

L3. Surveys of surveys

Olivier Godechot

Sciences Po – Campus Reims

Inquiries in Sociology

Building a questionnaire. Art or science

- Surveys (or Questionnaires) are one of the main tools of social sciences for inquiries:
 - sociology, political science, economics, psychology, ... But also, geography, epidemiology, and even history...
- Little research in France on the practice of questionnaire.
- Important research in the US on surveys in psychology/social psychology with tests.
 - Cf. Public Opinion Quarterly
 - mainly devoted to study differences in formulation. Not very normative.
- Questionnaire is the product of social science research
- Which needs to take into account « psychological factors » (interactions)
- ... and economic factors (price of the questionnaire)

1. History and examples of questionnaires

Historical elements on surveys. First survey (?)

1838.]

[11]

QUERIES OF THE STATISTICAL SOCIETY OF LONDON RELATING TO STRIKES.

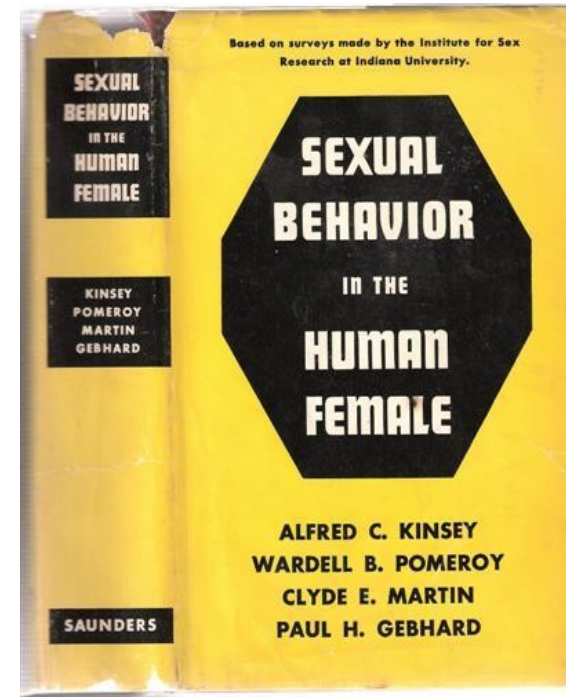
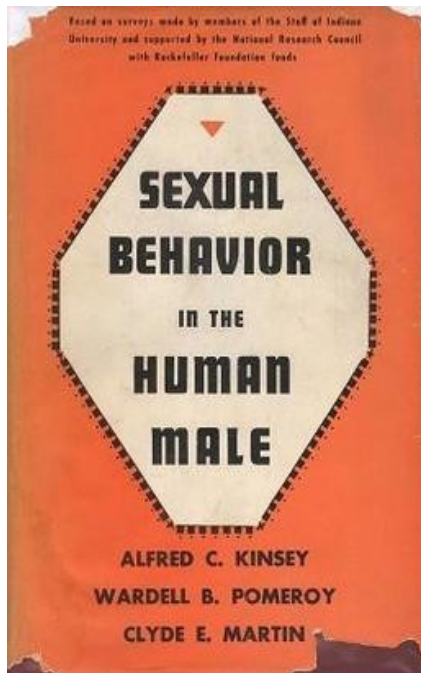
1. Name of town or district.
2. Population of ditto.
3. Characteristics of the population, and of their staple employ, with reference to the strike.
4. State the circumstances of the trade, whether flourishing or otherwise, at the time of the strike.
5. State the occupation of persons engaged in the strike or combination, shewing the particular class with which it originated.
6. What were the average weekly earnings immediately before the strike, and when in full work, of the classes with whom the strike originated?
7. Were the earnings, or other circumstances of the workmen, with whom the strike originated, inferior to those of persons employed in the same branch of manufacture, in the same or neighbouring towns or district?
8. Were they in any way, and in what way, superior?
9. Did any Trades' Union exist in the district before the strike, or was any formed in consequence?
Number and description of its members—its constitution—rules for management—remuneration of officers—its declared objects—mode of operation.
10. Were any, and what, compulsory measures taken to make workmen join the Union?
11. Did any association of masters exist in the district before the strike, for counteracting the operations of Trades' Unions, or for any other purpose?
12. Did the strike originate in the town or district, or was it part of a more extensive plan, and was it fomented by delegates from any other body of workmen?
13. Was the strike under the direction of any Committee; and, if so, what was its constitution and mode of operation; and had that Committee absolute and secret powers?
Endeavour to furnish a copy of the rules.
14. What was the ostensible cause of the strike or combination; state whether any and what notice was given, either by the masters or men, as the case might be, of the time of its commencement.
15. Were any, and what, proposals for accommodation made by either party for the acceptance of the other?
Send copies of any papers connected with such proposals.
16. On what day did the strike commence?
17. In how many establishments did the strike prevail? In how many did the workmen continue employed.
18. State the number of persons engaged in the strike or combination, and the number who continued at work.
19. State the number and description of workmen and other persons whose employments depended upon those of the persons engaged in the strike, and who were consequently deprived of occupation during its continuance.

- Early social statistics on administrative data mainly.
 - Life, Death, Crime, etc.
 - Exhaustive.
 - “Events” rather than accounts and perceptions
- 1838. “Fourth Annual Report of the Council of the Statistical Society of London”, *Journal of the Statistical Society of London*, 1(1): 5-13
 - Linked to the social question. On strikes
 - “to collect a statistical account of the various strikes and combinations which have existed in different parts of the United Kingdom for the purpose of altering the rate of wages, and of introducing new regulations between masters and men”
 - No sampling. Sent to a list of knowledgeable persons.
 - “the Council have printed, a numerous list of queries, designed to elicit the complete and impartial history of strikes. Copies of these queries, accompanied with a printed explanatory letter, have been transmitted to **many intelligent individuals connected with, or interested in the welfare of, manufactures and other industrial pursuits in which large bodies of operatives are employed; also to the Editors of all Newspapers in manufacturing districts, and to every Mechanics' Institute in the Kingdom.**”

Historical elements. Questionnaires in Sociology

- Early sampling techniques
- Agriculture inquiries under the New Deal (Didier, 2009)
- Gallup. *American Institute of Public Opinion*
- Random poll in order to predict election. Predicts 1936 presidential election.
- In sociology
 - Stouffer, Samuel A., et al. 1949. *The American soldier*
 - Survey (multiple surveys) on up 500,000 soldiers during WWII.
 - Questions ranging from food rations to promotions or attitudes towards the enemy
 - Attitudes and satisfaction

A scandalous questionnaire?



