Soziologisches Institut – Prof. Dr. Katja Rost

PROBLEM SPECIFICATION, STUDY RETRIEVAL AND CODING

MAER-Net Colloquium 2013

5 – 7 September 2013

Content

1. **Problem Specification**
2. Study Retrieval
3. Coding

Strenghts and Weaknesses

👍 More differentiated and sophisticated than qualitative "vote-counting"

👍 Hidden effects

👍 Large amount of information

👍 Moderator analysis

👎 Effort and expertise

👎 Social context

👎 Apples and oranges

👎 Study population

Problem Statement

o Needs to be straightforward and complete BUT, at this stage, not be highly detailed

o Guides

– the selection of studies

– the coding of information

– the analysis of the resulting data

o Example: (Under which conditions) Does pay-for-performance reduce work effort?

The positive research agenda Gap spotting

- Assumptions underlying existing literature remain unchallenged
    - "extend(ing) this literature"
    - to "address this gap in the literature"
    - to "fill this gap"
    - to point at themes that others "have not paid particular attention to"
    - or to "call for more empirical research"
- *under-problematize* existing literature
- incremental, footnote-on-footnote research
- reinforces already influential theories

Alvesson, Sandberg (2010)

That's Interesting!

o How do theories which are generally considered interesting differ from theories which are generally considered uninteresting?

o Interesting theories are those which *deny* certain assumptions of their audience, while non-interesting theories are those which *affirm* certain assumptions of their audience.

! Attack on the taken-for-granted world of their audience.

! A theorist is considered great, not because his theories are true, but because they are *interesting*.

Davis, M. S. (1971)
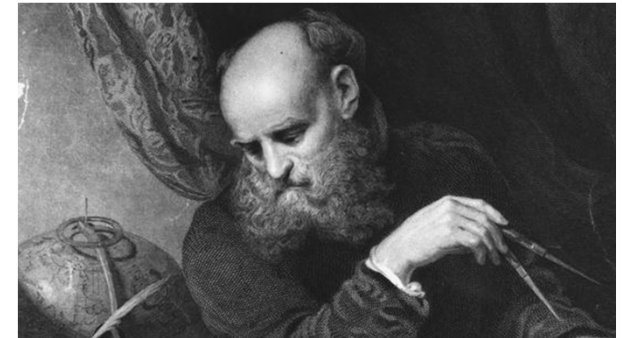
The Index of the Interesting

o What seems to be …
  – disorganized is organized
  – individual is holistic
  – local is general
  – ineffectively is effectively
  – bad is good
  – unrelated is correlated
  – exist together cannot exist together
  – heterogeneous is composed of a single element
  – stable/unchanging is unstable/changing
  – a positive co-variation is a negative co-variation
  – similar is opposite
  – independent is dependent

Davis, M. S. (1971)
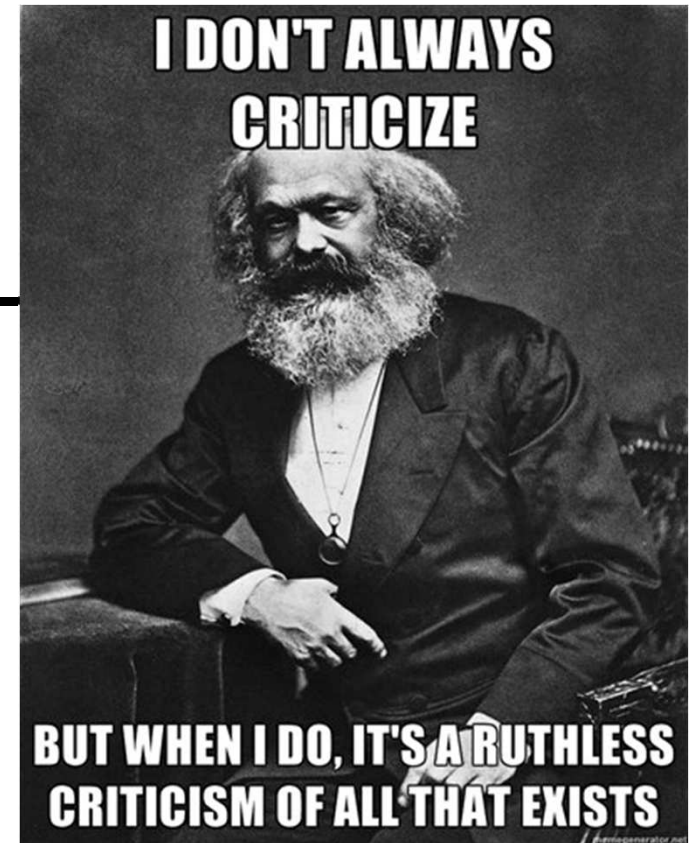
Examples for Abstraction

a. What seems to be an individual phenomenon is in reality a holistic phenomenon.

