Research Methods
Chapter 2
Clever Hans

- Clever Hans the horse could do simple math and spell out the answers to simple questions. He wasn’t always correct, but he was most of the time.

- While a team of scientists, veterinarians, zoologists and circus trainers could not figure out how Hans was correctly answering the questions, Oskar Pfungst, a psychologist did. What did he discover?
While Hans could not do math or correctly answer questions on his own, he was very perceptive.

Hans was picking up on subtle body language given off by his owner who asked the questions.

When the owner was hidden from view, suddenly Hans could not answer the questions correctly.

How does this story relate to methodology?
Emily Rosa was a 4th grader in Colorado in 1998. She entered a science competition with an experiment that challenged the legitimacy of therapeutic touch (TT).

We will use Rosa’s experiment to look at scientific method.
The scientific method is a 5 step process for empirical investigation of a hypothesis, under conditions designed to control biases and subjective judgments.

- Empirical investigation is the collecting of objective information, firsthand, by making careful measurements based on direct experience.
The goal of psychology is to develop explanations for behavior and mental processes...explain why we do what we do.

These explanations, based on solid empirical studies are called *theories*.

- A theory is a testable explanation for a set of facts or observations.
1. Developing a Hypothesis:

**Hypothesis**: A statement predicting the outcome of a scientific study or describing the relationship among variables in a study.

- A hypothesis literally means a little theory.
All hypotheses must be **testable** and **falsifiable**, or shown to be either supported or incorrect.

- **Falsifiability** is the possibility that an assertion can be shown false by an observation or experiment. That something is "falsifiable" does not mean it is false; rather, that if it is false, then this can be shown by observation or experiment.

  A theory cannot be considered scientific if it does not admit the possibility of being shown false.
All good hypotheses need an operational definition.

An operational definition is a specific description of the concepts involving the conditions of the scientific study.

- Operational definitions are stated in terms of how the concepts are to be measured or what the operations are being employed to produce them.
Rosa and Pfungst’s Hypotheses

- Rosa sought to prove that TT practitioners could not accurately sense the presence of her hand above theirs if they could not see it there.

- For Hans, Pfungst operationalized his hypothesis by stating the horse could not give the correct answer when it could not see its owner.
2. Performing a controlled test: A hypothesis must undergo rigorous tests before it will be accepted as a legitimate theory.

To make a test controlled, one must account for the independent variable.

- **Independent Variable**: A stimulus condition that the experimenter changes independently of all other carefully controlled conditions in the experiment.
In Rosa’s experiment, she tested 21 TT practitioners to see if they could sense which of their two hands was closest to her hand when they could not see it.

To do this, she made a cardboard screen with two holes in the bottom. The practitioners would put their hands through, palms up. Rosa would hold her hand, palm down a few inches from either of the practitioner’s hands.
Eliminating Patterns

- For both Rosa and Hans’ experiments, the presence of patterns in the experiment could have jeopardized the findings. To eliminate this, Rosa and Pfungst used random presentation.

- Random presentation is a process by which chance alone determines the order in which the stimulus is presented.
In Rosa’s experiment, randomization was achieved by a coin flip to determine whether she put her hand above the practitioner’s left or right hand.

In the Hans experiment, Pfungst made sure to ask math questions with random answers in which there were no predictable patterns (answers of 2, 4, 6...).
3. Gathering objective data: getting information by direct observation that relies only on the independent variable and not on the experimenter’s hopes. This data is called the dependent variable.

- **Dependent Variable**: The measured outcome of a study, or the response of the subjects in the study.
A good way to remember which is which:

- Independent Variable (IV) = stimulus or cause
- Dependent Variable (DV) = response or effect

- Both the IV and the DV must have an operation definition. That means, you must explain what each will look like and how it will be measured.
Rosa and Hans

- For Rosa, she simply recorded whether the TT practitioner said “left” or “right.”

- For Pfungst’s study of Hans, the DV was simply the horse’s hoof-taping in response to each question asked, making sure to keep track of the number of correct answers given.
4. **Analyzing the results:** This step consists of looking at the data collected and seeing if it supports or disproves the hypothesis.

- We will briefly discuss stats next class, but it is not a major part of our psychology class. No worries, this is not a math class.
In Rosa’s experiment, the analysis was pretty clear. By design, there was a 50% chance a practitioner could guess correctly. So in order to disprove her hypothesis, they would have to answer correctly significantly more than 50% of the time...they did not.