<https://www.youtube.com/watch?v=5Mx0h4xkoQQ>
(5'22 to 7'57)

Kinsey Report

- Kinsey, Alfred C., Wardell B. Pomeroy, and Clyde E. Martin, 1948. *Sexual Behavior in the Human Male*. Indiana University Press.
- A very large, marking, debated, controversial inquiry on sexual behavior

Total Number of Histories

Year	Increment	Total
1938 (6 months)	62	62
1939	671	733
1940	959	1692
1941	843	2535
1942	816	3351
1943	1510	4861
1944	2490	7351
1945	2668	10019
1946	1467	11486
1947 (part)	728	12214

Interviewers Involved

Interviewer	Histories taken	Percent of total histories
Kinsey	7036	57.6
Pomeroy	3808	31.2
Martin	890	7.3
Others	480	3.9

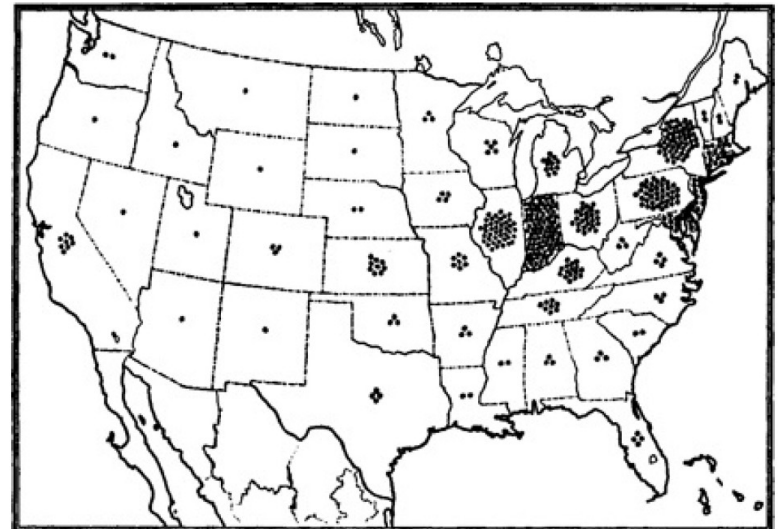


Figure 1. Sources of histories.
One dot represents 50 cases

Kinsey Report (2)

- Sampling.
 - Not random.
 - Based on volunteers found thanks to connections with community leaders
 - Over-samples students (departure population), prisoners, prostitutes and Indiana
 - Criticized for a selection bias. Challenge to produce such results with random samples
 - However, even with randomized survey → refusal to answer produce bias
- The questionnaire
 - No formal questionnaire
 - Adapting questions to the interviewee and the context
 - Cf. next slide
 - In fact a more or less formalized questionnaire existed and published after Kinsey's death.
 - Interviewers had to learn by heart the items and the coding (very demanding)
 - 570 items for coding sex history
- The principle of the when? questions
 - “When did you first masturbate”
 - “When did you have your first sexual intercourse with someone of the same sex?”

Standardizing the point /adapting the form of the question

“While the point of each question has been standardized, the form of each question has varied for the various social levels and for the various types of persons with whom the study has dealt. Standardized questions do not bring standardized answers, for the same question means different things to different people. In order to have questions mean the same thing to different people, they must be modified to fit the vocabulary, the educational background, and the comprehension of each subject. It is especially important to use a vocabulary with which the subject will feel at home, and which he will understand.”

- [illegible]

CL: $f \times w f 3-x$ inst $f \times$ ① Bk p-1st wr
REM: 18-2 w f $f \times$
6-12 inst.
ins
p q p
Pr $x \exp 9-16$ f thg 3-d

Items Covered on Sex Histories

I. SOCIAL AND ECONOMIC DATA

1. Sex
2. Age
3. Date of birth
4. Race
5. Geographic origin
 - State of subject's birth
 - Countries, states of residence for a year or more
 - Parents' place of birth
6. Rural-urban background
7. Religious background
 - Denominations involved
 - Degree of adherence
8. Occupational history
9. Economic status

* Items marked with asterisks (*) are checked only for selected series of individuals who are available for special study.

10. Educational history

- Years of schooling
- Colleges attended
- College majors
- Age upon leaving school
- Age while in high school

*11. Psychological test ratings

12. Recreational interests

- Extracurricular activities in school and college
- Moving pictures
- Dancing
- Cards
- Gambling
- Smoking
- Use of alcohol

I. SOCIAL AND ECONOMIC DATA (Cont'd)

12. Recreational interests (Cont'd)

- Use of narcotics
- Use of marihuana
- Hunting
- Fishing
- Reading
- Sewing
- Cooking
- Housework
- Special interest in music
- Special interest in sports
- Other special interests and sources of recreation

13. Athletic experience

- On high school and college organized teams

14. Fraternity or sorority membership in college

15. Home background

- Parents' occupational status
- Parents' economic status
- Parents' educational background
- Parental marital history
- Happiness of adjustment
- Separation or divorce
- Parent-child relationships
- Attachment to father
- Attachment to mother
- Brothers, sisters
- Number
- Ages
- Companions at 10 and at 16
- Number
- Relative numbers, male and female

16. Institutional history

- Prison, orphanage, etc.
- Army or Navy experience

17. Personality traits

II. MARITAL HISTORIES (FOR EACH MARRIAGE)

1. Marital status

2. Spouse's history

- Age
- Length of previous acquaintance
- Length of engagement
- Religious affiliations
- Educational history

3. Age of each spouse at marriage

4. Years married, divorced, separated, or widowed

5. Common law marriages

6. Offspring

- Sex
- Ages
- Mother's age at first childbirth

7. Abortions

- Spontaneous
- Induced

8. Marital adjustment

- Rating
- Sources of conflict

III. SEX EDUCATION

1. Sources of knowledge, ages when learned

- Pregnancy
- Coitus
- Fertilization
- Menstruation
- Veneral diseases
- Prostitution
- Contraception
- Abortions
- Male erection (in female histories)

