



# Effects of different translations of answer scales

Marlène Sapin, Dominique Joye, Sylvie Leuenberger-  
Zanetta Alexandre Pollien, & Nicole Schöbi

[Marlene.Sapin@fors.unil.ch](mailto:Marlene.Sapin@fors.unil.ch)

## Outline

- Context
- Research questions
- Method
- Some results and methodological issues
- Conclusion and outlook

## Context

- In multi-language survey, a necessary condition to obtain comparable measurement is to assure the equivalence of all the questions in all languages, both in meaning and in response scale (Smith, Mohler, Harkness & Onodera, 2005).
- In survey, a limited categories tend to be use.
- Attitudinal surveys seeks to arrange responses along an underlying continuum (ex: satisfaction / agreement)
- Each element of response scales can pose difficulties. Wording is one of those because of the structural and lexical differences accross languages.
- In mono-cultural contexts : 2 main approaches can be used to measure the strenght of answer categories:
  - Ask respondents to rate the strength of terms defining each point on the scale.
  - Measure distributions generated by different answer scales.

## Research questions

The goal of our experiment is to find out :

- Whether and how different translations influence the measurement of responses?
- which translation can provide the best comparability at a Swiss national level and accross countries?

In this presentation: exploratory results of analyses made at a national level

- Do different translation variants of response scale change the quality of measurement in terms of reliability and validity ?
- Do these translations variants modify the distribution of points along the supposed underlying continuum of the dimension measured by the question?

## Method

- data :
  - ESS, 2006 / additional national questions (F2F)
  - 3 groups-split balloting (variant A = identical; Variant B; Variant C)
  - 3 questions to test the 3 variants of the 11-points satisfaction scale
  - 3 questions to test the 3 variants of the 5-categories agreement scale
  - German (n=1326) / French (n=409) / Italian (n=69)
  
- Approaches
  - Two-groups split-ballot MTMM (Saris, Satorra & Coenders, 2004)
  
  - Log-multiplicative associative model (Clogg, 1982)

## The three wording variants of the 11-points satisfaction scale

	<b>Main questionnaire Variant A</b>	<b>Variant B</b>	<b>Variant C</b>
<b>German</b>	Sehr unzufrieden ... Sehr zufrieden	Äussert unzufrieden ... Äussert zufrieden	Überhaupt nicht zufrieden ... Äussert zufrieden
<b>French</b>	Très insatisfait... Très satisfait	Extrêmement insatisfait... Extrêmement satisfait	Pas du tout satisfait... Très satisfait

Extremely unsatisfied

Extremely satisfied

00      01      02      03      04      05      06      07      08      09      10

### Questions:

- 1) All things considered, how satisfied are you with your life as a whole nowadays?
- 2) On the whole how satisfied are you with the present state of the economy in Switzerland?
- 3) Now thinking about the Swiss government, how satisfied are you with the way it is doing its job?

## The three variants of the agreement scale

	<b>Variant A Main questionnaire</b>	<b>Variant B</b>	<b>Variant C</b>
<b>German</b>	Stimme stark zu	Stimme stark zu	Sehr einverstanden
	Stimme zu	Stimme eher zu	Einverstanden
	Weder noch	Weder noch	Weder noch
	Lehne ab	Lehne eher ab	Nicht einverstanden
	Lehne stark ab	Lehne stark ab	Überhaupt nicht einverstanden
<b>French</b>	Tout à fait d'accord	Tout à fait d'accord	Tout à fait d'accord
	Plutôt d'accord	D'accord	D'accord
	Ni d'accord, ni en désaccord	Ni d'accord, ni en désaccord	Ni d'accord, ni pas d'accord
	Plutôt en désaccord	En désaccord	Pas d'accord
	Tout à fait en désaccord	Tout à fait en désaccord	Pas du tout d'accord

Agree strongly

Agree

Neither agree nor disagree

Disagree

Disagree strongly

### Questions:

- 1) The government should take measures to reduce differences in income levels.
- 2) Gay men and lesbians should be free to live their own life as they wish.
- 3) Modern science can be relied on to solve our environmental problems.

