HOW PEOPLE FROM THE U.S. AND CHINA PERCEIVE SELF-RATED HEALTH QUESTIONS ?

&

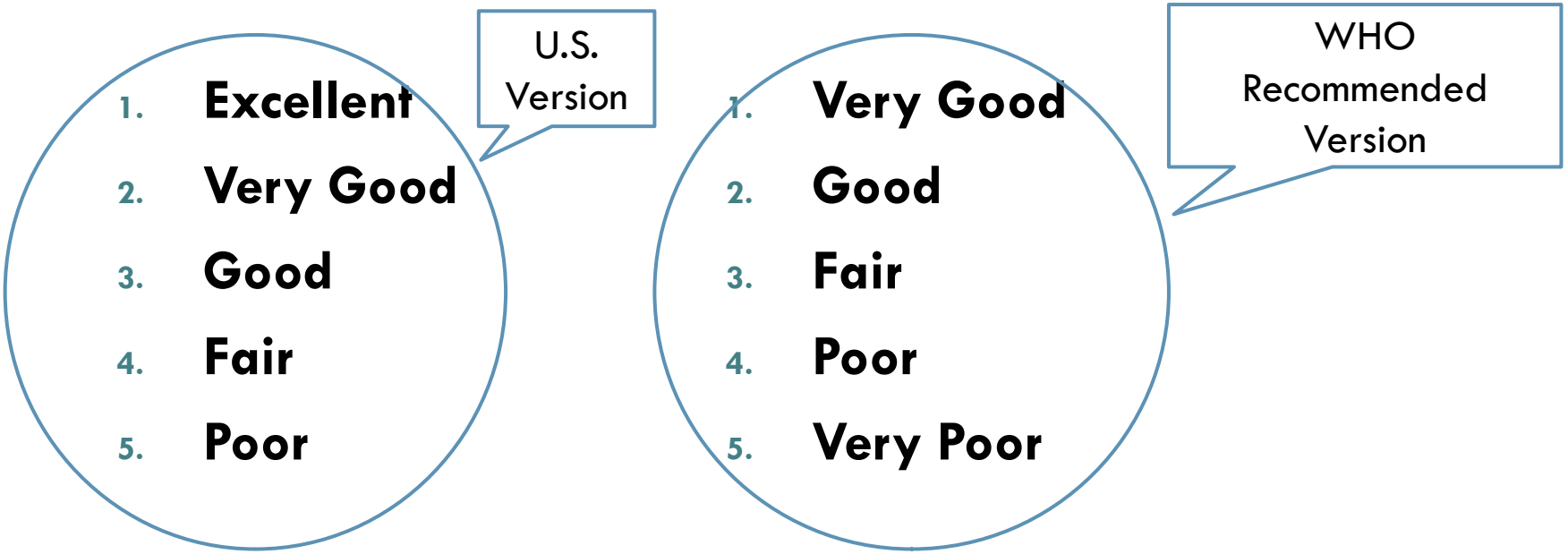
THE RESPONSE SCALE DESIGN AND QUESTION ORDER EFFECTS IN THE MEASURES OF SELF-RATED HEALTH STATUS IN CHINESE POPULATION