2. Parental contribution to sex education

3. Experience in observing sex behavior

4. Experience with graphic depictions of sexual activity

5. Formal sex education in school and college

6. Attitudes on nudity

- Of parents
- Of subject

IV. PHYSICAL AND PHYSIOLOGIC DATA

1. General development and health

- Height
- Weight, and maximum ever reached

*Pulse rate

*Blood pressure and BMR

- Thickness of lips
- Handedness
- History of chronic illnesses and handicaps

History of venereal disease

2. Adolescence: ages at onset of

- Erotic responsiveness
- First orgasm (and its source)

Pubic hair growth

Breast development (in females)

Breast knots (in adolescent males)

Menstruation

Voice change

Onset of rapid growth

Completion of growth

3. Genital characters: male

- Testes
- Descent
- Position (of right and left)
- *Size
- History of injury

VII. HETEROSEXUAL HISTORY (Cont'd)

12. Heterosexual prostitution (the subject as prostitute; males or females as heterosexual prostitutes) (Cont'd)

Love affairs
Percent married
Number of virgins
Occupations
Racial groups
Sources of contacts
Personal introductions by friends
In established houses of prostitution
Ages of subject
Number of houses involved
Size of houses
Geographic location
Physical or other force used to hold prostitutes
In own established house
Time involved
Geographic location
Number of persons employed
Street solicitation
Percent of rejects
Percent causing trouble
Petting techniques
Partners active in petting (in 9 techniques shown in VII-2)
Subject active (in same 9 techniques)
Coital positions (as shown in VII-9)
Orgasm, frequencies for subject
Prophylaxis and contraception
Examination of partners
Condoms
Antiseptic douche
Other prophylactics
Variant techniques
Anal coitus
Anilinctus
Flagellation
Other sadism
Masochism
Scatology
Voyeurism
Fetishism
Group activities
Age of subject
Frequency
Number of persons in group involved
Psychologic reactions

Exhibitionistic activity

Ages involved
Frequency
Sizes of audiences
Character of audiences
Psychologic reaction of subject
On first occasion
On subsequent occasions
Pay: range, average
Techniques of the exhibition
Income from prostitution
Range of payment
Average payment
Average earned per week, per month
Percent paid to houses
Percent paid to procurer
Rolling
Frequency
Range of intake
Average intake
Social involvements
With friends
With families
With police
Frequency of arrest and conviction
For prostitution
For rolling
For other concomitants
Legal penalties
Bases of dismissal
Amounts paid for protection
Socio-sexual background
Factors involved in beginning prostitution
Chief factors for continuing prostitution
Extent of pleasure derived from sexual relation
Plan for continuance of prostitution
Psychologic conflicts of prostitute
Effects on marriage
Willingness to recommend prostitution to others

VIII. HOMOSEXUAL HISTORY

1. Pre-adolescent play

Ages involved, frequencies
Companions
Ages
Number
Techniques
Exhibition
Manual manipulation
Vaginal or urethral insertion

VIII. HOMOSEXUAL HISTORY (Cont'd)

1. Pre-adolescent play (Cont'd)

Mouth-genital contact

Anal

2. Post-adolescent experience

Ages involved
First experience
Age
Partner
Age
Race
Relation to subject
Circumstances
Place of contact
Initiation of approach
Techniques employed, passive, active, or mutual
Financial arrangements
Satisfaction for subject
Age of first experience with each technique, passive and active
Manual
Oral
Anal
Breast (for female)
Femoral
Full body contact
Frequency
During first year
Maximum, ever, per day
Maximum, ever, per week
Average per week during each year
Total number of contacts
Partners
Total number
Age range
Comparisons with age of subject
Age preferences
Reasons for age preferences
Social position
Students in grade school
Students in high school
Students in college
Clergy
Teachers
Art groups
Professional persons
Business groups
Armed forces
Laboring groups
Law enforcement officers
Highest position held
Number married
Number without previous homosexual experience

Number of oncers

Duration of longest affairs
Relations involving love and affection

Percentage of approaches which are rejected
Races involved: white, Negro, others

Techniques

Petting, passive and active
Lip kissing
Tongue kissing
Body kissing
Breast manipulation, manual
Breast manipulation, oral
Genital manipulation, manual
Genital manipulation, oral
Flagellation on back, buttocks, genitalia
Urethral insertions
Anilinctus
Nudity
Positions involved (including 69)
Preference for light or dark
Places involved
Subject's orgasm
Frequency by each technique or by spontaneous ejaculation
Partner's orgasm
Frequency by each technique or by spontaneous ejaculation

3. Psychic reactions

Preferences for
Masculine or feminine type of partner
Partner of particular height
Partner of particular weight
Partner of particular complexion
Particular amount of body hair
Particular genital characteristics
Particular breast characteristics
Circumcised partner
Other physical qualities of partner
Reaction to odor and taste, genitalia and semen