EXAMPLE: Emile Durkheim's assertion in *Suicide* that suicide, which was considered at the time he wrote to be a behaviour characteristic of an individual, is in fact (more crucially) a process characteristic of a society.

b. What seems to be a holistic phenomenon is in reality an individual phenomenon.

EXAMPLE: Sigmund Freud's assertion in 'Thoughts for the Times on War and Death' that war, which was considered at the time he wrote to be a social phenomenon, is in fact (more crucially) a psychological phenomenon.

Davis, M. S. (1971: 316)

The negative research agenda
Over-problematization

o highlights what is 'wrong' with existing knowledge

  – Social constructionism

  – Postmodernism

  – Feminism

  – Critical theory

o *over-problematizes* existing literature

o disrupts, rather than builds upon, and extends an established body of literature

Alvesson, Sandberg (2010)

Principles for identifying and challenging assumptions

o Evaluate you own work.

- That's obvious!

- It's absurd!

- That's interesting!

Alvesson, Sandberg (2010)
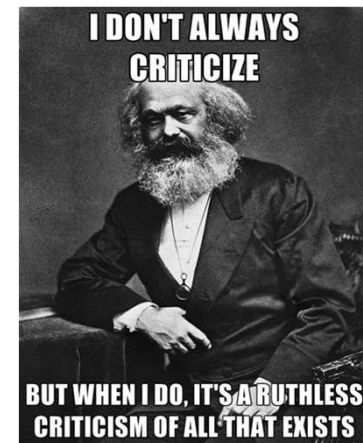
Content

1. Problem Specification
2. **Study Retrieval**
3. Coding

Identification of the Study Population

Who do you want to generalize to?

The Theoretical Population

What population can you get access to?

The Study Population

How can you get access to them?

The Sampling Frame

Who is in your study?

