Running An Experiment

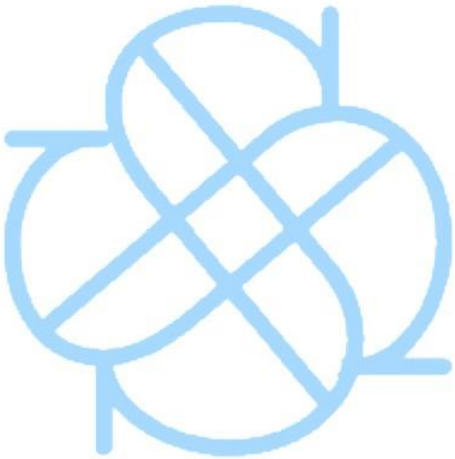
- Essentially: **follow the scientific method and make sure that your study is replicable**. What sets this section apart from the others is your *active intent in controlling the independent and other variables*, and often your inclusion of a control group. The best way to do this is to write a solid methods section and see if someone else who reads it fully understands how to conduct your study. Blakstad and Emerald Publishing offer more information on experimental research.
- When conducting a true experiment, it's extremely valuable to base your design off of previous well established researchers.

WRITING THE METHODS SECTION:

- This part of your writing should be relatively easy because you often don't have to cite many other researchers, you don't have to interpret any content or data, and you don't have to come up with anything new at this point. **The only thing you have to do is write out and justify exactly what you've done and how you plan to analyze your data.** This section must be clear and concise. Most researchers can put a decent rough draft of their methods section on paper in under an hour. The key is to stay focused on your **goal**: **clear, concise, justified, and replicable methods**.
    - VeryWell https://www.verywellmind.com/how-to-write-a-method-section-2795726 offers some guidance.
- The methods section is usually organized into several subsections. These often include Participants, Materials, Design, and Procedure.
    1. **Participants**. Explain (at very least) the basic characteristics of who was involved (age, sex, race, geographic region), how many were involved, and how participants were chosen. Ex: '30 eight to ten year old participants were randomly chosen from each of 10 elementary schools in the southern Arizona area, resulting in 300 total participants. 52% were female…'
    2. **Materials**. Describe any materials, measurements and equipment used. What was used in your study? You don't need to be super specific about ubiquitous items (for ex, you can say "a laptop" instead of "a grey 2009 macbook with a case"), but you should be specific enough that others can replicate your work. If you use the same stimuli or work from another study, be sure to reference and specify that. Ex: 'A website built by the researcher hosted the survey. Participants took the survey on their own devices, and data was stored in google sheets protected by a randomized 10+ digit password. The article used in group A was taken from Joe's (2017) study and was used to determine how participants would react to blatant racism in…'

3. **Design**. How was your study set up? Here, you clearly define your variables and the groups in your study. Explain if you used a within or between groups design, what kind of data you were gathering, and what type of study you chose. Ex: 'The study used a 4x2 between groups design to conduct a longitudinal correlational study. Quantitative data was gathered to determine the impact of the independent variable, milky way consumption, on the dependent variables of heart attack risk..' *Note: factorial design is essentially a description of how many independent variables and 'levels' you're assessing. Explained further here.*
4. **Procedure**. What are you going to actually *do*? Take it step by step and create a cohesive and concise narrative for how to conduct your study. Ex: 'The researcher gathered seawater samples at 8 different locations along the NE seaboard monthly for a period of 4 months. Samples were pulled from between 3 and 4 feet underwater by the niftydevicething and were then stored in sealed bottles for transport back to the lab. Once at the lab, each sample was..'

- Some studies, like Joe's might not need to include a participants section because he didn't actually have any participants in his study of water quality and paint fleck concentrations.
- If you're peer reviewing a friend's method section, or **if you want to check if you got the point of the methods section, ask yourself these questions.** (*Hint: the answers should all be yes!*)
    - Is it all written in past tense?
    - Did the researcher introduce the method they selected? Did they defend this choice?
    - Did the research describe the instruments used to conduct their research? Did they justify and explain the development and/or selection of these instruments?
    - Did the researcher provide detailed information regarding the participants/items studied? Did they justify those choices?
    - Did the researcher explain, in detail , the procedure taken within their research process?
    - Could the research be replicated without the aid of the researcher? If no, the procedure is not explicit enough.



