Module 1:
Data sources; data types; concepts of validity

MSIR 525

Monday, September 9, 2019

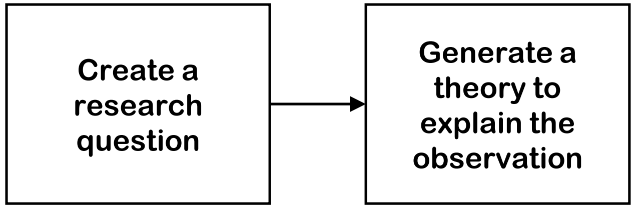
Agenda

- Textbook vs. actual HR analytics
- Null hypothesis significance testing
  - Test statistics
  - One-tailed vs. two-tailed tests
  - Type I error vs. Type II error
- Statistical power
- Research designs
  - experimental; correlational; quasi-experimental
  - Cross-sectional; longitudinal; episodic
- Data sources
- Data types
- XX types of validity

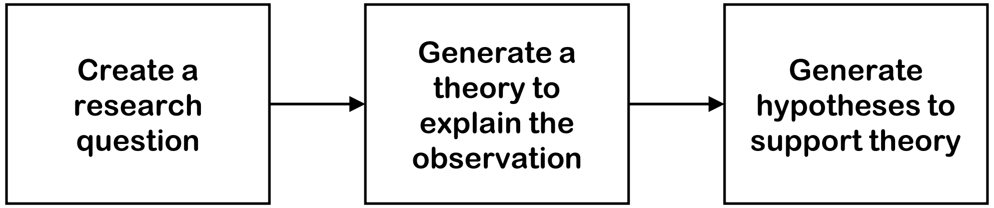
The textbook approach to HR analytics/evidence-based practice

Create a research question

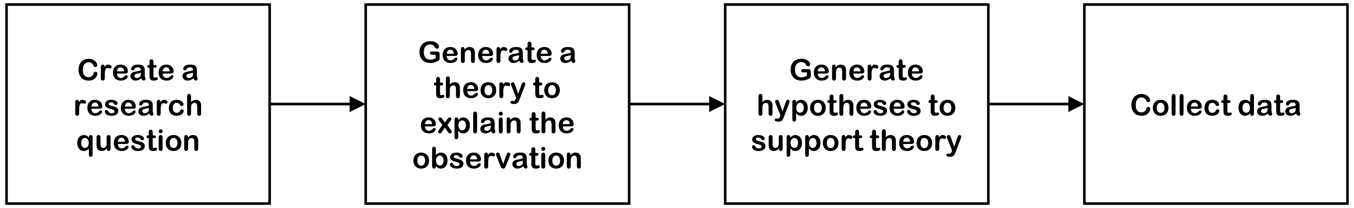
The textbook approach to HR analytics/evidence-based practice



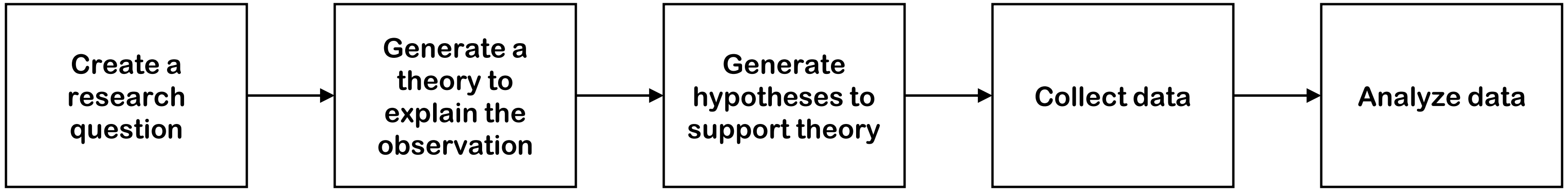
The textbook approach to HR analytics/evidence-based practice



The textbook approach to HR analytics/evidence-based practice



The textbook approach to HR analytics/evidence-based practice



The textbook approach to HR analytics/evidence-based practice

Create a research question → Generate a theory to explain the observation → Generate hypotheses to support theory → Collect data → Analyze data

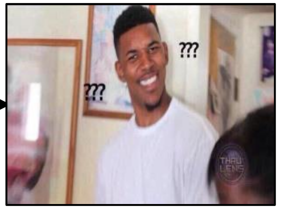
What HR analytics/evidence-based practice really looks like

The textbook approach to HR analytics/evidence-based practice

Create a research question → Generate a theory to explain the observation → Generate hypotheses to support theory → Collect data → Analyze data

What HR analytics/evidence-based practice really looks like

→



The textbook approach to HR analytics/evidence-based practice

Create a research question → Generate a theory to explain the observation → Generate hypotheses to support theory → Collect data → Analyze data

What HR analytics/evidence-based practice really looks like

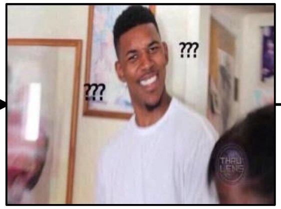
→

→

The textbook approach to HR analytics/evidence-based practice

Create a research question → Generate a theory to explain the observation → Generate hypotheses to support theory → Collect data → Analyze data

What HR analytics/evidence-based practice really looks like

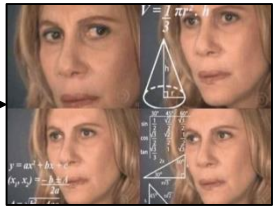


The textbook approach to HR analytics/evidence-based practice

Create a research question → Generate a theory to explain the observation → Generate hypotheses to support theory → Collect data → Analyze data

What HR analytics/evidence-based practice really looks like

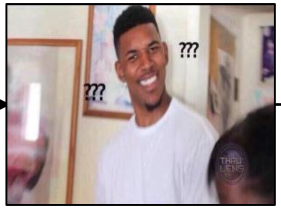


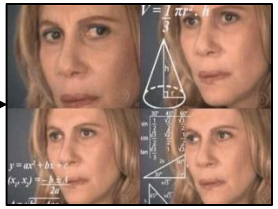



The textbook approach to HR analytics/evidence-based practice

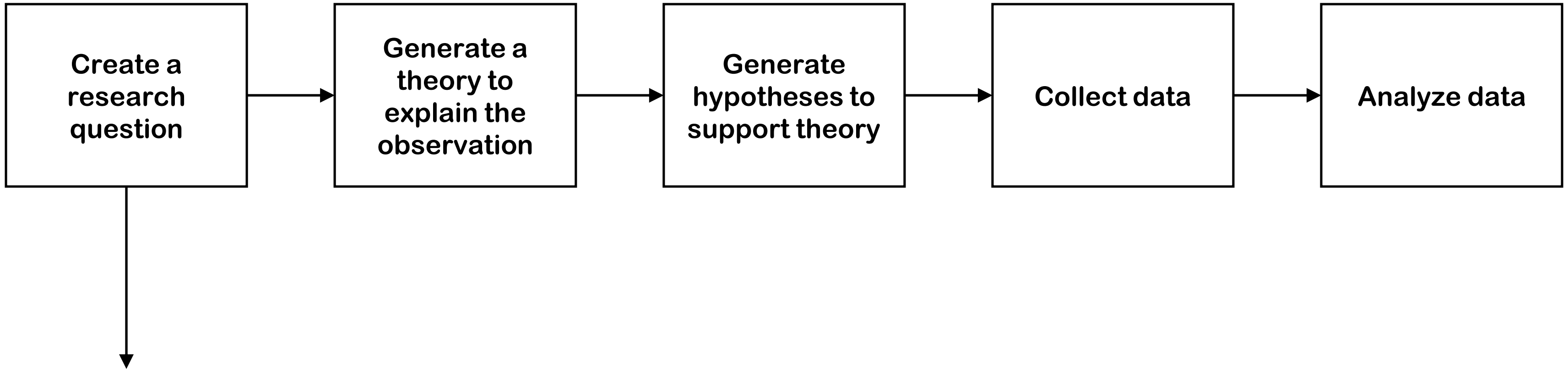
Create a research question → Generate a theory to explain the observation → Generate hypotheses to support theory → Collect data → Analyze data

