



Evidence-Based Survey Design Principles

Presented at:

- ISPI Atlanta Chapter
- Wednesday, March 31, 2021
- Noon - 1 pm (ET)

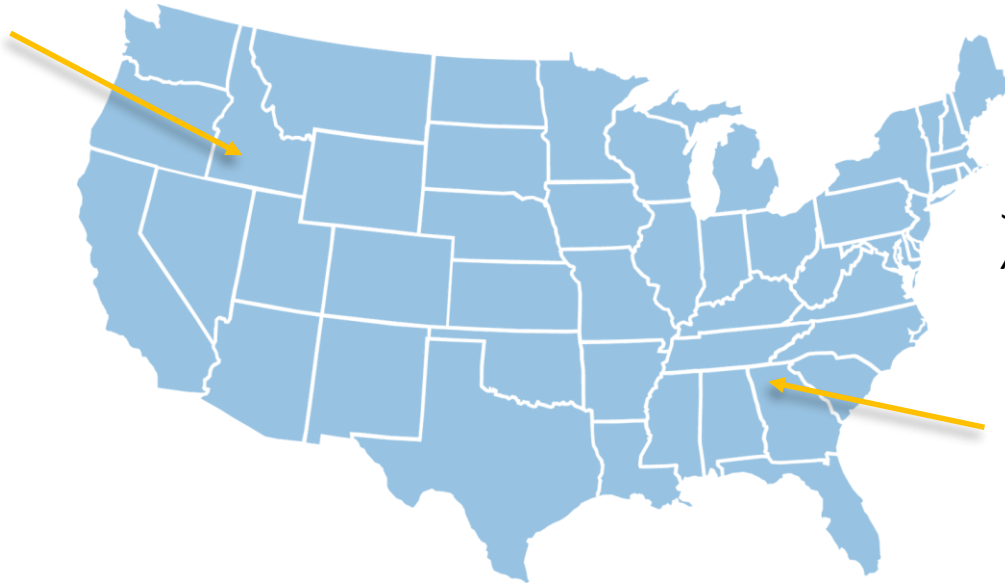
Presented by:

- Yonnie Chyung, Ed.D.
- Jennifer Shamsy, M.S.
- