An Introduction to, and Overview of Some Basic Research Design and Methods Concepts and Principles

October 5th, 2019

Winfred Arthur, Jr.
Department of Psychological and Brain Sciences
and Department of Management

Research Methods

- means of discovering truth

Research Methods

- means of discovering truth
- what **is** truth?

Research Methods

- means of discovering truth
- what **is** truth?
  - Riveda Sandhyavandanam ⇒ "There is only one truth [but] people often see it in different ways"
  - see also Ludwig Wittenstein (1889-1952) ⇒ *The metaphysics of space and motion and the wave structure of matter (WSM) gives absolute truth and meaning to language*

Research Methods

- means of discovering truth
- what **is** truth?
  - Riveda Sandhyavandanam ⇒ "There is only one truth [but] people often see it in different ways"
  - see also Ludwig Wittenstein (1889-1952) ⇒ *The metaphysics of space and motion and the wave structure of matter (WSM) gives absolute truth and meaning to language*

- **realism**
- rationality
- regularity
- causality/determinism
- discoverability

Research Methods

- means of discovering truth
- what **is** truth?
  - VALIDITY?
  - a conclusion based on a piece of research is valid when it corresponds to the actual or true state of the world

Research Methods

	Quantitative	Qualitative
Conceptual	Concerned w/ facts about phenomena Assumes a fixed and measureable reality	Concerned w/ understanding behavioral/social phenomena from informant's perspective Assumes a dynamic and negotiated reality Peculiar to behavioral & social "sciences"
Methodological	Data are collected through measuring things Data are analyzed through numerical comparisons and statistical inferences Data are reported through statistical analyses	Data are collected through participant observation and interviews Data are analyzed by themes from descriptions by informants Data are reported in the language of the informant

Adapted from Minchiello et al. (1990, p.5)

Validity

- a key (maybe THE key) criterion in the evaluation of any piece of research or test (measure)
- the appropriateness of **inferences** drawn from **data**
  - data = results of research study ➡ research validity
  - data = test scores ➡ test and measurement validity

Validity

- a key (maybe THE key) criterion in the evaluation of any piece of research or test (measure)
- the appropriateness of **inferences** drawn from **data**
  - data = results of research study ➡ research validity
  - data = test scores ➡ test and measurement validity

Validity

research validity

- validity → appropriateness of inferences drawn from the results of a research study
- a conclusion based on research is valid when it corresponds to the true state of the world
- dimensions of research validity
  - internal
  - external
  - statistical-conclusion
  - construct

test & measurement validity

- validity → appropriateness of inferences drawn from test scores
- validation → process by which we determine/establish the validity of test scores
- different validation approaches → simply different means (i.e., sources of validity evidence) by which we establish construct validity
  - criterion-related
  - content-related
  - construct-related
  - meta-analysis & validity generalization
  - synthetic validity/j-coefficients
  - face validity

Validity

research validity

- validity → appropriateness of inferences drawn from the results of a research study
- a conclusion based on research is valid when it corresponds to the true state of the world
- dimensions of research validity
  - internal
  - external
  - statistical-conclusion
  - construct

test & measurement validity

- validity → appropriateness of inferences drawn from test scores
- validation → process by which we determine/establish the validity of test scores
- different validation approaches → simply different means (i.e., sources of validity evidence) by which we establish construct validity
  - criterion-related
  - content-related
  - construct-related
  - meta-analysis & validity generalization
  - synthetic validity/j-coefficients
  - face validity

Validity

- a key (maybe THE key) criterion in the evaluation of any piece of research or test (measure)
- the appropriateness of **inferences** drawn from **data**
  - data = results of research study ➡ research validity
  - data = test scores ➡ test and measurement validity

Research Validity

- two philosophies, world views, or approaches to truth discovery
  - Donald Campbell and Donald Rubin
    - Shadish (2010; Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods, 15,* 3-17)
- Campbell's causal model $\Rightarrow$ methodologically and design driven
- Rubin's causal model $\Rightarrow$ quantitatively and statistically driven

Research Validity

- two philosophies, world views, or approaches to truth discovery
  - Donald Campbell and Donald Rubin
    - Shadish (2010; Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods, 15,* 3-17)
- Campbell's causal model ⇒ methodologically and design driven
- Rubin's causal model ⇒ quantitatively and statistically driven

Research Validity ⇒ re Campbell's causal model [CCM]

- facets of research validity
  - internal
  - external
  - statistical-conclusion
  - construct
- threats to research validity
- controls

- internal validity ⇒ the validity of inferences about whether observed covariation b/n *A* (presumed treatment) and *B* (presumed outcome) reflects a causal relationship from *A* to *B*, as those variables were manipulated or measured

Threats

- history
- maturation
- testing
- attrition or mortality
- selection
- regression

Controls?

- history effects (events outside the lab)
  - observed effect between IV and DV might be due to events occurring between the pretest and posttest when these events are not the treatment of research interest

- maturation effects
  - source of error in a study related to the amount of time between measurements

- testing effects
  - effects due to the number of times particular responses are measured
  - resulting from familiarity with the measurement instrument

- attrition or mortality effects
  - the dropping out of some participants before a study is completed, causing a threat to validity

- selection effects
  - result from biases associated with the selection of, and assignment of research participants into groups

- regression effects
  - tendency of participants with extreme scores on first measure to score closer to the mean on a second testing
  - a statistical threat

- internal validity
  - extent to which we can infer that a relationship between two variables is causal or that absence of a relationship implies absence of cause
  - is the observed effect real or artifactual?
  - threats

- correction $\Rightarrow$ random assignment/ randomization

- external validity ⇒ the validity of inferences about whether the cause-effect relationship holds over variations in persons, settings, time, treatment variables, and measurement variables

Threats

- population validity
- ecological validity
- temporal validity

Controls?

- other participants
  - interaction of selection and treatment
  - population validity
- other settings
  - interaction of setting and treatment
  - ecological validity
- other times
  - interaction of history and treatment
  - temporal validity

- external validity
  - generalizability
- enhanced or increased $\Rightarrow$ random sampling for representativeness
- trade-offs between internal and external validity?

- statistical conclusion validity ⇒ the validity of inferences about the correlation (covariation) b/n treatment and outcome

Threats

- low statistical power
- violation of test assumptions
- poor test reliability

Controls?