University of Michigan
March, 2013

MENGYAO (Maggie) HU

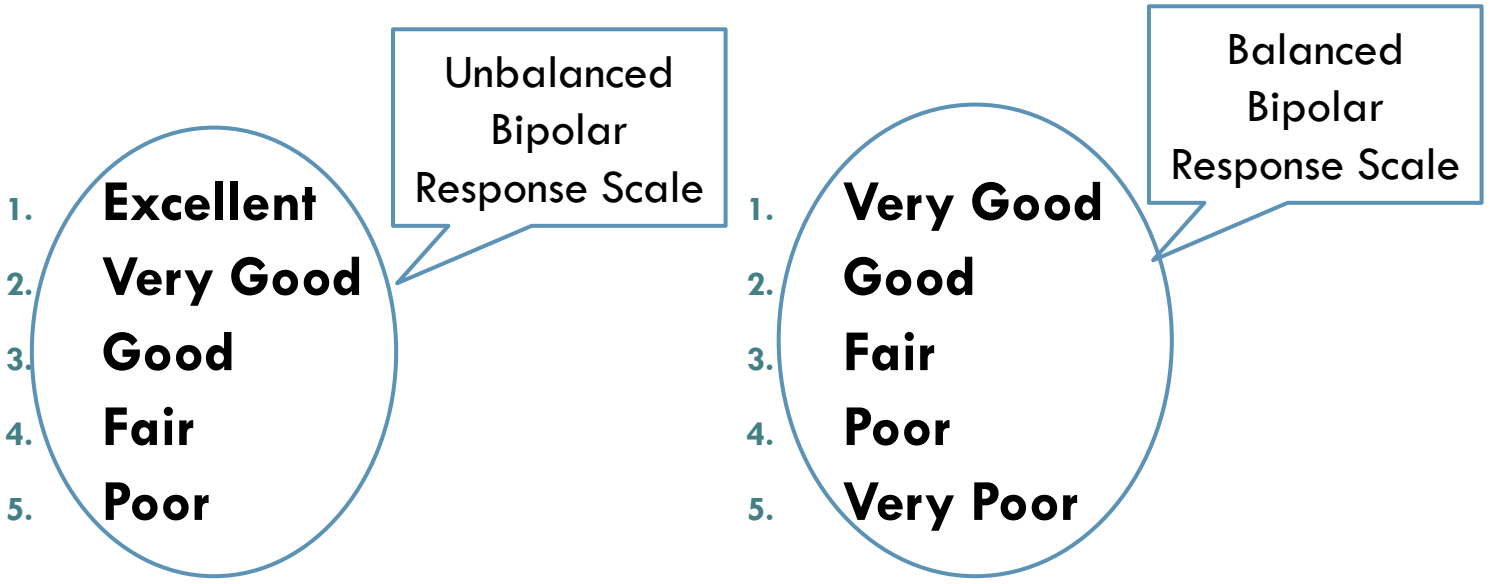
Introduction

- Why is Self-Rated Health Important?
  - An independent predictor of mortality
  - Most widely used comprehensive health measurement
- Different Versions of Response Scales for Self-Rated Health Questions



Background

- *Response Scale Designs*
  - Bipolar scales vs. unipolar scales
  - *Balanced vs. Unbalanced*

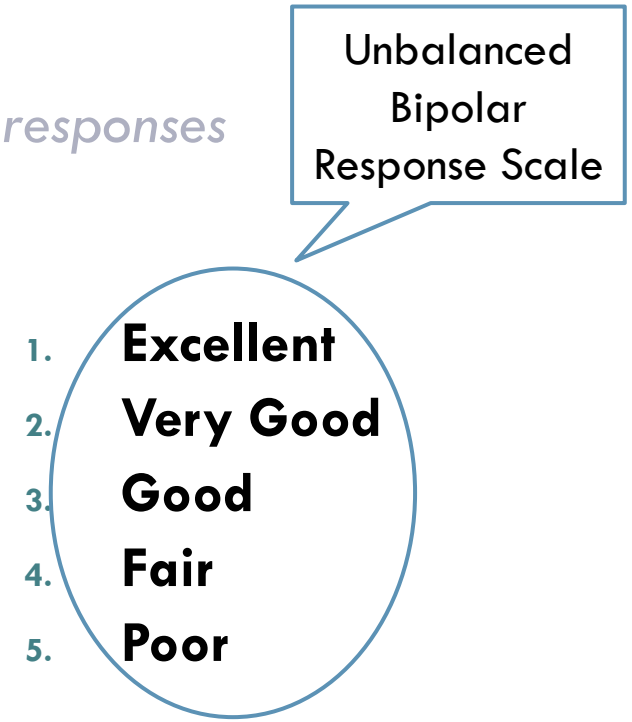


Background

- *Response Scale Designs*
  - Bipolar scales vs. unipolar scales
  - *Balanced vs. Unbalanced*
- *Question Order Effects*
  - *Impact of the order of questions upon survey responses*
    - *Beginning of the questionnaire / health section*
    - *End of the questionnaire / health section*

Background

- *Response Scale Designs*
  - Bipolar scales vs. unipolar scales
  - *Balanced vs. Unbalanced*
- *Question Order Effects*
  - *Impact of question orders upon survey responses*
- *Comparison across cultures*
  - Hofstede's Six Dimensions Model:
    - Individualistic (e.g. U.S.)
    - Collectivistic (e.g. China)



Research Questions

- In the Chinese population only:
  - Do *question order effects* have an influence on people's answers to self-reported health questions?
  - Do *different response scale designs* have an influence on people's answers to self-reported health questions?
- Is the reporting pattern for self-reported overall health for the U.S. population different from that for the Chinese population, even when holding the actual health status constant?

Data

- China Health and Retirement Longitudinal Study (CHARLS)
  - Baseline wave (2011)
  - Aged 45 years or above
  - Sample size: 12,652