What HR analytics/evidence-based practice really looks like





The textbook approach to HR analytics/evidence-based practice

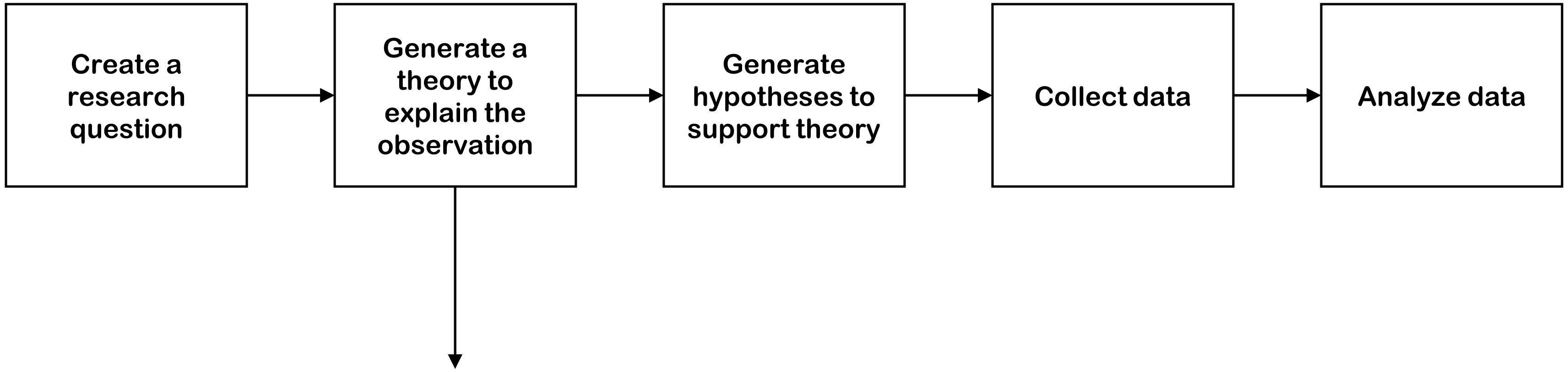


Problem identification –or- future state target:

Example: What is driving our star employees to turnover (i.e., quit)?

Example: What characteristics should we be (not) looking for in a job candidate?

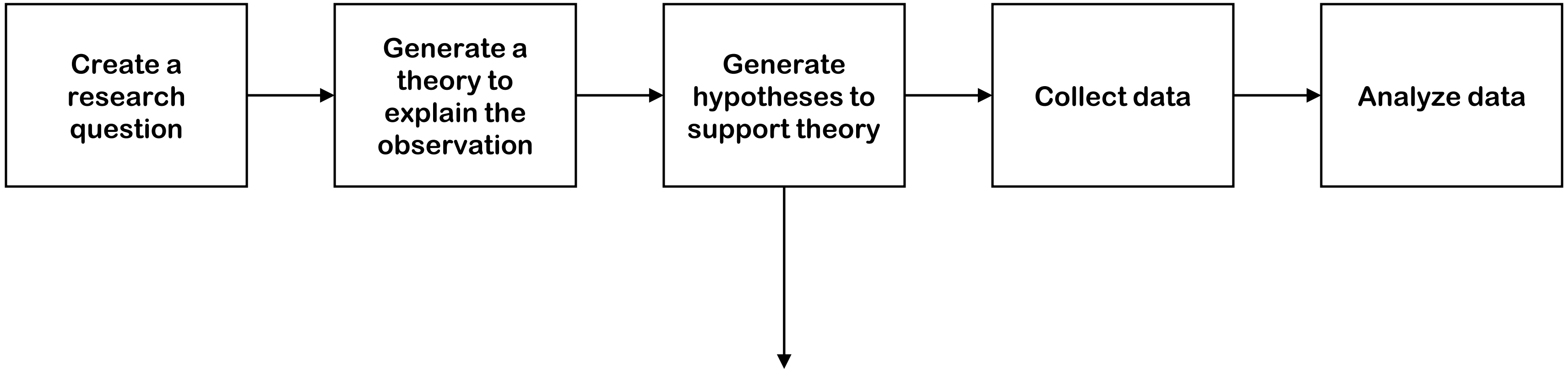
The textbook approach to HR analytics/evidence-based practice



Example: What is driving our star employees to turnover (i.e., quit)?

According to personal resources theory, high levels of demands at work require one to focus personal resources in this area leaving fewer resources to tackle demands in other areas, such as those in the family domain.

The textbook approach to HR analytics/evidence-based practice

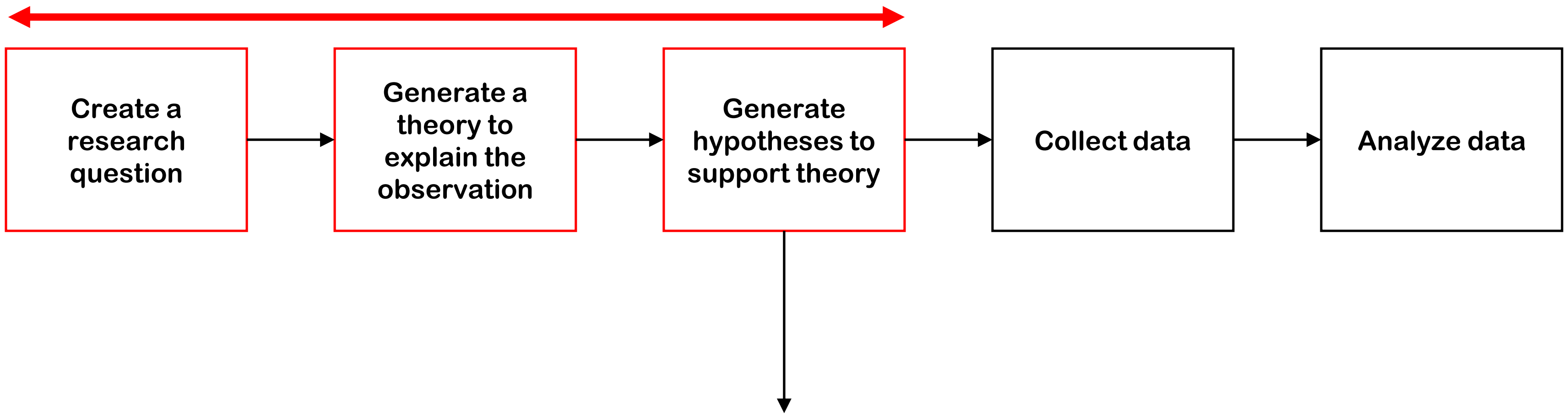


Example: What is driving our star employees to turnover (i.e., quit)?

Theory: Personal resources theory

Hypothesis: Individuals who are emotionally exhausted are more prone to quitting.

The textbook approach to HR analytics/evidence-based practice



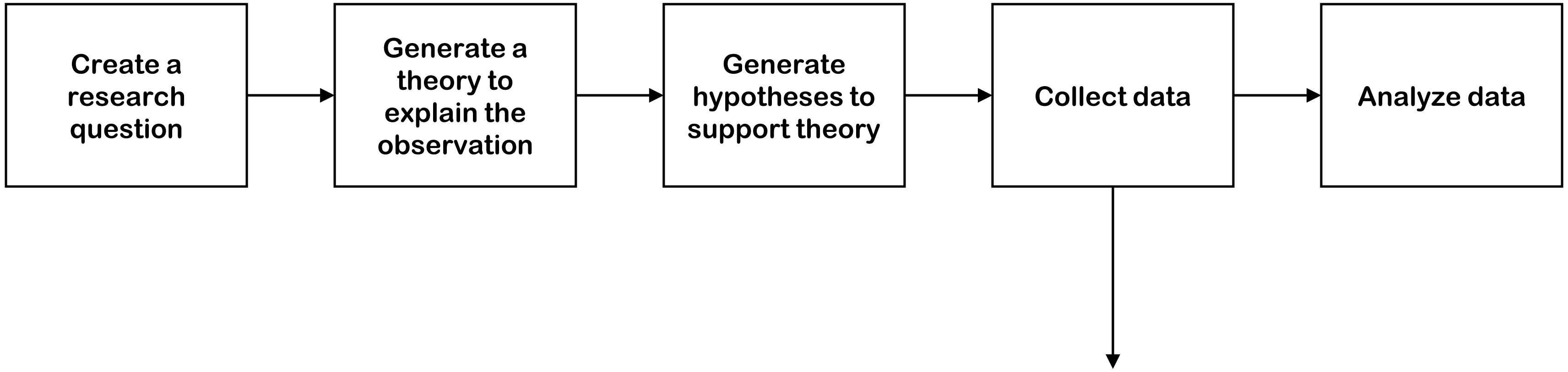
Example: What is driving our star employees to turnover (i.e., quit)?

