Numbers for policy: Practical problems in quantification

Casteldefells 26/08/2018-01/09 /2018

Bruna De Marchi

The licence not to quantify. Qualitative versus quantitative analyses

Alternative Title

From the old days: An old social scientist's memories

Tuesday, August 28, 2018

Content and Purpose

Lecture:

Sharing some hints and some lessons from my professional experience as a sociologist (starting 1973), practicing both "qualitative and quantitative *research*". Today I will concentrate on the latter

Purpose:

Invitation to reflect whether such "memories" may be useful/inspiring/applicable at the time of BIG DATA

Quantitative vs Qualitative *Research* in Sociology (Social Sciences)

Often a case of "Epistemological Tribalism" (not sure when it started)

A tendency (need, craze) to define, distinguish and **take sides** …

Each side is an "epistemic bubble", reinforcing its own views and ignoring contrary ones

… On the basis of an inappropriately labelled distinction

The Physics Envy and its Redemption

In order to be "scientific" you have to speak the language of numbers

The task was immensely facilitated by:

- The advent of computers and, later, personal computers
- The availability of easy to use softwares (such as SPSS - Statistical Packages for the Social Sciences)

MORE or LESS of WHAT?

Quantity: more questions, more data, more statistical analyses, more tests, more diagrams, more and more "colorful" graphic representations, …

Quality: more chances for GIGO (garbage in, garbage out)

SURVEYS (but not only)

- WHO ASKS
- WHAT
- TO WHOM
- WHY (WHAT FOR)
- HOW
- WHERE
- WHEN

META-QUESTIONS
WHO, WHY (WHAT FOR)

Who are the investigators

Who funds their work

Which are their declared objectives

Which are their tacit objectives

Which are their assumptions

Which are the possible implications of their findings

….

RESEARCH DESIGN
TO WHOM

Representative samples?

Random samples?

Inquiry mode

Occasional vs Professional (paid) respondents

…

RESEARCH DESIGN
WHERE, WHEN

- Prevalence of WEIRD (Western, educated, industrialized, rich, and democratic)

- Who is available/willing when (time of day/week/year)

- Comparative samples

- Local situation and circumstances

- Translations

- …

Selection Bias
(Postal, Telephone, Internet Surveys)

- 1936 US presidential election: Republican Alfred Landon, against the current President, Franklin D. Roosevelt.

- On the basis of a (very expensive) **poll** with a **sample size of some 2.4 million people**, the prestigious magazine *Literary Digest* **predicted** a victory for **Landon** with **57%** of votes against **Roosevelt** with **43%**

- **The actual results** of the election were **62% for Roosevelt** against **38% for Landon**

- At the same time, George Gallup was able to correctly **predict a victory for Roosevelt** based on a **sample of about 50,000** (potential) voters

RESEARCH DESIGN
WHAT, HOW

- Topic (familiar, pertinent, significant, artificial, hypothetical, …)
- Question phrasing and format
- Response phrasing, format and instructions
- Number of questions
- Attitude of interviewee
- Interaction investigator/interviewer – interviewee
- Methodological assumption and technical artefacts
- …..

A SIMPLE QUESTION

Do you know Andrea Saltelli?

      1. Yes

      2. No

Numbers just labels (nominal *"scale of measurement")*

Not strictly necessary to provide an instruction such as: "Select one of the two" as answers are clearly mutually exclusive.

If there is no response, it is likely out of distraction forgetfulness (must be coded anyway)

BUT, there is a (possibly unrecognized) **assumption**: "knowing" means the same to everybody

IT IS REALLY A SIMPLE QUESTION?

Do you know Andrea Saltelli?

Possible new format after **realization** that "knowing" has no univocal meaning

1. I know him personally

2. I know about him

3. I never heard of him

4. ….

Instruction "Select one of the following" is advisable, as answers are not clearly mutually exclusive.

A LESS and LESS SIMPLE QUESTION

Please indicate your gender
 1. Female
 2. Male

FINISHED?

How do you interpret lack of answer in this case?

If you need/want to be more accurate, you may insert a generic
3. Other …
or be even more specific
BUT there are considerations other than technical ones in phrasing questions and answers

YES-NO Scales –
An Example of Oversimplification

- The so-called C-scale developed by the researchers Glenn Wilson and John Patterson in the 1960s is widely used to measure conservatism.