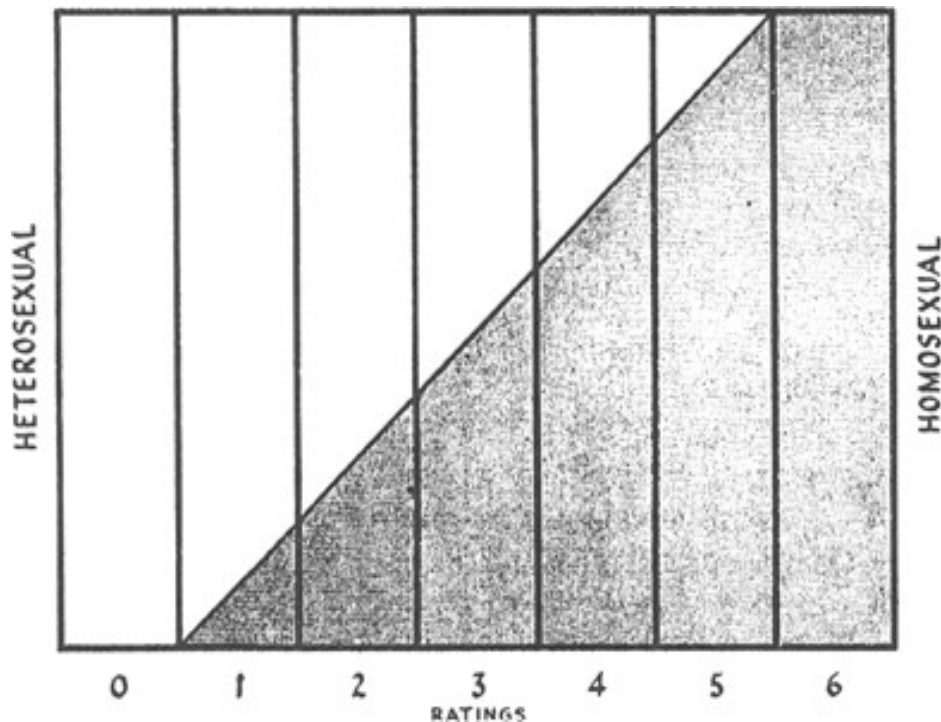
4. Sources of contacts

Personal friends
Pick-ups
Street

An exploration of the core and (outrageous) margins

- III. Sources of sexual outlet
- 14. Masturbation
- 15. Nocturnal Emissions
- 16. Heterosexual Petting
- 17. Premarital intercourse
- 18. Marital intercourse
- 19. Extra-marital intercourse
- 20. Intercourse with prostitutes
- 21. Homosexual outlet
- 22. Animal contacts

A few results on homosexuality



AGE	CASES	HETEROSEXUAL-HOMOSEXUAL RATING: ACTIVE INCIDENCE SINGLE WHITE MALES—EDUC. LEVEL 0-8							
		X	0	1	2	3	4	5	6
5	820	%	%	%	%	%	%	%	%
6	820	92.6	2.3	0.1	0.4	1.5	0.2	0.1	2.8
7	819	87.1	2.9	0.5	1.2	2.8	0.2	0.1	5.2
8	819	81.8	4.2	1.5	2.2	3.2	0.9	0.2	6.0
9	819	73.7	6.2	1.8	3.5	4.3	0.7	0.4	9.4
10	819	70.5	6.6	1.8	4.3	4.6	0.7	0.5	11.0
11	819	62.9	8.5	2.2	5.1	5.9	1.0	0.5	13.9
12	818	59.9	10.0	2.3	5.4	6.4	1.1	0.5	14.4
13	815	54.1	13.3	2.8	6.5	7.1	1.2	0.6	14.4
14	814	49.8	20.8	2.3	5.7	6.0	1.6	0.9	12.9
15	811	41.2	31.4	1.7	5.3	6.3	2.1	1.4	10.6
16	803	32.0	42.1	2.7	6.2	5.1	2.7	1.5	7.7
17	774	24.1	50.1	3.7	7.6	5.0	3.1	1.4	5.0
18	708	17.5	56.4	4.0	8.1	5.1	3.1	1.4	4.4
19	635	11.5	61.6	4.6	7.9	5.7	2.4	1.7	4.6
20	550	10.7	62.5	4.5	8.0	6.0	2.5	1.3	4.5
21	473	9.3	65.1	3.2	7.6	5.7	2.5	1.5	5.1
22	386	8.8	63.0	4.1	7.8	6.7	2.9	2.3	4.4
23	306	7.8	63.9	4.2	6.9	7.8	2.6	2.6	4.2
24	269	8.2	61.7	4.5	8.2	7.4	3.3	2.6	4.1
25	231	6.1	62.3	4.3	9.1	8.2	3.5	1.3	5.2
26	187	4.8	59.9	7.0	9.1	9.1	3.2	1.6	5.3
27	172	2.3	60.5	5.8	9.9	10.5	3.5	1.7	5.8
28	152	2.6	57.3	7.2	10.5	11.2	3.3	2.6	5.3
29	134	2.2	55.3	6.0	11.2	13.4	3.7	1.5	6.7
30	121	2.5	54.5	5.8	12.4	13.2	4.1	1.7	5.8
31	107	1.9	53.2	6.5	11.2	13.1	4.7	1.9	7.5
32	95	2.1	54.7	5.3	8.4	11.6	7.4	4.2	6.3
33	87	2.3	54.2	5.7	8.0	11.5	5.7	4.6	8.0
34	78	2.6	56.4	5.1	10.3	11.5	2.6	5.1	6.4
35	72	2.8	58.2	1.4	11.1	12.5	2.8	5.6	5.6
36	68	2.9	58.8	1.5	11.8	13.2	1.5	5.9	4.4
37	65	3.1	58.5	1.5	10.8	13.8	1.5	4.6	6.2
38	59	1.7	57.5	1.7	11.9	15.3	1.7	3.4	6.8
39	57	1.8	57.8	1.8	12.3	15.8	0.0	3.5	7.0
40	50	2.0	64.0	2.0	8.0	18.0	0.0	0.0	6.0

Estimation of homosexual practices

- Kinsey Report
 - 37% of males,
 - 20% of females
 - involved at least once in their life in homosexual relations
- Many think massive overestimation due to sampling
 - Prisoners / Homosexual community
- After adjusting for bias, Van Wyk and Geist (1984; see also Gagnon & Simon, 1973) found
 - 1.8% of women and 4.7% of men in scored 4.01-6.00 on the Kinsey scale,
 - whereas 4.1% of women and 9.6% of men scored $K = 0.005-4.000$
 - bisexual experience is more common than predominantly homosexual experience.