Theory: Personal resources theory

Hypothesis: Individuals who are emotionally exhausted are more prone to quitting.

Not this class!

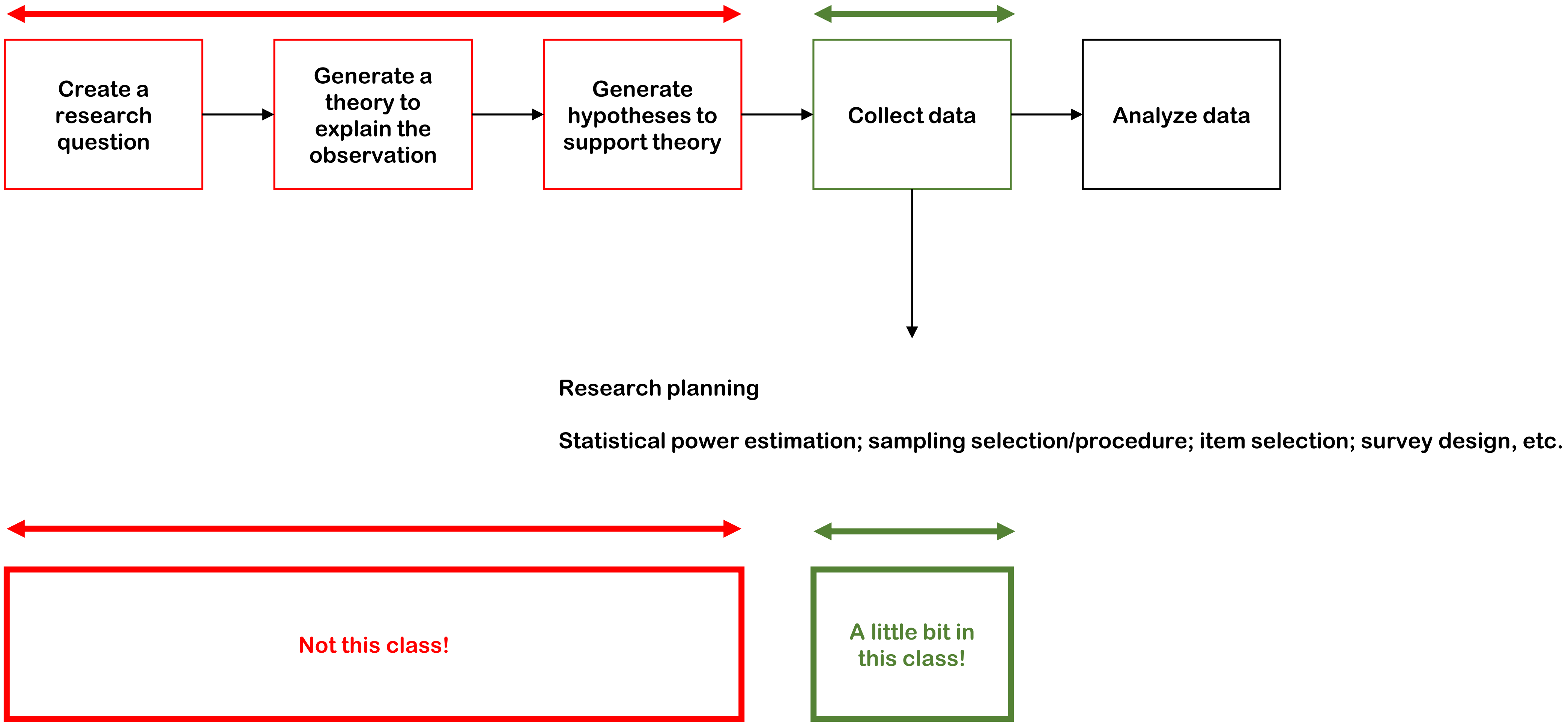
The textbook approach to HR analytics/evidence-based practice



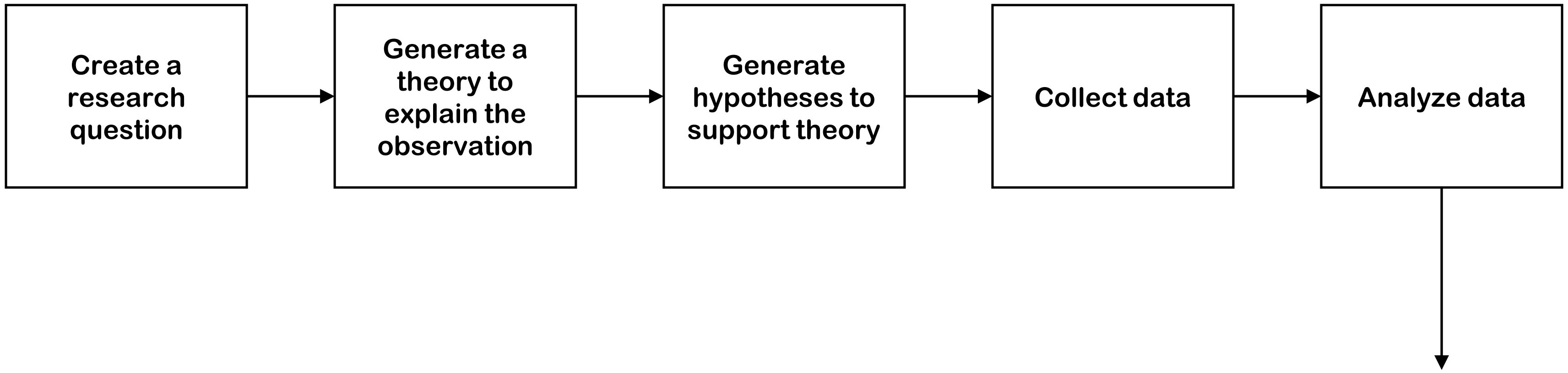
Research planning

Statistical power estimation; sampling selection/procedure; item selection; survey design, etc.

The textbook approach to HR analytics/evidence-based practice



The textbook approach to HR analytics/evidence-based practice



Statistical analysis (depends on what the objective is)

Descriptive vs. prescriptive

Correlation (i.e., association); linear regression; group differences and moderation; non-linear relations

The textbook approach to HR analytics/evidence-based practice

Create a research question → Generate a theory to explain the observation → Generate hypotheses to support theory → Collect data → Analyze data

Statistical analysis (depends on what the objective is)

Descriptive vs. prescriptive

Correlation (i.e., association); linear regression; group differences and moderation; non-linear relations

Not this class!

A little bit in this class!

A lot in this class!

The textbook approach to HR analytics/evidence-based practice

Create a research question → Generate a theory to explain the observation → **Generate hypotheses to support theory** → Collect data → Analyze data

Statistical analysis (depends on what the objective is)

Descriptive vs. prescriptive

Correlation (i.e., association); linear regression; group differences and moderation; non-linear relations

Not this class!

A little bit in this class!

A lot in this class!

Hypothesis Testing Framework

Generate hypotheses to support theory

- Null hypothesis significance testing (Fisher, 1925)
  - Null hypothesis vs. alternative hypothesis

Hypothesis Testing Framework

Generate hypotheses to support theory

- Null hypothesis significance testing (Fisher, 1925)

Null hypothesis:

A statistical test of the hypothesis that suggests that there is no difference between specified populations (or no relation between constructs) and that any observed difference is due to sampling or experimenter error.

$r = 0$

Alternative hypothesis:

A statistical test of the hypothesis that suggests that there is a difference between specified populations (or relation between constructs).

