

# **A Primer on How to Read a Scientific Paper for Substance Misuse Prevention Professionals**



Pacific Southwest (HHS Region 9)

**PTTC**

**Prevention Technology Transfer Center Network**

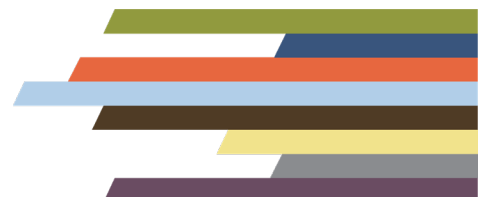
Funded by Substance Abuse and Mental Health Services Administration

### DISCLAIMER

This document was prepared by the [Pacific Southwest Prevention Technology Transfer Center \(PTTC\)](#) under a cooperative agreement from the Substance Abuse and Mental Health Services Administration (SAMHSA). All material appearing in this document, except that taken directly from copyrighted sources, is in the public domain and may be reproduced or copied without permission from SAMHSA or the authors. Citation of the source is appreciated. Do not reproduce or distribute this presentation for a fee without specific, written authorization from the Pacific Southwest PTTC. At the time of this presentation, Elinore F. McCance-Katz, served as SAMHSA Assistant Secretary. The opinions expressed herein are the views of the authors and do not reflect the official position of the Department of Health and Human Services (DHHS), SAMHSA. No official support or endorsement of DHHS, SAMHSA, for the opinions described in this document is intended or should be inferred.

This work is supported by the following cooperative agreement from the Substance Abuse and Mental Health Services Administration: Pacific Southwest PTTC: H79SP081015-01

Revised on December 22, 2020.



# Primer on How to Read a Scientific Paper for Substance Misuse Prevention Professionals

Reading scientific articles (sometimes referred to as journal articles or scientific papers) can be both intimidating and challenging. However, by approaching articles strategically, prevention professionals can glean a great deal of information from scientific papers or journal articles. Learning how to decipher a journal article is crucial for substance misuse prevention professionals, especially during their work on Step 3 (Planning) of the Strategic Prevention Framework. Actions to complete in this step include identifying evidence-based interventions for implementation that are a good conceptual and practical fit. At times, prevention professionals will need to read journal articles to review the research completed on interventions in order to determine whether or not an intervention is an appropriate fit for their community.

This document is a compilation of resources to assist prevention professionals in enhancing their reading and understanding of scientific papers. Topics covered include: what “peer-reviewed” means, the purpose of each section of an article, step-by-step instructions on how to read an article, and types of study designs.

## *What does peer-reviewed mean?*

Before reading a scientific paper, it is important to determine if it underwent a peer-reviewed approval process before being published. This is important for substance misuse prevention professionals because the peer review process provides some quality assurance to consumers of research. The following is a description of the peer-review process:

### **PEER-REVIEWED**

A process in which two or more expert reviewers read an article and make suggestions for revisions before it is accepted for publication.

“Peer review is a process in which two or more experts in a field read an article and make suggestions for revisions before it is accepted for publication. To pass peer review, the article must give an original perspective on an important topic. It also must be thoroughly researched, logically argued, and well written. Most articles that are submitted to journals do not make it through peer review. The articles that do get published usually have extensive revisions. Most articles you see are the result of years of work by many people. To learn if a journal requires articles to be peer reviewed, locate journal’s information in the database you are using or the journal’s website.”

---

Excerpt from [Lone Star College: The CyFair Writing Center, Scholarly Sources](#)

## *What is the purpose of each section of an article?*

Journal articles are often organized into particular sections. These sections can include an abstract, an introduction, methods, results, a discussion section, and sometimes a brief conclusion. Journal articles will also include a references section. Confused as to what each section of a journal article contains? Here's a quick primer:

### **ABSTRACT**

A brief summary of the whole article—it's main points and findings

### **INTRODUCTION**

The purpose of the study and the main question(s) to be answered—what's already known and what the authors are trying to find out

### **LITERATURE REVIEW**

A review of existing research on a topic that frames how authors arrived at the research question/topic

### **METHODOLOGY**

How the research was performed—subjects, testing conditions, controls, etc.

### **RESULTS / DATA**

The numbers and the outcomes—often presented visually with charts, graphs, and tables

### **DISCUSSION**

An extended summary of the findings—delineates new insights gained from the results and anything unexpected that arose

### **CONCLUSION**

Restates the findings and results—what was discovered and what future research still needs to be done

### **REFERENCES / BIBLIOGRAPHY**

Previous research the authors reviewed to formulate their study design and research questions—what has already been published on the topic

---

Excerpt from [\*Brandeis University: Tips for Reading Scholarly and Journal Articles\*](#)

## *How to Read a Journal Article*

Follow the steps below to assist you in gleaning the most information possible from the articles you read.

