برنامج مهارات البحث العلمي (10)

محاضرة «تحليل البيانات وتفسير المعطيات»

مركز بحوث الدراسات العلمية والطبية

جامعة الملك سعود

Analysis Of Data, Breaking Problems Down Into Manageable Units

Outline

- Data Types
- Data Analysis
- Positive vs Negative results
- Examples

Outcomes

You will be able to:

- Distinguish between Quantities and Qualitative data.
- Dealing with the three Cs: Coding, Categorizing, and Concepts
- Dealing with the negative results similar to positive one .

Bogdan and Biklen (1982)

"Working with data, organizing it, breaking it into manageable units, synthesizing it, searching for patterns, discovering what is important and what is to be learned, and deciding what you will tell others"

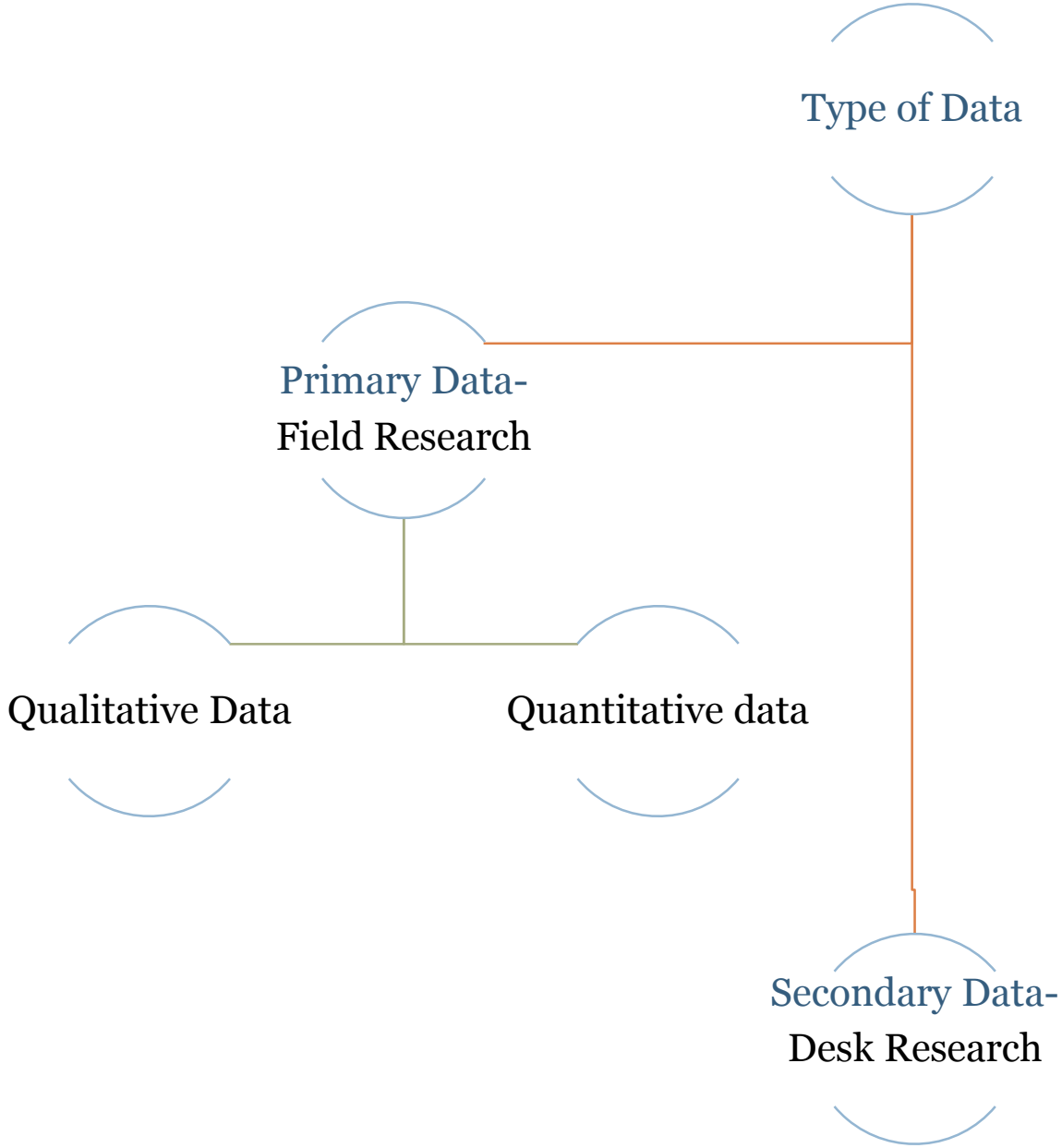
Challenge

- ✓ to place the raw data into logical, meaningful categories;
- ✓ to examine them in a holistic approach;
- ✓ to communicate this interpretation to others.