The Sample
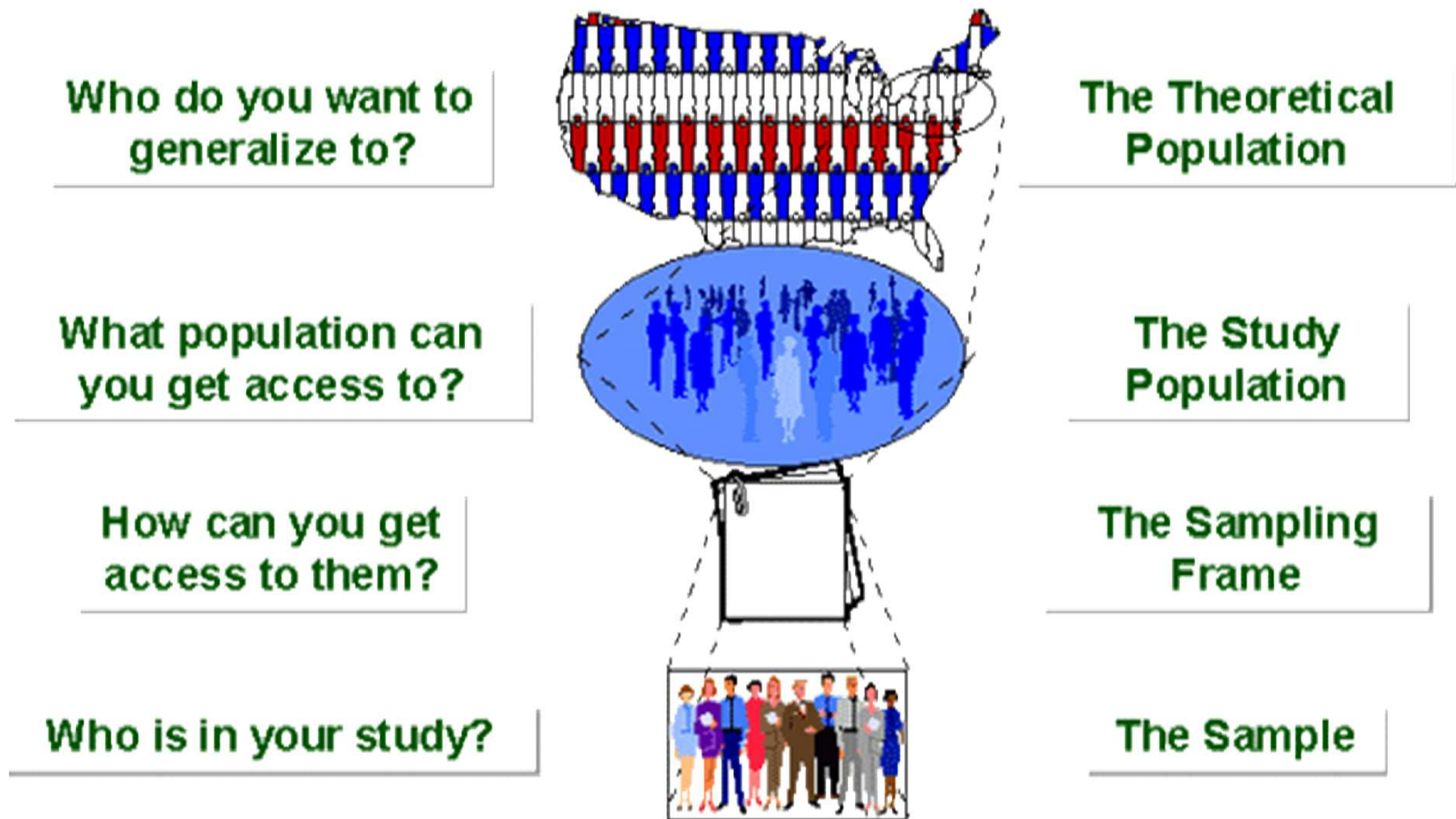
Predefinition of
Study Eligibility Criteria
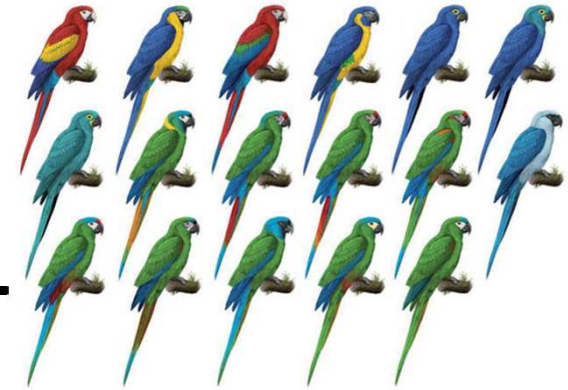
o Distinguishing features, e.g. topic, constructs

o Research respondents, e.g. SMEs, adults

o Key variables, e.g. control variables

o Research method, e.g. experimental design, survey

o Cultural and linguistic range, e.g. countries, only English reports

o Time frame, e.g. only recent studies

o Publication type, e.g. articles, books, dissertations, technical reports, unpublished manuscripts, conference presentations

Identification of Studies

o More! than one computerized bibliographic database (Web of Science, SSRN, Google Scholar, Research Gate)

  1. Effective keyword search

  2. Review articles

  3. Backward- and forward citation analyses

  4. Manual search of relevant publication outlets

  5. Experts

o Retrieve everything feasible; any omissions may create a selection bias!

Recording of identified Studies

o Identification numbers

o Bibliography

o Abstract

o Publication type

o Availability (ordered from...),

o Eligibility (dropped, abandoned...)