## Results

- Descriptions (test-retest, mean)
  
- Reliability and validity
  - 11-points satisfaction scale
  
- Estimates of the distances between categories
  - 11-points satisfaction : satisfaction with life
    - instrumental variable: happiness
  - Discussion : the role and the difficulties of the instrumental variable choice
  - 11-points satisfaction scale: satisfaction with life
    - comparaison with another instrumental variable

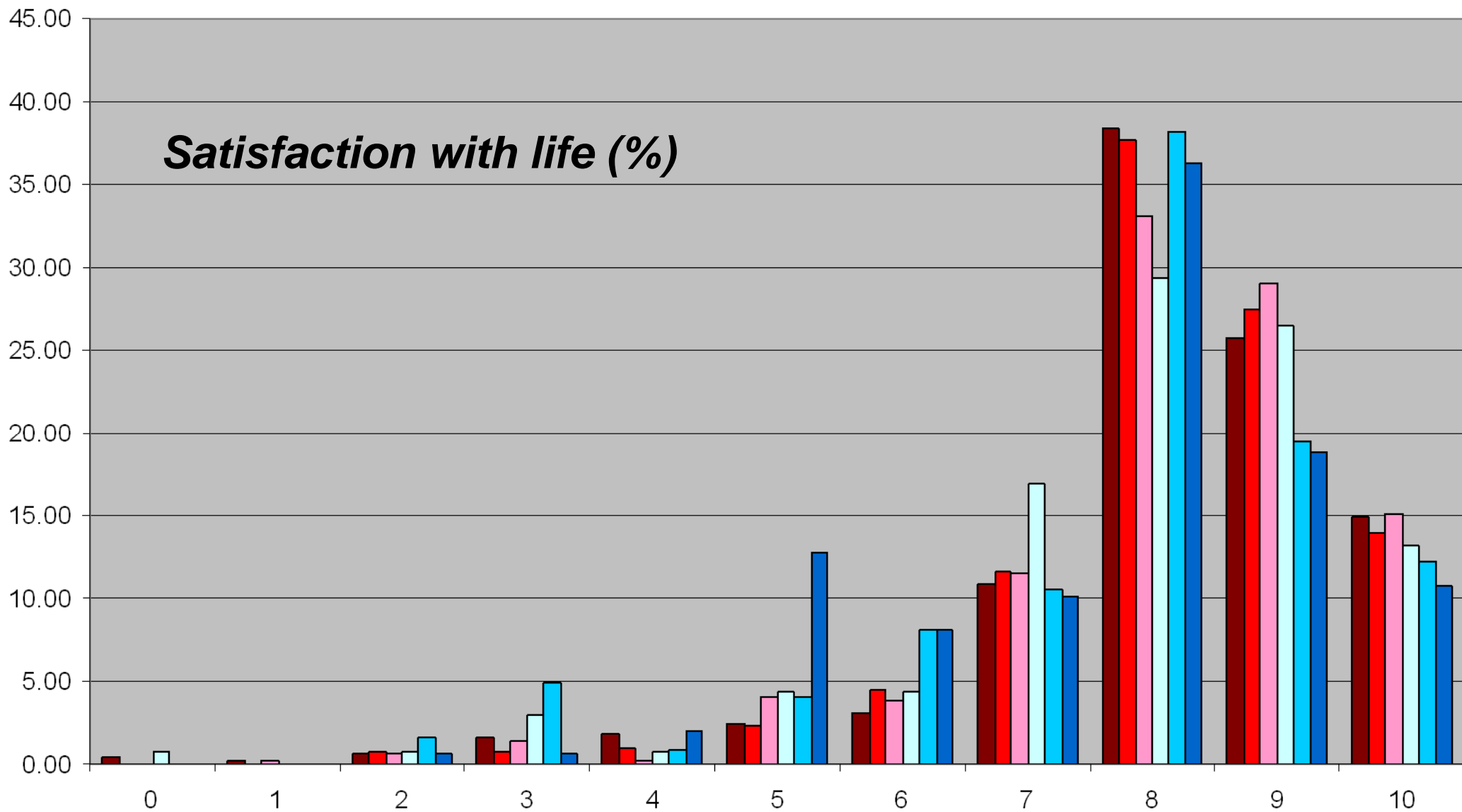


## *Test-retest reliability scores and means of the 3 items testing the 11- points satisfaction scale*

<b>Test-retest scores</b>	<b>Variant A</b>		<b>Variant B</b>		<b>Variant C</b>		
	<b>Satisfaction with</b>	<b>German</b>	<b>French</b>	<b>German</b>	<b>French</b>	<b>German</b>	<b>French</b>
life		0.71	0.77	0.68	0.62	0.69	0.78
economy		0.69	0.67	0.72	0.65	0.67	0.74
Gouvernement		0.74	0.77	0.74	0.67	0.71	0.76