- low statistical power
  - power analysis
    - sample size (*n*)
    - effect size (magnitude of effect)
    - power (.80)
    - alpha (*p*-value, .05)

- low statistical power
  - power analysis
    - sample size (*n*)
    - effect size (magnitude of effect)
    - power (.80)
    - alpha (*p*-value, .05)

Parameter	Study 1	Study 2
power	.80	**?? higher or lower**
effect size (*d*)	0.30	0.30
alpha (*p*)	.05	.05
N	500	125

- low statistical power
  - power analysis
    - sample size (*n*)
    - effect size (magnitude of effect)
    - power (.80)
    - alpha (*p*-value, .05)

Parameter	Study 1	Study 2
power	0.80	0.60
effect size (*r*)	.25	.25
alpha (*p*)	.05	.05
N	600	**?? higher or lower**

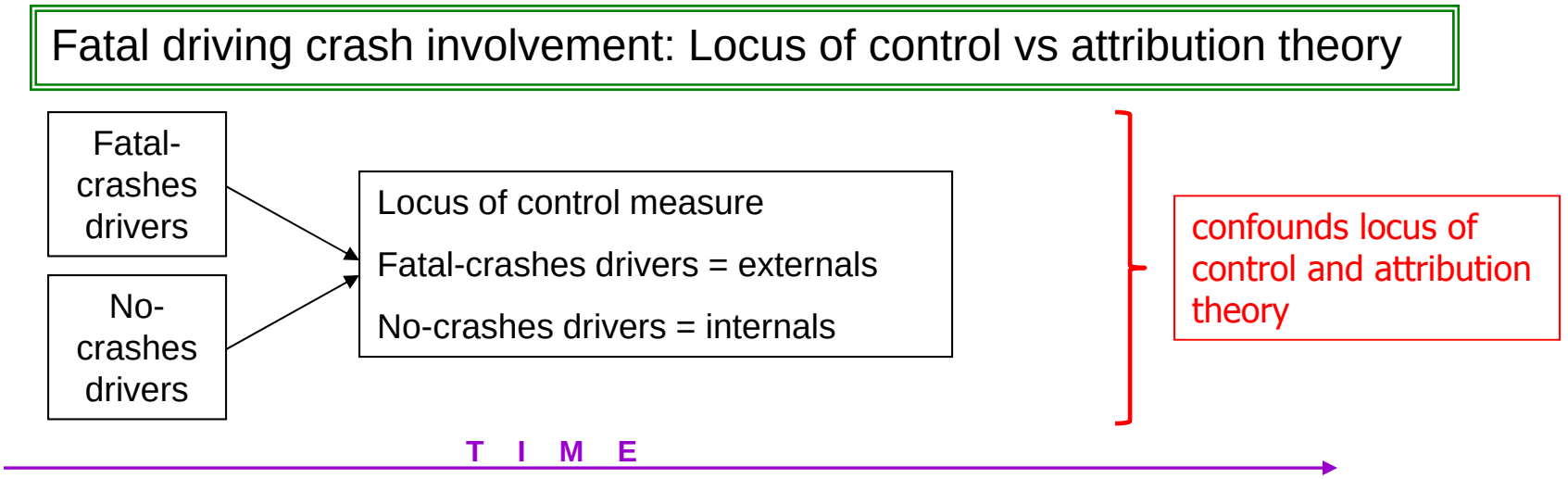
- low statistical power
  - power analysis
    - sample size (*n*)
    - effect size (magnitude of effect)
    - power (.80)
    - alpha (*p*-value, .05)
- violations of statistical test assumptions
- poor reliability of measures

- address threats $\Rightarrow$ adequate power, meet test assumptions, and use reliable measures

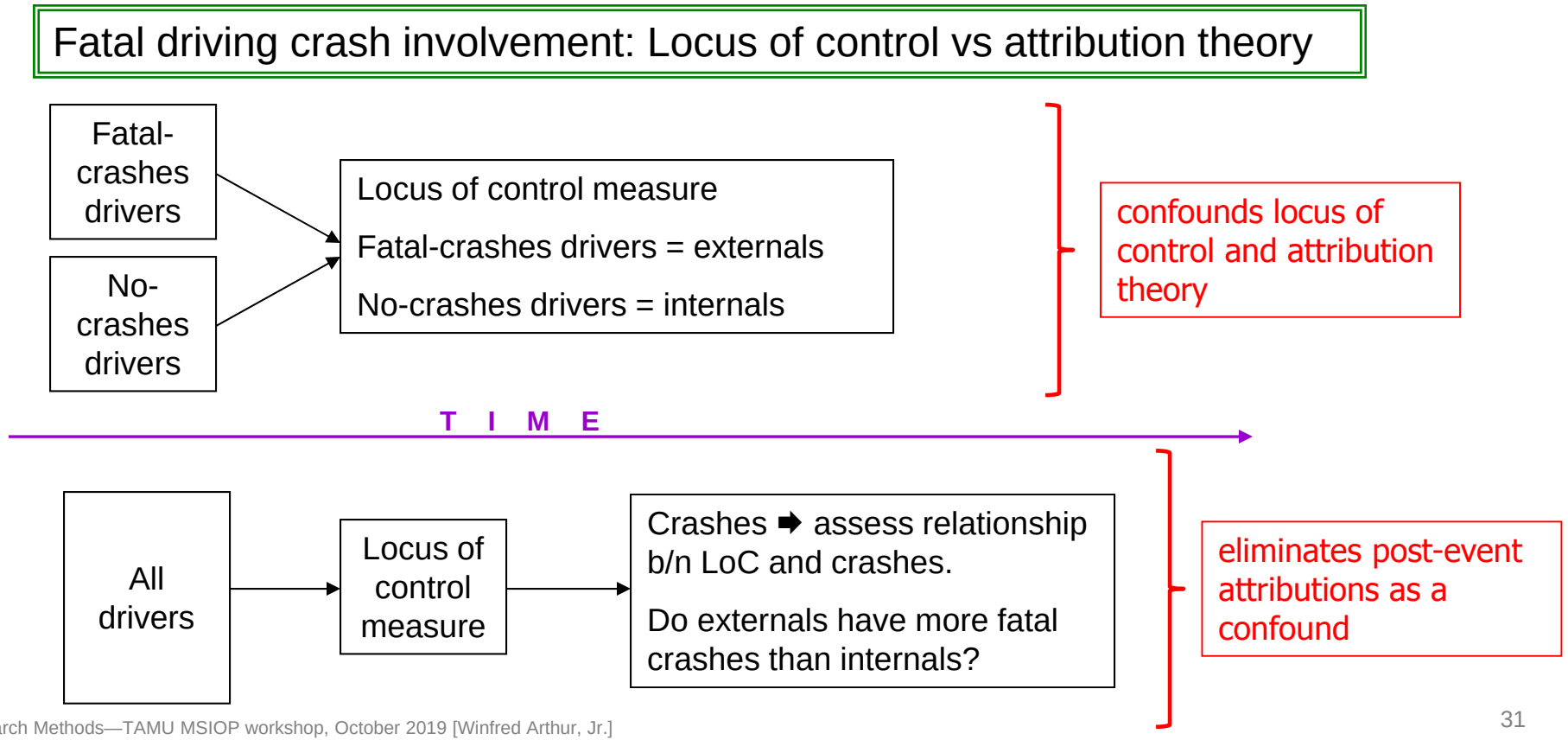
- construct validity ⇒ the validity with which inferences can be made from the operations in a study to the theoretical constructs those operations are intended to represent

Threats	Controls?
• loose connection between theory and study • evaluation apprehension • experimenter expectancies ("good-subject" response)	

- construct validity
  - use of appropriate theories, theoretical labels, or models to describe and explain phenomenon being studied



- construct validity
  - use of appropriate theories, theoretical labels, or models to describe and explain phenomenon being studied



- threats
  - loose connection between theory and study
  - changes in participants' behavior as a result of being studied
    - Hawthorne effect
    - "good subject" response
    - social desirability responding
    - evaluation apprehension
    - responses to experimenter expectancies, and experimenter effects

- control or minimize threats
  - double-blind procedures
  - single-blind procedures
  - deception

- interrelatedness of different facets or dimensions of research validity
  - statistical conclusion
  - internal
  - external
  - construct

- interrelatedness of different facets or dimensions of research validity

Which of the following dimensions of research validity would have to be demonstrated or established before one could meaningfully engage in a discussion of any of the other dimensions of research validity?

