Reading a Research Article Quickly and Efficiently

Summary: In short, a fast way to review an article is to read

1. **Title**
2. **Abstract**
3. Last paragraph in the **Background** section (purpose statement)
4. Scan the **Methods** section for data source and design
5. Review tables and charts in **Results** section. If these are unclear, go to 6.
6. Read first few paragraphs of the **Discussion** section, often summarizing the results and providing implications of the research.
7. Then, go back and check the Methods and Results section (if unclear), or re-read the entire article.

Be an active reader. Identify the "what", "who", "when", "where", "how", and "why" as you read. Take notes in the margins (if possible) and look up what you don't know or understand to fill in gaps.

To review a paper efficiently and quickly, it helps to understand the structure of a journal article and the common elements found in most articles.

I. Structure of an Article

Most articles have five sections. This structure gives readers an <u>organized description</u> of the research. These sections are, in general,

1. An introduction with the **purpose** of the research along with objectives or aims,
2. a **background** section with a literature review a description of the gap in knowledge that the research answers,
3. **methods** used with the research design, selection of subjects and acquisition of data, and analysis plan,
4. **results**, including tables and graphs, and a
5. general **discussion** of the research findings, comparing the results to other studies, the impact of the research and, sometimes, hinting at new areas of research needed.

These five sections tend to be found in nearly all research journal articles and abstracts. Each of these sections may vary depending on the preferences of the journal editors, the field of study, and the intent of the research. The author has to follow these journal instructions in layout, content and style.

<u>Order:</u> The order of these five structure elements (purpose, background, methods, results, discussion) may vary depending on the journal's preferences and the area of study. The order above works for most fields. What differs most often are in sections 2 and 5, the background and general discussion sections. In some journals and fields, like in medicine, nursing, public health and general sciences, the background section can be quite long. It provides the reader with a description of the problem or knowledge gap (what isn't known), a review of the literature of research that has preceded this submission and sets up the final paragraphs, the most important part of the background section.

In other fields, the background section is short while discussion section is very long, containing the literature review in parallel to the research findings. However, the overall five-part structure remains the same.

<u>Purpose or objectives (Background section):</u>

All of the literature review provided in the background section leads to a final paragraph. **In general, this final paragraph states the purpose and objectives of the research.**

For example, suppose the final paragraph in the Background section reads as follows:

> *"In this paper, we will determine the associations between cigarette and tobacco use among women and the frequency of miscarriages in a population of women attending a regional system of federally qualified health centers across five states collected 2004-2008. We will use data from the …. data source… which contains outpatient information from surveys provided by a random selection of women attending FQHC clinics. A wider description of this data system is described by Smith et al (2004). We will present population estimates of miscarriages as well as other demographic and potentially confounding factors which may also be associated with miscarriages. By identifying women with and without a history of miscarriages, and matching them based on race and ethnicity, age groups, parity and zip code, we hope to better identify how tobacco use may be associated with the occurrence of miscarriages after adjusting for additional individual and contextual factors."*

Notice how the first two sentences list the purpose (what), the target population (who), data source (from where), the dates when the data were collected (when). Later sentences provide the methods (how). The final sentence often provides the main objective of the research. This paragraph must be concise and interesting, inviting the reader to read further.

The last paragraph(s) in the background section of most papers should be one of the first paragraphs you need to review before performing a quick review, or before (re-)reading the entire paper.

In medicine, nursing and most natural sciences, the background section contains the literature review. The discussion section refers to specific findings from the Results section in comparison to similar or dissimilar results found in similar studies. This provides the reader with evidence of consistency, or in opposition to previous results.

In psychology, sociology and other sciences, papers may have shorter background section and a longer discussion that includes much of the literature review.

II. Methods

In general, the methods section provides details on the design of the study, the data source(s), population or selection of study subjects, the data elements (variables) collected from the data

source(s), and statistical methods used for analysis and presentation. Methods used to validate the findings are described.

The design of the study is one of the most important things you should find in any review.