- She **concluded** that TT practitioners could not detect the “human energy field.”
Pfungst’s Analysis

- For Han’s the chance level of simply guessing the correct response was near zero, so any consistent level of correct responses would support the hypothesis that Hans could do math.

- That hypothesis was rejected, however, as Hans was unable to correctly answer any questions in the absence of his owner.
5. Publishing, criticizing and replicating the results: The last step of the scientific method is to have the results withstand the criticism and scrutiny of the science community.

- Critics check each others’ work by replicating the study, sometimes under slightly different circumstances to see if the same results can be duplicated.

Replicate: To do a study over to see if the same results are obtained. To control for bias, the replication is most often done by someone other than the original researcher.
Experimental Method: A kind of research in which the researcher controls and manipulates the conditions including the IV.

Experimental method must account for independent variables, dependent variables and confounding or extraneous variables.
A good way to remember which is which:

- Independent Variable (IV) = stimulus or cause
- Dependent Variable (DV) = response or effect

- Both the IV and the DV must have an operation definition. That means, you must explain what each will look like and how it will be measured.
**Confounding Variables**

- **Confounding Variables**: Variables that have unwanted influence on the outcome of an experiment.
  - Or, other possible explanations for the dependent variable (result).
The Challenges of Experiments

- There are many challenges with conducting experiments. First one has to make sure that all groups being tested have the same conditions. This is called control.

- Second, for an experiment to be valid, one has to make sure the subjects are drawn from a population which consists of everyone who fits the description of your test group.
To ensure we have a group which represents the demographic we want, we must use random selection.

**Random Selection:** Each subject of the sample has an equal likelihood of being chosen for the *experimental group*, or the group which gets the *independent variable*.

- Ex. Names drawn out of a hat.
Sampling

- To have confidence in results, they need to be taken from a sample of participants chosen in an unbiased manner.

- **Random Sample:** A sample group of subjects selected by chance, or without biased selection techniques.
Representative Sample: A sample obtained in such a way that it reflects the distribution of important variables in the larger population in which the researcher are interested—variables such as age, SES, ethnicity, education....
On the sheet of paper...

Please write the following things:

- Gender
- Shoe size
- Height in cm
- Hair length in cm

\[ \text{in} \times 2.54 = \text{cm} \]

- Do not write your name on the sheet of paper
Non-Experiment Designs

- Sometimes we are unable to do experiments for ethical or practical reasons. In this case we must do another kind of research.

- **Ex post facto:** Research in which we choose subjects based on a pre-existing condition.
  - Ex: Cancer research.
A correlation study is one where researchers try to show the relationship (or correlation) between two variables.

Correlation studies are largely based in statistics.
- It is important to remember that correlation does not necessarily mean causation.
A **survey** is a research method where questions are asked to subjects who report their own answers.

What are some dangers of using a survey?
Naturalistic observations are a method where subjects are observed in their natural environment.

Why would it be important for subjects to not know they are being observed?
In a **longitudinal study**, one group or subject is studied for an extended period of time to observe changes in the long term.

- Same subjects for the entire study
- Time and expense
These studies are designed to cut down on time and expense.

- **Cross-sectional studies** look at a cross section of the population and studies them at one point in time.
  - Ex: No child left behind

- **Cohort-sequential studies** look at a cross section of population and then studies them over a short period of time.
Sources of Bias

- **Personal Bias**: When the researcher allows his or her personal beliefs affect the outcome of the study.

- **Expectancy Bias**: When the researcher allows his or her expectations to affect the outcome of the study.
Reducing Bias

- **Double Blind Study:** An experiment where both the subject and the person administering the experiment do not know the nature of the independent variable being administered.
Ethics in Research

- Each university or group doing research must have an Institutional Review Board which is responsible for making sure research is performed in an ethical manner.

- The APA says deception is to be avoided whenever possible. However, when deception must be used, the subjects are to be debriefed as soon as possible after the study.
**Frequency Distribution:** A summary chart which shows how frequently each of the various scores in a set of data occur.