Data analysis can be said to be:

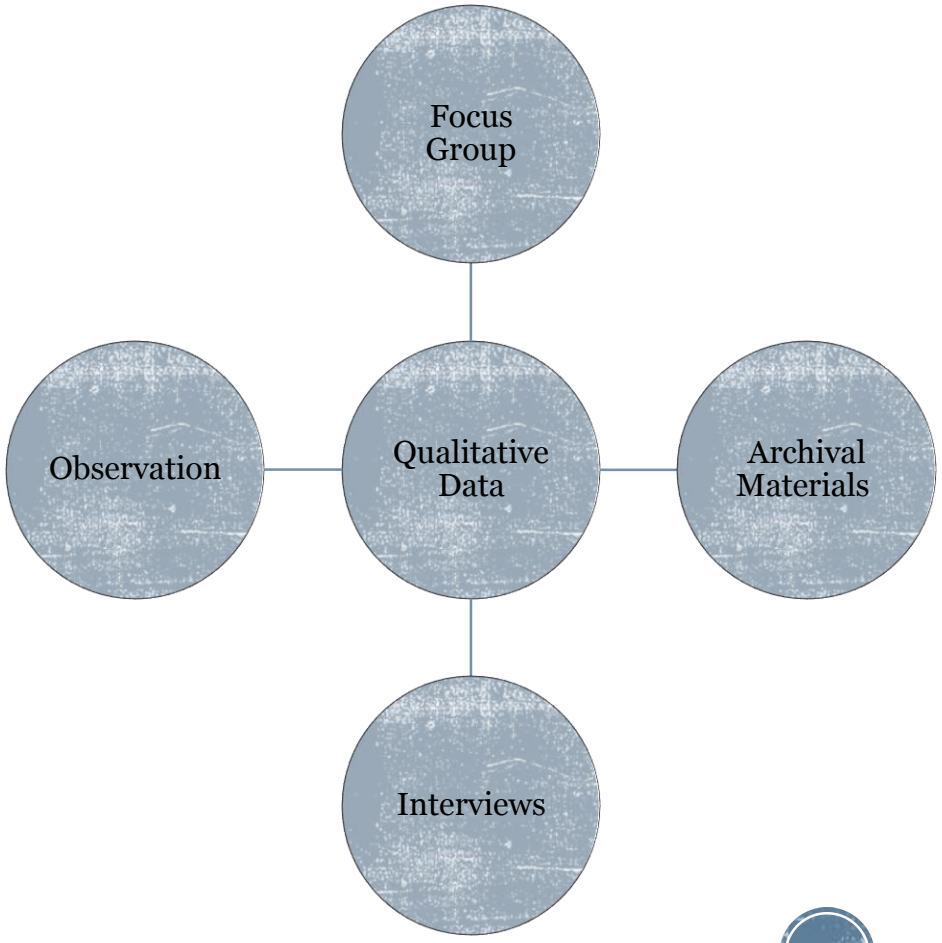
"The process of breaking down, examining, comparing, conceptualizing, and categorizing data" (Strauss & Corbin, 1998).

As data analysis proceeds, the researcher moves back and forth between data analysis and data collection in order to create and explain the findings.


Type of Data
Primary Data-
Field Research
Qualitative Data
Quantitative data
Secondary Data-
Desk Research