A. Construct validity
B. Internal validity
C. Statistical-conclusion validity
D. External

- good research design or study
  - free from threats
  - no alternative explanations
  - permits robust conclusions about relationships between study's variables
- if research is conceptualized as the "discovery of truth", it then becomes blatantly obvious that **no single [primary] study will be sufficient to permit a claim of truth discovery**
  - at best, it can only be claimed to be another datum in the cumulative body of knowledge that slowly moves us closer to the discovery of truth

Philosophies of Causal Inference ⇒ Conditions of Causality

- contiguity
- temporal precedence
- constant conjunction

Philosophies of Causal Inference ⇒ Conditions of Causality

- contiguity
- temporal precedence
- constant conjunction

Constant conjunction stipulates that (A) if X is observed, then Y will occur; and conversely, (B) if Y is observed, then X must have been present in the past.

A number of noteworthy points pertaining to the above:

1. An effect can obviously have multiple causes.
2. This makes Stipulation B tenuous.
3. Which is ok and indeed reflects the elegance of the inference system since after all it entails a post-dictive paradigm and we know that post-dictive inferences are very susceptible to alternative explanations, as is the case here.
4. Thus, in summary, consistent with the predicates of predicitve and post-dictive paradigms, one would expect Stipulation A to permit stronger causal inferences than Stipulation B.

Philosophies of Causal Inference ⇒ Conditions of Causality

- contiguity
- temporal precedence
- constant conjunction

- summary
  - cause (X) must precede consequence (Y) in time; that is, X is manipulated and then Y is measured **[temporal precedence and constant conjunction]**
  - statistical covariation b/n X and Y must be present **[contiguity]**
  - alternative causes of Y must be controlled via random assignment to groups **[experimental dsgn]**

Philosophies of Causal Inference ⇒ Conditions of Causality

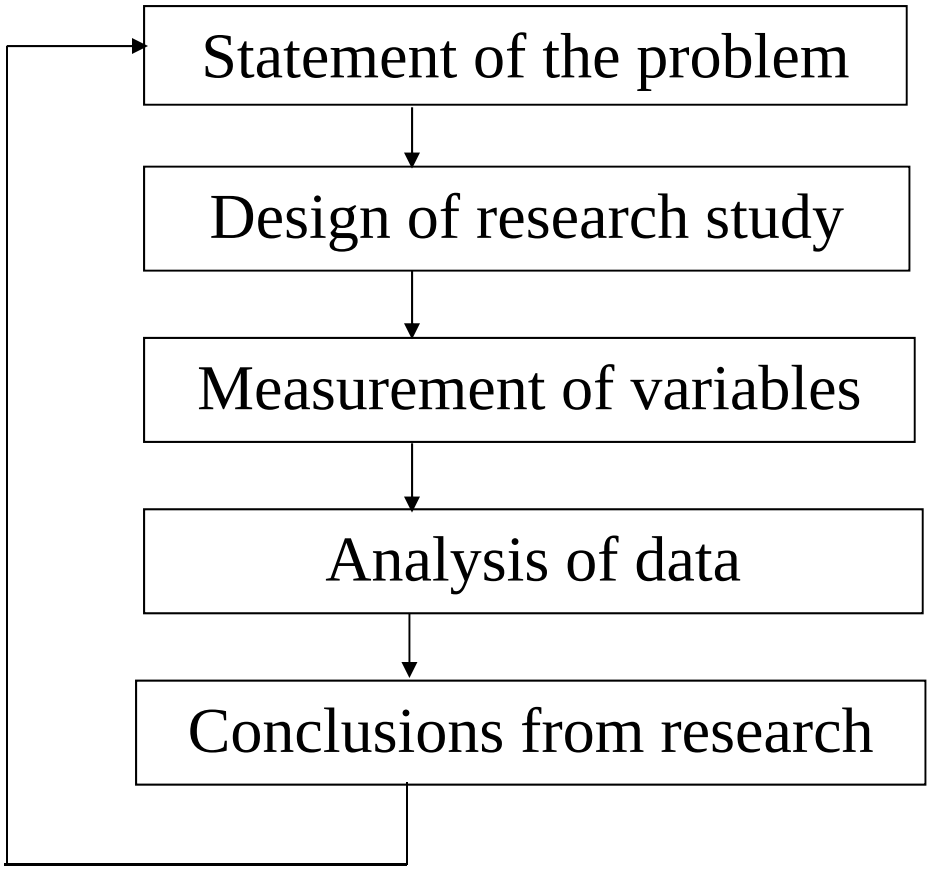
- contiguity
- temporal precedence
- constant conjunction

A rooster crows each dawn and observes that shortly after (30 minutes or so), the sun rises. After nine (9) months of this, the rooster concludes that he is responsible for making the sun rise. The farmer disagrees with this assertion and in an attempt to disprove this claim, asks the rooster to crow at 10:00 pm one evening.

By doing this, which condition of causality is the farmer trying to assess?