Reading a journal article from beginning to end is not always the best approach

### Step 1: Literature Search

Begin your search by finding journal articles on a topic you want to learn about. See “How to Conduct a Literature Search” for searchable databases that might be of interest to you.

### Step 2: Understand the Big Picture

Start with the Abstract to get an idea of the population that was studied, the intervention that was introduced, and the bottom-line of the results. Next, move to the introduction, which tends to be less arduous to follow than other sections. In the introduction, the authors provide big picture context for the study and review previous research in the field. The introduction is where you'll gain an understanding of the relevant research that has been conducted to address or inform a particular research question or topic and what specifically the authors are interested in studying (e.g., the hypotheses or research questions).

You want to come away from the introduction with an understanding of the central problem the researchers are trying to address and any related research questions. The authors will often offer hypotheses about the answers to those questions and review the ways in which the hypothesis(es) or research questions were examined.

### Step 3: Read the Discussion/Conclusion

The discussion section summarizes and evaluates the findings of the research in a digestible way. In this section, the authors offer an interpretation of the data from their results and whether their data supports or refutes their hypotheses and related research questions. The authors will also speak to how their findings relate to the bigger picture. Keep in mind that all scientific studies have limitations, such as limitations related to the study's design. These limitations are often acknowledged in the discussion section of a paper and should be given ample consideration when deciding an intervention's feasibility and appropriateness for a particular community.

### Step 4: Read the Methods Section

Before reviewing the results section of an article, it is a good idea to review the methods section to get an idea of how the study was conducted. The methods section will usually review the population that was studied, how the population was recruited for participation, the design of the intervention, and any materials or protocols to which the population under study was exposed. This information helps prevention practitioners to interpret the data they will review in the results section of an article. Some of the most important aspects of the methods section to pay

## SAMHSA'S Pacific Southwest Prevention Technology Transfer Center

attention to are the study's design (e.g., was it a blind experiment with randomization or a quasi-experiment?), the use of a control or comparison group, sample size, and sample demographics. Researchers use many types of designs when studying interventions. See table below for a review of the types of study designs and Appendix A for a glossary of scientific terms.

### Step 5: Understand the Results

The results section presents the data objectively (i.e., without interpretation), often making it the most difficult part of the paper to digest.

Being familiar with a few statistical terms can help you navigate the data, particularly “significant” and “non-significant.” This basic statistical terminology is used by scientists to describe statistical outcomes that could be the result of random chance (“non-significant”) versus statistical outcomes that could represent meaningful discoveries (“significant”). If you're not familiar with statistics, consider reading a primer like [this one from SAS](#). Be aware that the results section can be challenging to understand, even for other scientists who are reading outside of their field of expertise.

Pay attention to the figures and charts, which pack a lot of information. Read the captions closely, since they tend to explain the results with simple, clear language. Also, ask yourself: *What is being measured? How did they measure their variables (e.g., what does it mean when the authors said they measured “recidivism” or “overdose”)? Given what you read in the introduction and discussion sections, what do the data show?*

### Step 6: Review the References Section

The scientific paper or journal article you are reading could be the first of its kind, examining an intervention, population, or factors not many researchers have investigated previously. It could also be one paper in a very long line of others that explore similar interventions, populations, or factors. If you're interested in learning more about a particular topic, you can turn to the references section for further reading or even to aid contacting one of the authors with questions. The references section can be extremely useful to find other research that has explored related topics.

---

*These steps are excerpts from [The Non-Scientist's Guide to Reading and Understanding a Scientific Paper](#)*

## Types of Study Designs

As described above, an important step when reading journal articles is to understand the study design. Below are some designs often used in research.

Study Design	Description
Meta-analysis	A way of combining data from many different research studies. A meta-analysis is a statistical process that combines the findings from individual studies.
Systematic Review	A summary of the clinical literature. A systematic review is a critical assessment and evaluation of all research studies that address a particular clinical issue. The researchers use an organized method of locating, assembling, and evaluating a body of literature on a particular topic using a set of specific criteria. A systematic review typically includes a description of the findings of the collection of research studies. The systematic review may also include a quantitative pooling of data, called a meta-analysis.
Randomized Controlled Trial	A controlled clinical trial that randomly (by chance) assigns participants to two or more groups. There are various methods to randomize study participants to their groups.
Matched Comparison Group Studies (Quasi-Experimental)	A quasi-experimental study design is one that compares the outcomes of groups but does not include random assignment. If a randomly controlled study is not feasible (e.g., the investigator has no control over group assignment), researchers will often compare the outcomes between groups who received an intervention and those who did not by matching these groups on key characteristics (e.g., baseline scores). How well identified the key characteristics are and how closely matched the groups are can determine the validity of the study's results. One common type of quasi-experimental comparison group design is the non-equivalent groups design. For this design, investigators compare similar groups (no random assignment) on some outcome using pretests and posttests.
Non-equivalent groups design	This design is often used when there is no control by the investigator over assignment to groups. If the groups are not randomly assigned, they may not be similar to each other, ergo, non-equivalent.