$r \neq 0$

Hypothesis Testing Framework

Generate hypotheses to support theory

- Null hypothesis significance testing (Fisher, 1925)

Null hypothesis:

A statistical test of the hypothesis that suggests that there is no difference between specified populations (or no relation between constructs) and that any observed difference is due to sampling or experimenter error.

$r = 0$

Alternative hypothesis:

A statistical test of the hypothesis that suggests that there is a difference between specified populations (or relation between constructs).

$r \neq 0$

Null hypothesis:

There is no relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover

$r = 0$

Alternative hypothesis:

There is a positive relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover

$r > 0$

Hypothesis Testing Framework

Generate hypotheses to support theory

- Null hypothesis significance testing (Fisher, 1925)

Null hypothesis:

A statistical test of the hypothesis that suggests that there is no difference between specified populations (or no relation between constructs) and that any observed difference is due to sampling or experimenter error.

$r = 0$

Alternative hypothesis:

A statistical test of the hypothesis that suggests that there is a difference between specified populations (or relation between constructs).

$r \neq 0$

Null hypothesis:

There is no relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover

$r = 0$

Alternative hypothesis:

There is a positive relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover

$r > 0$

Hypothesis Testing Framework

Generate hypotheses to support theory

- Null hypothesis significance testing (Fisher, 1925)

Null hypothesis:

A statistical test of the hypothesis that suggests that there is no difference between specified populations (or no relation between constructs) and that any observed difference is due to sampling or experimenter error.

$r = 0$

Alternative hypothesis:

A statistical test of the hypothesis that suggests that there is a difference between specified populations (or relation between constructs).

$r \neq 0$

BUT HOW DO WE DETERMINE WHETHER OR NOT WE CAN TRUST THE RESULT?

(SPECIFICALLY, WHAT MUST OCCUR FOR US TO "BELIEVE" THAT EMOTIONAL EXHAUSTION IS ASSOCIATED WITH AND, THUS, MAY CAUSE AN INDIVIDUAL TO QUIT THEIR JOB?)

Null hypothesis:

There is no relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover
$r = 0$

Alternative hypothesis:

There is a positive relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover
$r > 0$

Hypothesis Testing Framework

Generate hypotheses to support theory

- Null hypothesis significance testing (Fisher, 1925)

Null hypothesis:

A statistical test of the hypothesis that suggests that there is no difference between specified populations (or no relation between constructs) and that any observed difference is due to sampling or experimenter error.

$r = 0$

Alternative hypothesis:

A statistical test of the hypothesis that suggests that there is a difference between specified populations (or relation between constructs).

$r \neq 0$

BUT HOW DO WE DETERMINE WHETHER OR NOT WE CAN TRUST THE RESULT?

(SPECIFICALLY, WHAT MUST OCCUR FOR US TO "BELIEVE" THAT EMOTIONAL EXHAUSTION IS ASSOCIATED WITH AND, THUS, MAY CAUSE AN INDIVIDUAL TO QUIT THEIR JOB?)

Null hypothesis:

There is no relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover
$r = 0$

Alternative hypothesis:

There is a positive relation between emotional exhaustion and turnover behavior

Emotional exhaustion → Turnover
$r > 0$

IN A NUTSHELL, THE P-VALUE MUST BE LESS THAN .05.

BUT, WHAT DOES THAT MEAN?

Hypothesis Testing Framework

Generate hypotheses to support theory

- $p < .05$
  - The Lady Tasting Tea (Fisher, 1956)
  - The take-home point…
    - When there was a very small probability that the woman could complete the tea-task by luck alone would we conclude that she had a genuine skill in detecting whether milk was poured into a cup before or after the tea.
    - If there is a very small probability that *emotional exhaustion* is associated with *turnover behavior* (i.e., by chance; 5%), we begin to “believe” that they are related to each other

Hypothesis Testing Framework

Generate hypotheses to support theory

- *p* < .05

  - If there is a very small probability that *emotional exhaustion* is associated with *turnover behavior* (i.e., by chance; 5%), we begin to “believe” that they are related to each other

  - In this case, we would *reject the null hypothesis* (basically, we say that it is wrong to say that the two constructs are unrelated)

  - Furthermore, we conclude that the relation between EE and TO is *statistically significant*

Hypothesis Testing Framework

Generate hypotheses to support theory

- $p > .05$
  - If we observed no relation between emotional exhaustion and turnover behavior (i.e., $p > .05$), we would *fail to reject the null hypothesis*
  - Note that we <u>do not</u> accept the null hypothesis
  - Furthermore, we could state that the relation between EE and TOI is *not statistically significant* because the corresponding *p*-value is greater than .05

Hypothesis Testing Framework

Generate hypotheses to support theory

- *p* > .05 (diving deeper!)
  - What does a *p*-value ***<u>really</u>*** tell us?
  - In a nutshell, the *p*-value tells us how well the independent variable predictors the dependent variable
  - In other words, the *p*-value tells us how well the independent variable explains the behavior of the dependent variable

Hypothesis Testing Framework

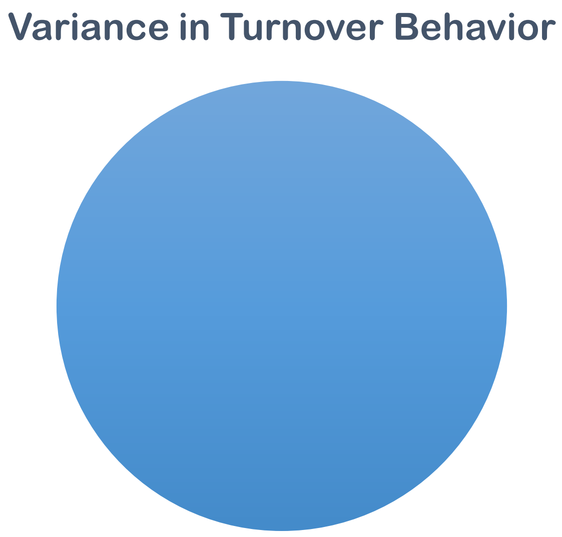
Generate hypotheses to support theory

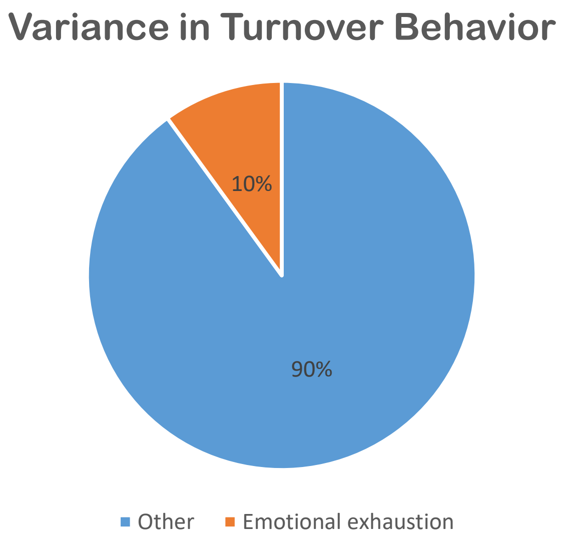
- $p > .05$ (diving deeper!)
  - What does a *p*-value ***really*** tell us?
  - In a nutshell, the *p*-value tells us how well the independent variable predictors the dependent variable
  - In other words, the *p*-value tells us how well the independent variable explains the variance (i.e., change in) of the dependent variable
    - We use a metric called R^2 (r-squared) to capture this

Hypothesis Testing Framework

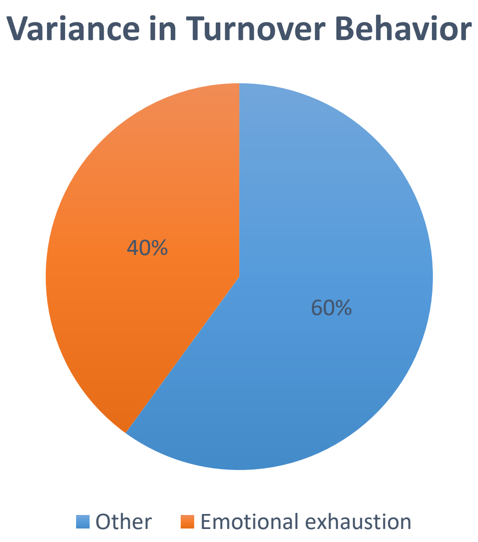
Generate hypotheses to support theory

- p > .05 (diving deeper!)
    - What does a p-value ***<u>really</u>*** tell us?