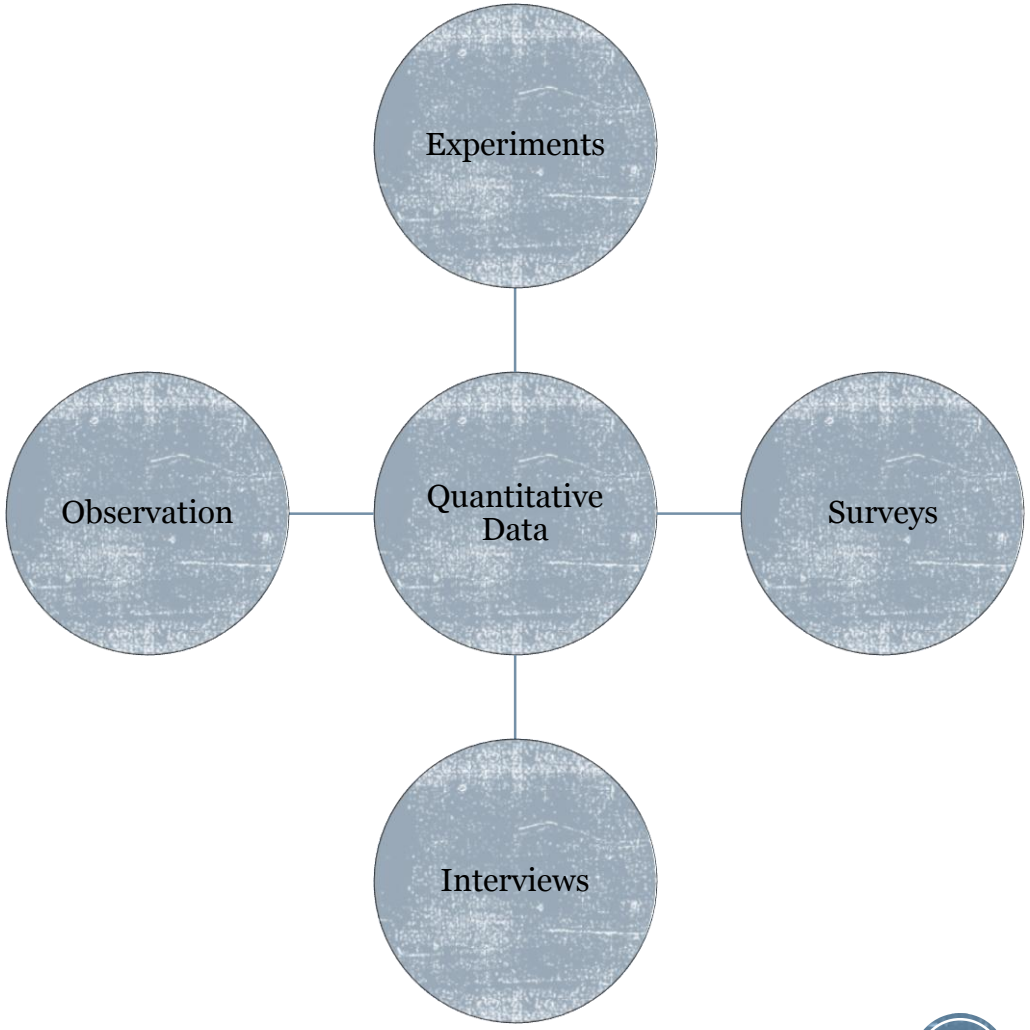
Qualitative Data

- Qualitative Research is primarily exploratory research. It provides insights and understanding about a particular problem to develop ideas or hypotheses for potential quantitative research.
- The researcher's ability to interpret the data and to present the findings clearly makes a qualitative research study useful.
- Since the nature of data is descriptive, it can be approximated with but cannot be computed.
- Qualitative data collection methods vary using unstructured or semi-structured techniques.



Quantitative Data

- Quantitative Research is used to quantify the problem by way of generating numerical data or data that can be transformed into useable statistics.
- Quantitative data collection methods are much more structured than Qualitative data collection methods.



“A little bit of data and a lot of right brain- Agar, 1991”

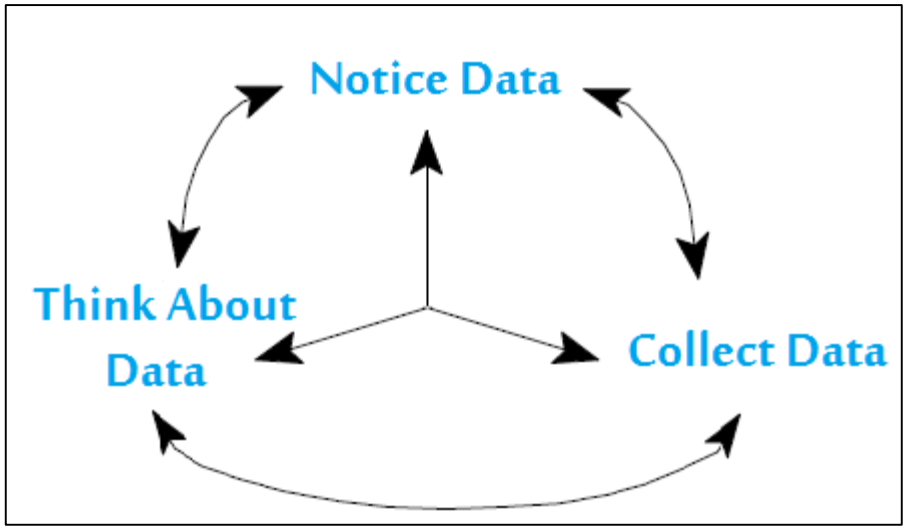
- The question is, how do you come up with that “little bit of data?”
- Obviously you start by reading and rereading the data record. In the process you *notice* a few interesting things. You then *collect* one or more of these things and intensively *think* about them.

The process of Data analysis

- Once data collection has begun, it is time for the researcher to begin data analysis. Data analysis continues after data collection has been completed.
- The researcher explores the data to answer the questions:

- What is going on here?
- What does this mean?
- Why do the participants behave this way? And so on.