<table>
<thead>
<tr>
<th>Battery life, minutes</th>
<th>Frequency (f)</th>
<th>Relative frequency</th>
<th>Percent frequency</th>
</tr>
</thead>
<tbody>
<tr>
<td>360–369</td>
<td>2</td>
<td>0.07</td>
<td>7</td>
</tr>
<tr>
<td>370–379</td>
<td>3</td>
<td>0.10</td>
<td>10</td>
</tr>
<tr>
<td>380–389</td>
<td>5</td>
<td>0.17</td>
<td>17</td>
</tr>
<tr>
<td>390–399</td>
<td>7</td>
<td>0.23</td>
<td>23</td>
</tr>
<tr>
<td>400–409</td>
<td>5</td>
<td>0.17</td>
<td>17</td>
</tr>
<tr>
<td>410–419</td>
<td>4</td>
<td>0.13</td>
<td>13</td>
</tr>
<tr>
<td>420–429</td>
<td>3</td>
<td>0.10</td>
<td>10</td>
</tr>
<tr>
<td>430–439</td>
<td>1</td>
<td>0.03</td>
<td>3</td>
</tr>
</tbody>
</table>

**Table: Life of AA batteries, in minutes**
Using Data for Analysis

- **Histogram**: A bar graph depicting a frequency distribution. The height of the bars indicates the frequency of a group of scores.
Mean (average): The measure of central tendency most often used to describe a set of data.

- To calculate mean, simply add all the scores and divide by the number of scores.
- *While the mean is easy to calculate, it has a big downside. It can easily be influenced by extreme scores.*
Median: A measure of central tendency represented by the score that separates the upper half of the scores in a distribution from the lower half.

- The big advantage of this is the median is not effected by extreme scores.

Mode: A measure of central tendency which represents the score that occurs most often.
The weekly salaries of six employees at McDonalds are $140, $220, $90, $180, $140, $200.

For these six salaries, find:
- (a) the mean
- (b) the median
- (c) the mode
Mean, Median, Mode

**Answers**

**Mean:**  \[
\frac{90 + 140 + 140 + 180 + 200 + 220}{6} = 161.67
\]

**Median:**  
90, 140, 140, 180, 200, 220

The two numbers that fall in the middle need to be averaged.

\[
\frac{140 + 180}{2} = 160
\]

**Mode:**  
90, 140, 140, 180, 200, 220

The number that appears the most is 140
Standard Deviation (SD): A measure of variability that indicates the average distance between the scores and their mean.

- A low standard deviation indicates that the data points tend to be very close to the mean, whereas high standard deviation indicates that the data are spread out over a large range of values.
Normal Distribution

- The standard deviation and mean together tell us a lot about the distribution of scores.

A data set with a mean of 50 (shown in blue) and a standard deviation of 20.
Normal Distribution

A normal distribution is a bell shaped curve.

A standard deviation of 15 accounts for about 68% of responses.
A distribution is skewed if one of its tails is longer than the other.

- The first distribution shown has a **positive skew**. This means that it has a long tail in the positive direction.

- The second distribution has a **negative skew** since it has a long tail in the negative direction.

- Finally, the third distribution is symmetric and has no skew (**normal distribution**).
A Skewed Distribution

- Are the results positively or negatively skewed?
**Correlation: Negative and Positive**

- **Correlation**: A relationship between two variables in which change in one variable are reflected in the changes in the other variable.
  - **Correlation Coefficient**: A number between $-1$ and $+1$ expressing the degree of relationship between two variables.
## Comparing Correlation Coefficients

- If the correlation coefficient is a positive number, there is a positive correlation (connection) between the variables.
  - SAT scores and college achievement—among college students, those with higher SAT scores also have higher grades

- If the correlation coefficient is a negative number, there is a negative correlation (connection) between variables.
  - Education and years in jail—people who have more years of education tend to have fewer years in jail

- If the correlation coefficient is 0, there is no correlation between variables.
Positive Correlation Coefficients

- Positive correlation coefficients indicate a stronger connection as they get closer to 1.
# Comparing Research Methods

<table>
<thead>
<tr>
<th>Research Method</th>
<th>Basic Purpose</th>
<th>How Conducted</th>
<th>What Is Manipulated</th>
</tr>
</thead>
<tbody>
<tr>
<td>Descriptive</td>
<td>To observe and record behavior</td>
<td>Case studies, surveys, and naturalistic observations</td>
<td>Nothing</td>
</tr>
<tr>
<td>Correlational</td>
<td>To detect naturally occurring relationships; to assess how well one variable predicts another</td>
<td>Computing statistical association, sometimes among survey responses</td>
<td>Nothing</td>
</tr>
<tr>
<td>Experimental</td>
<td>To explore cause and effect</td>
<td>Manipulating one or more factors and using random assignment to eliminate preexisting differences among subjects</td>
<td>The independent variable(s)</td>
</tr>
</tbody>
</table>