A design can be either

1. Cross-sectional: taken at one time only, similar to a snapshot in time.
2. Retrospective: data that looks back in time to follow subjects back for their exposure or disease history
3. Prospective: data that looks forward in time, following subjects forward for their exposure or disease history. It can be longitudinal, collecting similar data from time points prospectively. There are many forms of analysis for longitudinal data.

Key words describing the study design include:

- case study
- observational study
    - case-control study (or identifying cases and/or controls by disease)
    - cohort study (identifying those exposure and unexposed to determine the incidence of disease(s)).
    - ecological study, where a natural experiment or event occurs (e.g. earthquake, tornado, flood, etc.) and the population affected are monitored for the outcome in question. These can be very weak studies, but profoundly useful when the event occurs that frames the population.
- Treatment studies where the selection of subjects is randomized.
    - Randomized trial, where subject selection was limited to persons having specific eligibility criteria and then systematically selected to some treatment, intervention or exposure. Subjects, researchers and even the analysts overseeing the randomization of subjects may be blinded to the assignment of subjects to treatment (or non-treatment) groups.
    - non-randomized studies, often selecting subjects by convenience or convenient sampling: not random; biased in nature; with any inferences limited to the study population.

The design of the study reflects how the data are collected and limits the statistical tests that can be applied and the inferences that you can make from the study subjects to a larger target population. *Every design has inherent biases.* A bias is the error in the research measurements caused by inherent features in the design. Bias may be systematic, unintentional, or even random. It is worth your time to look up each of these study designs and read about their advantages and disadvantages, such as bias.

Some comments on some study designs:

Some bias, like selection bias, can invalidate an entire study because the choice of a comparison group, say, control patients who do not have a disease, may not compare well to a set of patients (cases) who have a disease. The validity of such case-control studies, one type of study design, rests in the proper selection of control subjects, and how well the disease is defined, what is called the 'case definition'. Bias can occur on cohort studies, another type of study design. In cohort studies, the study identifies persons who were "exposed" to an intervention, natural occurrence, agent, chemical, medication, education (etc.) and those not-exposed. Ideally, those exposed should differ from those unexposed only in their exposure history.

From there, the study subjects can be followed either forward or backward in time.

If they are followed forward in time, (prospective cohort study), then the exposed and unexposed subjects are observed for the incidence of disease or some specific outcome.

In a retrospective cohort study, the exposed and non-exposed are identified at the start of the study or in a series of records or documents, like in medical or occupational records. For example, suppose you were interested in the association between exposure and disease. Then, these records are reviewed for all subjects to determine their exposure history in the beginning of some time period. Also, the records are reviewed for the incidence of disease. In short, people exposed and unexposed are identified in the past and followed in time to determine the incidence of disease or an outcome. But both the exposure and disease happened in the past.

Let's talk about another study called a convenience survey. This is a heavily biased study since there are many sources of bias.

For example, suppose a researcher stands in a shopping mall. As shoppers walk by, s/he offers them the opportunity to take a survey. Not all shoppers are given equal opportunity to participate. In fact, selection may be biased based on how the researcher recruits shoppers, chooses persons, the location of where he/she stands, etc. If data are collected using convenience surveys that anyone can take (or refuse or ignore), then the findings represent, at a minimum, only those respondents completing the survey. Since not everyone had an opportunity to take the survey, the probability of selection is not the same for everyone in the target population (all shoppers that day). A researcher would have to make a strong argument to extend the study findings to a larger population (e.g. all shoppers that go to this mall in the year.) But that kind of thing happens all of the time, where the survey sponsor assumes that the survey respondents reflect everyone who shops in the store or mall where the survey was given. Convenient samples are known for their bias, but few people realize just how biased they can be.

Randomized designs trials will randomize subjects to an intervention and follow subjects up for outcome(s) of interest. These designs can come in many forms and often make the strongest inferences, but only on the ideal subjects similar to those in the study. Randomization should balance subjects. That is, subjects should be similar in demographics and characteristics across each study arm but differ only in the intervention. Clinical trials often use this design.

Other items in the methods section include descriptions of how data are analyzed (i.e. specific statistical tests or methods, including software used) and presented (i.e. percentages, rates, etc.).