R^2 = .10



R^2 = .40

Hypothesis Testing Framework

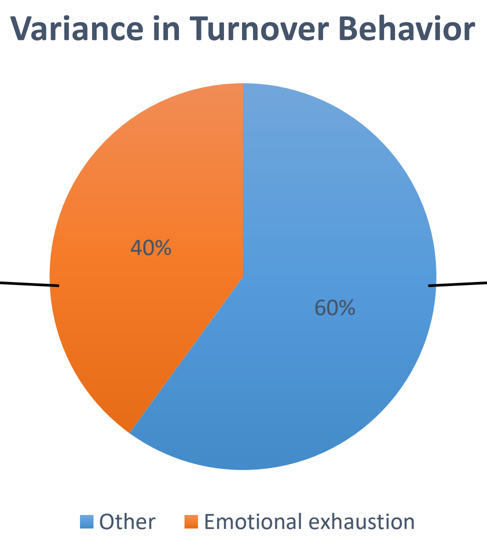
Generate hypotheses to support theory

- *p* > .05 (diving deeper!)
  - What does a *p*-value ***really*** tell us?

Systematic variance

This variation in the outcome variable (i.e., turnover behavior) is due to the independent variable of interest (i.e., emotional exhaustion)

In this example, we can conclude that 40% of the change in turnover behavior is due to emotional exhaustion alone.



Unsystematic variance

This variance in the outcome variable (i.e., turnover behavior) results from random factors.

In this example, we can conclude that 60% of the change in turnover behavior is due to everything else not including emotional exhaustion.

Hypothesis Testing Framework

Generate hypotheses to support theory

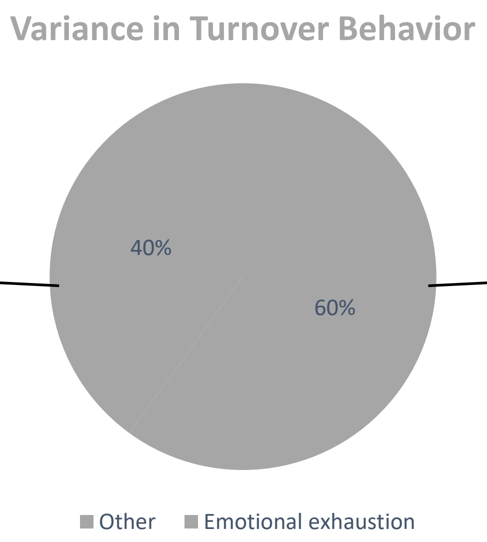
- *p* > .05 (diving deeper!)
  - What does a *p*-value ***really*** tell us?

$$\text{Test statistic} = \frac{\text{Systematic variance}}{\text{Unsystematic variance}} = \frac{\text{Effect}}{\text{Error}}$$

Systematic variance

This variation in the outcome variable (i.e., turnover behavior) is due to the independent variable of interest (i.e., emotional exhaustion)

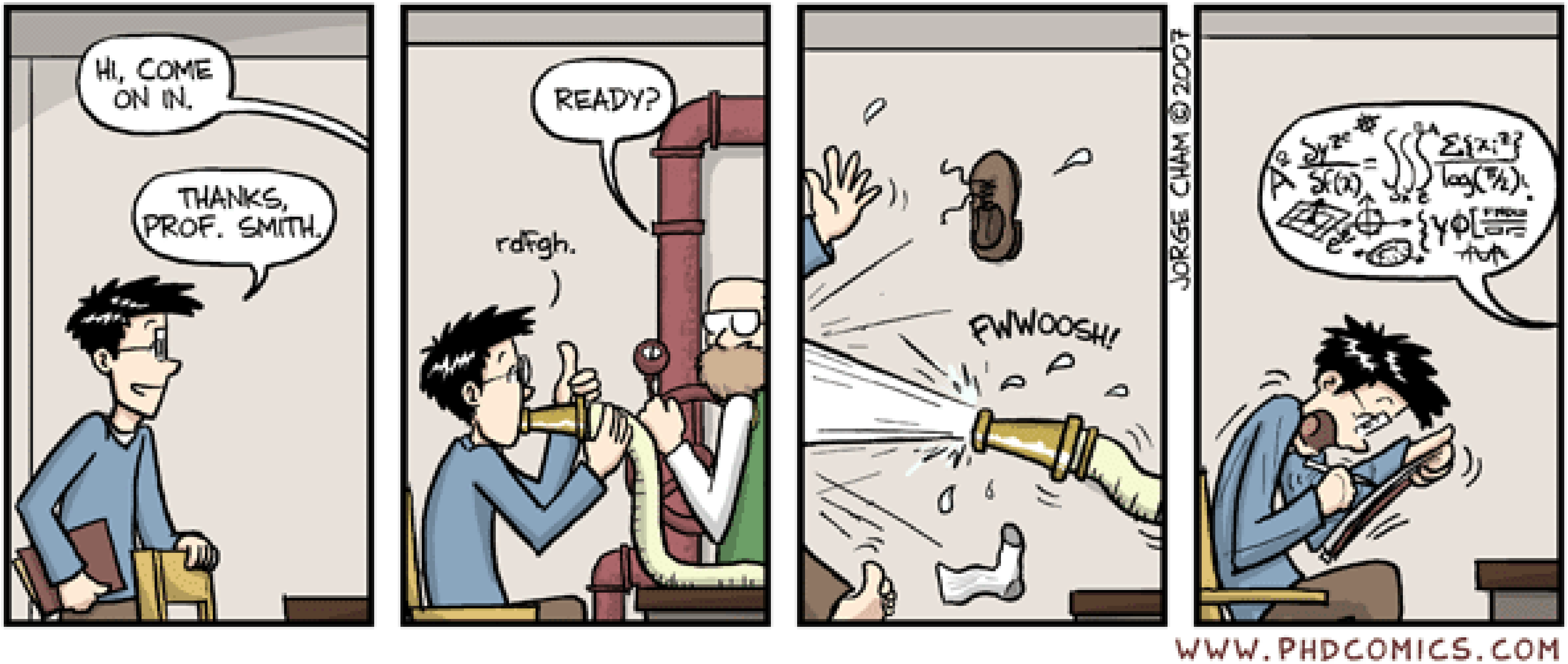
In this example, we can conclude that 40% of the change in turnover behavior is due to emotional exhaustion alone.



Unsystematic variance

This variance in the outcome variable (i.e., turnover behavior) results from random factors.

In this example, we can conclude that 60% of the change in turnover behavior is due to everything else not including emotional exhaustion.


HI, COME ON IN.
THANKS, PROF. SMITH.
READY?
rdfgh.
FWWOOSH!
JORGE CHAM © 2007
WWW.PHDCOMICS.COM

Hypothesis Testing Framework

Generate hypotheses to support theory

- Summary

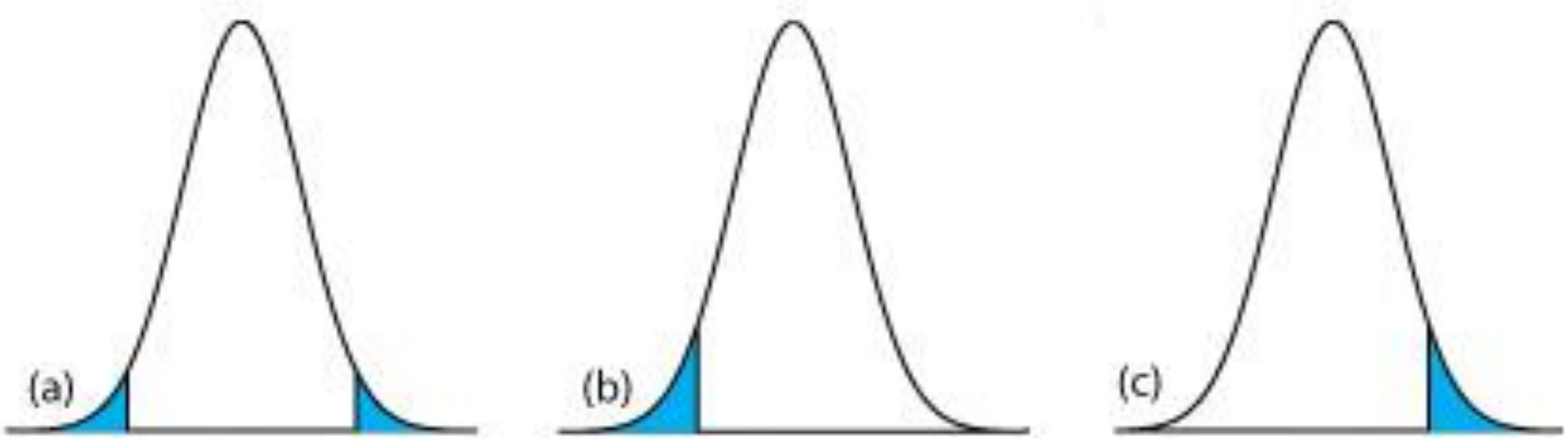
	When the *p*-value is…	
	> .05	< .05
	Fail to reject the null	Reject the null
	Claim that the effect is likely not to be present	Claim that the effect is likely to be present
We…	(Informally, we are saying that is likely wrong to say that an effect is present)	(Informally, we are saying that is likely wrong to say that there is no effect)
	State that the observed result is not statistically significant	State that the observed result is statistically significant
	Should not use propose an evidence-based practice recommendation	Have grounds to make an evidence-based practice recommendation

- NOTE: We never claim to "accept the null" or to "accept the alternative"

Hypothesis Testing Framework

Generate hypotheses to support theory

- Null hypothesis significance testing (Fisher, 1925)
  - One- vs. two-tailed tests



Two-tailed test	**One-tailed test**	**One-tailed test**
There is a relation between emotional exhaustion and turnover behavior.	**There is a negative relation between emotional exhaustion and turnover behavior.**	**There is a positive relation between emotional exhaustion and turnover behavior.**

Hypothesis Testing Framework

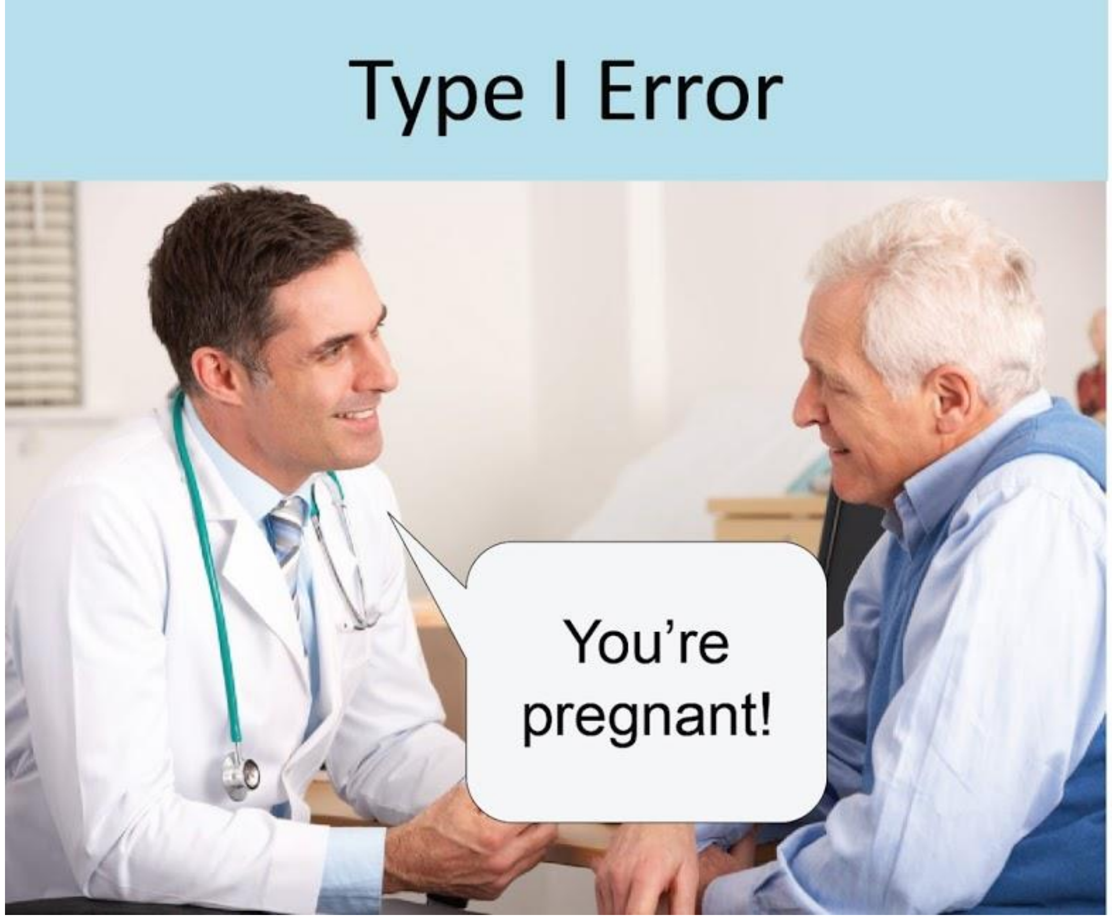
Generate hypotheses to support theory

- Type I vs. Type II error

Hypothesis Testing Framework

Generate hypotheses to support theory

- Type I vs. Type II error



Hypothesis Testing Framework

Generate hypotheses to support theory

- Type I vs. Type II error

Hypothesis Testing Framework

Generate hypotheses to support theory

- Type I vs. Type II error

Hypothesis Testing Framework

Generate hypotheses to support theory

- Type I vs. Type II error

Hypothesis Testing Framework

Generate hypotheses to support theory

- Type I vs. Type II error

Hypothesis Testing Framework

Generate hypotheses to support theory

- Type I vs. Type II error

Table of errors		Null hypothesis is… "True"	Null hypothesis is… "False"
Decision about the null hypothesis	Fail to reject	Correct inference *(true negative)*	Type II error *(false negative)*
	Reject	Type I error *(false positive)*	Correct inference *(true positive)*

Hypothesis Testing Framework

Generate hypotheses to support theory

- How to avoid Type I and Type II errors?
    - Full disclosure, the relationship between TI and TII is *very* complex. Although we know that as the probability of making a TI error decreases, the probability of making a TII error increases, the exact nature of the relationship between the two errors is unknown.

Hypothesis Testing Framework

Generate hypotheses to support theory

- How to avoid Type I and Type II errors?
  - A priori power analysis vs. post hoc power analysis