Estimation of homosexual practices (2)

- France ; CSF (2006)
 - 4% of females, 4.1% of males declare having sex with same sex partners once in their life – often during teenage years (13% of them)
 - Comparable with ACSF (1992)
 - 2.6% of females, 4.1% of males
 - 1.0% of females, 1.6% of males during the last 12 months.
 - Rare are people exclusively homosexual: 0.3%
 - 0.5% of females and 1.1% of males define themselves as homosexuals
 - 0.8% of females and 1.1% of males say they are bisexuals.
- Comparable figures for other countries US, UK.
 - Arch Sex Behav. 1995 Jun;24(3):235-48.
 - HOPE survey. Larger estimates :
 - Attraction : 8.7, 7.9, and 8.5% of males and 11.1, 8.6, and 11.7% of females in the US, the UK, and France report some homosexual attraction.
 - Behavior: 6.2, 4.5, and 10.7% of males and 3.6, 2.1, and 3.3% of females in the US, the UK, and France, had sexual contact with someone of the same sex in the previous 5 years

A problematic survey: *Analyse des comportements sexuels des français (1992)*

Ref: Spira, Bajos et ACSF, 1993 ; Giami et al., 1998

- Not the only major inquiry on sexual behavior in France
 - 1970. Simon ; 2006. CSF
 - Strong methodological reflexion
- Context : Peak of AIDS epidemic
- Questionnaire done by a multidisciplinary team : sociology, demography, economics, psychology, psychoanalysis, epidemiology).
- Choice of the mean of interview
 - Phone interview
 - Strong procedure of anonymization through separation of phoning and interviewing
 - One respondent complained of sexual phone harassment after questionnaire → after investigation, came from relatives rather than interviewer.
 - Protects both the interviewer and the interviewee

A problematic survey: Analyse des comportements sexuels des français (1992) - 2

- Interviewers
 - 90 interviewers.
 - Strong emphasis on the training of the interviewers
 - Devereux, 1980 : “An interview on sexuality, even if scientific, is a kind of sexual interaction”.
 - Risk of counter-transfer
 - Training on : AIDS/ Questionnaire structure / Interviewer – interviewee interaction (tone of voice)
 - Presence of researchers and psychologists in the phone call center
 - Frequent collective “debriefing” of interviewers
 - Difficult cases : interviewee with AIDS
 - Interview of older persons by young interviewers (image of parental sexuality)

The filter

- Sampling and survey
 - Two questionnaires
 - A short one (15 minutes) 20 000 males and females from 18 to 65 years
 - A longer one (45 minutes) : person with a HIV risk profile (2200) + a random sampling (2600).
 - Selection question➔
 - And also a questionnaire for refusal to answer (age, gender, occupation, reason)

As I told you, this is a survey of 20,000 people. It's not useful to give the same questionnaire to everyone. In order to know which questionnaire I should ask you, I will read you a list of situations and you will just tell me if one of these situations concerns you, without telling me which one. I mean without telling me which one:

- *You were born on the 4th, 17th, or 20th of any month of the year.*
- *You have had sex with at least two different people in the last 12 months.*
- *In the past 5 years, you have had sex at least once with a person of the same sex as you.*
- *In the past 5 years, you have paid for sex at least once.*
- *In the past 12 months, you have used a soft (hash, marijuana...) or hard (cocaine, heroin...) drug at least once.*
- *You have hemophilia*

(Translated with www.DeepL.com)

The design of ACSF questionnaire

- Problem of Kinsey report
 - Adapting the vocabulary to the understanding of the interviewee or standardize
 - Only 9 (high skilled) interviewers
- Close questionnaire. Choice of a descriptive vocabulary.
 - Avoid both scientific vocabulary (ex. fellatio, sodomy, cunnilingus) or slang :
 - Choice of descriptive vocabulary
 - “You sucked the sex of your partner”
- Answers are designed in a way not to let other people understand to what question you are answering
 - Yes/No ; Number; Frequency, etc.
- Questions parallel to sex life
 - Questions on attitudes towards risks
 - Security belt in a car
 - Car accident, cancer, etc.
 - Questions on death
 - Frequency of thoughts of death
 - Emotions related to death
- Progression of the questionnaire (400 questions)
 - Start with socio-economic characteristics, childhood, opinions related to sex and HIV, sexual practices, sexual diseases, ends with fantasies
- Sensibility of responses to the wording (Béjin, 1993)
 - Did you masturbate ?
 - 42% of female at least once in their life time
 - Do you reach orgasm when you masturbate?
 - 51% say yes.

2. Why questionnaire?

Why questionnaire?

- Measuring
 - Setting equivalences
 - Quantifying
 - We don't have an idea of figures on large population
- Exploring
 - Discovering new facts
 - Links in the data
 - Inductive
- Confirming theories
- Set up the questions according to these goals?
 - With the statistical exploitation in mind

Prioritize key questions (confirmatory approach)

- Shape the questionnaire accordingly to what you want to demonstrate
 - Questions on your Y
 - Questions on your X
 - Not always easy. Think in terms of proxy
 - Don't forget
 - Control variables
 - Instruments (advanced)

Godechot 2014. "Getting a job in finance. The strength of collaboration ties." *European Journal of Sociology*.

- Finance pay through hold up
 - On line questionnaire with efinancialcareers
- Key variables
 - Y: Description of pay
 - X: Asset moves
 - Description of the *capacity to move assets*
 - Description of the give value to assets through *collaboration ties*

Finding a proxy for moveable assets and collaboration ties

What was at stake during this recruitment?

(multiple choice question ; n=489)

Replacing someone 27%

Reinforcing a team 55%

Bringing new techniques
21%

Bringing new clients 7%

Providing new strategies
11%

Developing a new business
25%

Table 3. Ties and collaboration ties

Questions	Items	%
Who was at the origin of the last move? (n=532)	A head hunter contacted me	20%
	An employee of another firm contacted me	18%
	I contacted a head hunter	7%
	I contacted directly a firm	27%
	I replied to a firm	27%
Did you know employees in the service where you were hired? (n=531)	Former colleagues	22%
	Business partners	13%
	Former classmates	13%
	Friends	8%
	Others	15%
Once in your new job, did you help to hire some former colleagues? (n=469)	No, I did not try	76%
	I tried with no success	10%
	1 or 2 colleagues	12%
	3 and more	2%
When you changed jobs, did you ever move with other colleagues in an another firm? (n=469)	No	85%
	With 1 or 2 colleagues	12%
	With 3 and more	3%
If your boss moves to another firm and invites you to come along, would you follow him or her? (n=233)	Very likely	15%
	If other members of the team would go	2%
	If conditions are interesting enough	66%
	No	18%

3. *Asking questions*

Standardizing, that's the question!