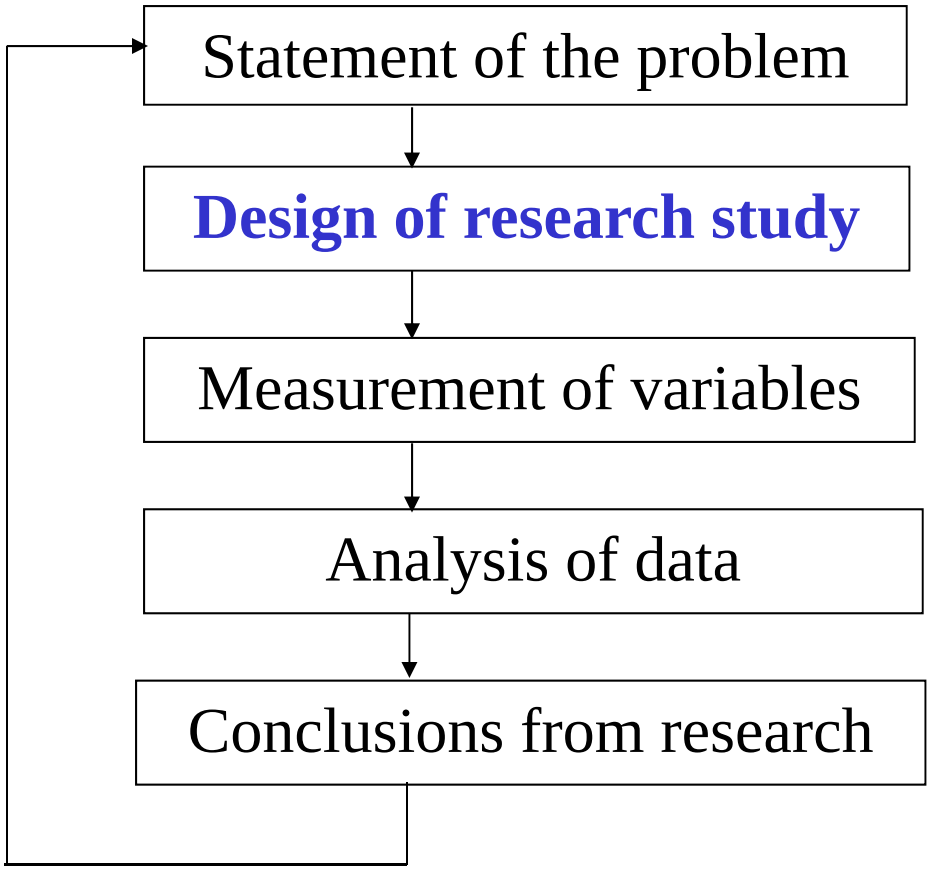
The Empirical Research Cycle

Research process ⇒ summarized as 5-step sequence



The Empirical Research Cycle

Research process ⇒ summarized as 5-step sequence

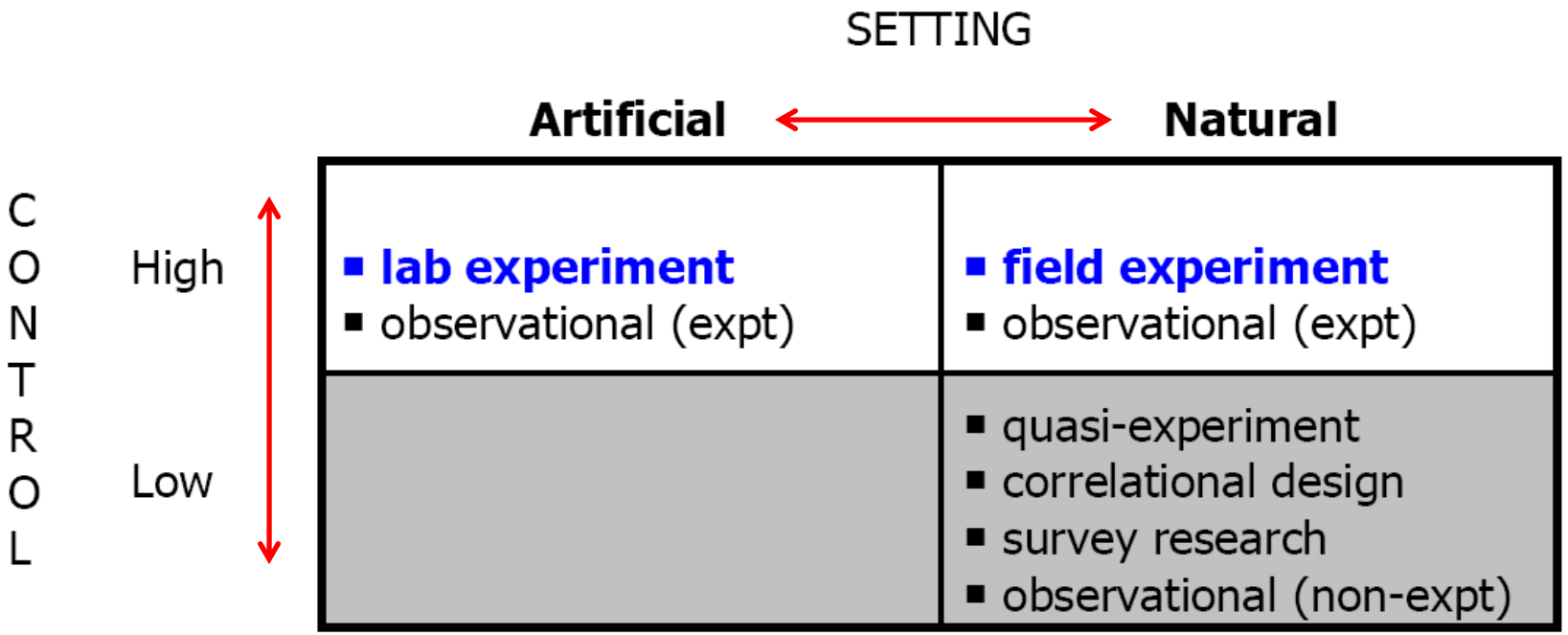


Research Design

- naturalness of the research setting
  - lab or field
  - "naturalness" or "artificiality" of the setting
  - field research typically employs a real-life setting, while lab research is more contrived and artificial

Research Design

- degree of control
  - confounding and extraneous variables
  - manipulation $\Rightarrow$ this is reflective of a high degree of control
  - research designs that permit random assignment (and manipulation) are technically referred to as "experiments"


SETTING
Artificial
Natural
CONTROL
High
Low
lab experiment
observational (expt)
field experiment
observational (expt)
quasi-experiment
correlational design
survey research
observational (non-expt)

Experimental Designs

- **experiment**
  - a research method in which the investigator manipulates a variable under carefully controlled conditions and observes whether changes occur in a second variable
  - experiments are used in order to detect *cause-and-effect* relationships
- **conditions that make a <u>true</u> experiment**
  - **<u>random assignment</u>** into experimental conditions (experimental & control conditions)
  - **<u>manipulation</u>** of independent variables

resultant effect $\Rightarrow$ high levels of control over the **<u>who</u>**, what, when, where, and how, etc.