ANALYZING QUANTITATIVE DATA:

- The first step to analyzing your data is to get into a **format** you can work with. Spreadsheets from google or word will likely be powerful enough to use. Once you have a sheet with all of your data compiled, it might also be necessary to create sheets for any manipulated subgroups (for ex, a sheet where the IV was applied and a control sheet). That way, whatever analysis you might perform on the control you can compare with the

same analyses on the sheet(s) where the independent variable was applied. If there's significant differences in the results of your two groups, then it appears your independent variable did have an impact.

- Some students might find it valuable to also create sheets of other **natural subgroups** that aren't evenly spread throughout the population or that you suspect impact results (a male and female sheet, a 'rich' and 'poor' sheet, a boss and employee sheet, etc).
- For example, if a snowball survey ended up gathering data from 70% males and 30% women, it would be important to find out the impact of gender on the results. If the same analyses found that men and women didn't noticeably differ from each other, then your composite results would be stronger because you would be able to say that 'although a disproportionate amount of men were surveyed, gender didn't appear to have a significant impact on responses'. Or, if your survey is on what kinds of clothing should be allowed in the workplace, and you survey 50% males and 50% females, it's probably reasonable to assume that men and women would in general have different opinions. Check that assumption by **performing the analyses** on the male/female sheets and seeing how/if the results differ. **The point of these subgroup analyses is to determine how strong your overall results are if you've over/under represented a population and to see if there's another variable that might have had a noteworthy impact on your results** (which you could then talk about in your discussion section, and likely your limitation section).
- **Stay organized throughout!** Keep a piece of paper next to you or open up another document just to keep notes on the data you're finding and what trends/results it might suggest as you go. This'll be a lifesaver later, especially if you're working with lots of data or numbers. Be sure to clearly label everything!
- Before, during *and* after your analyses it can be helpful to create **graphs**. Sometimes, creating a graph of data before actually analyzing it can help you decide whether or not it is necessary to perform an in depth analysis. Afterwards, graphs provide good visual aids to both you and readers, and serve as an excellent organization tool.
- As for what analyses you're actually performing, **it's usually best to model after what other researchers in your field have done.** That lends credibility to your own methods and makes your results more easily comparable with others. Of course, if you're trying to

attack a new angle, this isn't always possible, but try and find a study from *somewhere* that does something similar to yours (and reference it in the methods section!).

- In general, the first analysis you'll perform will determine at least some of the descriptive statistics, and many of your studies will require crosstab and chi square analyses, which is why they are in depth discussed below.

Data Measurement Level and Descriptive Statistics

- There are four levels of data measurement you must understand before you can calculate your descriptive stats. They are:
    1. **Nominal**. Data is basic classification with no order/hierarchy. Ex: male/female.
    2. **Ordinal**. Data has a logical order but no consistent interval between values. Ex: leadership position in a business.
    3. **Interval**. Data is continuous with a logical order and consistent intervals between values. *There is no natural zero, so ratios are meaningless: cannot say that somebody agrees 2x as much as somebody else because there is no absolute zero*. Ex: the Likert Scale, or degrees fahrenheit.
    4. **Ratio (scale)**. Data is continuous, in a logical order, has consistent intervals, and has a natural zero. *You can meaningfully compare ratios and say that something is 2x as tall as something else because a natural zero exists.* Ex: height, weight, age.
- Descriptive statistics are the **mean, median, mode, frequency, min/max values**, and **percentages**. Useful resource Pell explains a mean should only be calculated from interval and ratio data and a meaningful median should only be calculated from ordinal, interval, and ratio data. Mode and min/max values can be calculated for all levels of measurement.
- For studies that aren't meant to be generalized to the population, these can sometimes be enough. Otherwise, read through these subsections describing and explaining some common forms of quantitative analysis:

Cross Tabulation and Chi-Square

- Also known as **crosstab**, and is used to analyze categorical data. Useful in trying to determine if there's a correlation between variables. It is a table that records the

intersection of how many people have characteristic(s). (Qualtrics elaborates.) A crosstab might look like this:

	'Poor'	'Middle'	'Rich'	
Male	18	2	8	28
Female	12	12	4	28
	30	14	12	56

- To analyze, you should use a **Chi square analysis**, a correlation test for categorical values. If the two variables are independent, then the results will be "non-significant", and if they are related, then the results will be "statistically significant". Watch Khan Academy's video on chi square analysis before reading the following.
- **Essentially: the null hypothesis is that there is no relationship between variables. Rejecting the null hypothesis means that you can state there is a relationship between variables** (and then you can go on to look at what kind of relationship etc). To reject the null hypothesis, your results have to have a very low chance of occurring naturally (for example, flipping a coin 50 times and getting heads only twice has a low enough probability of occurring by chance that it would probably reject the null hypothesis). Many fields/studies use **5%** as their threshold, or alpha level: if your study had a less than 5% chance of occurring randomly, then it rejects the null hypothesis; if your study had over 5% chance of occurring randomly, then we fail to reject the null hypothesis (and it remains possible that there is no relationship between the variables). The lower the percent your study has of occurring randomly, the more it suggests that your variables have a relationship. Use Illinois as a resource for chi squares.
- **To actually perform a chi square analysis, you must:**
    1. State your null hypothesis and an alpha level. In this case, it would be "Wealth is independent of gender". We'll set the alpha level to 0.05, or 5%.
    2. Calculate your degrees of freedom. This equals (rows-1)*(columns-1), or in our example, (2-1)*(3-1) = 2 degrees of freedom.

3. Look up a chi square table and use your degrees of freedom and alpha level to determine your critical chi-squared value. Our example's critical value is 5.99.
4. Compute your test statistic. To do this, you need to have calculated your marginals (the sums of the different categories on the edges of your table). Then, you need to calculate the expected frequencies of every cell. (Expected frequencies are explained well here.) Expected frequencies can be calculated using the formula of ((row total)*(column total))/total. The expected frequency of the top left corner of our example would be calculated as ((28)*(30))/56 = 15. Our table now looks like this, with expected frequencies listed below each value:

	'Poor'	'Middle'	'Rich'	
Male	18 15	2 7	8 6	28
Female	12 15	12 7	4 6	28
	30	14	12	56

5. Now, calculate the chi squared statistic doing these operations to each cell and then adding all of the final values together.
    a. Subtract the actual value of each cell- its expected value
    b. Square that answer
    c. Divide the above by the expected frequency

    For the top left corner, the value that would be added to all of the other cells would be calculated $(18-15)^2/15 = 0.6$. The bottom left corner would be calculated (12-15)/15 = 0.6. The table would now look like this, with said value put under the other actual and expected value:

	'Poor'	'Middle'	'Rich'	
Male	18 15	2 7	8 6	28

	0.6	3.57	0.666	
Female	12 15 0.6	12 7 3.57	4 6 0.666	28
	30	14	12	56

Adding all of these values together (0.6+3.57+0.666+0.6+3.57+0.666) gives us a sum and thus chi squared statistic of roughly 9.672.

6. Finally, compare your critical value against your chi squared statistic. If your chi squared stat is larger than your critical value, you must reject the null hypothesis. If it is equal to or smaller than your critical value, you "Fail to reject the null hypothesis". In this case, 9.672 > 5.99, so we reject the null hypothesis and conclude that there *is* a relationship between gender and wealth.

Other Methods

- There are **4 more commonly used analysis methods**. You should probably only use these if you have a strong grasp of statistics and access to the computer programs often used to carry them out (for ex, SPSS, SAS, STATA, and MINITAB). This page links to more thorough descriptions of the first three methods (and itself links to more resources for each).
    - **Correlation**. Describes the nature of relationship. Strong/weak, positive/negative.
    - **Analysis of Variance (ANOVA)**. Determines whether the differences in means of 3+ groups is statistically significant. Must later perform a post hoc test to see which 2 groups differed. Laerd elaborates.
    - **Regression**. Determines if your manipulated variables are predictors of outcome. Statistically controls every variable in your experiment. Statistician blogger Jim writes a good (relatively basic) explanation, and the Harvard Review has a good initial (very basic) introduction. **Very helpful for many kinds of research, but best with experience and/or software**

- **Quantitative Text Analysis (QTA)**. Automated systematic interpretation of large quantities of text. MethodSpace explains.