<b>Means</b>	<b>Variant A</b>		<b>Variant B</b>		<b>Variant C</b>		
	<b>Satisfaction with</b>	<b>German</b>	<b>French</b>	<b>German</b>	<b>French</b>	<b>German</b>	<b>French</b>
life		8.07 (1.62)	7.85 (1.80)	8.17 (1.37)	7.65 (1.83)	8.11 (1.54)	7.60 (1.66)
economy		6.69 (1.58)	5.96 (1.69)	6.55 (1.59)	6.19 (1.77)	6.55 (1.59)	6.06 (1.69)
Gouvernement		6.04 (1.65)	5.63 (1.89)	5.88 (1.65)	5.63 (1.60)	5.77 (1.68)	5.40 (1.97)

# *Satisfaction with life (%)*



■ German variant A ■ German variant B ■ German variant C ■ French variant A ■ French variant B ■ French variant C

## *Test-retest reliability scores and means of the 3 items testing the 5-categories agreement scale*

Test-retest scores	Variant A		Variant B		Variant C	
	German	French	German	French	German	French
Measure for income difference reduction	0.77	0.77	0.83	0.80	0.81	0.80
Lifestyle freedom for homosexuals	0.84	0.81	0.86	0.83	0.86	0.82
Science reliable to solve environmental problems	0.76	0.57	0.76	0.66	0.75	0.74

Means	Variant A		Variant B		Variant C	
	German	French	German	French	German	French
Measure for income difference reduction	2.28 (.93)	2.15 (1.06)	2.35 (1.01)	2.20 (1.15)	2.27 (1.00)	2.22 (1.10)
Lifestyle freedom for homosexuals	2.04 (.97)	1.85 (1.06)	2.00 (.96)	2.15 (1.15)	1.92 (.92)	2.12 (1.08)
Science reliable to solve environmental problems	2.54 (.86)	3.25 (.88)	2.68 (.96)	3.21 (.93)	2.49 (.94)	3.33 (.86)

## *Eleven points satisfaction scale: 2-groups split-ballot-MTMM*

	German			French		
	Variant A ESS	Variant B	Variant C	Variant A ESS	Variant B	Variant C
<b>Reliability</b>						
Satisfaction with life	.94	.78	.79	.96	.75	.87
Satisfaction with economy	.84	.90	.87	.84	.88	.96
Satisfaction with gouvernement	.87	.89	.88	.90	.88	.88
<b>Validity</b>						
Satisfaction with life	.98	.96	.93	.99	.86	.95
Satisfaction with economy	.98	.98	.95	.98	.88	.95
Satisfaction with gouvernement	.98	.98	.95	.98	.87	.96
<b>Method variance</b>	.10 <sup>a</sup>	.09	.19	.10 <sup>a</sup>	.52	.24