Lastly, keep track of any <u>inclusion or exclusion criteria</u> for subject selection. These criteria may limit how the author can make wider inferences to a population beyond the study subjects. Inclusion and exclusion criteria are critical to understanding subject selection, potential biases due to the selection, and inferential range of the findings.

Randomized studies often have many exclusion criteria. These criteria are needed so that the outcomes observed should derive only from the intervention and not reasons related to the exclusion criteria.

III. Abstract

The abstract is a concise structured description of the research. Abstracts should contain the five structural elements listed in Section I. The abstract tends to limited from as few as 250 words to as many as 500. It must be submitted with a journal article, poster, or submission for review publication.

The first line of an abstract describes the **literature gap**, "what is unknown or not understood." For example,

> *"Little is known about the association of XXXXXXX, a new class of psychoactive medications, to the incidence of suicide ideation among youth and young adults prescribed these medications."*

The second or third line provides the study purpose and objectives. It can refer to the what/who/when of the study.

This is followed by 1 -3 sentences describing the methods, including the data source (from what, who, when) and how the subjects were selected. Most important, the abstract should include the study design.

The remainder of the abstract is written to keep the word count as low as possible. The methods section is often just one or two sentences with the data source and design. The results section provide the number of subjects (study size), and a quick review of the major results. Lots of details are missed here. The key is to provide the results making the most impact. A conclusion section summarizes the findings and hints at the potential impact of the research. This impact may be clinical, ecological or population-based. It may suggest additional implications or areas of potential research or funding. You often find the "hook" in the last like. This hook should reflect the impact of the study, often reaching beyond the population studied.

IV. Results

Results sections will first describe the sample population (size) and whether there were some subjects lost in the selection or exclusion process. Often, in a design that involves random assignment of an intervention, an author will compare those subjects across intervention groups to determine random differences that can still occur when randomizing eligible subjects.

General demographic descriptions of the subjects or populations are provided early. As results are provided, pay attention when and where figures are noted and placed in the article.

In the figures, pay attention to:

- the labels and units on axes in charts or graphs
- the ranges of values in charts
- In histograms, look for frequency or percents in the axes. In the other axis, data grouped into bars may show only the value ranges. Check the midpoints between the histogram bars to determine the actual value being noted as a group frequency or percent.
- Look for the zero (0, 0) value where axes cross. When the zero value is not the value where the axes cross, pay close attention to any distortion that can occur in the diagram due to the values in one or both of the axes.
- The legend and footnotes, both very important.

Results should be logically presented, flowing through the purpose of the research to representation of the data. Conclusions and inferences from the data should be in the Discussion section. However, some authors like to extend their inferences in the Results section.

V. Discussion Section

The first paragraph of the discussion section tends to summarize the overall findings in the study.

Look for implications of the research here, in addition to paragraphs that describe the limitations and strengths of the research.

The discussion may end with a "hook" that extends the implications from the research well beyond the study population. The "hook" may be an impactful statement or proposition, often having clinical or sociological impact beyond the statistical findings. You have to consider these carefully, since the author can go "out on a limb" here and exaggerate perceived implications. Some articles are remembered only for the hook or last paragraph. One example is the original paper by Crick and Watson, the researchers that discovered the double helix shape of DNA.

> *"It has not escaped our notice that the specific pairing we have postulated immediately suggests a possible copying mechanism for the genetic material."*

Watson JD, Crick FH (April 1953). "Molecular structure of nucleic acids; a structure for deoxyribose nucleic acid" (PDF). Nature 171 (4356): 737–738.
http://www.nature.com/nature/dna50/watsoncrick.pdf

Pretty powerful stuff, eh?

VI. Some final notes

Finally, read slowly and actively. That is, identify the message by dissecting each sentence and section. Remind yourself the "what", "who", "when", "where", "how", and "why" as you read. Don't forget that the abstract leaves a lot of information out, and you often need to review the last paragraph in the Background and first few in the Discussion to understand the purpose and conclusions quickly. The methods section will clue you to the design. Knowing study designs help you to anticipate sources of bias. When you know what biases to look for, you can review the Results with more confidence and a more critical eye.