ANALYZING QUALITATIVE DATA:

- Analyzing qualitative data generally is done in the following 3 steps.
    1. Initial review and notes.
    2. Initial categorizations of data.
    3. Revised & finalized coding.
- Many researches describe analyzing qualitative data as a **cyclical process**. It's okay to go back to steps, and it's possible to be in more than one step at a time.
- Throughout, it's important to keep in mind what question(s) you're trying to answer. Staying focused will save you a lot of time and result in a far clearer and more useful paper in the end.

(Basic) Codes and Themes

- **'Coding' qualitative data** doesn't mean signing numerical codes to different pieces of data. It just **refers to organizing your data based off of various themes or categories that consistently appear**.
    - Pell Institute offers a relatively short but still valuable guide to coding and analyzing data. Renner is an excellent source to use when thinking about qualitative analysis and personally coding data. Erlingsson & Brsyiewicz provide

an invaluable step by step guide for finding and creating codes, categories, and themes. **Students should read all of these carefully before beginning their own analysis.**

- Reading through your information and coming up with codes based on what is there and not current theory is known as **open coding** (**closed coding** uses a pre-established coding scheme).
    - Open codes are developed after you have a solid understanding of all of the data you've collected, and often change as you become more and more familiar with it. The point of them is to organize your data so that you can better draw conclusions about it and how it interacts with itself and other research.
    - For example, consider the question asking "How did you feel about NYC's firefighters in the immediate aftermath of 9/11?". Participants answered:
        - "Strongly", "Not good- I was furious that they let my son die", "Hurt. It seemed like they couldn't do anything", "Proud of their heroic efforts", "Disappointed, in a word", "Amazed at their strength and courage", "I felt a duty to be incredibly supportive no matter their results", "Sad in solidarity with their losses", "I hurt with them", and "I didn't know what I could ever do to thank them".
        - How would you sort these? You could sort them based off of the different emotions people seemed to express: things like sad, proud, angry, grateful. Or, you could code them into more complex themes- such as 'Firefighters inspired'. You could even try to sort them based off of repeated phrases or words.
        - Whatever you choose, it's helpful to **create an abbreviation to refer to each category**. Angry feelings might be coded as Ang, grateful could be Grf. Within categories, you might even have **subcategories.** Those who are angry could be split into groups that are more specific, such as angry at firefighter leadership ('Ang L') or angry at their efforts ('Ang E'). You can also create supercategories that combine multiple large categories (for ex: creating a supercategory called 'Upset' that would include angry, disappointed, etc).

- **It's vital to allow yourself to update and revise codes and categories as you work and as new or more refined themes and ideas appear.**
- Beyond creating and revising these codes, you also need to be able to **understand and describe how they relate to each other.** Keep notes on their 'interactions' so you can expound on that in your discussion/results sections (which can sometimes be combined for qualitative papers). An easy way to notice and write their interactions with each other is to think about it a little bit like your lit review or 'professional conversation' section. You might also want to keep notes on how codes/data relate to your lit review itself.

(Further) Code Analysis

- Once you've coded your data into different categories, what can you do with that? If you're not satisfied with grouping the codes into categories and themes as explained by Erlingsson & Brsyiewicz, you can also perform these (and more!) analyses.
- Often, researchers will perform **axial coding**. Axial coding is the process of exploring and defining the relationships between your open codes. Sage elaborates.
- A third step might be **selective coding**, or figuring out the core variable that impacts all of your data and then going back through and 'selectively' coding any data that relates to the core variable. Gallicano provides fantastic explanations and a thorough example that will help you understand the relationship between open codes, axial coding, and selective coding.
- Axial and selective coding are sometimes associated with **grounded theory** (of which there are two main types), which is it's own extremely well established and useful analysis process. Grounded theory is an attempt to discover/construct theory from data that was comparatively analyzed. This NCBI article offers a much more detailed explanation.
- Although the above analyses will be more than sufficient for the majority of you, it's worth nothing that **there are forms of qualitative analysis not listed here**. Feel free to further research other analysis to find one that best fits your goals and research question. You might start by searching 'narrative analysis', 'framework analysis', and 'discourse analysis' or looking through this Research Methodology page.

Avoiding Bias

- There are of course **issues with all of these categorization methods/styles**. Going back to our 9/11 views of firefighters example, emotions or sentiment can be hard to accurately interpret from that little detail. Some data might defy categorization altogether (what emotion does "Strongly" fit?). Themes can be difficult because bias from the researcher can heavily influence what they see no matter what was actually said (does "I felt a duty to be incredibly supportive no matter their results" really belong under the theme 'Firefighters inspired'?) . And categorizing solely based on repeated words or phrases isn't always accurate (both responses that included the word "hurt" were quite arguably describing different sentiments). So, **choosing an appropriate categorization is of course an important way to avoid bias and misinterpretation.**
- But even beyond that, you should be wary of the fact that your own thoughts and opinions might misconstrue what your participants really intended. To try and avoid that, use at least one of the following three methods (preferably all of them!).
    1. **Get someone to check your codes.** For the purposes of AP Research, it isn't possible to have a full other partner that would generate their own codes you could then check against yours, but you should at least be able to get a friend to either a) quickly look over some of your data and come up with a few of their own codes or b) check your codes against the data and see if they agree.
    2. **Have your participants check your codes to see if they agree.** This is an *awesome* way to help ensure your own biases aren't clouding your participants' intended messages.
    3. **Triangulation!** This is when you try to verify what you think your results are saying with other sources. If they match up, you should have more confidence in your interpretations.

WRITING THE RESULTS SECTION:

- Woah. You've made it this far! Go ahead and give yourself a congratulatory pat on the back. Seriously, you should be proud. Anyways, back to your regularly scheduled guide:
- Writing the results section is just **reporting what data you found**. You should not be offering any commentary here (unless you're one of the few people doing a qualitative study whose results would 100% be best presented alongside analysis). T**his is all about laying out your data, *not interpreting* it.**
- You don't need to include all of your data, instead you can **summarize** important points and interesting findings (*including data that may 'disagree' with your conclusion or hypothesis!!!*). This is an excellent place in your paper to include self-created graphs *if* they will *enhance* your readers' understanding of your data. **The golden rule of graphs is to only use them when words cannot clearly and succinctly explain your information. Refer to your graphs, don't repeat them**. In addition:
    - Use a consistent, clean style among your graph and figures with clear labels.
    - Never use a green/red comparison because of the colorblind.
    - If you have a lengthy amount of data you need to reference, consider creating and using an appendix.
    - Check your discipline's guidelines on how to refer to figures/graphs/tables/etc.
    - Use leading zeros (0.7 vs .7).
- Other than that, the most important thing about writing your results section is to **organize it in a logical manner**. It should be easy to follow, and if not a fun read, at least a sensible one. It might be helpful to lay out your results similar to how you analyzed them: if you did different spreadsheets for variables outside of your control, you could have a 'Uncontrolled Variables Impact' section.
    - Swaim offers more advice. Here's what part of Joe's results section might look like:

"Sites D, E, and F regularly reported a paint fleck level in the water of 6.4 ppm. This was on average 3.2 ppm less than samples from sites A-C, which were located on national park waterfronts.

All sites' concentration increased by an average 17% in January in conjunction with rising wind speeds of 8%."

- **Qualitative research results** often begins with a *concise* summary of major themes, findings, patterns, and categories. Then they might go more in depth and review their findings under the headings of the different categories. (Some papers might even put the participants and data analysis methodology here). Here's what part of a qualitative results section could look like:

"What remained consistent across both the A/B control/experimental groups was that when participants knew the emoji sender was an authority figure, their comfort was at its lowest point. In fact, of those surveyed, 83% of the authority-figure group used the term "uncomfortable" to describe their experience, as opposed to 32% when participants didn't know of the sender's authority.