- It consists simply of 50 items where the respondent circles "yes" or "no" to indicate whether they "favor or believe in" the item in question.

- For scoring purposes, saying you dislike jazz contributes just as much to your designation as a "conservative" than saying you are in favor of the death penalty.

From: **Jesse Singal:** How Social Science Might Be Misunderstanding Conservatives

http://nymag.com/daily/intelligencer/2018/07/how-social-science-might-be-misunderstanding-conservatives.html?utm_source=undefined&utm_medium=undefined&utm_campaign=feed-part

More Elaborated Scales
Likert items and scales (origin 1932) - Ex. 1

Please state your agreement/disagreement with each of the following statements:

1. Bruna's lecture was interesting

2. Bruna's lecture was superficial

3. Bruna's lecture was funny

4. Bruna's lecture was incomplete

n. ...

1 Strongly Disagree	2 (Partially) Disagree	3 Neutral	4 (Partially) Agree	5 Strongly Agree

More Elaborated Scales
Likert type item and scales (origin 1932) - Ex. 2

Please state how often you preform the following actions:

1. Brushing your teeth after eating

2. Praying before going to bed

3. Washing your hands before touching food

4. Watching TV

n. …

Never	Rarely	Often	Always

"Scales of Measurement" ASSUMPTIONS and DOUBTS

- Words are translated into numbers

- Numbers are not just labels but represent a ranking. From nominal/qualitative to numerical/quantitative)

- 5>4>3>2>1 (or vice versa). Ranking → ordinal scale;

- The semantic distance between the available choices (5, 4, 3, 2, 1) is the same → interval-scale

- The numbers obtained can be subjected to the application of statistical techniques (including very sophisticated ones)

GOOD: as far as you recognize that these are artefacts for allowing statistical treatment and do not take numbers at face value …

A BIT OF EVERYTHING and TOO MUCH OVERALL

Investigator:

The job of analysis is easy and quick with the computer software.

So, let's get as much "data" as possible!

Respondent:

Do you think I don't have anything to do but fill your questionnaire?

57a. Lei pensa che i progressi scientifici che stiamo facendo, alla lunga, aiuteranno l'umanità o le nuoceranno?

A	Aiuteranno	1
B	Nuoceranno	2
C	Saranno in parte utili e in parte dannosi	3
	Non so	-1
	Non risponde	-2

o25

MOSTRARE IL CARTELLINO 58

58. Guardi questo cartellino e mi dica, per ogni istituzione elencata, quanta fiducia Lei ha in essa: molta, abbastanza, poca o nessuna.

(REGISTRARE UNA RISPOSTA PER OGNI ISTITUZIONE)

		Molta	Abbast.	Poca	Nessuna	Non so	Non risp.	
A	La Chiesa	1	2	3	4	-1	-2	v200
B	Le Forze armate	1	2	3	4	-1	-2	v201
C	Il sistema scolastico	1	2	3	4	-1	-2	v202
D	La stampa	1	2	3	4	-1	-2	v203
E	I sindacati	1	2	3	4	-1	-2	v204
F	La polizia	1	2	3	4	-1	-2	v205
G	Il parlamento	1	2	3	4	-1	-2	v206
H	La pubblica amministrazione	1	2	3	4	-1	-2	v207
I	Il sistema di previdenza sociale	1	2	3	4	-1	-2	v208
J	L'Unione Europea	1	2	3	4	-1	-2	v209
K	La NATO	1	2	3	4	-1	-2	o26
L	l'Organizzazione delle Nazioni Unite (ONU)	1	2	3	4	-1	-2	v210
M	Il sistema sanitario	1	2	3	4	-1	-2	v211
N	Il sistema giudiziario	1	2	3	4	-1	-2	v212
O	Le grandi aziende	1	2	3	4	-1	-2	o27

MOSTRARE IL CARTELLINO 58a

58a. E' d'accordo o contrario sulle seguenti affermazioni riguardanti la cooperazione internazionale?