Data Analysis: A Model of the Process



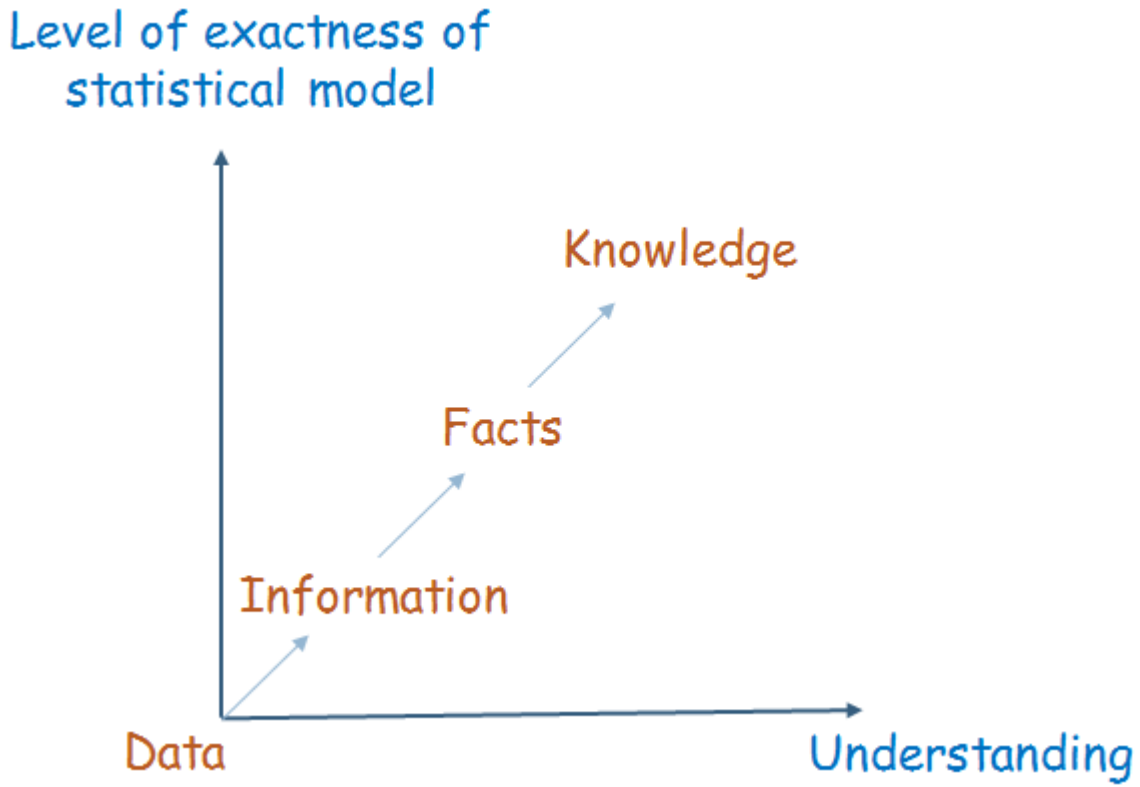
- Data Analysis based on three notes: Noticing, Collecting, and Thinking about interesting things (Seidel, 1998).

Interpreting the Data

- Once the data are sorted into manageable chunks through coding, the process of interpretation begins.
- It is important to note that this phase of the research process overlaps with data collection and coding, although it often extends long after the data collection has been completed.

After data have been sorted into codes, the codes themselves must be organized. Coffey and Atkinson (1996) note the importance of being able to see all of the data from a particular code at the same time. If the researcher is using QDAS, he or she may print out all of the data from a particular code. Some researchers create tables or written summaries of each code. Whatever the strategy, it is important to be able to see each piece of data in relation to other like data.

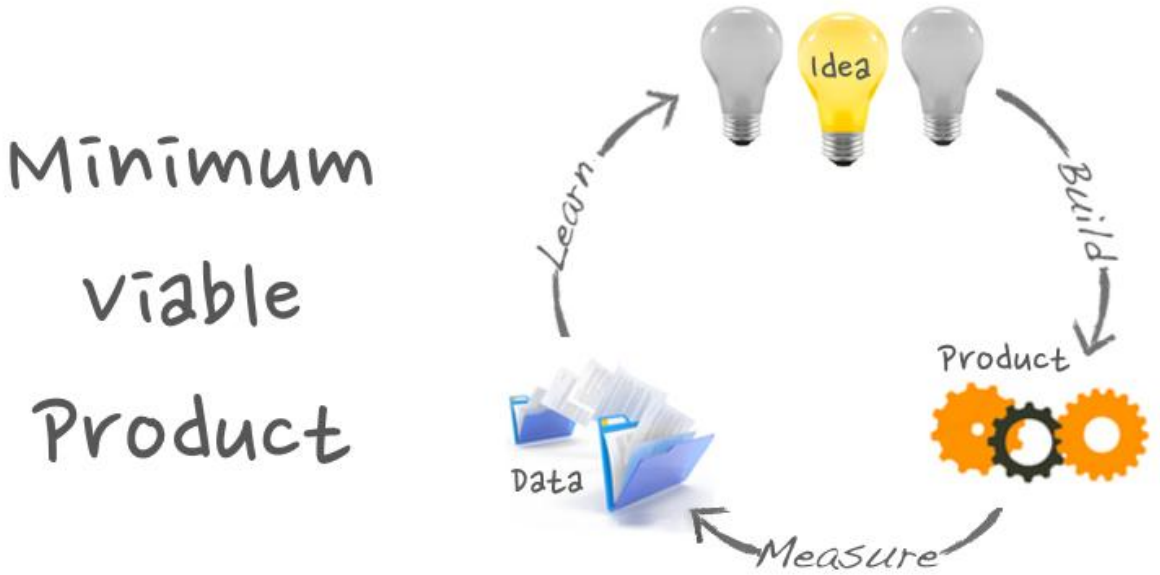
“Data is known to be crude information and not knowledge by itself ”