Remove effects of respondents' familiarity with the instruments

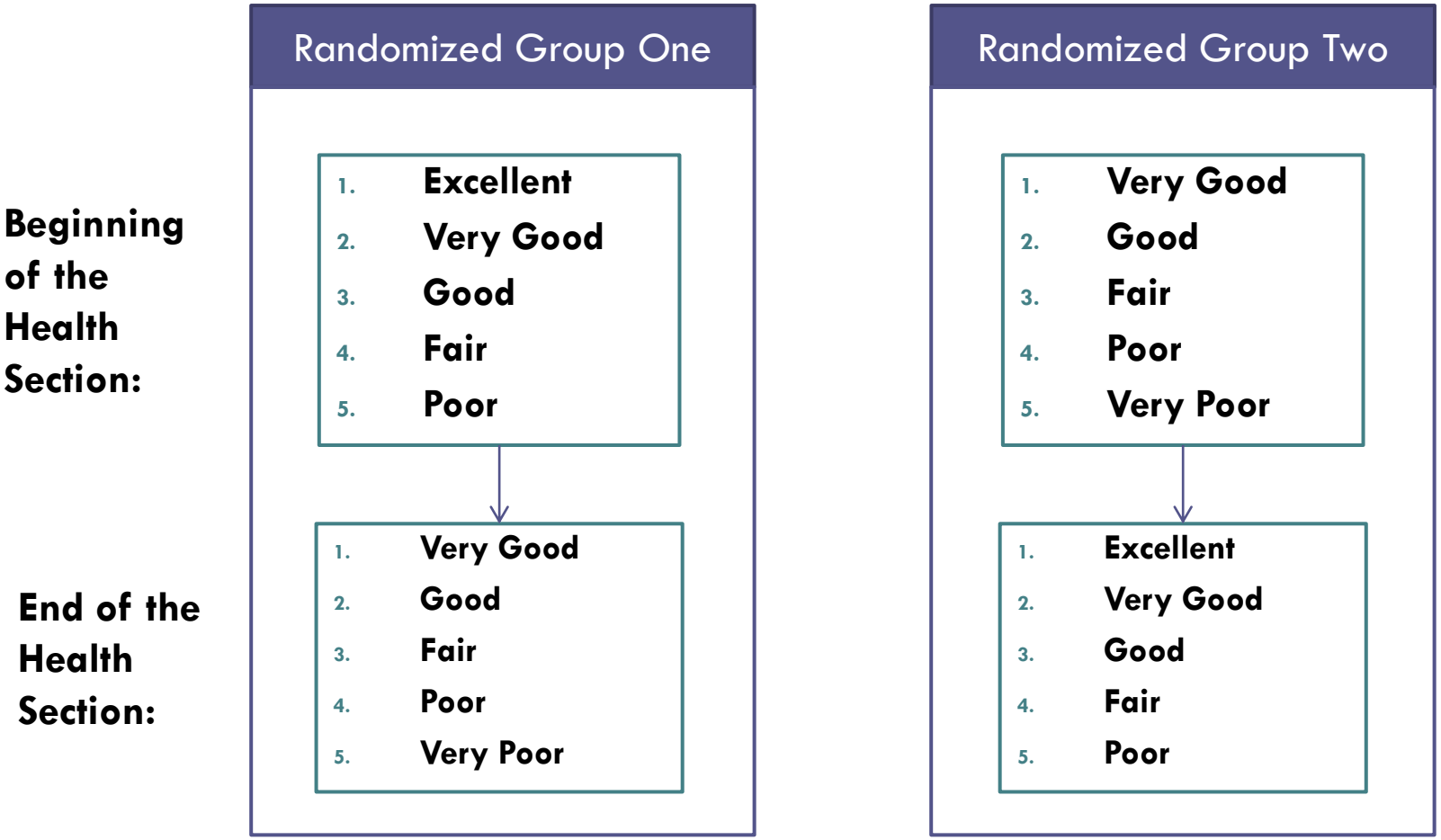
Remove the effects of age-differences with the very first cohort

- U.S. Health and Retirement Study (HRS)
  - Baseline wave (1992)
  - Over the age of 50
  - Sample size:17,587