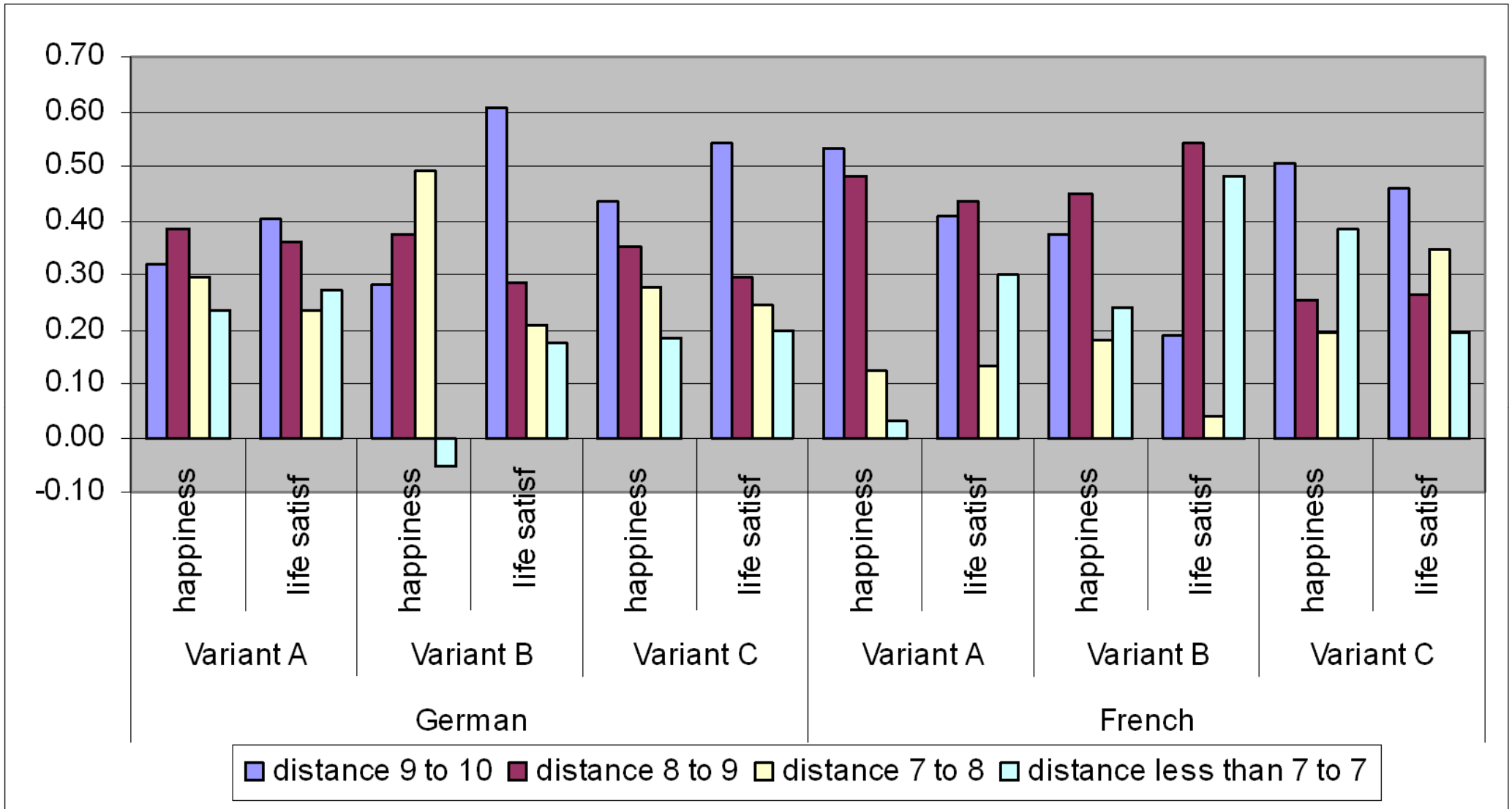
## 11-point satisfaction scale: the model of association for the satisfaction with life – Overall Happiness

	German			French		
	Variant A	Variant B	Variant C	Variant A	Variant B	Variant C
<b>Extremely satisfied – 10</b>	0.65	0.62	0.72	0.77	0.71	0.77
<b>9</b>	0.33	0.34	0.28	0.24	0.34	0.26
<b>8</b>	-0.05	-0.03	-0.07	-0.24	-0.11	0.00
<b>7</b>	-0.35	-0.52	-0.35	-0.37	-0.29	-0.19
<b>Less than point 7</b>	-0.58	-0.47	-0.53	-0.40	-0.53	-0.57
<b>Weighted mean</b>	.06	.06	.06	-.02	-.02	-.02
<b>Range of the scores</b>	1.23	1.09	1.24	1.17	1.25	1.34
<b>Distance between 10 - 9</b>	0.32	0.28	0.44	0.53	0.38	0.51
<b>Distance between 9 - 8</b>	0.39	0.37	0.35	0.48	0.45	0.25
<b>Distance between 8 - 7</b>	0.30	0.49	0.28	0.12	0.18	0.19
	$\chi^2 = 380.98^{***}$ , $L2 = 178.98^{***}$ , BIC = 20.29, $df = 22$			$\chi^2 = 73.24^{***}$ ; $L2 = 49.88^{***}$ , BIC = -82.36, $df = 22$		

## ***Requierements for the choice of the instrumental variable*** ***(Clogg, 1982)***

- There must be association between R and C.
- The model must produce estimated category scores which do not violate known “ordinality” requirements.
- The model must fit the data to an acceptable degree .
- The instrumental variable must be chosen with the best substantive and/or theoretical available information about how the R variable is associated with it.
- It may be advisable to exploit information from more than one instrumental variable or to exploit a group of in order to reinforce the inferences.