Analytical Thinking: Involves additional processes

- Examining it in detail from many angles
- Looking for possible flaws in the reasoning, the evidence, or the way that conclusions are drawn
- Comparing the same issues from the point of view of other writers
- Being able to see and explain why different people arrived at different conclusions
- Being able to argue why one set of opinions, results or conclusions is preferable to another
- Checking for hidden assumptions
- Checking for attempts to lure the reader into agreements

For most, the end products of research are books, papers, presentations, or plans for action. Data analysis moves you from the rambling pages of description to those products.



Many experimental results never see the light of publication day. For a large number of these, it comes down to the data being “negative”, i.e. the expected and/or wanted effect was not observed.

Negative Results: The Dark Matter of Research

- The estimate for 2014 was 2.5 *million articles*, and that number is sure to be higher today. Still, the publication record is only a tiny slice of all the research data in existence around the world. Results that are inconclusive or challenge our assumptions are frequently hidden in lab notebooks, never to be shared. These data represent the "dark matter" of our research universe. *But is the current system really best for science?*

Reducing the positive bias in the scientific literature

- One recurring misunderstanding among scientists is that **negative results** are equivalent to **bad results** and are products of flawed or ill-designed science.
- This quality argument can be easily challenged, because even positive results are not exempted from having their quality questioned. Furthermore, some of the most well-known examples of fraud and data fabrication in science were associated with studies claiming positive correlations or supporting researchers' anticipated hypotheses.
- Instead of suggesting poorly conducted science, negative results can indicate novel findings or unexpected outcomes of rigorous scientific investigations, directly or indirectly contributing to scientific discovery.

There are several consequences to keeping negative results hidden.

1. sharing a failed experiment may prevent a number of other research groups from wasting time and money on the same idea. Even if another lab wanted to try a similar experiment, they could make slight changes based on the previous attempt.

1. Second, the current tendency to focus on splashy results may be leading to false conclusions. In a famous 2005 paper in PLOS Medicine, John Ioannides even argued that most published findings are false.

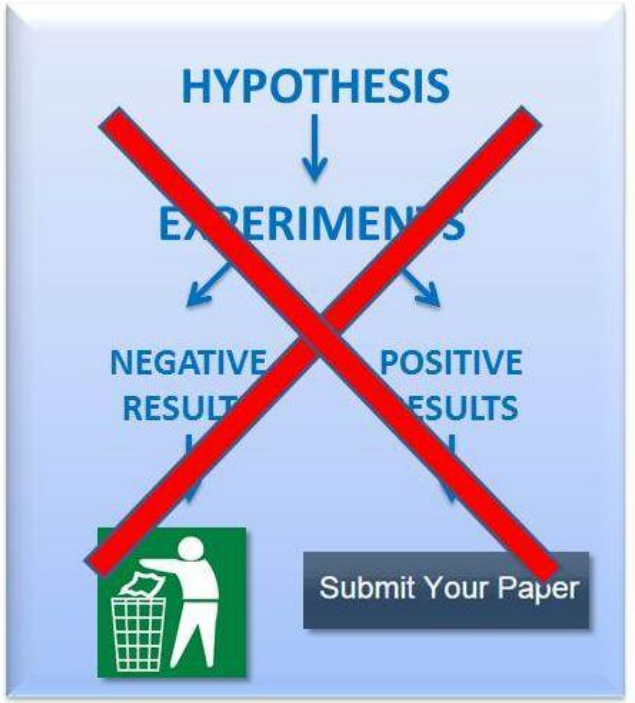


Figure 1: The most common approach taken by journals, in which only those experiments yielding positive results end up as publication material.

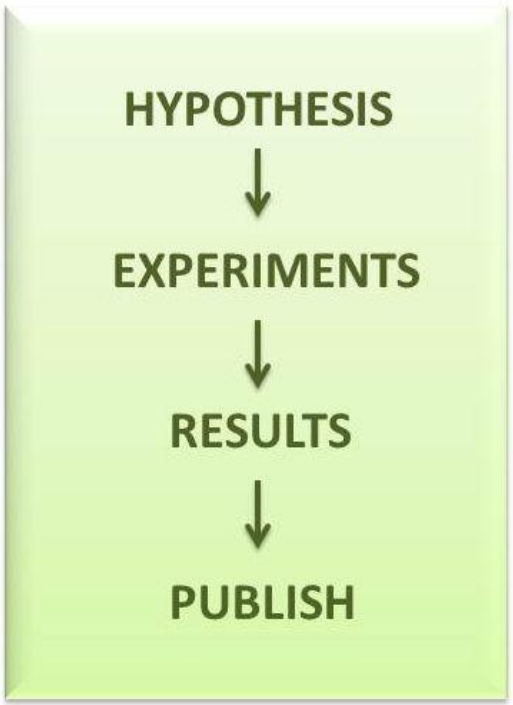


Figure 2: A more neutral approach, in which all results are published, as long as they are generated by well-carried out experiments based on sound hypotheses.

Why Science Needs to Publish Negative Results

The traditional journals are highly biased toward positive findings. As a consequence

1. Important negative findings often do not get reported.
2. Not knowing someone else has done the same experiment, the scientific community is at risk of spending time and money replicating failure rather than, as we do not do enough, replicating positive findings.

Should negative results be treated with the same rigor as positive results?

A negative finding may be important as it

i. Argues against a specific hypothesis (these can and are sometimes published).

ii. May support the null hypothesis.

iii. But a negative finding can just as easily arise because the experiment was not done correctly, because it was poorly designed, because it was simply a bad idea in the first place.

However, interpretation of negative findings can be subject to the same problem that interpretation of positive findings are: they can be over-interpreted, over-generalized, etc.

No result is worthless: the value of negative results in science

i. Fanelli (2012) demonstrates that negative results have been gradually disappearing from academic literature over the past two decades.

ii. Meanwhile, articles primarily and clearly stating positive results have grown 22% between 1990 and 2007.

iii. As positive results are more likely to lead to prestigious publications, discarding odd and unexpected findings is common in the scientific publishing system that privileges these “successful” results.

iv. Traditionally, it is expected that successful studies will obtain research findings in alignment with well-established literature or expected outcomes.

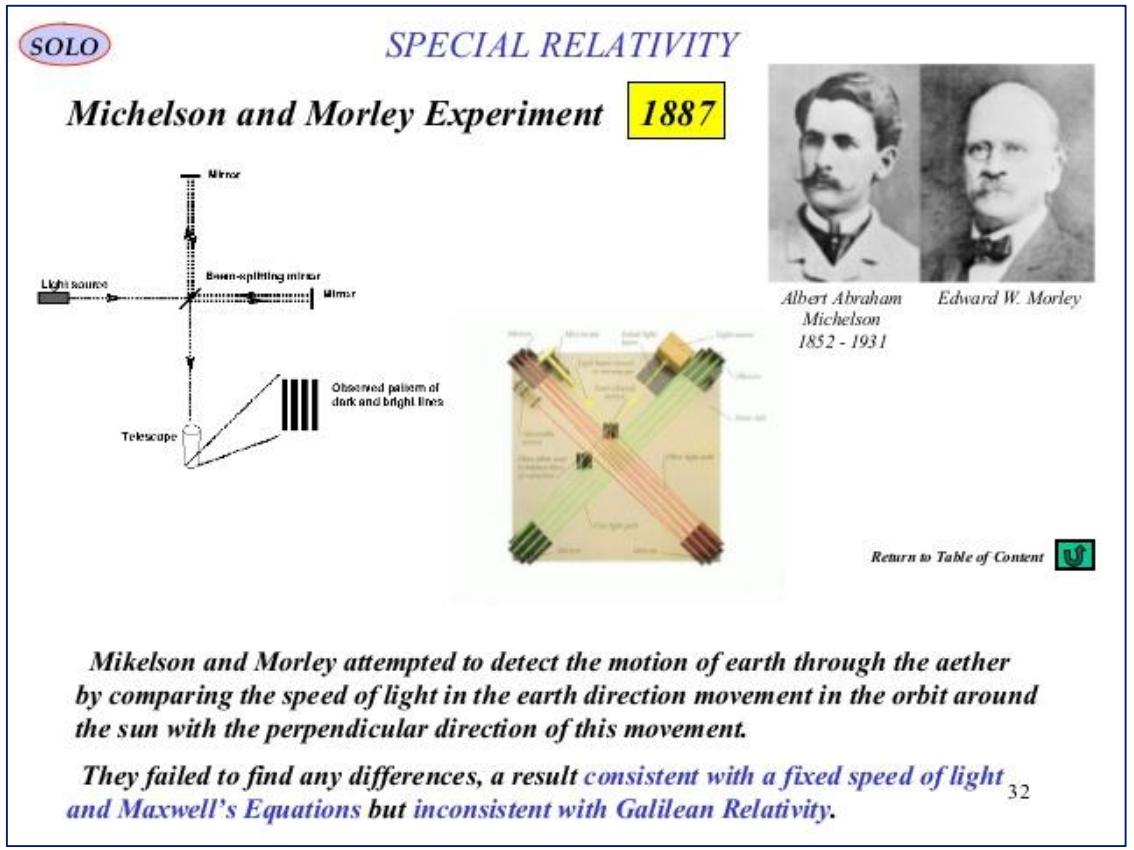
Are Negative Results Indeed Meaningless? Or Is There Potential Value In Sharing Negative Results With A Broader Academic Community?

- Reasons for the low profile of negative results publications, one common assumption is that publishing negative results might harm scientists' reputations.
- Along these lines, negative results are believed to indirectly communicate to the scientific community that a study was poorly designed and researchers were either unknowledgeable about the phenomenon or incapable of tailoring more robust research hypotheses.
- Moreover, the discouragement to submit publications reporting negative results is due to a higher likelihood that these papers will be filtered by the peer-review firewall, given their perceived lack of soundness in comparison to studies with "successful" results.

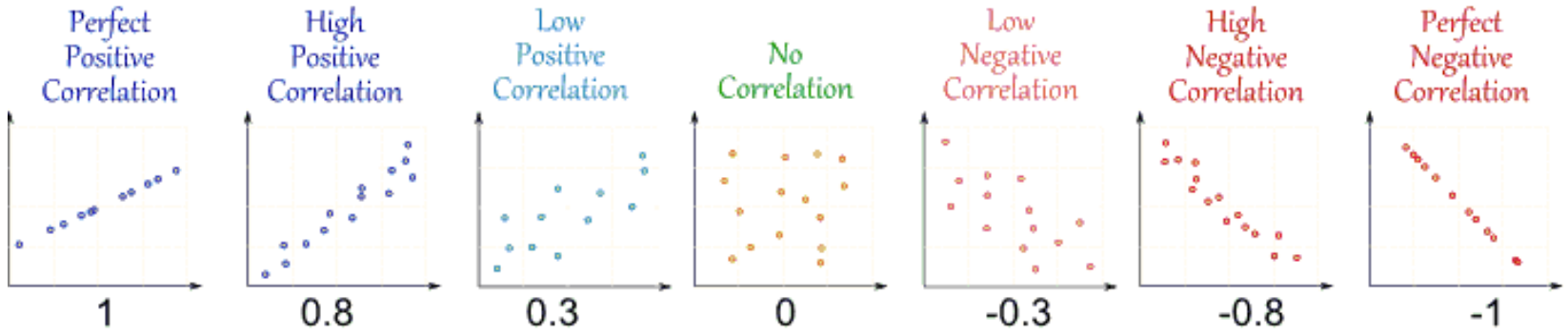

JOURNAL OF NEGATIVE RESULTS IN BIOMEDICINE
IMPACT FACTOR 1,47
The Null Journal
Negative Results, Failures, and Voids
JOURNAL`S OF NEGATIVE RESULTS
JOURNAL OF NEGATIVE RESULTS
- ECOLOGY & EVOLUTIONARY BIOLOGY -
NEGATIVE RESULTS
Journal of Pharmaceutical Negative Results
Journal of Articles in Support of the Null Hypothesis

Example Of Negative Results

- A classic example of negative results being recognized by researchers as scientific and ushering in a paradigm shift in science took place in the 17th century. All of their efforts to advance the theory led to a continual rejection of their research hypotheses. A few years later, these null results were published in the ***American Journal of Science*** and played an important role in inspiring new experiments, including a well known one that confirmed a major physical theory proposed by Albert Einstein in 1905: the special theory of relativity.



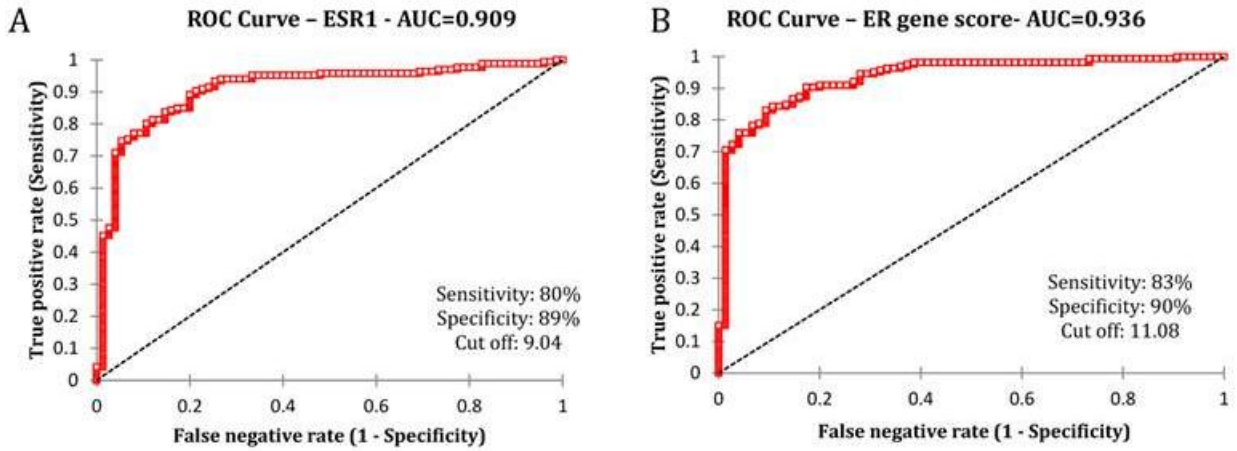
Correlation



- Correlation can have a value:
- **1** is a perfect positive correlation
- **0** is no correlation (the values don't seem linked at all)
- **-1** is a perfect negative correlation

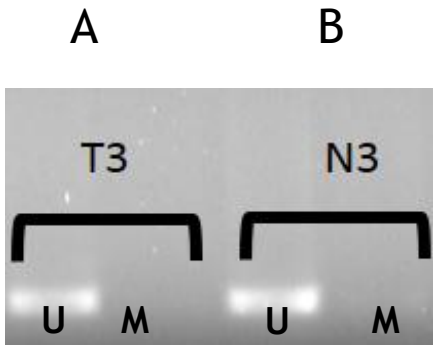
Absence of correlation as negative results between any 2 variables is very important as +ve and –ve correlation and must be interpreted and is worth publishing.

Receiver operating characteristic analysis



- Estrogen receptor (ER) is an important prognostic and predictive marker in human breast cancer. Patients with tumors that are positive for ER are known to respond to endocrine therapy and have improved disease specific survival and overall survival compared to those with tumors that are ER negative.

Methylation of hTERT promoter CTCF binding region in CRC



The methylation status of CTCF binding region of 16 samples showed that there was no difference between tumor tissue and normal mucosa, in which 13 showed unmethylated region, the rest were partially methylated both for tumor and normal mucosa as it is shown in Figure.
A. Majority of samples showed unmethylation, B. some are partially methylated

Association studies of ACE gene polymorphism and recurrent miscarriage in different populations

Reference	Ethnicity	Genotype frequencies				Associations (main results)
			DD n (%)	ID n (%)	II n (%)	
Buchholz et al., 2003	German	RM Case = 184 Control= 127	59 (32.1) 30 (23.6)	83 (45.1) 71 (55.9)	42 (22.8) 26 (20.5)	Significant association
Goodmanet al., 2009	American	RM Case= 120 Control=48	34 (28.3) 28 (33.3)	55 (45.8) 34 (40.8)	31 (25.8) 22 (26.2)	No significant differences
Sallout et al, 2010	Palestinians	RM Case=100 Control=100	49 49.0) 54 (54.0)	42 (42.0) 34 (34.0)	9 (9.0) 12 (12.0)	No significant differences
Zhang et al., 2011	Chinese	RM Case=127 Control=132	21 (16.5) 8 (6.1)	49 (38.6) 34 (25.8)	57 (44.9) 90 (68.2)	Strong association
Mello et al., 2003 42	Italian	RM cases = 48	25 (52)	20 (42)	3 (6)	Significant association.

Frequency of rs861539 (C/T) in XRCC3 different types of cancer in different populations.

rs861539 (C/T) Association of T allele				
Breast cancer	Protective role	100	Saudis	This study
Colorectal cancer	Increased Risk (association)	100	Polish	40
Breast Cancer	Increased Risk (association)	1826	UK	2
Bladder Cancer	Protective Role	214	Italian	42
Colerectal Cancer	Protective Role	128	British	4
Lung Cancer	No association	272	Danish	43
Breast Cancer	No association with cancer. Association with cancer progression and grading	700	Polish	3

Their Opinion,,,

Why science needs to publish negative results

Manuela Battaglia • 3 years ago

I'm not sure I agree in general. To me there is no such a thing as negative result in science. A result is a result. it is what you expected from the experiments that affects how you see your data. A well planned experiment never gives negative data. It gives a RESULT. I don't recall - in my 20-year old career as immunologist- a single data not published because considered "negative". Rather, many data were not published because the experiments were wrongly planned to address a specific question. However, you can then re-look at your data and give new-different interpretation.

Personally, I don't think I will ever submit my data to a journal called "Negatives on..." It goes against the way I see Science.
Thanks a lot for the article.

Fabiano Fernandes • 3 years ago

I agree that a so called negative result (or better: a result that do not show good improvement on a theme) should be published, especially if it is a new result. As said in the article, it will save time of many authors, and will point out where to go by showing where not to go. Unfortunatelly most of the journals and reviewers to not agree with it and it is almost impossible to publish such

“We need a better record to learn how well science distinguishes truth from fallacy.”


Thank
you