- Experimental model as an ideal
 - Question in a questionnaire \approx experimental protocol
 - Ask the same question with the same response items to different persons
 - Put into equivalence persons who respond the same way to the same question.
- Consequences (contradictory)
 - Questions both simple and precise
 - Answers simple, few and precise but sufficiently numerous to capture the response the interviewee wants to give.
 - Two goals:
 - standardizing without distorting.
 - Asking questions without imposing research problem (questions that interviewee never ask themselves)

The correct level of standardization

- Very imprecise
 - + Short and simple
 - + Standardized
 - + Easy to use in statistics
 - - Inadequate
 - - Respondent unsatisfied
- Very precise
 - + Adequacy
 - + Respondent satisfaction
 - + Good descriptions
 - - Complexity and length
 - - Loosing standardization

Social variation in interpretations of questions

Bessière, C. and F. Houseaux. 1997.
“Suivre des enquêteurs”, *Genèses*
29: 100-114.

- Followed interviewers and see that in interaction the understanding of questions does not correspond to the one intended by the designers
- *Do you ever feel during your work:*
 - *The impression that what you do, anybody could do it ?*
 - *Yes*
 - *No*

- Housekeeper. Says yes but as a source of pride.
 - ➔ Housekeeping is not exclusively specific to one deprived strata of population

Starting from fieldwork... (but beware)

- Common advice is to do exploratory qualitative open interviews
- Generally good advice
 - Avoid questions, answers, wording which do not correspond to people's experience
- but might lead to impose one specific fieldwork experience to all interviewees.
- Ex : Contacts survey (1983). Friendship part
 - Do you eventually know a few couples or persons who you would consider as your best friends ?
 - Yes/ No
 - How many persons or couples?
 - 1, 2 or 3
- Imposition of the notion of « couple of friends ». 83% of friends who live in couple are therefore declared as couples of friends.
- Leads to much more gender mix than when we ask to describe the three best friends without proposing/imposing the “couple” item (Godechot, 2000)

Finding the right level of generality

- Risk of asking irrelevant questions to interviewees
- ... Or of not proposing the relevant answer
 - Interviewee becomes angry
 - Might abandon
- To overcome this risk temptation of making an exhaustive nomenclature of reality
 - Ex: Survey on job mobility in Finance
 - Object : finance → finance
 - But other possibilities
 - Finance -> other sector
 - Other sector -> finance
 - Finance -> Unemployment -> Finance
 - Finance -> Unemployment -> Other sector
 - Other sector -> Unemployment -> finance
 - Ex. Description of the family
 - Classical description : Father, Mother, Brother 1-2 ; Sister 1-2
 - Now step father ; step mother ; same sex parents ; large families, etc.
- To capture all the possible cases
 - Costly in terms of questions / time
 - A lot of questions for rare cases...
 - Costly in terms of cognitive tiredness
- In order to match to reality, temptation to set filters and have nested questions for different people
 - Advantage to ask nested questions : adaptation to different groups.
 - Only ask to the different groups the question that are relevant to them.
 - Limits: sample on subgroup might get small.
 - If pushed too far : different questionnaire for all individuals.
 - We would lose the logic of the experimental nature of the questionnaire, where all individuals are submitted to the same research protocol,
 - Selection bias

The wording of questions

Grémy, Jean-Paul. 1987. “Les expériences françaises sur la formulation des questions d'enquête: Résultats d'un premier inventaire,” *Revue française de sociologie*, 28 (4): 567-599

Difference between positive or negative wordings

- Should have accepted the invitation
Versus
- Was right to refuse the invitation

- Form A. Do you think General de Gaulle **should have accepted** President Roosevelt’s invitation to Algiers?
- Form B. Do you think General de Gaulle **was right to refuse** President Roosevelt’s invitation to Algiers?
- Form A:
 - Yes 27%.
 - No 58%.
 - NA 15%
- Form B:
 - No 15%.
 - Yes 63%.
 - NA 22%

Practice versus opinion

- Objective quantification versus subjective quantification
- Social sciences tend to prefer objective questions
 - Easier to use in a causal framework
- Limit: illusion of objectivity.
 - Quantification might not be the correct one.
 - Quantification and memory effort might be too important.
 - Interviewees don't count precisely but rather answer broadly.
- Ex: Travail et modes de vie, 1997/
 - Do you work at night ?
 - Yes, more than 50 nights per year
 - Yes, less than 50 nights per year
 - No
 - Do you have to work at a fast pace?
 - All the time
 - More than half of the time
 - Less than half ...
 - How much are involved in your work?
 - Little
 - Just the right amount
 - A lot

Different types of question

- Closed questions (single answer)
- Closed questions (multiple answers)
- Open questions (short answer)
- Open questions (long answer)

Comparison between open question and closed question

- What to do with agricultural surpluses?
 - People don't think spontaneously of the market option "Sell for less"
 - When proposed it becomes obvious
- Is it or not an "imposition of a problem"?
- Form A. For several agricultural products, such as wheat or milk, France produces more than it needs. In your opinion, what should be done with the surplus? (open question)
- Form B. For several agricultural products, such as wheat or milk, France produces more than it needs. Among the following measures to reduce the surplus, which one(s) would you prefer: produce less, sell for less, sell abroad, give to underdeveloped countries?

	Form A	Form B
- sell abroad	52%	40%
- sell for less	10%	34%
- give to underdeveloped countries	16%	17%
- produce less	1%	3%
- other measures 12% -	12% /	
- don't answer	9%	6%

There is a capital of knowledge accumulated in questions

- French occupation (PCS) in French Labor Force Survey. (Enquête *Emploi*)
 - Job's title
 - Job's status
 - Job's position (in collective agreements)

A1 a - Profession principale. Intitulé précis de profession.

.....

b - Si M.. est agent de l'Etat, d'une collectivité locale, d'un hôpital public, d'un service (EDF, SNCF, etc.), militaire de carrière, **préciser sa classification : corps grade, etc.**

.....

c - SI NECESSAIRE, éléments complémentaires précisant l'intitulé de profession.