**Positive side of the 11-point satisfaction scale: Distances found in using two different instrumental variable - happiness vs life satisfaction (Q: life satisfaction)**



## Conclusion & Outlook

- wording scale variants are closed, but they are not similar.
- The impact of wording variants seems to be stronger in the case of hard questions or question with more social desirability. Is the impact of wording variants stronger in the case of hard questions or question with more social desirability?
- The MTMM approach:
  - the scores of reliability and validity of the 11-points satisfaction scale can be modified by the wording variants.
  - Alternative method for agreement scale?
- Clogg's approach:
  - Even if we manage to find the good instrumental variable, can we postulate that this variable is invariant across the cultural context?
- Alternative methodological approaches?



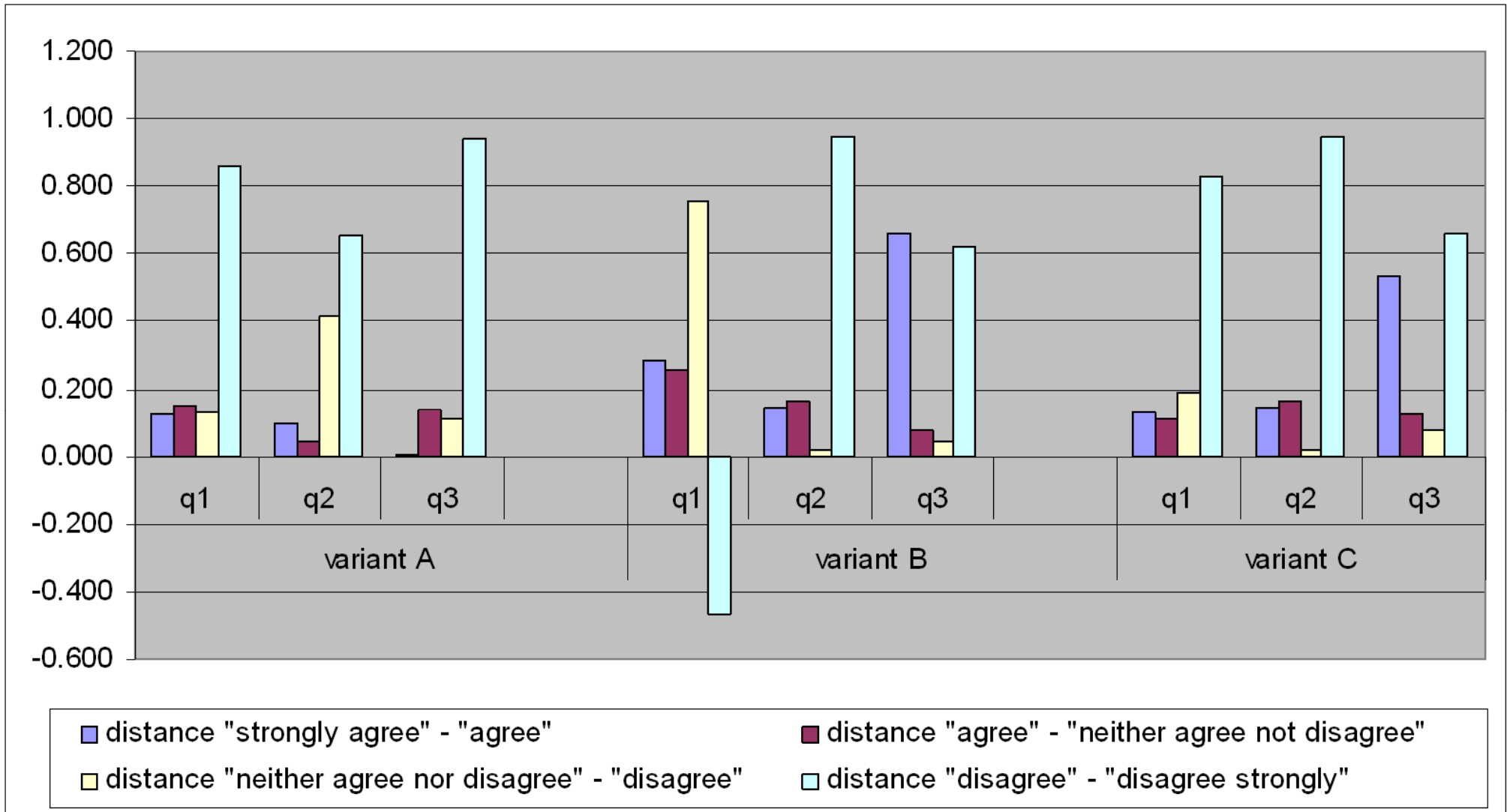


Thank you

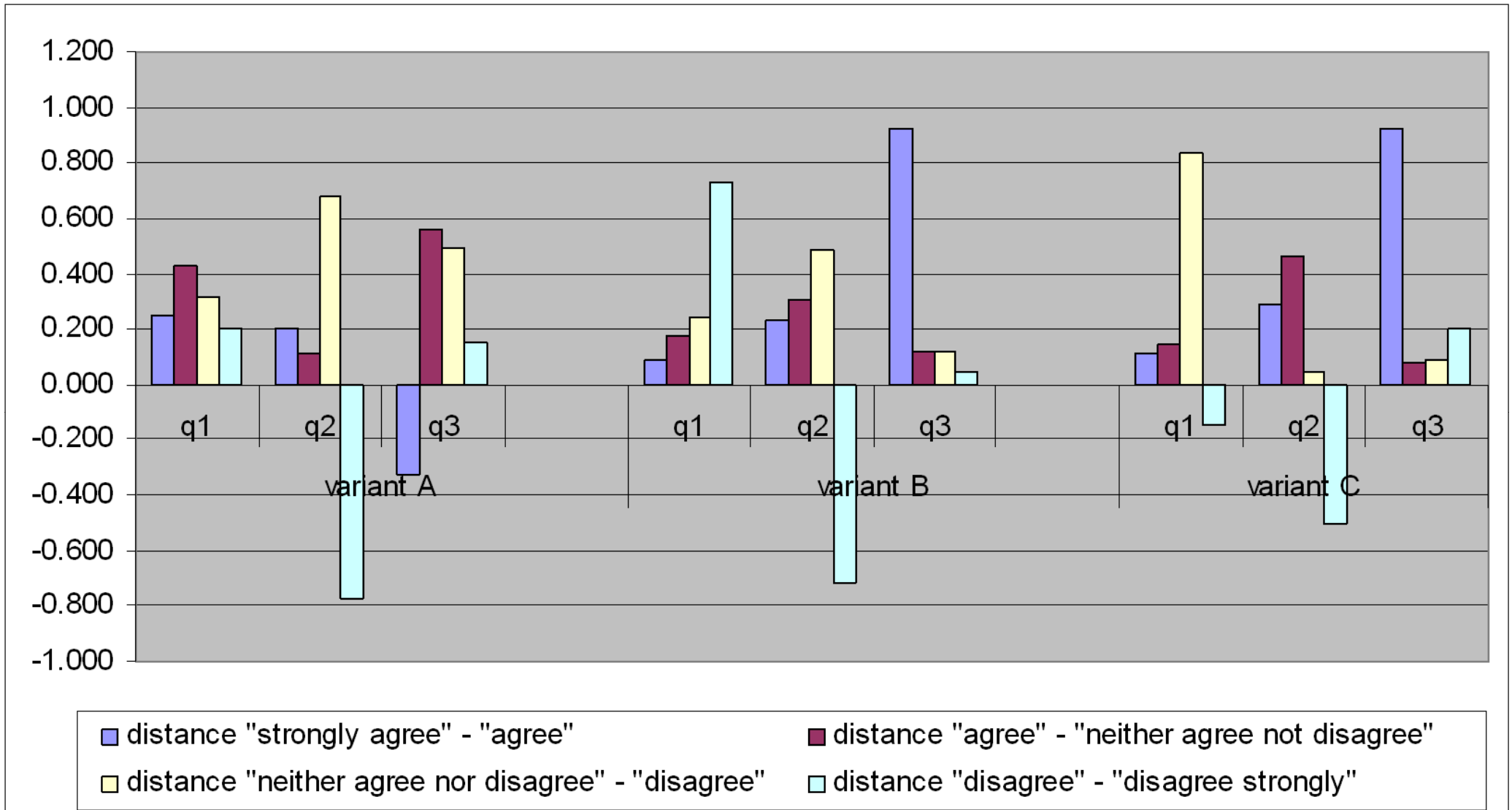
## Distances between the answer categories for the wording variants - **lifestyles freedom for homosexuals** (same question as instrument)