One: Response Scale Design Effects

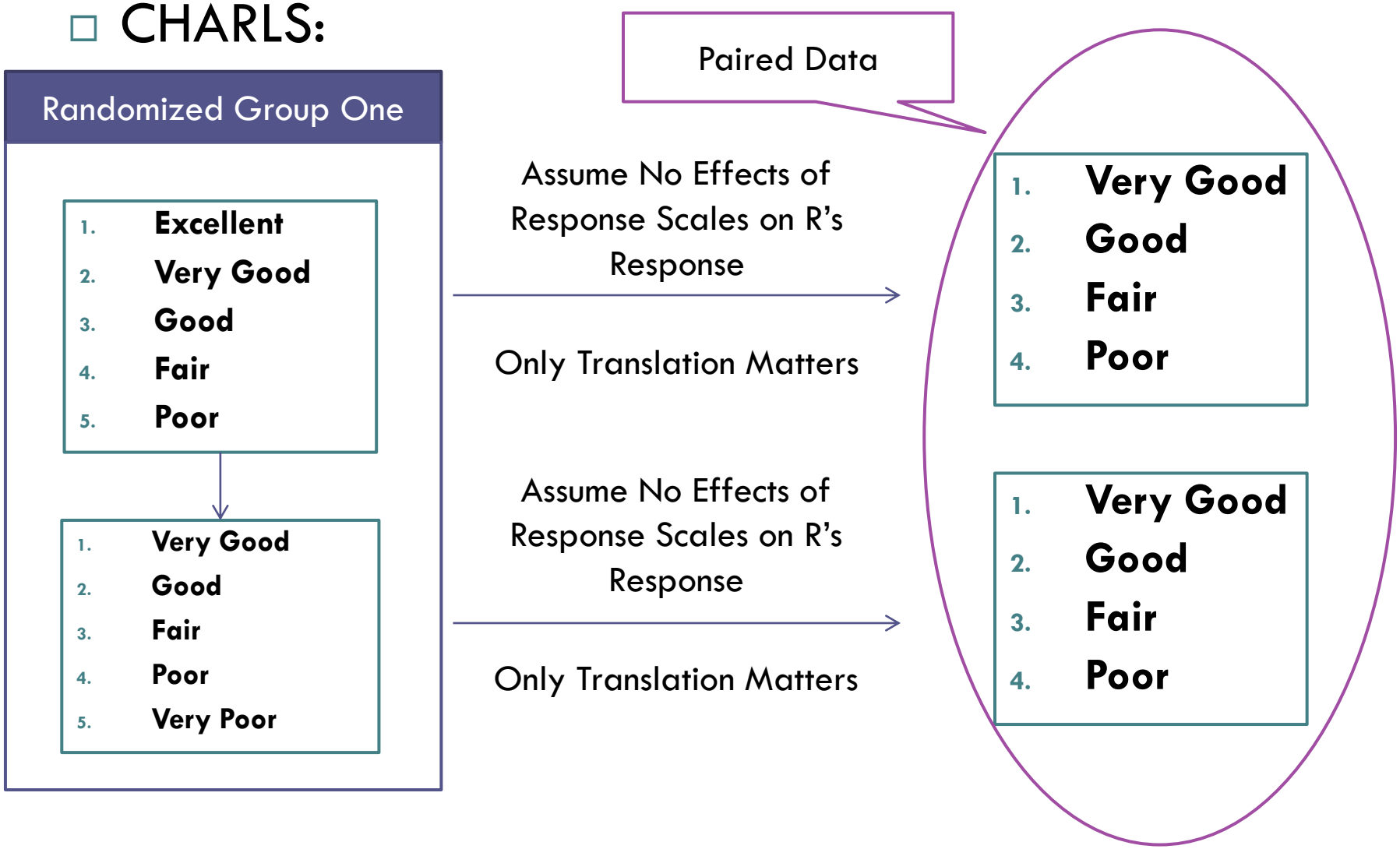
Design of CHARLS Study

- Respondents are randomized into two groups:



How do people answer this question? Do Response Scale Designs Matter?

- CHARLS:



How do people answer this question? Do Response Scale Designs Matter?

Table. Bivariate Distribution of Observed Sample Frequencies - "Very Good" to "Poor" Response Scales (Paired Test)

Self-Rated Overall Health (Beginning)	Self-Rated Overall Health (End)				
	Very Good	Good	Fair	Poor	Total
Very Good	371 (45.80%)	303 (37.41%)	126 (15.56%)	10 (1.23%)	810 (100.00%)
Good	110 (7.72%)	827 (58.04%)	460 (32.28%)	28 (1.96%)	1425 (100.00%)
Fair	80 (1.99%)	392 (9.75%)	3228 (80.30%)	320 (7.96%)	4020 (100.00%)
Poor	11 (0.45%)	69 (2.83%)	597 (24.51%)	1759 (72.21%)	2436 (100.00%)
Total	572 (6.58%)	1591 (18.31%)	4411 (50.75%)	2117 (24.36%)	21365 (100.00%)

Source: China Health and Retirement Longitudinal Study (CHARLS) 2011 & Health and Retirement Study (HRS) 1992