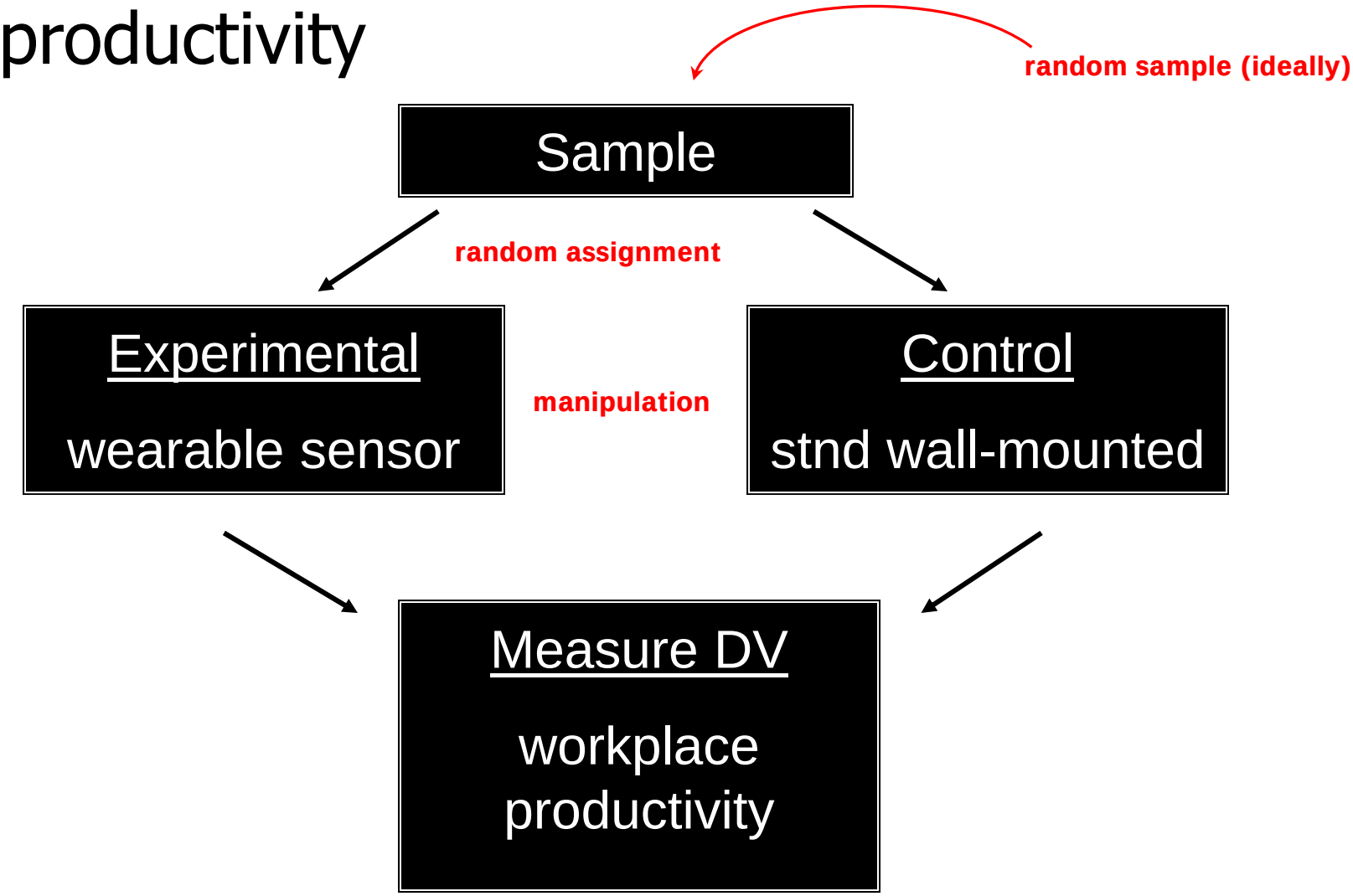
Experimental Designs

- random assignment
- manipulation
- high levels of control over the **who**, what, when, where, and how, etc.

Experimental Designs

- experimental and control groups
  - logic of the experimental approach
    - if the two groups are identical in all respects **[both known and unknown factors]** except for the variation created by the manipulation of independent variable, then any differences between the experimental and control groups on the dependent variable **must be** due to the manipulation of the independent variable

Effect of sensory-integrated building micro-adjustments on workplace productivity



Experimental Designs

- random assignment and manipulation
  - pretest-posttest with a control group
  - Solomon four-group
  - posttest only with a control group
    - but prudent and preferable to have a pretest

o within- and between-subjects designs

Quasi-Experimental designs

- self-selection groups
- pre-existing groups

  - nonequivalent control group designs
    - delayed control group
    - mixed factorials
  - designs without control groups
    - interrupted time series
    - repeated treatment

It doesn't take more than a few episodes of the Discovery Channel's Deadliest Catch to get the idea that commercial fishing can be a career path rife with risk, making it one of the most dangerous occupations in the U.S.

Sometimes the danger stems from how fish are harvested. Rules for catching fish can vary by region and species. About a third of U.S. fisheries operate under what's known as **derby-style fishing** - a season opens for a few weeks or months, and fishermen race to land their catch before it's closed again.

But derby-style fishing means commercial fishermen are sometimes forced onto the water in stormy weather or before their boats can be properly maintained. Miss the window to catch fish, and for some, it can mean the difference between keeping their businesses afloat or not.

But over the last decade or so, a different kind of fisheries management program known as **catch shares** has been gaining ground. The idea here is to allot fishermen a portion of the catch ahead of time, and allow them to fish until they reach it. Proponents claim catch shares create incentives for fishers to slow down, eliminating the need to race each other to fish.

A new study published online Wednesday in the journal *Nature* proves they're right. "This is the first time we see broad systematic evidence that catch shares are slowing the race to fish," says study co-author Martin Smith, professor of environmental economics at the Nicholas School of the Environment at Duke University.

Researchers looked at monthly data from 39 federally managed commercial fisheries (worth a combined $402 million) such as Pacific halibut, Atlantic cod, New England haddock, Gulf of Mexico red snapper and more, which operate under catch share programs. The researchers then compared that to data from a similar control fishery that did not operate under a catch share program. They found that under a catch share program, harvesters - on average - took an extra month to fish.

Although many studies have associated spanking with increases in aggression and externalizing problems, these studies tend to be nonexperimental in nature. The difficulty with moving beyond nonexperimental research in this field is that spanking is difficult to study in a causal manner; it would be unethical, for example, to randomly assign a child to be spanked or not spanked. The authors of the current study used propensity score matching (PSM) to approximate conditions found in randomized studies. PSM takes two groups (in this case spanked and not-spanked children) and matches them on a series of observed covariates. The outcome of this method is that children in both groups end up differing only in the variable of interest—whether they were spanked or not spanked. The researchers used this technique to study the impact of spanking on children who were part of the Early Childhood Longitudinal Study—Kindergarten Cohort. Parents reported how often they had spanked their 5-year-old child in the past week, and teachers reported children's externalizing behavior problems at ages 5, 6, and 8. The researchers found that spanking at age 5 predicted externalizing problems at age 6 and 8, suggesting that spanking leads to increased externalizing behavior.

Quasi-Experimental designs

	Allocation to Groups	Treatment	Posttest
Group I	Nonequivalent naturally occurring groups	Yes [A_1]	Yes
Group II		No [A_2]	Yes

Quasi-Experimental designs

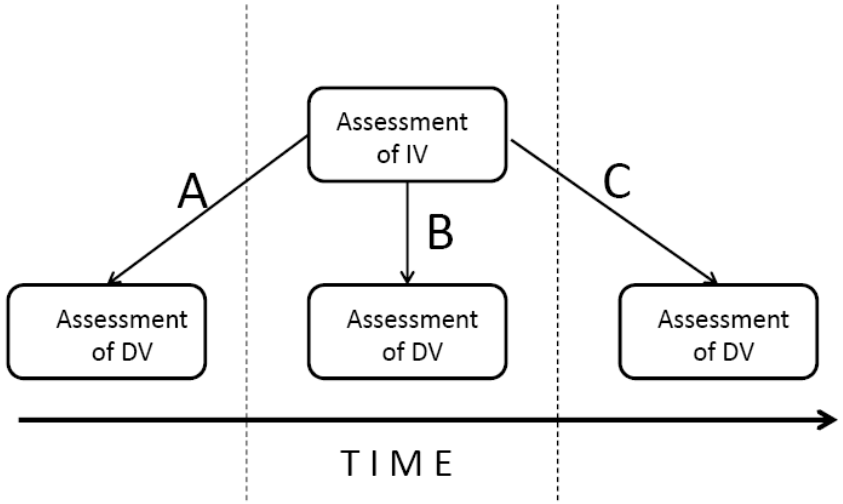
	Allocation to Groups	Treatment	Posttest
Group I	Nonequivalent naturally occurring groups	Yes [A_1]	Yes
Group II		No [A_2]	Yes