An interesting finding was that within the code of 'uncomfortable' there appeared to be 3 distinct subsets. These were based on the reason the participant cited as to why they were uncomfortable.

The first was because of fear of misunderstanding the authority figure and not appropriately following their direction. This subset comprised about 23% of.."

- Some students may find it helpful to use this PDF (http://dl.icdst.org/pdfs/files/22e390b2eb0c8e951f3a742fda5b2d1d.pdf) as a guide for writing parts of their qualitative research section, although I wouldn't recommend using the entire thing as a cohesive guide.

WRITING THE LIMITATIONS SECTION:

- The limitations section is simple: **you need to describe all of the things**

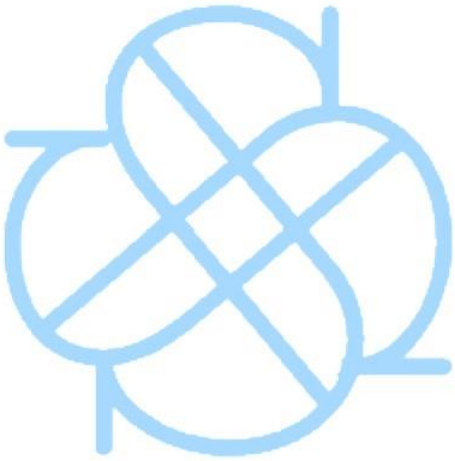
- Often the hardest part about it is deciding whether or not to make its own section or to add it to the discussion/conclusion. *That* decision will simply come down to the style of your paper and whatever you think will flow best.
- **Limitations might include** threats to your internal or external validity such as the population you surveyed or the replicability of your study. Or, you might simply have limitations in your methodology (did you forget to ask a question you meant to? Was your stimulus appropriate if you used one? etc). It's entirely possible that an event or sudden shift in your environment created limitations (for example, 9/11 would have a major impact on a foreign policy views survey), especially if said incident was out of your control. If the execution of your study was difficult (testing equipment wasn't as accurate as you wanted, it wasn't possible to access a certain data set, etc) that would also be a limitation.
- **Essentially: if anything *might* have impacted the intensity or integrity of your results, now's the time to list and explain it.**
    - In the end, people trust a study with clearly stated limitations far more than one which doesn't acknowledge its own. Be clear and upfront and you'll do better on the exam and earn more respect for your work.
- **Example:** Joe's limitations section might look like:
    - "An unusually strong winter weather season might have exaggerated the impacts of paint erosion in seaside areas. In addition, the construction of a nearby coal plant muddied the waters and made the lab tests less reliable...."

WRITING THE DISCUSSION SECTION:

- This section is where you (finally!) are the main force **providing your own commentary and conclusions.** What do you think about any trends or patterns you've noted in the results section (or, 'what do my findings mean?')? How do they relate to your lit review or other previous commentary? What are the implications of your findings? And what further research needs to be done? McCombes provides valuable guidance for writing

your discussion section, and explains a good discussion section includes Interpretations, Implications, Limitations, and Recommendations (IILR).

- A **mistake** many students make in their discussion section is to simply restate their data and the implications of it without actually explaining *how their data actually indicates those implications*.
    - An example of this mistake could be Joe's study with lots of data on water pollution. He might write something like:
        - "Paint fleck concentrations were highest in areas with many painted oceanfront properties. This suggests that animals there die quicker as storms continue to worsen."
        - The issue with this is that he never explained *why* paint fleck concentrations actually lead to animal deaths. Instead, writing this…
        - "Paint fleck concentrations were highest in areas with many painted oceanfront properties. This is unfortunate, as the rising rates of damaging stores suggest paint concentrations will only increase, and Bob (2016) writes the paint flecks are damaging to sealife's respiratory systems. This suggests that the increasing paint fleck concentration will cause quicker wildlife death from the respiratory issues caused by increased paint fleck pollution."
        - ...is far more clear about the potential causality of his findings. The key? *Don't ever make your readers make an inference.* You should draw all conclusions clearly *for them* in your discussion section, and then explain why these conclusions matter and what impact they'll have.
- **Other common mistakes** are simply reiterating your results section, and even worse, bringing up new data in this section. Avoid these mistakes, and include IILR, and you'll do just fine.