## SAMHSA'S Pacific Southwest Prevention Technology Transfer Center

Study Design	Description
Cohort Study (Prospective Observational Study)	A clinical research study in which people who presently have a certain condition or receive a particular treatment are followed over time and compared with another group of people who are not affected by the condition.
Case-control Study	Case-control studies begin with the outcomes and do not follow people over time. Researchers choose people with a particular result (the cases) and interview the groups or check their records to ascertain what different experiences they had. They compare the odds of having an experience with the outcome to the odds of having an experience without the outcome.
Cross-sectional Study	The observation of a defined population at a single point in time or time interval. Exposure and outcome are determined simultaneously.
Case Reports and Series	A report on a series of patients with an outcome of interest. No control group is involved.

---

Excerpt from [Literature Reviews: Types of Clinical Study Designs](#)



## *Definitions of Basic Statistical Terms*

The following are some of the most common statistical terms in scientific papers. Understanding these terms will assist you in understanding the results and conclusions sections.

### **N**

"N" is usually used to indicate the number of subjects in a study. For example: If you have 76 participants in a study,  $N=76$ . A lowercase n is often used to indicate the number of subjects in a subsample.

## Measures of Central Tendency: The Three Ms

### **MEAN**

The average result of a test, survey, or experiment.

### **MEDIAN**

The score that divides the results in half - the middle value.

### **MODE**

The most common result (the most frequent value) of a test, survey, or experiment.

## Significant Difference

### **SIGNIFICANCE**

The measure of whether the results of research were due to chance. The more statistical significance assigned to an observation, the less likely the observation occurred by chance.

### **P-VALUE**

The way in which significance is reported statistically (i.e.  $p<.01$  means that there is a less than 1% chance that the results of a study are due to random chance). Note that in general p-values need to be fairly low (.01 and .05 are common) in order for a study to make any strong claims based on the results.

## Correlation

### **CORRELATION**

The degree to which two factors appear to be related. Correlation should not be confused with causation. Just because two factors are reported as being correlated, you cannot say that one factor causes the other. Establishing causation requires certain conditions to be met. First, the

## SAMHSA'S Pacific Southwest Prevention Technology Transfer Center

factors/variables must be correlated. Second, the cause must always precede the effect (temporal precedence). Third, ruling out a confounding variable (sometimes called a “third” variable or mediator) that actually causes both and is the real reason the variables or factors are related to each other. For example, in the summertime, as ice cream sales increase, crime rates also increase (i.e., a positive correlation between ice cream and crime). Should we conclude that eating cream causes people to commit crime? Probably not. The more likely cause, or third variable, is hot weather, which can explain both an increase in ice cream sales and crime. Correlation, therefore, is just one step in determining causality.

### R-VALUE

The way in which correlation is reported statistically (a number between -1 and +1). Generally, r-values should be  $\geq \pm .3$  in order to report a significant correlation.

- An r-value of -1 indicates an extreme negative correlation between two variables - as one variable's value tends to increase, the other variable's value tends to decrease.
- An r-value of +1 indicates an extreme positive correlation between two variables - as one variable's value tends to increase, the other variable's value also tends to increase.
- An r-value of 0 means there is no correlation at all between the elements being studied.

---

Excerpt from [\*iStudy for Success!\*](#)

## *Additional Resources*

Are you interested in learning about p-values and other statistical analyses? Begin your journey by watching the following videos:

- [Understanding the P-Value – Stastics Help \(YouTube\)](#)
- [Statistical Significance, the Null Hypothesis and P-Values Defined & Explained in One Minute \(YouTube\)](#)
- Need help understanding what a statistical term means in plain English? [Visit Statistic How To: Statistics for the rest of us!](#)
- Looking for other approaches to reading scientific papers? [Review this useful infographic.](#)

## References

- Larsen, A. (2013). Scholarly sources. Lone Star College CyFair Writing Center. Downloaded on September 13, 2018 from [http://www.lonestar.edu/departments/tutoring/scholarly\\_sources.pdf](http://www.lonestar.edu/departments/tutoring/scholarly_sources.pdf)
- Rosenblatt, J. & Smith, E. (2014). Tips for reading scholarly and journal articles. Brandeis University. Downloaded on September 13, 2018 from <https://www.brandeis.edu/acserv/studyskills/ReadingScholarlyandJournalArticles.pdf>
- Endpoints. (2017). The non-scientist's guide to reading and understanding a scientific paper. Downloaded on September 13, 2018 from <https://endpoints.elysiumhealth.com/how-to-read-a-scientific-paper-695188037080>
- Georgia State University Library. (2018). Literature reviews: Types of clinical study designs. Downloaded on September 13, 2018 from <http://research.library.gsu.edu/c.php?g=115595&p=755213>
- Penn State University. (2017). Definitions of basic statistical terms. Downloaded on September 13, 2018 from <http://tutorials.istudy.psu.edu/basicstatistics/basicstatistics2.html>

## *Appendix A: Study design terminology*

### **BIAS**

Can be understood as any deviation in the results from the truth (e.g., the true effectiveness of a treatment). Randomized controlled studies tend to reduce the likelihood of introducing bias but there are several potential sources of bias beyond study design including the ways in which a study is carried out (procedure), the analysis, and reporting.

### **CASE CONTROL STUDIES**

Studies which start with the identification of persons with a disease of interest and a **control** (comparison, referent) group without the disease. The relationship of an attribute to the disease is examined by comparing diseased and non-diseased persons with regard to the frequency or levels of the attribute in each group.

### **CAUSALITY**

The relating of causes to the effects they produce. Causes are termed necessary when they must always precede an effect and sufficient when they initiate or produce an effect. Several factors may be associated with the potential disease causation or outcome, including predisposing factors, enabling factors, precipitating factors, reinforcing factors, and risk factors.

### **CONTROL GROUPS**

Groups that serve as a standard for comparison in experimental studies. They are similar in relevant characteristics to the experimental group but do not receive the experimental intervention.

### **CONTROLLED CLINICAL TRIALS**

Clinical trials involving one or more test treatments, at least one control treatment, specified outcome measures for evaluating the studied intervention, and a bias-free method for assigning patients to the test treatment. The treatment may be drugs, devices, or procedures studied for diagnostic, therapeutic, or prophylactic effectiveness. Control measures include placebos, active medicines, no-treatment, dosage forms and regimens, historical comparisons, etc. When randomization using mathematical techniques, such as the use of a random numbers table, is employed to assign patients to test or control treatments, the trials are characterized as Randomized Controlled Trials.

### **COST-BENEFIT ANALYSIS**

A method of comparing the cost of a program with its expected benefits in dollars (or other currency). The benefit-to-cost ratio is a measure of total return expected per unit of money spent. This analysis generally excludes consideration of factors that are not measured ultimately in economic terms. Cost effectiveness compares alternative ways to achieve a specific set of results.

### **CROSS-OVER STUDIES**

Studies comparing two or more treatments or interventions in which the subjects or patients, upon completion of the course of one treatment, are switched to another. In the case of two treatments, A and B, half the subjects are randomly allocated to receive these in the order A, B and half to receive them in the order B, A. A criticism of this design is that effects of the first treatment may carry over into the period when the second is given.

### **CROSS-SECTIONAL STUDIES**

Studies in which the presence or absence of disease or other health-related variables are determined in each member of the study population or in a representative sample at one particular time. This contrasts with LONGITUDINAL STUDIES which are followed over a period of time.

### **DOUBLE-BLIND METHOD**

A method of studying a drug or procedure in which both the subjects and investigators are kept unaware of who is actually getting which specific treatment.

### **EMPIRICAL RESEARCH**

The study, based on direct observation, use of statistical records, interviews, or experimental methods, of actual practices or the actual impact of practices or policies.

### **EVALUATION STUDIES**

Works consisting of studies determining the effectiveness or utility of processes, personnel, and equipment.

### **GENOME-WIDE ASSOCIATION STUDY**

An analysis comparing the allele frequencies of all available (or a whole genome representative set of) polymorphic markers in unrelated patients with a specific symptom or disease condition, and those of healthy controls to identify markers associated with a specific disease or condition.

### **INTENTION TO TREAT ANALYSIS**

Strategy for the analysis of Randomized Controlled Trial that compares patients in the groups to which they were originally randomly assigned.

### **LOGISTIC MODELS**

Statistical models which describe the relationship between a qualitative dependent variable (that is, one which can take only certain discrete values, such as the presence or absence of a disease) and an independent variable. A common application is in epidemiology for estimating an individual's risk (probability of a disease) as a function of a given risk factor.

### **LONGITUDINAL STUDIES**

Studies in which variables relating to an individual or group of individuals are assessed over a period of time.

### **LOST TO FOLLOW-UP**

Study subjects in cohort studies whose outcomes are unknown e.g., because they could not or did not wish to attend follow-up visits.

### **MATCHED-PAIR ANALYSIS**

A type of analysis in which subjects in a study group and a comparison group are made comparable with respect to extraneous factors by individually pairing study subjects with the comparison group subjects (e.g., age-matched controls).

### **META-ANALYSIS**

Works consisting of studies using a quantitative method of combining the results of independent studies (usually drawn from the published literature) and synthesizing summaries and conclusions which may be used to evaluate therapeutic effectiveness, plan new studies, etc. It is often an overview of clinical trials. It is usually called a meta-analysis by the author or sponsoring body and should be differentiated from reviews of literature.

### **NUMBERS NEEDED TO TREAT**

Number of patients who need to be treated in order to prevent one additional bad outcome. It is the inverse of Absolute Risk Reduction.

### **ODDS RATIO**

The ratio of two odds. The exposure-odds ratio for case control data is the ratio of the odds in favor of exposure among cases to the odds in favor of exposure among noncases. The disease-odds ratio for a cohort or cross section is the ratio of the odds in favor of disease among the exposed to the odds in favor of disease among the unexposed. The prevalence-odds ratio refers to an odds ratio derived cross-sectionally from studies of prevalent cases.

### **PATIENT SELECTION**

Criteria and standards used for the determination of the appropriateness of the inclusion of patients with specific conditions in proposed treatment plans and the criteria used for the inclusion of subjects in various clinical trials and other research protocols.

### **PREDICTIVE VALUE OF TESTS**

In screening and diagnostic tests, the probability that a person with a positive test is a true positive (i.e., has the disease), is referred to as the predictive value of a positive test; whereas, the predictive value of a negative test is the probability that the person with a negative test does not have the disease. Predictive value is related to the sensitivity and specificity of the test.

### **PROSPECTIVE STUDIES**

Observation of a population for a sufficient number of persons over a sufficient number of years to generate incidence or mortality rates subsequent to the selection of the study group.

### **QUALITATIVE STUDIES**

Research that derives data from observation, interviews, or verbal interactions and focuses on the meanings and interpretations of the participants.



## **QUANTITATIVE STUDIES**

Quantitative research is research that uses numerical analysis.

### **RANDOM ALLOCATION**

A process involving chance used in therapeutic trials or other research endeavor for allocating experimental subjects, human or animal, between treatment and control groups, or among treatment groups. It may also apply to experiments on inanimate objects.

### **RANDOMIZED CONTROLLED TRIAL**

Clinical trials that involve at least one test treatment and one control treatment, concurrent enrollment and follow-up of the test- and control-treated groups, and in which the treatments to be administered are selected by a random process, such as the use of a random-numbers table.

### **REPRODUCIBILITY OF RESULTS**

The statistical reproducibility of measurements (often in a clinical context), including the testing of instrumentation or techniques to obtain reproducible results. The concept includes reproducibility of physiological measurements, which may be used to develop rules to assess probability or prognosis, or response to a stimulus; reproducibility of occurrence of a condition; and reproducibility of experimental results.

### **RETROSPECTIVE STUDIES**

Studies used to test etiologic hypotheses in which inferences about an exposure to putative causal factors are derived from data relating to characteristics of persons under study or to events or experiences in their past. The essential feature is that some of the persons under study have the disease or outcome of interest and their characteristics are compared with those of unaffected persons.

### **SAMPLE SIZE**

The number of units (persons, animals, patients, specified circumstances, etc.) in a population to be studied. The sample size should be big enough to have a high likelihood of detecting a true difference between two groups.

### **SENSITIVITY AND SPECIFICITY**

Binary classification measures to assess test results. Sensitivity or recall rate is the proportion of true positives. Put differently, it is a measure of a test's ability to correctly identify true positives. Specificity is the probability of correctly determining the absence of a condition (true negatives). Single-Blind Method - A method in which either the observer(s) or the subject(s) is kept ignorant of the group to which the subjects are assigned.

### **TIME FACTORS**

Elements of limited time intervals, contributing to particular results or situations.

---

Excerpt from [\*Literature Reviews: Types of Clinical Study Designs\*](#)



Pacific Southwest (HHS Region 9)

PTTC

Prevention Technology Transfer Center Network

Funded by Substance Abuse and Mental Health Services Administration