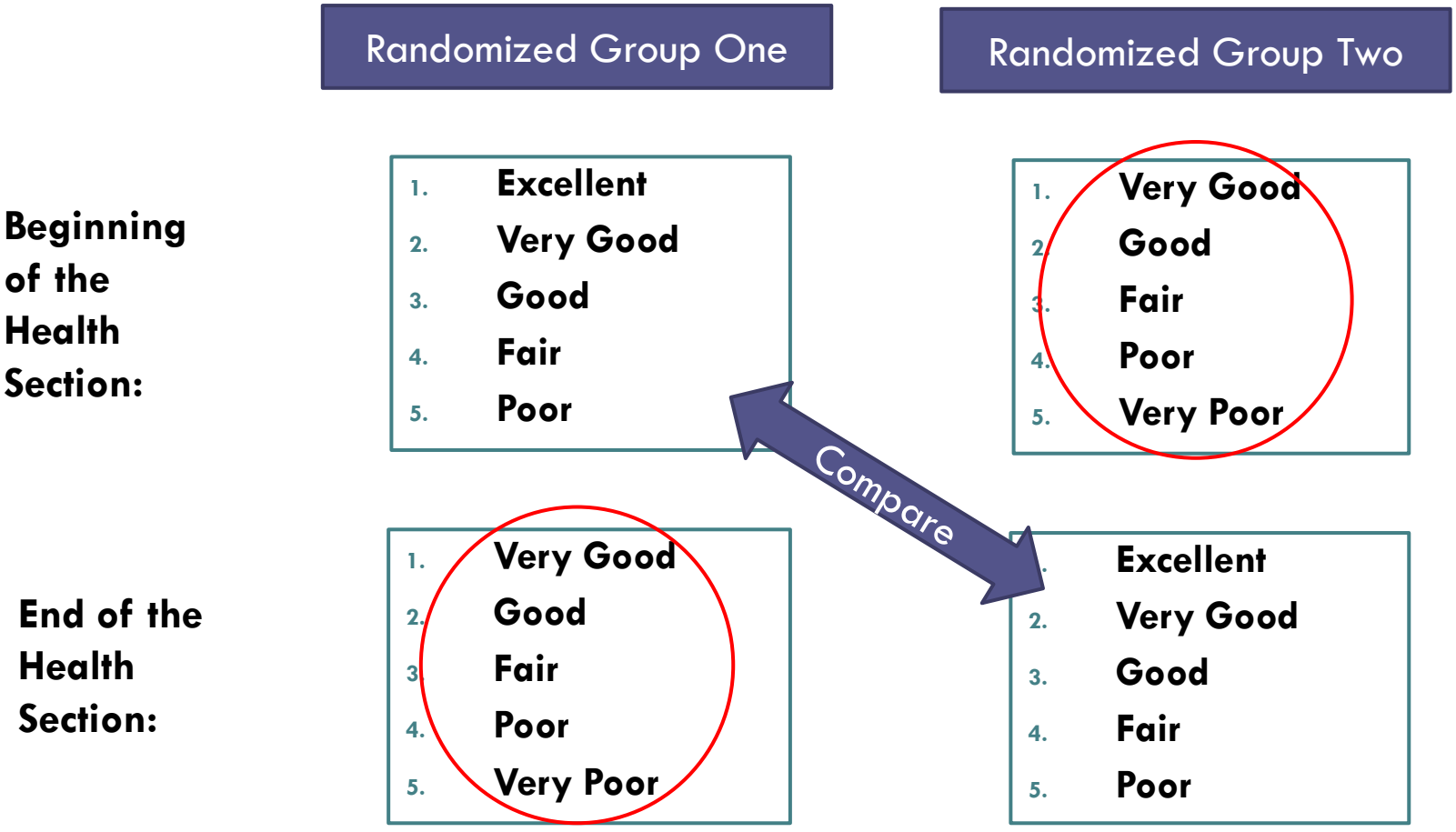
Weighted Kappa Coefficient = 0.58 ~ (0.56, 0.60): a moderate level of agreement

Bowker's test for internal symmetry: S (df=6)=206.95; P < 0.0001

Two: Question Order Effects

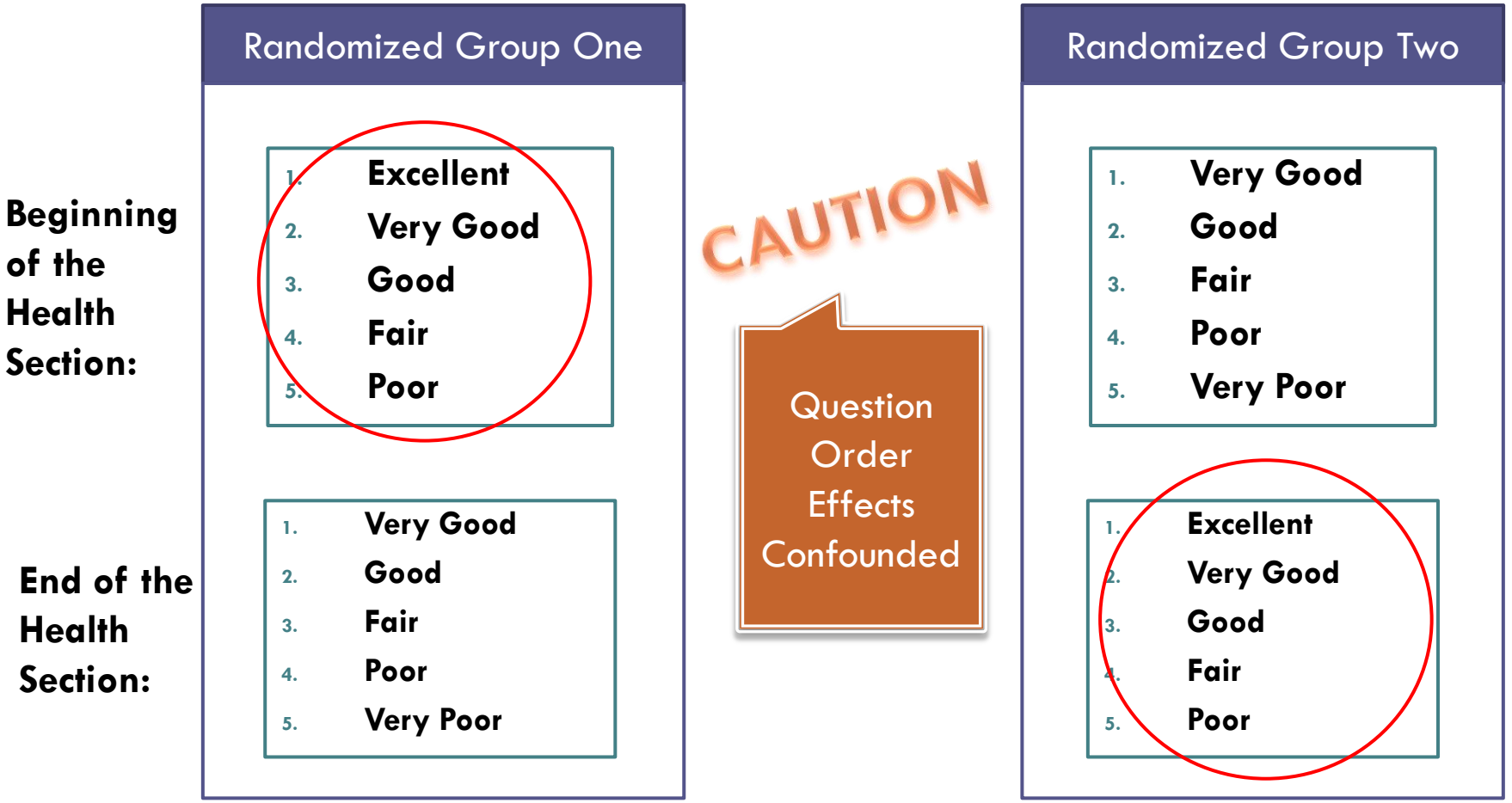
Do Question Orders Matter?