Quasi-Experimental designs

- major/primary threat
  - nonequivalency $\Rightarrow$ alternative/competing explanations
- enhancing interpretation
  - matching
  - pretesting
  - statistical control
  - moderator variables

Correlational Designs

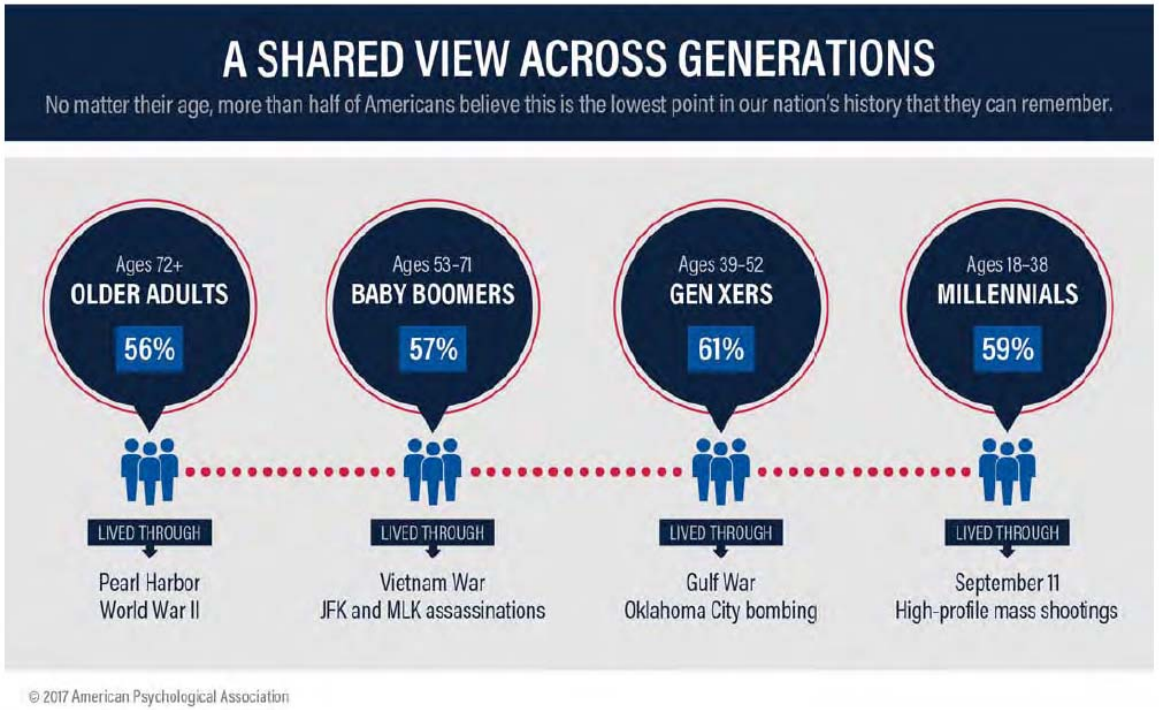
- measurement of two of more variables and assessing relationship/association b/n them
  - predictive
  - concurrent
  - postdictive



Survey Research

- measurement and assessment of opinions, attitudes, and such, usually by means of questionnaires and sampling methods

STRESS IN AMERICA™:
THE STATE OF OUR NATION



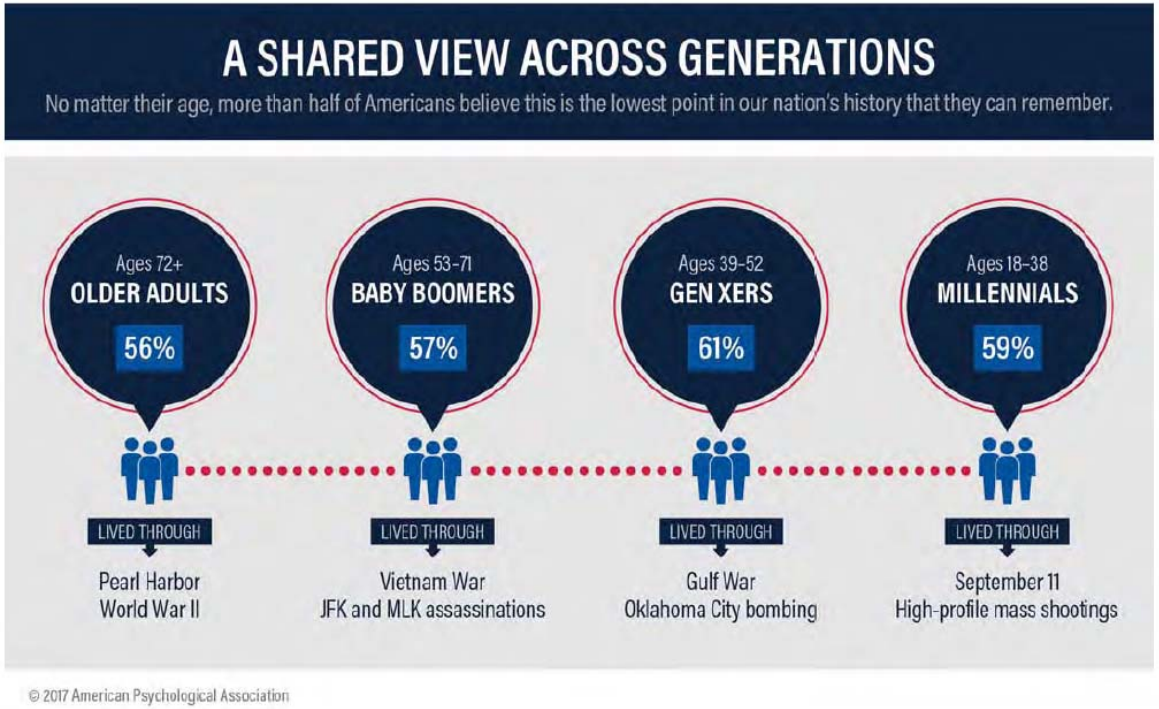
Survey Research

- measurement and assessment of opinions, attitudes, and such, usually by means of questionnaires and sampling methods

STRESS IN AMERICA™:
THE STATE OF OUR NATION

"Do you believe in killing unwanted babies?"

"Should a woman be forced to bear unwanted children?"