Study selection according to methodological criteria?

o Methodological quality exists in the eye of the beholder

– Research findings are not robust to methodological differences among studies

– Few schemes for assessing methodological quality

o Full representation of available research on a topic

Study selection according to form of research findings?
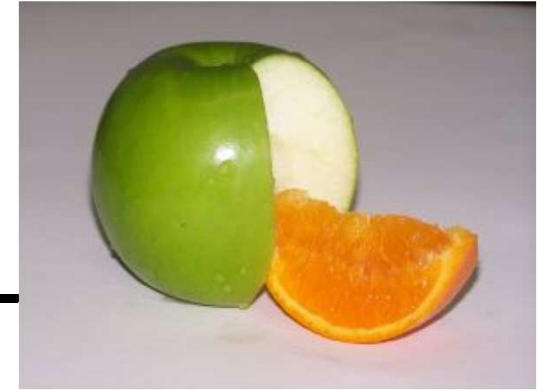
o Bivariate correlations

o Measurement research (reliability)

o Std. vs. unstd. regression coefficients

o Proportion of observations

o Pre-Post contrasts

o Group contrast , mean differences

⚠️ You can mix different forms but the same effect size statistics should be used for analyzing all the findings in a given meta-analysis and you have to control for the form of research in the meta-regressions

Content

1. Problem Specification
2. Study Retrieval
3. **Coding**

Developing a Coding Protocol

o Study descriptors (x)

– Study-level: time, sample and source descriptors, method, theory, researcher etc.

– Sub-effect-level: measurements, type of statistical test, control variables, construct reliability, missing data etc.

o Effect sizes (y)

– Std. size of the effect

– Sample size

– T-statistic

Example of a Data Base

	B1			f_x	ID_Metaanl															
	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S	T	U	V	W	X
1	ID_Met	ID	SubID	Verkett	Article	PubJahr	Theory	discussi	discussi	CSR-CFI	main re	Number	Other H	DepVar	Random	Sample	hard CS	relative	Kind of	Kind of	relative	Account indu
2	1763	1	1	1_1	Adams,	1998.0	Stakeh	yes	no	positiv	ressou	3.0	leverag	CSR	yes	1994.0	yes	no	hard	Cash gi	yes	Accoun no
3	1378	2	2	2_2	Alexan	1987.0	Shareh	no	no	0.0	positiv	0.0		CFP	no	1971-19	no	no	own su	CSR rar	yes	Event n yes
4	264	3	3	3_3	Amato,	2007.0	Busine	yes	no	positiv	due to	3.0	size, a	CSR	no	1999.0	yes	yes	hard	Cash gi	yes	Accoun yes
5	1580	5	4	5_4	Bansal	2005.0	Resour	no	no	0.0	/	7.0	interna	CSR	no	1986-19	no	no	own co	CSR rar	yes	Accoun no
6	777	6	5	6_5	Barnett	2006.0	Portfoli	yes	yes	u-shap	weak C	3.0	higher	CFP	no	1972-20	no	no	SRI	SRI vs.	yes	Market no
7	32	7	6	7_6	Bechet	2008.0	Shareh	yes	yes	0.0	positiv	0.0		CFP	no	k.A.	no	no	KDL	CSR rar	no	Accoun no
8	1964	7	7	7_7	Bechet	2008.0	Shareh	yes	yes	0.0	positiv	0.0		CFP	no	k.A.	no	no	KDL	CSR rar	yes	Accoun no
9	41	7	8	7_8	Bechet	2008.0	Shareh	yes	yes	0.0	positiv	0.0		CFP	no	k.A.	no	no	KDL	CSR rar	yes	Accoun no
10	141	7	9	7_9	Bechet	2008.0	Shareh	yes	yes	0.0	positiv	0.0		CFP	no	k.A.	no	no	KDL	CSR rar	yes	Accoun no
11	1319	8	10	8_10	Berron	2007.0	Norma	yes	no	positiv	stakeh	2.0	Corpor	CFP	no	2000-20	no	no	SiRi (cc	CSR rar	no	Accoun yes
12	1350	8	11	8_11	Berron	2007.0	Norma	yes	no	positiv	stakeh	2.0	Corpor	CFP	no	2000-20	no	no	SiRi (cc	CSR rar	no	Market yes
13	158	11	12	11_12	Bramm	2006.0	Efficier	yes	yes	0.0	positiv	0.0		CFP	no	2002.0	no	no	EIRIS	CSR rar	yes	Event n no
14	2001	12	13	12_13	Bramm	2004.0	Stakeh	no	yes	positiv	manag	10.0	industr	CSR	no	1989-19	yes	yes	hard	Cash gi	yes	Accoun no
15	2007	12	14	12_14	Bramm	2004.0	Stakeh	no	yes	positiv	manag	10.0	industr	CSR	no	1989-19	yes	yes	hard	Cash gi	yes	Accoun no
16	2008	12	15	12_15	Bramm	2004.0	Stakeh	no	yes	positiv	manag	10.0	industr	CSR	no	1989-19	yes	yes	hard	Cash gi	yes	Accoun no
17	2005	12	16	12_16	Bramm	2004.0	Stakeh	no	yes	positiv	manag	10.0	industr	CSR	no	1989-19	yes	yes	hard	Cash gi	yes	Accoun no
18	1611	13	17	13_17	Bramm	2005.0	Stakeh	no	no	0.0	/	12.0	philan	CSR	no	2002.0	yes	no	hard	Cash gi	yes	Accoun no
19	1936	14	18	14_18	Bramm	2008.0	empiri	yes	yes	0.0	diversi	1.0	cost-be	CSR	yes	1990-19	yes	yes	hard	Cash gi	yes	Accoun no
20	2566	14	19	14_19	Bramm	2008.0	empiri	yes	yes	0.0	diversi	1.0	cost-be	CFP	yes	1999.0	yes	yes	hard	Cash gi	yes	Market yes
21	2582	14	20	14_20	Bramm	2008.0	empiri	yes	yes	0.0	diversi	1.0	cost-be	CFP	yes	1995-19	yes	yes	hard	Cash gi	yes	Market yes
22	2580	14	21	14_21	Bramm	2008.0	empiri	yes	yes	0.0	diversi	1.0	cost-be	CFP	yes	1990-19	yes	yes	hard	Cash gi	yes	Market yes
23	1333	15	22	15_22	Bramm	2009.0	Stakeh	yes	no	0.0	positiv	3.0	firm-le	CSR	no	2002-20	yes	no	hard	Cash gi	yes	Accoun no
24	1311	16	23	16_23	Branco,	2008.0	Resour	yes	yes	0.0	legitim	4.0	interna	CSR	no	2003.0	no	no	own co	CSR rar	yes	Accoun no
25	1308	16	24	16_24	Branco	2008.0	Resour	yes	yes	0.0	legitim	4.0	interna	CSR	no	2003.0	no	no	own co	CSR rar	yes	Accoun no