- Respondents are randomized into two groups:



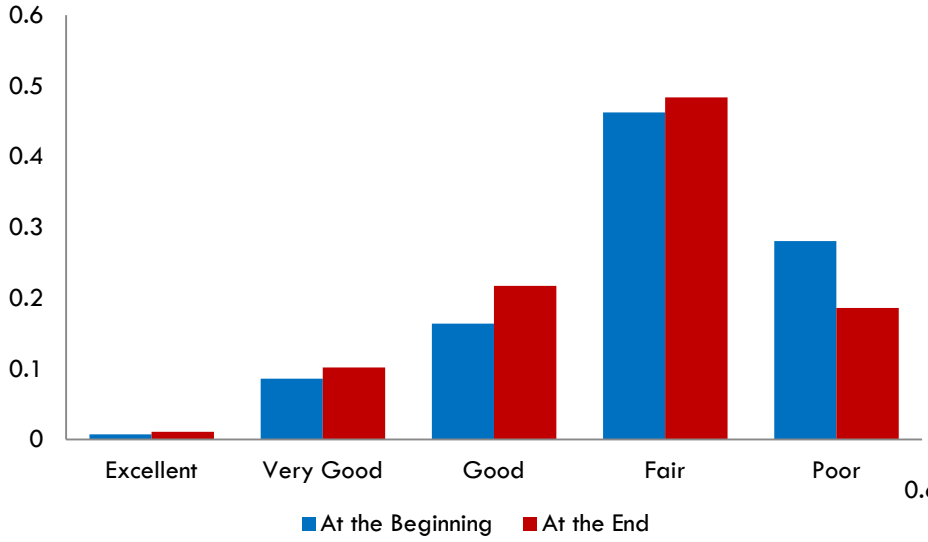
Design of CHARLS Study

- Respondents are randomized into two groups:



Do Question Orders Matter?

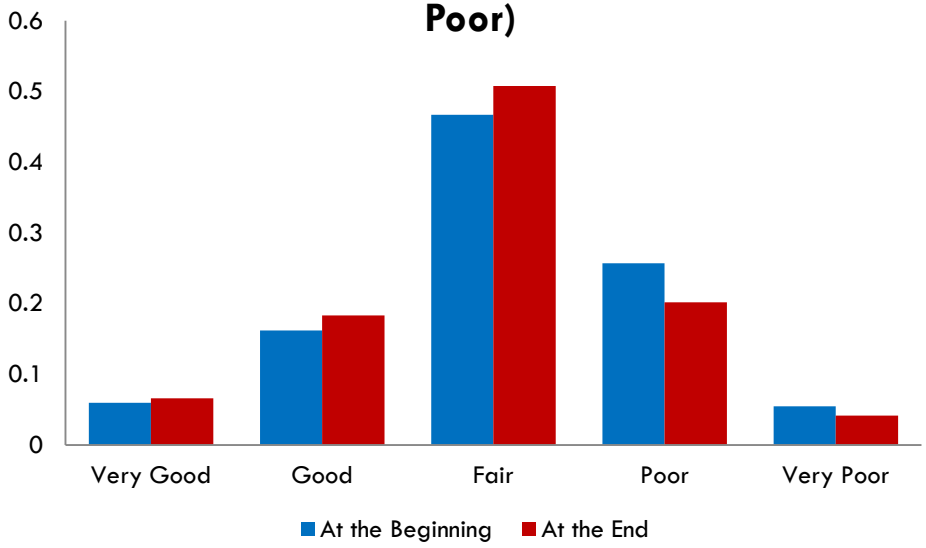
SRH by Question Order (Excellent - Poor)



$$\chi^2(df = 4) = 153.56\ (\mathrm{p} < 0.0001)$$

$$\chi^2(df = 4) = 101.75\ (\mathrm{p} < 0.0001)$$

SRH by Question Order (Very Good - Very Poor)