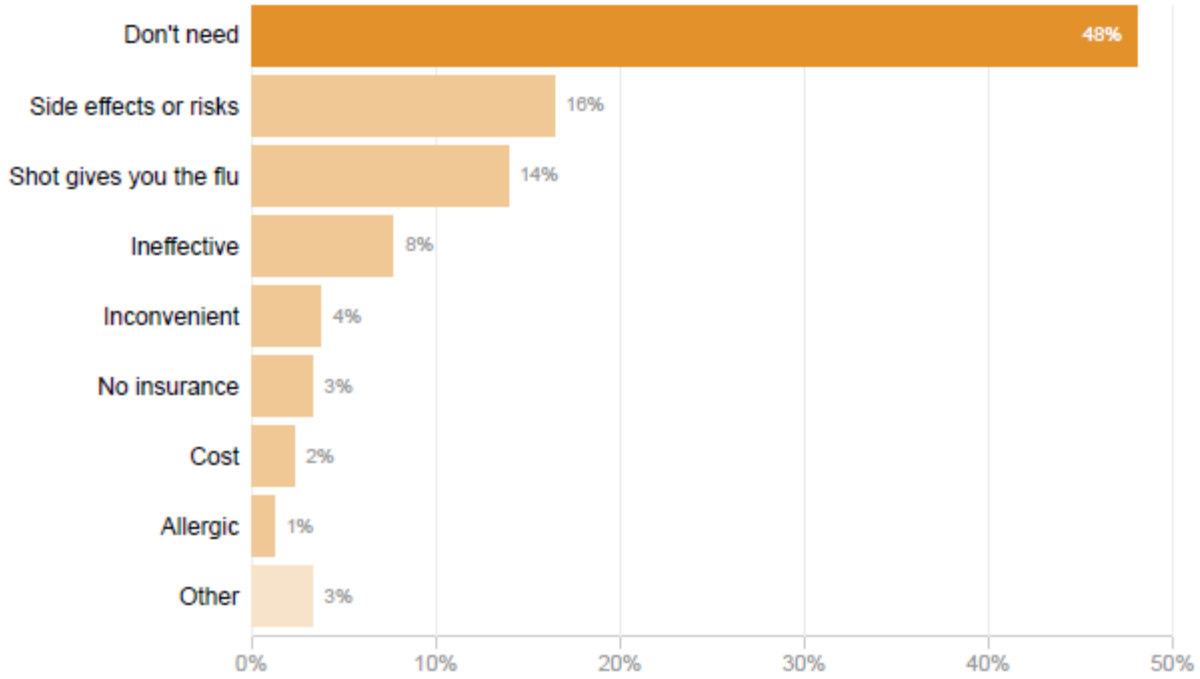


Survey Research

Many Americans Believe They Don't Need The Flu Vaccine

Poll: 'Why are you not planning on getting a flu vaccine this year?'

Asked of the respondents who said they were not planning to get a flu vaccine



Source: The Truven Health Analytics-NPR Health Poll of 3,008 adults interviewed from October 1–15, 2015. The margin of error is +/-1.8 percent.

Credit: Alyson Hurt/NPR

Survey Research

- margin of error ➡ extent/amount of sampling error
- if sample = population, then margin of error = 0
- "what is the sample size for a specified margin of error (*m*)?"
  - $N = 1/m^2$
  - if want 2% margin of error, then sample size required is $1/(.02)^2 = 2{,}500$

Additional Design and Methods Issues

- time $\Rightarrow$ longitudinal and cross-sectional designs
- data collection protocols $\Rightarrow$ observational designs?
- secondary research designs $\Rightarrow$ meta-analysis

Additional Design and Methods Issues

- time ⇒ longitudinal and cross-sectional designs

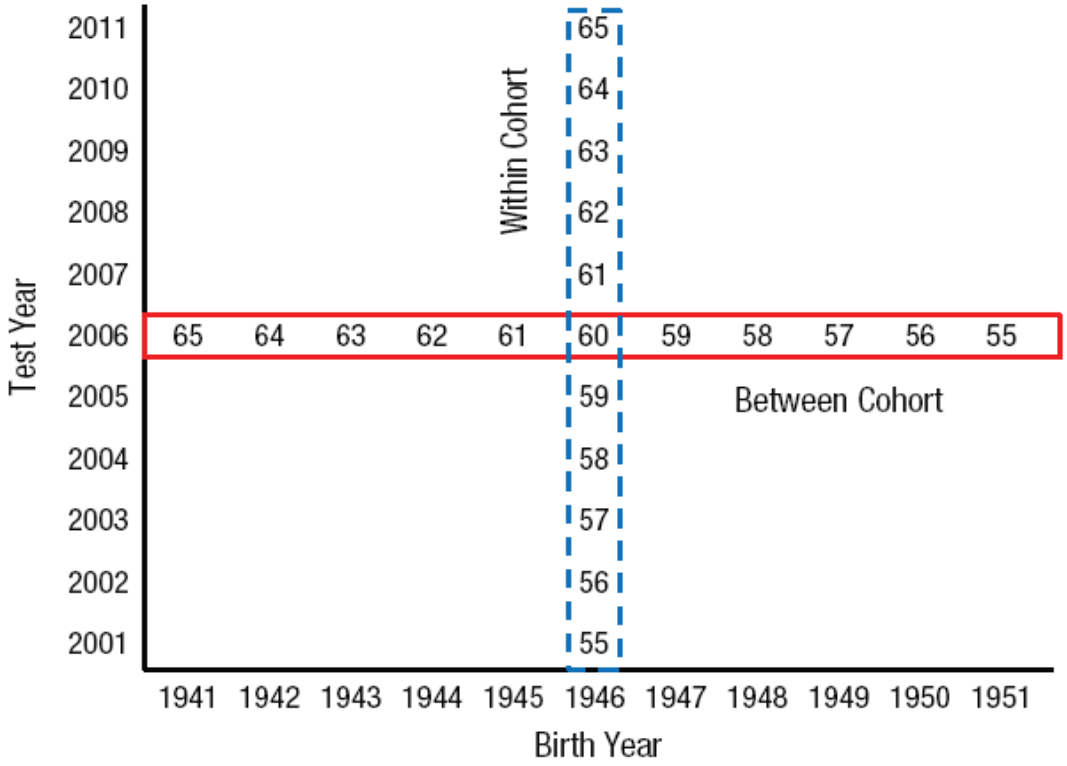


Fig. 1. Illustration of the structure of the data in the current project. Although the birth years ranged from 1907 to 1989, only a limited number of birth years are illustrated for the sake of clarity. Each cell in the matrix consists of data from different people. Thus, comparisons along a row are between cohort, because they involve people who were of different ages (and thus born in different years) but who were tested in the same year; comparisons along a column are within cohort, because they involve people who were of different ages (but born in the same year) and who were tested in different years. The numbers in the cells correspond to the ages of individuals from the indicated birth and test years.