EDITING:

- **First steps:**
    1. Use an **in-document search function** and **delete contractions**. Use the same to look for any placeholders you might have used while writing.
    2. **Skim** through your entire document, quickly, to look for **informal writing** and bad grammar. Correct.
    3. **Check and fix your word count**. Remember that you have a 10% buffer on your required 4,000-5,000 word paper (as of 2020, check this for yourself).
    a. **If you need more words**, try and beef up your lit review and/or discussion by adding more perspectives or sources. You could also perform an additional analysis of your data and include that in your results and discussion. (These are

usually preferable to just subbing in more complex phrases for simple statements, which you can also do: for example, instead of "therefore," you might write "in light of that fact,".)

b. **If you need less words**, *first* go through delete overused words. These include "even", "thus", and especially "that". Then, go through and start reading your whole paper out loud. If you can't get through a sentence/paragraph and keep track of your flow, it's too wordy. Rewrite it so that 7th-grade-you could understand it. (This step is super tedious but it *works* and is a great editing tip anyhow.) If you still have too many words, see if there's any sections in your lit review or discussion that don't need to be included at all, or that can be summed up in one sentence.

4. The **main goal (and one of *the most important parts to AP Research*)** is to **have an *aligned* research paper from start to finish**. Read through your paper and consider deleting now-irrelevant sections from your lit review, or adding a discussion of your methods.
5. **Check for 'flow'**. This will help support having a readable, synthesized, and aligned paper.

a. **If you want to rearrange parts**, *stop typing and changing things*. Pull out a blank sheet of printer paper and make flow-charts. Doodle. Find a pattern of organization that makes sense and then go back and rearrange or write the parts the way you want them. This is usually the best way to keep track of what changes you want to make and how you're going to do them.

b. If specific sentences are throwing you off, try and imagine how you would explain the concept within it to 7th-grade-you. If you have to split a sentence up, do so. Spend at least 5 minutes with it before you show it to a friend. (Easy to hear a suggested revision and get it 'stuck' in your head, blocking any possible better revisions.)

c. **If the flow of a paragraph or a smaller section isn't great, think about what your goal is within it.** Are you achieving that goal? If not, what needs to be deleted or moved to another section (so that it can support a different goal)?

Usually, the issue is that you have extra information that only sort-of supports your goal.

6. **Make your own version of the capstone rubric.** Write out the different scores in your own words so that you can be sure you more clearly understand what it takes to get each score. That makes it easier to check if you're where you want to be. Check to see if you're meeting the emphasized points and nuances in your paper.
7. **Check your references against what you've actually cited**. Read through your paper and write down each citation as you come across it (with the year in case you cite an author multiple times). Make a new document and copy-paste your references section onto it. Then, check your new list against the document and highlight each source you actually did use. When you're done, go back through and delete any sources you didn't highlight, and add any sources that were never actually referenced in your paper.
8. **Now's when you get your trusted critic** (that hopefully hasn't seen your paper at all this year, as suggested at the top of this sheet!!) **to look over your paper** in its entirety. Go back through and address all of their comments/questions/concerns.
   a. **Ask your teacher if they have an editing packet** or sheet of specific things to look for when editing a peer's paper.
   b. If you delete or add any sources, be sure to **keep your references up to date!**
9. Go through steps 1-7 again as necessary.

- The above is a good basic format for editing your paper, although everyone's process will be a bit unique. Be sure to take your time!!!

Creator's Note: *GREAT JOB!!! Getting discouraged at this stage is super easy, but makes your editing much harder and less effective. Your efforts are worthwhile, and you should be proud.*

******All used sources in this study guide are cited throughout with their appropriate links.***