Question Orders Effects: Excellent - Poor

$\chi^2(df = 4) = 153.56\ (\mathrm{p} < 0.0001)$

Table. Cumulative Regression Models for SRH with the Scales Ranging from *Excellent* to *Poor*

Predictor	$\widehat{\beta}$	Cumulative Odds Ratio $\varphi_{y=<k:j}$	95% CI for φ
Question Order			
Beginning of the Health Section	0.46***	1.59	(1.41, 1.79)
Health Index (Disease)	0.41***	1.51	(1.45, 1.58)
Health Index (Body Function)	0.14***	1.16	(1.15, 1.17)
Emotional Health			
Unhealthy	0.71**	2.03	(1.35, 3.06)
Age			
65 and over	-0.40***	0.67	(0.60, 0.76)
Education			
(Below High School)			
High School Graduates	-0.11*	0.89	(0.80, 1.00)
College or Above	-0.45**	0.64	(0.49, 0.83)

Source: China Health and Retirement Longitudinal Study (CHARLS) 2011.

Notes: N= 11530. Wald chi-square statistic (df) = 1478.66 (7); -2 Log L = 793121480; Weighted to national level;

Question Orders Effects: Very Good - Very Poor

$\chi^2(df = 4) = 101.75\ (\mathrm{p} < 0.0001)$

Table. Cumulative Regression Models for SRH with the Scales from *Very Good* to *Very Poor*

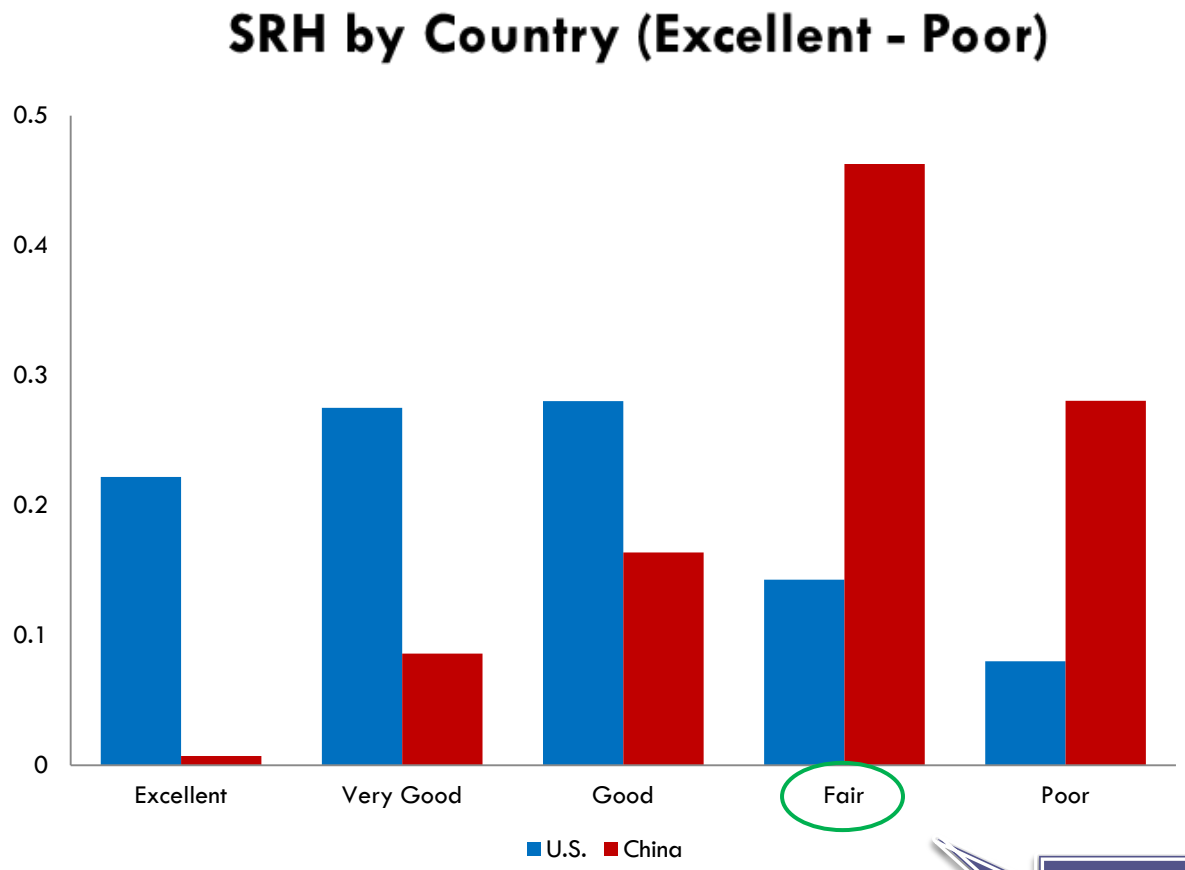
Predictor	$\widehat{\beta}$	Cumulative Odds Ratio $\varphi_{y=<k:j}$	95% CI for φ
Question Order			
Beginning of the Health Section	0.23***	1.25	(1.15, 1.37)
Health Index (Disease)	0.39***	1.48	(1.44, 1.53)
Health Index (Body Function)	0.15***	1.16	(1.15, 1.17)
Emotional Health			
Unhealthy	0.56**	1.75	(1.18, 2.60)
Age			
65 and over	-0.31***	0.74	(0.67, 0.81)
Education			
(Below High School)			
High School Graduates	-0.10*	0.90	(0.83, 0.99)
College or Above	-0.41***	0.66	(0.54, 0.81)