Additional Design and Methods Issues

- time ⇒ longitudinal and cross-sectional designs

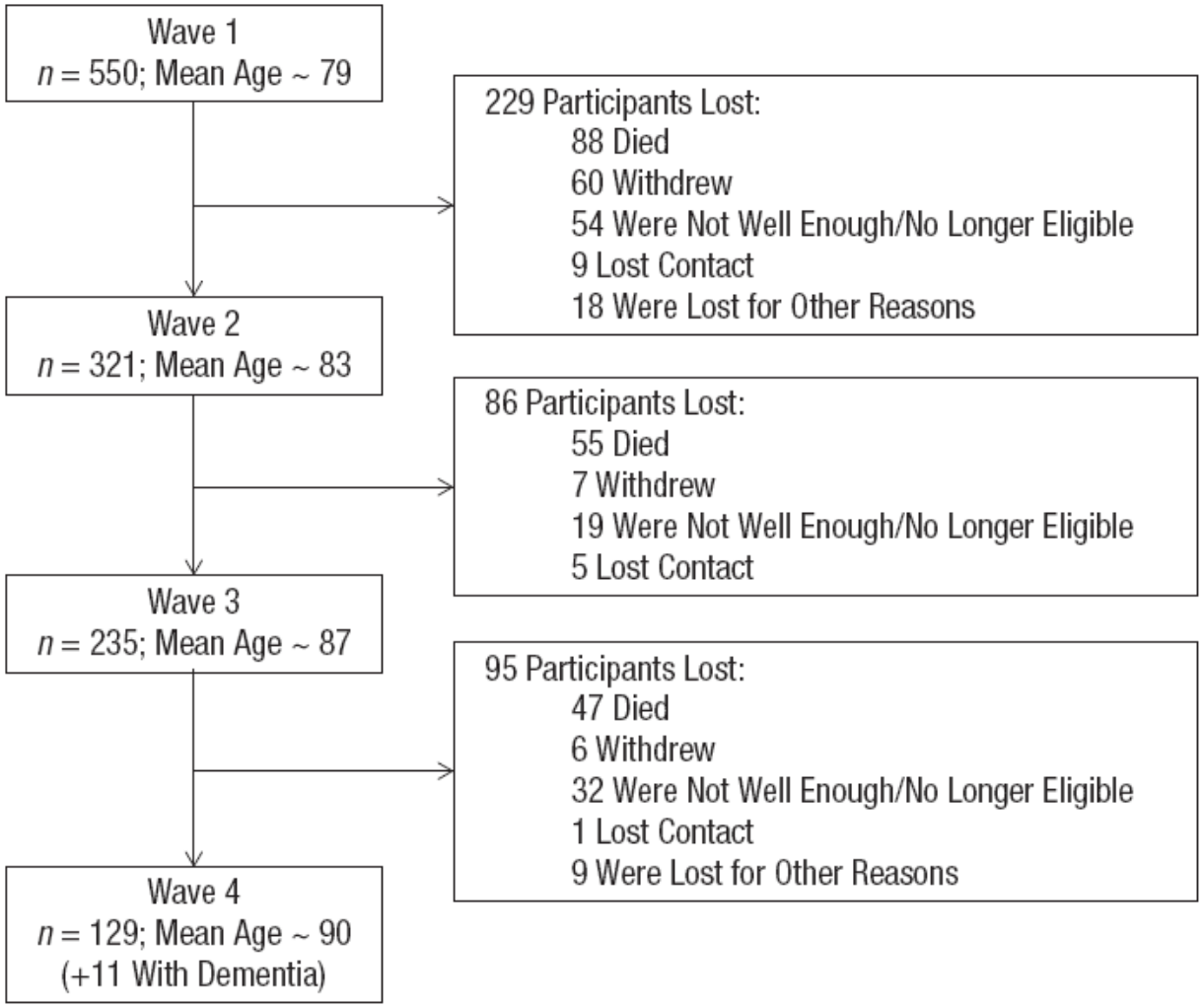


Fig. 1. Number and mean age of participants and causes of attrition at each wave of testing in the Lothian Birth Cohort of 1921. "Withdrew" refers to participants who did not wish to continue to participate. "Not well enough/no longer eligible" refers to participants who withdrew for health reasons, including being too frail to travel, and participants who had severe visual impairments or dementia (except at Wave 4, when 11 people with dementia were examined at home by a physician, but not given tests of cognitive functioning). "Other reasons" for leaving the study include, for example, living too far away to travel to the testing site, caring for a spouse, refusal by the participant's general medical practitioner, and missing an appointment.

Additional Design and Methods Issues

- time $\Rightarrow$ longitudinal and cross-sectional designs
- data collection protocols $\Rightarrow$ observational designs?
- secondary research designs $\Rightarrow$ meta-analysis

Additional Design and Methods Issues

- control techniques and strategies
  - random assignment to groups
  - matching
  - within-subject designs
    - order and carry-over effects
      - counter-balancing
  - manipulation checks
  - instrumentation of response
    - common method variance
  - building nuisance variables into the study [moderators]
  - statistical control
    - control variables
  - experimenter effects and bias reduction
    - double-blind procedures
    - automation

equivalency of conditions

Research Validity

- two philosophies, world views, or approaches to truth discovery
  - Donald Campbell and Donald Rubin
    - Shadish (2010; Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods, 15,* 3-17)
- Campbell's causal model $\Rightarrow$ methodologically and design driven
- Rubin's causal model $\Rightarrow$ quantitatively and statistically driven

Research Validity

- two philosophies, world views, or approaches to truth discovery
  - Donald Campbell and Donald Rubin
    - Shadish (2010; Campbell and Rubin: A primer and comparison of their approaches to causal inference in field settings. *Psychological Methods, 15,* 3-17)
- Campbell's causal model ⇒ methodologically and design driven
- Rubin's causal model ⇒ quantitatively and statistically driven

Research Validity ⇒ re Rubin's causal model [RCM]

- reliance on statistical procedures to discover truth
- can we reclaim/salvage truth from poor designs and data?
- can we recreate or recover truth statistically?
  - SEM (path analysis) to make causal inferences?
    - test for mediation ⇒ "full mediation"; "fully mediated models"?
      - what does that really mean?
  - missing values and data imputation? [RMNET]
    - "we have no information as to why subjects' fail to respond to items; to assume we do is total unwarranted ignorance" (Romie Littrell, 04/17/10)

Research Validity ⇒ re Rubin's causal model [RCM]

- can we recreate or recover truth statistically?
  - meta-analysis?
    - Dieckmann, Malle, and Bodner (2009; An empirical assessment of meta-analytic practice. *Review of General Psychology, 13,* 101-115)
    - psychometric meta-analysis ⇒ controls for statistical artifacts
  - control variables?
  - corrections for faking?
    - "once test users take a wrong course, there is no going back to the choice point" (Cronbach, 1990, p. 521)

a? **the**? take away?

- good research design or study
  - free from threats
  - no alternative explanations
  - permits robust conclusions about relationships between study's variables
- if research is conceptualized as the "discovery of truth", it then becomes blatantly obvious that **no single [primary] study will be sufficient to permit a claim of truth discovery**
  - at best, it can only be claimed to be another datum in the cumulative body of knowledge that slowly moves us closer to the discovery of truth
- SCIENCE IS A MARATHON **NOT** A SPRINT!


QUESTIONS ?

An Introduction to, and Overview of Some Basic Research Design and Methods Concepts and Principles

October 5th, 2019

Winfred Arthur, Jr.
Department of Psychological and Brain Sciences
and Department of Management