GR

A3 a - M.. exerce-t-il sa profession :

1. A son compte (*agriculteur, artisan, commerçant, industriel, profession libérale, gérant majoritaire de SARL, libre ou en location gérance, etc.*).....

2. Comme salarié

b - M... est-il :

A5 - POUR LES SALARIES

Quelle est la position professionnelle de l'emploi actuel de M... ?

1. Manoeuvre ou ouvrier spécialisé (*OS1, OS2, OS3, etc.*).....

2. Ouvrier qualifié ou hautement qualifié (*P1, P2, P3, TA, OQ, etc.*).....

3. Agent de maîtrise dirigeant des ouvriers, maîtrise administrative ou commerciale.....

4. Agent de maîtrise dirigeant des techniciens ou d'autres agents de maîtrise.....

5. Technicien, dessinateur, VRP (*non cadre*).....

6. Instituteur, assistant(e) social(e), infirmier(e) et autre personnel de catégorie B de la fonction publique.....

7. Ingénieur ou cadre (*les employés, techniciens, agents de maîtrise n'ayant pas la qualité de cadre ne pas se classer ici, même s'ils cotisent à une caisse de retraite*).....

8. Professeur et personnel de catégorie A de la fonction publique.....

9. Employé de bureau, employé de commerce, agent de service, aide soignant(e), gar

Ex. Wealth

- Wealth is difficult to measure
 - Several methods: complete nomenclature (securities, real estate, furniture, debts, etc.). Very long. Value
 - Global question.
 - In French Wealth survey (*Enquête Patrimoine*):
- *« A votre avis, si vous aviez à liquider la totalité de ce que les personnes de votre ménage possèdent, sans oublier les actifs professionnels et les biens durables, combien le ménage pourrait-il retirer de la vente ? »*
- “If you were to sell everything you own, without forgetting professional assets and durable goods, how much would you get from the sale?”
 - Presentation of a card with tranches.

Some questions in the French LFS worsen with time

Quel est votre matrimonial status ?

Sans objet (moins de 15 ans) ou non renseigné.

- 1 Célibataire.
- 2 Marié(e) ou remarié(e).
- 3 Veuf(ve).
- 4 Divorcé(e).

Vivez-vous en couple (avec un membre du ménage) ?

1. Oui
2. Non

Legal question

What is your matrimonial status ?

Without object (less than 15 years) or non answers.

- 1 Bachelor.
- 2 Married or remarried.
- 3 Widow.
- 4 Divorced.

New forms of family led to use mainly the following

Do you live with a partner (with someone from the household) ?

1. Yes
2. No

A supplementary question on the PACS only introduced in 2013

Legality issue (France). Law Informatique et liberté (ART 8)

Forbidden to collect race, religion, political opinion

“I. - Il est interdit de collecter ou de traiter des données à caractère personnel qui font apparaître, directement ou indirectement, les origines raciales ou ethniques, les opinions politiques, philosophiques ou religieuses ou l'appartenance syndicale des personnes, ou qui sont relatives à la santé ou à la vie sexuelle de celles-ci.”

But with many exceptions

II. - Dans la mesure où la finalité du traitement l'exige pour certaines catégories de données, ne sont pas soumis à l'interdiction prévue au I :

1° Les traitements pour lesquels la personne concernée a donné son **consentement exprès**, sauf dans le cas où la loi prévoit que l'interdiction visée au I ne peut être levée par le consentement de la personne concernée ;

(...)

7° Les **traitements statistiques réalisés par l'Institut national de la statistique et des études économiques** ou l'un des services statistiques ministériels dans le respect de la [loi n° 51-711](#) du 7 juin 1951 sur l'obligation, la coordination et le secret en matière de statistiques, après avis du Conseil national de l'information statistique et dans les conditions prévues à l'article 25 de la présente loi ;

8° Les **traitements nécessaires à la recherche**, aux études et évaluations dans le domaine de la santé selon les modalités prévues au chapitre IX.”

Contradictory demands. Privacy issues / Knowledge

- France
 - CNIL will judge the questionnaire regarding the goal
 - More constraints on public survey (Insee/Ined) than on private ones (Sofres, etc.)
 - Political opinion
 - Bread and butter of electoral survey
- Questions on religion
 - Religious practice (easy)
 - Religion in specific survey
- Race
 - Very difficult to ask questions on race (not necessarily impossible)
 - Question planned in TEO1 (2008), but last minute suppressed
 - Roundabouts (TEO1) => Country of birth of parents, grand parents. Language spoken at birth
- US
 - Race in the census
 - Religion not in the census

4. Conducting a questionnaire

Conducting a questionnaire

- First questionnaire
 - (Questions. Cf. part 3) ; Size ; Order
- Testing first draft
 - Verify the good understanding of questions and answers
 - Estimate length and cost
 - Avoid questions with Stalinist scores (95% yes).
- Revision of the questionnaire
- The layout of the questionnaire as a *technical device*
 - Ex. Questionnaire for French LFS *Emploi*
 - Helps the collection
- Sample of respondents
- Training of interviewers
- Interviews
 - Feedback from interviewers
- Data collection.
- Data production
 - Description of the inquiry (feedback); Codebook ; Sample weights ; Data cleaning ;
- Statistical exploitation
- Eventual post-inquiry interviews

Size

- It's really imperative to limit the number of questions
 - Limits the cost of a questionnaire
 - Limits the fatigue or abandons of interviewees.
 - Statistical illusion that all questions will be exploited.
Generally only a limited set of questions are exploited
- Don't neglect control questions
 - Difference between a questionnaire and randomized experience.
 - Important to control for a set of given variables

Size. Prioritize key questions

- Building a questionnaire is the occasion of an analytical decomposition of the object studied
- Good questionnaires speak to people.
 - They don't seem artificial.
 - In line with the problems they face
- Temptation of too much details
- Think already in terms of statistical exploitation
 - What will you learn ?
 - What type of statistical analysis
- Value-added of one question relative to another one