Daten_codiert | Tabelle9

Bereit

Example of Study Coding

Publication Year
- 1975-1995
- 1996-2010

Journal Impact Factor
- Low Impact Factor (<.60)
- High Impact Factor (>.61)

Working Paper
- Published in a Journal
- Working Paper

Underlying Theory
- No Theory
- Social Science Theory
- Finance/ Economic Theory

H0 hypothesis
- No
- Yes

CSP-CFP Pros and Cons Discussion
- no discussion, only pro discussion
- balanced discussion

Industry-fixed effects
- No
- Yes

Firm-fixed effects
- No
- Yes

Time-lagged-effects
- No
- Yes

Kind of Analysis
- Regression
- *T*-test, mean comparison
- Correlation

CSP Measure (CSP ranking)
- SRI vs. Non-SRI Portfolio/Funds
- Crime/recall/incidents
- Cash giving/contributions
- Social disclosure
- CSP regulations/principles

CFP Measure: Accounting based
- Event based
- Market based

Analyzed Time Period
- 1960–69
- 1970–79
- 1980–89
- 1990–99
- 2000–09

Number of year analyzed

Important Aspects of Coding Protocols

o Use theoretical consideration to decide which variable should be included and which not

o Use close-ended items (open-ended only if categories for variables cannot predetermined)

o Cluster items that deal with similar topics

o Include not too many variables (time!)

o Include only variables which are available among all studies

o Develop coder guidelines, i.e. full definition of each item and coding of ambiguous, unusual, or borderline cases

Training of Coders

o Must understand the coding protocol

o Must have knowledge and skills to read and interpret research

o Must understand complex statistical reporting

o Mostly at least Master student!

Tip: Do the coding on your own and involve a master student for reliability checks

Reliability Checks

o Internal validity

– Consistency of a single coder: Draw a subsample of 20 or more studies and code it again (time lag!)

– Consistency of different coders: Compare the coding of two independent coders for a subsample of 20 or more studies

– You may include a confident scale for difficult items

o External validity

– Include missing data studies and analyze them separately for systematic differences

– Differentiate between „missing" and „not applicable"