Hypothesis Testing Framework

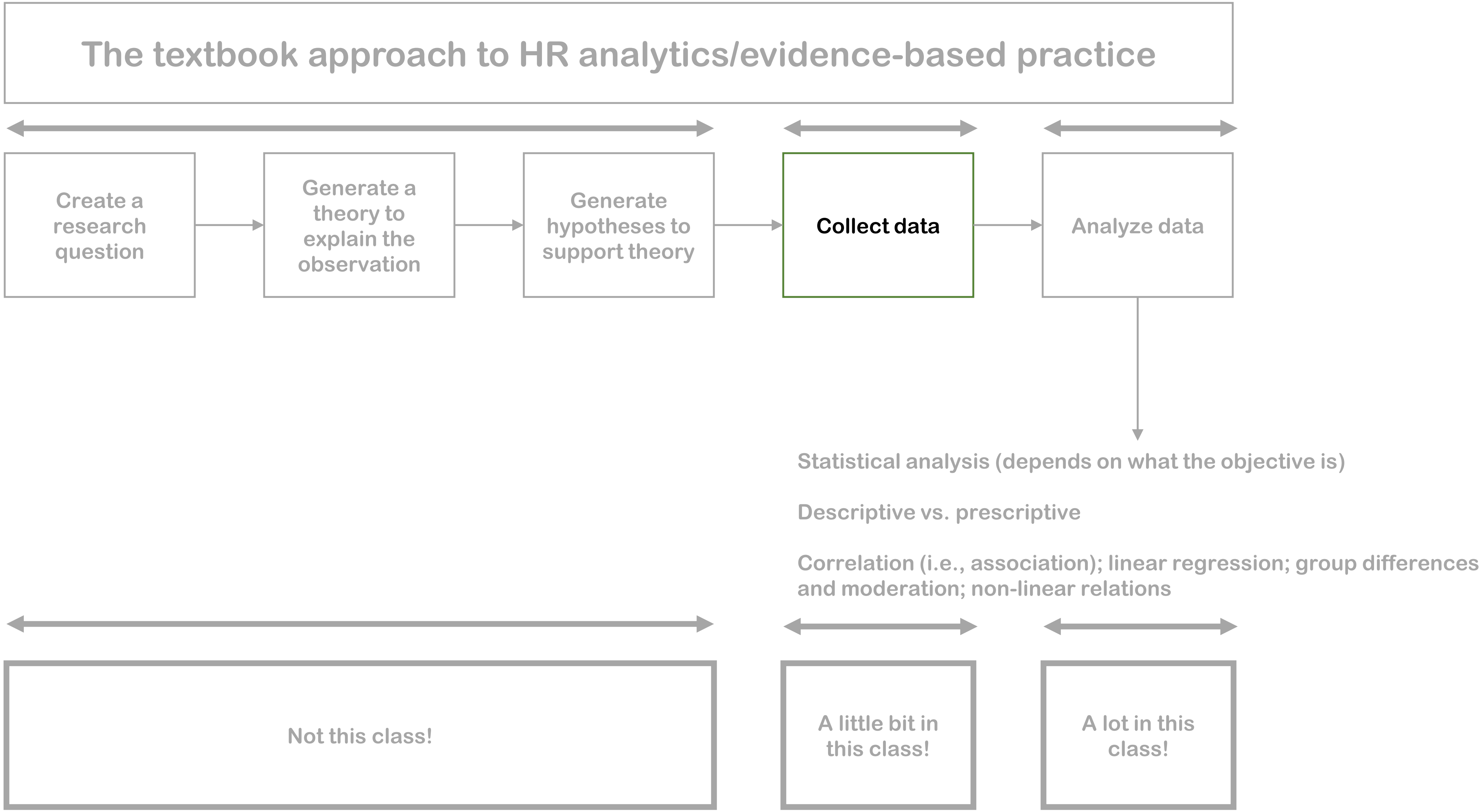
Generate hypotheses to support theory

- How to avoid Type I and Type II errors?
  - A priori power analysis vs. post hoc power analysis
  - Make sure that your study has sufficient statistical power
    - Effectively, we need to make sure that our study/projects include a sufficient number of individuals
    - What would you trust more, a result drawn from 100 vs. 1000 individuals?

Hypothesis Testing Framework

Generate hypotheses to support theory

- How to avoid Type I and Type II errors?
  - A priori power analysis vs. post hoc power analysis
  - Make sure that your study has sufficient statistical power
    - Effectively, we need to make sure that our study/projects include a sufficient number of individuals
    - What would you trust more, a result drawn from 100 vs. 1000 individuals?
  - Statistical power ranges from 0 to 1.0
  - We should aim to achieve a power of .8, or an 80% of detecting an effect if one genuinely exists.


The textbook approach to HR analytics/evidence-based practice
Create a research question
Generate a theory to explain the observation
Generate hypotheses to support theory
Collect data
Analyze data
Statistical analysis (depends on what the objective is)
Descriptive vs. prescriptive
Correlation (i.e., association); linear regression; group differences and moderation; non-linear relations
Not this class!
A little bit in this class!
A lot in this class!

Research Designs

Collect data

- Experimental
- Observational
- Quasi-experimental

Research Designs

Collect data

- Experimental
- Observational
- Quasi-experimental

- Cross-sectional design vs. longitudinal design

Data sources

Collect data

- Convenience samples
- Non-convenience samples
- Archival sources and organizational records

Data sources

Collect data

- **Convenience samples**
- Non-convenience samples
- Archival sources and organizational records

Convenience sampling (also known as availability sampling) is a specific type of non-probability sampling method that relies on data collection from population members who are conveniently available to participate in study.

PROS:

1. Ease of data collection
2. Cheap
3. Timely

CONS

1. Vulnerable to selection bias
2. High level of sampling error
3. Low credibility (sometimes)

Data sources

Collect data

- **Convenience samples**
- Non-convenience samples
- Archival sources and organizational records

Convenience sampling (also known as availability sampling) is a specific type of non-probability sampling method that relies on data collection from population members who are conveniently available to participate in study.

PROS:

1. Ease of data collection
2. Cheap
3. Timely

CONS

1. Vulnerable to selection bias
2. High level of **sampling error**
3. Low credibility (sometimes)

Data sources

Collect data

- **Convenience samples**
- Non-convenience samples
- Archival sources and organizational records

Convenience sampling (also known as availability sampling) is a specific type of non-probability sampling method that relies on data collection from population members who are conveniently available to participate in study.

PROS:

1. Ease of data collection
2. Cheap
3. Timely

CONS

1. Vulnerable to selection bias
2. High level of **sampling error**
3. Subjective
4. Low credibility (sometimes)

Sampling error vs. sampling variation vs. sampling distribution

Data sources

Collect data

- Convenience samples
- **Non-convenience samples**
- Archival sources and organizational records

Non-convenience sampling is a specific type of non-probability sampling method that relies on data collection from population members who are not conveniently available to participate in study.

PROS:

1. Population specific
2. Informed/Experienced
3. Credible
4. Motivated

CONS

1. Sample size concerns
2. Access
3. Turnover
4. Subjective

Data sources

Collect data

- Convenience samples
- Non-convenience samples
- Archival sources and organizational records

At times, researchers may be able to collect data from organizational records that can assist them in their HR analytic efforts. For example, it may be prudent for performance data to be curated from org records instead of asking the respondent to self-report their performance.

PROS:

1. Accurate/Objective

CONS

1. Availability
2. Out dated

Data types

Collect data

- Observed vs. latent variables

Data types

Collect data

- Observed vs. latent variables

"The many, as we say, are seen but not known, and the ideas are known but not seen" (Plato, The Republic)