Order. Helping people to memorize

- Battery effect.
 - Battery of questions with all the same wording and style. => boring. Repetition of the same answers.
- Questionnaire tries to help people to memorize without tiring or boring
- Not too much answer items. Problems of memory (if >7). No instantaneous memorization:
 - Some interviewees don't wait for reading/listening the complete list
 - Others get the complete list of items but forget elements
 - Presentation of a card

Order. Similarity/diversity

- Boredom, fatigue, length

Herzog A. Regula and Jerald G. Bachman. 1981. “Effects of Questionnaire Length on Response Quality,” *The Public Opinion Quarterly*, 45(4): 549-559

- Problem of abandon (especially on internet/or via the phone)

Table 2. Percentages of Respondents Answering with Identical or Nearly Identical Response Categories to an Entire Set of Items in the Long and the Short MtF Questionnaires^a

	<i>Long Form</i>	<i>Short Form</i>
All identical response categories	4.0%	0.7%
All-but-one identical response categories	2.9	0.6
All-but-two identical response categories	2.6	0.8
All-but-three identical response categories	3.1	1.0
All-but-four identical response categories	4.0	1.1

^a The set consists of 23 items on the importance of various job characteristics.

Order. Coherence effect

Howard Schuman, Graham Kalton and Jacob Ludwig, « Context and Contiguity in Survey Questionnaires », *The Public Opinion Quarterly*, Vol. 47, No. 1 (Spring, 1983), pp. 112-115

- Communist reporter item: Do you think the United States should let Communist newspaper reporters from other countries come in here and send back to their papers the news as they see it?
- American reporter item: Do you think a Communist country like Russia should let American newspaper reporters come in and send back to America the news as they see it?

Table 1. | Percent Yes to the Communist Reporter Question by Context and Contiguity

<i>Order</i>	<i>A</i> <i>Com./Amer.</i>	<i>B</i> <i>Amer./Com.</i>	<i>C</i> <i>Amer./(17 items)/Com.</i>
Percent	44.4%	70.1%	66.4%
(Sample size)	(117)	(107)	(107)
	A vs. B: $\chi^2_1 = 15.20$, $p < 0.001^a$		
	B vs. C: $\chi^2_1 = 0.34$, n.s.		
	A vs. C: $\chi^2_1 = 10.95$, $p < 0.001^a$		

^a The overall χ^2 for the table is $\chi^2_2 = 18.11$; $p < 0.001$. For an approximate multiple comparisons test based on the Bonferroni inequality, the p -values may be increased by a factor of 3; the A vs. B contrast remains significant at the $p < 0.001$ level and the A vs. C contrast is significant at the $p < 0.005$ level.

Order. Framing of memory

- Surveys on sociability :
 - 1983: 78% of interviewee talked at least once a week with a friend
 - 1997: 66% of interviewee talked at least once a week with a friend
- Collect method
 - 1983: collection of a full week schedule where all contacts are recorded
 - 1997: pre-coded grid of contacts (friends 1 to 5) during one week.
- Difference in figures might owe much to the difference in recording

Order. Manage the order

- Random order.
 - Psycho-sociology. Capture more reflex answers.
 - Does not stimulate memory
 - Fatigue and boredom for the interviewee
 - Less motivation
 - No memory work
- Structuring around a history
 - Increases the confidence of the interviewee
 - Stimulates memory
 - Organization of the questionnaire with difficult/boring questions at the end (building on confidence).

Techniques of collection

- Face to face
- Telephone
- Self-administered
 - Traditional email (T envelop)
 - Internet
- Incentives for respondents?
- Cost
 - 200 000 euros for *salaire et justice* (phone interviews: 3000 respondents/ 25 to 40 minutes) in 2008, done by Insee
 - 100 000 euros for *salaire et justice fonction publique* in 2010 (1000 respondents), done by a private firm.

Sampling

- Exhaustive
 - Small population
 - Census
 - Cost
- Random sample
 - Best way of generating a representative sample
 - Central limit theorem insure that the mean estimated on a sample is an unbiased estimator of the mean in the population
 - Possible to calculate confidence intervals
 - Limit: under the condition that non-responses do not bias the sample
 - Limit: you need an initial sampling database (exhaustive or representative) to do the sampling.
 - Addresses file ; Phone book ; Other possibilities (based on date of birth...)

Sample. Confidence interval of a mean

- The “true mean” is in the following interval

$$\left[\bar{X} - 1.96 \frac{s}{\sqrt{n}}, \bar{X} + 1.96 \frac{s}{\sqrt{n}} \right]$$

- When proportion = 50% in the sample, CI of the “true proportion” in a sample of :
 - 10 : CI of [20%,80%] (Gaussian proxy discussable)
 - 100: CI of [40%,60%]
 - 1000: CI of [47%,53%]
 - 10,000: CI of [49%,51%]

Sample. Complex sampling

- Stratified random sample
 - If we know in advance our population, we can draw directly subsamples in subgroups.
 - Ex: Rich males, poor males, rich females, poor females
 - Improves statistical properties of simple random sample: decreases variance in estimation.
- Clustered sample
 - First step random sample of clusters (geographic or other sort).
Than a sample of individuals within the cluster.
 - If the cluster is a territory, decreases the cost.

Sample. When random sample are impossible

- Quota sample
 - According to a set of given variables important to represent properly (ex: gender, age, occupation).
 - Set up of quotas
 - We ask the interviewer to find n interviewees for each intersection
 - Find 10 blue-collar males aged 25 to 35 years.
 - No necessity to have a sampling database
 - Problem of finding a sheep with five legs for the interviewer
 - We don't know exactly the bias and this makes it more difficult to calculate confidence intervals...
- Snowball sample
 - The interviewer asks the interviewee to recruit, or provide names of, other interviewees.
 - It might be difficult to have a representative sample
 - Good for small milieu ; network studies

Sample. How representative is a representative survey?

- Levels of analysis
 - Relation between a preliminary sample and a population
 - Relation between preliminary sample and responding sample
 - For each question some people don't answer. Missings
 - Different techniques for correcting for missing... More or less based on missing at random
 - Ask yourself more: what kind of population does the survey in the end represent
- Can a non representative survey be helpful?
 - Differentials (ex. male vs female) in a non-representative sample may represent true differentials
 - Hard to prove