Source: China Health and Retirement Longitudinal Study (CHARLS) 2011.

Notes: N= 11530. Wald chi-square statistic (df) = 1478.66 (7); -2 Log L = 793121480; Weighted to national level;

Three: U.S. & China Comparisons

U.S. & China Comparisons



"Fair" was translated as "一般", which means similar as others

↓

"Fair" in Chinese questionnaire seem to be a better neural category.

People from collectivistic cultures are mote likely to "Fitting in"

$$\chi^2(df = 4) = 6215.96\ (p < 0.0001)$$

U.S. & China Comparisons – Cumulative Logistic Model

Table. Cumulative Logistic Regression Model Results for Self-Rated Overall Health (with the Response Scales Ranging from *Excellent* to *Poor)*

Predictor	$\hat{\beta}$	Cumulative Odds Ratio $\varphi_{y=<k:j}$	95% CI for φ
Intercept			
Cut 5	-4.61***		
Cut 4	-2.30***		
Cut 3	-0.73***		
Cut 2	0.59***		
Country:			
U.S. Population	-0.14*	0.87	(0.78, 0.98)
Health Index – Disease	0.58***	1.79	(1.75, 1.82)
Health Index – Body Function	0.15***	1.16	(1.15, 1.17)
Emotional Health - Unhealthy	0.99***	2.69	(2.41, 3.01)
Age: > = 65 years old	-0,23***	0.79	(0.73, 0.87)
Education:			
(Below High School)			
High School Graduates	-0.31*	0.74	(0.68, 0.80)
College or Above	-1.07***	0.34	(0.31, 0.38)

Source: CHARLS 2011 and HRS 1992; Notes: * - significant at 95% confidence level; ** - significant at 99% confidence level; *** - significant at 99.9% confidence level. [a] Reference categories in parentheses.

Conclusions

- U.S. population are more likely to select the highly-optimistic choices about their health than Chinese population with similar health status
- Response scale designs influence people's response to self-rated health questions
- Question order – at the beginning of the health section or at the end of the health section – have affected responses to self-rated health questions
- For Chinese population, "Very Good" to "Very Poor" response scales provide better differentiation at the negative end

Future Research

- Study Design:
  - Randomize respondents into four groups to prevent confounded question order effects

	Group 1	Group 2	Group 3	Group 4
Beginning of the Health Section:	**Excellent** **Very Good** **Good** **Fair** **Poor**		**Very Good** **Good** **Fair** **Poor** **Very Poor**	
End of the Health Section:		**Excellent** **Very Good** **Good** **Fair** **Poor**		**Very Good** **Good** **Fair** **Poor** **Very Poor**

Future Research

- Study Design:
  - Randomize respondents into four groups to prevent confounded question order effects
  - Or, follow-up in longitudinal study and ask respondents SRH using the same response scales with a different question order (Paired Tests)

Future Research

- Study Design:
- Analysis:
  - Use Body Examination Data such as Blood Test Results as Control Variables

Future Research

- Study Design:
- Analysis:
  - Use Body Examination Data such as Blood Test Results as Control Variables
  - Compare HRS “Excellent - Poor” Scale with CHARLS “Excellent - Poor” Scale with different question orders (Suggestions from Sunghee Lee)
    - HRS vs. CHARLS (Beginning)
    - HRS vs. CHARLS (End)

Future Research

- Study Design:
- Analysis:
  - Use Body Examination Data such as Blood Test Results as Control Variables
  - Compare HRS "Excellent - Poor" Scale with CHARLS "Excellent - Poor" Scale with different question orders
  - Analyze different translations of the response scales in Chinese

Future Research

- Study Design:
- Analysis:
  - Use Body Examination Data such as Blood Test Results as Control Variables
  - Compare HRS “Excellent - Poor” Scale with CHARLS “Excellent - Poor” Scale with different question orders
  - Analyze different translations of the response scales in Chinese
  - Compare CHARLS with SHARE

Future Research

- Study Design:
- Analysis:
  - Use Body Examination Data such as Blood Test Results as Control Variables
  - Compare HRS “Excellent - Poor” Scale with CHARLS “Excellent - Poor” Scale with different question orders
  - Analyze different translations of the response scales in Chinese
  - Compare CHARLS with SHARE
- Translate “Fair” as “尚可” instead of “一般”

Thank you ☺

maggiehu@umich.edu