Data types

Collect data

- Observed vs. latent variables

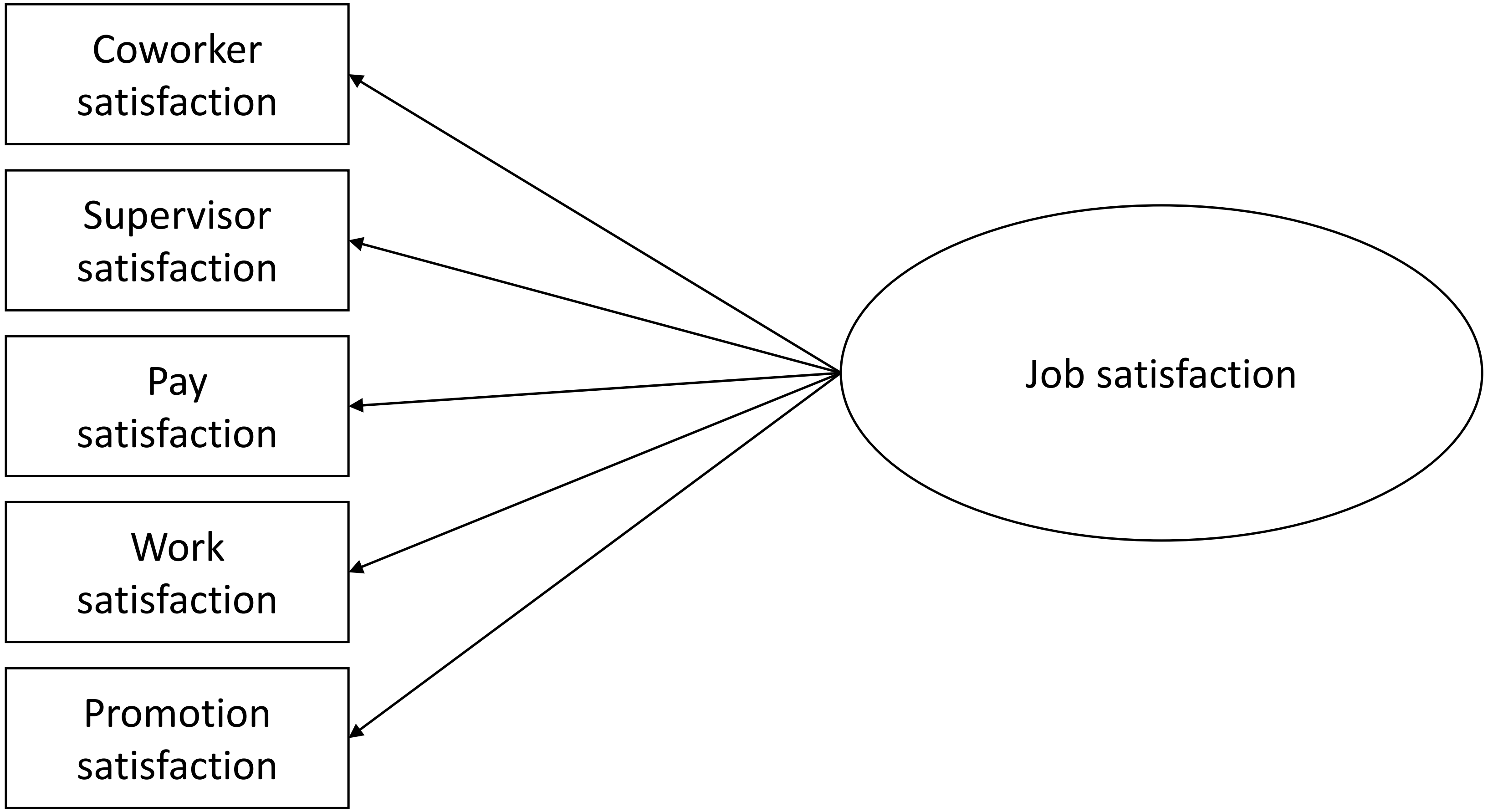
Observed variables (sometimes called *observable* variables or *measured variables*) are actually measured by the researcher.

Latent variables are variables that are not directly observed but are rather inferred (through a mathematical model) from other variables that are observed (directly measured).

Data types

Collect data

- Observed vs. latent variables



Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

Nominal variables are used to scale labels used to identify *groups* of people who share a common attribute that is not shared by people in other groups.

Characteristics: Cannot be ordered; cannot be measured

Examples: Sex; race; etc.

Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

Can be analyzed using the grouping method. The variables can be grouped together into categories, and for each category, the frequency or percentage can be calculated.

Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

Can be analyzed using the grouping method. The variables can be grouped together into categories, and for each category, the frequency or percentage can be calculated.

Nominal data can be used in statistical analysis after they are "dummy coded."

A	B
Sex	Sex_dummy
Male	1
Female	2
Male	1
Male	1
Male	1
Female	2
Female	2
Female	2

Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

Ordinal scales produce ranks in which people can be ordered according to the amounts of some attribute that the possess.

Ordinal variables represent labels that indicate the relative position of people with regard to the relative levels of the attribute being measured.

However, there is no attempt to ascertain how of that attribute is actually possessed by each person.

Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

Imagine that three teams (a pro team, a college team, and a high-school team) were asked to rank their three best soccer players according to athleticism.

Rank	Pro team	College team	HS team
1	"Best" player	"Best" player	"Best" player
2	Second "best" player	Second "best" player	Second "best" player
3	Third "best" player	Third "best" player	Third "best" player

The most athletic player from each team is ranked as "1"

But, are all three players equally athletic?

Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

An interval variable is based on numbers that represent quantitative differences between people in terms of the attribute being measured.

Importantly, the size of the unit of measurement is constant and additive, but the scale does not allow multiplicative interpretations.

This is because interval scales have an ***arbitrary*** zero value.

Example: Temperature; IQ; SAT scores

Data types

Collect data

- Nominal variable (categorical)
- Ordinal variable (categorical)
- Interval variable (continuous)
- Ratio variable (continuous)

A ratio scale has exactly the same characteristics as the interval scale except that the zero on the scale means: does not exist.

Ratio scales are considered a “higher” level of measurement than interval, ordinal, and nominal scales, because they provide more information and allow sophisticated inferences.

Specifically, ratio scales allow additivity as well as multiplicative interpretations in terms of ratios (i.e., 40 miles is twice as far as 80 miles).

Data levels

Collect data

- Individual-level data
- Group-level data
- Organizational-level data

The textbook approach to HR analytics/evidence-based practice

Create a research question

Generate a theory to explain the observation

Generate hypotheses to support theory

Collect data

Analyze data

Statistical analysis (depends on what the objective is)

Descriptive vs. prescriptive

Correlation (i.e., association); linear regression; group differences and moderation; non-linear relations

Not this class!

A little bit in this class!

A lot in this class!

Types of Validity

Analyze data

- Criterion validity
- Content validity
- Convergent validity
- Discriminant validity

NOT THE SAME AS RELIABILITY!!

Types of Validity

Analyze data

- Criterion validity
- Content validity
- Convergent validity
- Discriminant validity

NOT THE SAME AS RELIABILITY!!

Reliability does not imply validity

That is, a reliable measure that is measuring something consistently is not necessarily measuring what you want to be measured. For example, while there are many reliable tests of specific abilities, not all of them would be valid for predicting, say, job performance.

Types of Validity

Analyze data

- Criterion validity
- Content validity
- Convergent validity
- Discriminant validity

Criterion validity is the extent to which a measure is related to an outcome.

Two types:

(1) Concurrent validity
(2) Predictive validity

Types of Validity

Analyze data

- Criterion validity
- Content validity
- Convergent validity
- Discriminant validity

Content validity is a non-statistical type of validity that involves "the systematic examination of the test content to determine whether it covers a representative sample of the behavior domain to be measured" (Anastasi & Urbina, 1997 p. 114).

Specifically, content validity evidence involves the degree to which the content of the test matches a content domain associated with the construct.

Types of Validity

Analyze data

- Criterion validity
- Content validity
- Convergent validity
- Discriminant validity

Both convergent validity and discriminant validity are subtypes of *construct validity*.

The definition for construct validity is *very* similar to the definition for content validity.

Construct validity examines the question: Does the measure behave like the theory says a measure of that construct should behave?

Importantly, one needs to demonstrate convergent validity *and* discriminant validity in order to establish construct validity.

Types of Validity

Analyze data

- Criterion validity
- Content validity
- Convergent validity
- Discriminant validity

Convergent validity refers to the degree to which two measures of constructs that theoretically should be related, are in fact related.

Types of Validity

Analyze data

- Criterion validity
- Content validity
- Convergent validity
- Discriminant validity

Discriminant validity or divergent validity tests whether concepts or measurements that are not supposed to be related are actually unrelated.

Other issues

Analyze data

- Measurement error
- Sampling error
- Common method bias

Other issues

Analyze data

- Measurement error
- Sampling error
- Common method bias

Measurement error (also called observational error) is the difference between a measured quantity and its true value.

It includes random error (naturally occurring errors that are to be expected with any observation/experiment) and systematic error (caused by a mis-calibrated instrument that affects all measurements).

Other issues

Analyze data

- Measurement error
- Sampling error
- Common method bias

Since the sample does not include all members of the population, statistics on the sample, such as means and quartiles, generally differ from the characteristics of the entire population, which are known as parameters. These differences are known as sampling error.

Other issues

Analyze data

- Measurement error
- Sampling error
- Common method bias

Common-method variance (CMV) is the spurious "variance that is attributable to the measurement method rather than to the constructs the measures are assumed to represent“ (Podsakoff et al., 2003).

In other words, CMV is a "systematic error variance shared among variables measured with and introduced as a function of the same method and/or source".

There are several ex ante remedies to help avoid CMV/CMB.