		Molto d'accordo	D'accordo	Contrario	Molto contrario	Non so	Non risponde	
A	L'Italia non può risolvere da sola i suoi problemi ambientali, ma ha la necessità di collaborare con le organizzazioni internazionali che si occupano di protezione dell'ambiente	1	2	3	4	-1	-2	o28
B	L'Italia non può risolvere da sola i suoi problemi di criminalità, ma ha la necessità di collaborare con le organizzazioni internazionali di polizia	1	2	3	4	-1	-2	o29
C	L'Italia non può risolvere da sola i suoi problemi della disoccupazione, ma ha la necessità di collaborare con le organizzazioni internazionali che si occupano di sviluppo economico	1	2	3	4	-1	-2	o30

64. *Quanto sono rispettati i diritti umani attualmente in Italia? Secondo Lei c'è:*

A	molto rispetto per i diritti umani	1
B	abbastanza rispetto	2
C	poco rispetto	3
D	nessun rispetto	4

Non so	-1	
n risponde	-2	v224

MOSTRARE IL CARTELLINO 65

65. *Usando questo cartellino per favore mi dica per ciascuno dei seguenti comportamenti se Lei ritiene che possa essere giustificato mai oppure sempre oppure qualcosa di intermedio secondo le circostanze.* **(LEGGERE INVERTENDO L'ORDINE ALTERNATIVAMENTE DA UN'INTERVISTA ALLA SEGUENTE. REGISTRARE UNA RISPOSTA PER CIASCUNA AZIONE O CONDIZIONE)**

A   Cercare di ottenere dallo Stato benefici a cui non si ha diritto

1   2   3   4   5   6   7   8   9   10   -1   -2   v225
Mai giustificato   Sempre giustificato   Non so   Non risp.

B   Non pagare le tasse (o pagarle meno del dovuto) se si ha la possibilità

1   2   3   4   5   6   7   8   9   10   -1   -2   v226
Mai giustificato   Sempre giustificato   Non so   Non risp.

C   Prendere senza permesso e guidare un'automobile di altri (per fare un giro)

1   2   3   4   5   6   7   8   9   10   -1   -2   v227
Mai giustificato   Sempre giustificato   Non so   Non risp.

D   Usare marijuana o hashish

1   2   3   4   5   6   7   8   9   10   -1   -2   v228
Mai giustificato   Sempre giustificato   Non so   Non risp.

E   Dire il falso nel proprio interesse

1   2   3   4   5   6   7   8   9   10   -1   -2   v229
Mai giustificato   Sempre giustificato   Non so   Non risp.

F   Un uomo o una donna sposati che hanno una relazione fuori del matrimonio

1   2   3   4   5   6   7   8   9   10   -1   -2   v230
Mai giustificato   Sempre giustificato   Non so   Non risp.

G   Accettare denaro non dovuto (una bustarella) nell'adempimento del proprio dovere

1   2   3   4   5   6   7   8   9   10   -1   -2   v231
Mai giustificato   Sempre giustificato   Non so   Non risp.

H   Omosessualità

1   2   3   4   5   6   7   8   9   10   -1   -2   v232
Mai giustificato   Sempre giustificato   Non so   Non risp.

I   Aborto

1   2   3   4   5   6   7   8   9   10   -1   -2   v233
Mai giustificato   Sempre giustificato   Non so   Non risp.

J   Divorzio

1   2   3   4   5   6   7   8   9   10   -1   -2   v234
Mai giustificato   Sempre giustificato   Non so   Non risp.

K   Eutanasia (porre termine alla vita di un malato incurabile)

1   2   3   4   5   6   7   8   9   10   -1   -2   v235
Mai giustificato   Sempre giustificato   Non so   Non risp.

THE PRESENT

The general gist: we get much of our news from Facebook feeds and similar sorts of social media. Our Facebook feed consists mostly of our friends and colleagues, the majority of whom share our own political and cultural views. We visit our favourite like-minded blogs and websites. At the same time, various algorithms behind the scenes, such as those inside Google search, invisibly personalise our searches, making it more likely that we'll see only what we want to see. These processes all impose filters on information.

From: **C. Thi Nguyen**, Escape the echo chamber, AEON Newsletter

https://aeon.co/essays/why-its-as-hard-to-escape-an-echo-chamber-as-it-is-to-flee-a-cult

THE FUTURE
From Epistemic Bubbles to Echo Chambers?

- Remember I started with *"epistemological tribalism"* and *"epistemic bubbles"*?

- Where an **epistemic bubble** merely omits contrary views

- An **echo chamber** brings its members to actively distrust outsiders. Dissenting voices are not only banned, but actively discredited.

I AM NO LONGER TALKING ABOUT RESEARCH ALONE

SEE YOU THIS AFTERNOON FOR THE PRACTICUM