	German			French		
	Variant A	Variant B	Variant C	Variant A	Variant B	Variant C
<b>Agree strongly</b>	-0.24	-0.27	-0.28	-0.40	-0.46	-0.44
<b>Agree</b>	-0.03	-0.04	0.01	-0.30	-0.32	-0.37
<b>Neither agree nor disagree</b>	0.09	0.27	0.47	-0.26	-0.16	-0.28
<b>disagree</b>	0.77	0.75	0.52	0.15	-0.14	0.34
<b>disagree strongly</b>	<b>0.00</b>	<b>0.03</b>	<b>0.01</b>	0.81	0.81	0.69
<b>Weighted mean</b>	-0.03	-0.03	-0.03	-0.28	-0.28	-0.28
<b>Range of the scores</b>	1.01	1.02	0.80	1.21	1.27	1.13
<b>Distance between « agree strongly » and « agree »</b>	0.21	0.24	0.29	0.09	0.14	0.07
<b>Distance between « agree » and « neither agree nor disagree »</b>	0.12	0.31	0.46	0.04	0.17	0.10
	$\chi^2 = 550735.34^{***}$ , L2 = 112.03 <sup>***</sup> , BIC = -45.65, dl = 22			$\chi^2 = 38.09^{**}$ L2 = 37.22 <sup>**</sup> , BIC = -94.92, dl = 22		

**5-categories agreement scale: Distances found for the 3 French variants of the 3 items, in taking the similar questions as instrumental variables for the 3 items.**



**5-categories agreement scale: Distances found for the 3 German variants of the 3 items, in taking the similar questions as instrumental variables for the 3 items.**



**Satisfaction with life**

**Instrument:  
happiness**

	German			French		
	Variant A	Variant B	Variant C	Variant A	Variant B	Variant C
<b>Extremely unsatisfied</b>	0.25	0.02	0.01	0.07	-0.04	-0.04
<b>1</b>	-0.59	0.02	0.14	-0.81	-0.04	-0.04
<b>2</b>	0.02	-0.04	-0.82	0.00	-0.63	-0.39
<b>3</b>	0.00	0.18	-0.13	-0.06	0.06	-0.39
<b>4</b>	-0.59	-0.83	0.06	0.07	-0.63	-0.39
<b>5</b>	0.00	0.10	-0.09	-0.06	0.03	-0.39
<b>6</b>	0.00	-0.05	0.02	-0.01	-0.02	0.03
<b>7</b>	0.08	-0.02	0.04	-0.02	0.07	0.12
<b>8</b>	0.17	0.17	0.15	0.04	0.13	0.20
<b>9</b>	0.28	0.29	0.28	0.27	0.25	0.31
<b>Extremely satisfied</b>	0.37	0.39	0.42	0.51	0.36	0.51
<b>Weighted mean</b>	.19	.19	.19	.14	.14	.14
<b>Range of the scores</b>	0.95	1.22	1.24	1.32	.99	.89
<b>Range of score from “point 6” thru “point 10”</b>	0.37	0.44	0.40	.52	.38	.48
<b>Distance between “point 10” and “point 9”</b>	0.09	0.10	0.15	0.24	0.11	0.20
<b>Distance between “point 9” and “point 8”</b>	0.11	0.13	0.13	0.23	0.13	0.11
<b>Distance between “point 8” and “point 7”</b>	0.09	0.18	0.10	0.06	0.05	0.08
	$\chi^2=330.97^{***}$ , $L2 = 192.21^{***}$ , BIC = -4987.91, dl= 58			$\chi^2= 92.22^{***}$ ; $L2 = 66.80$ (p=.20), BIC = -281.85, dl=58		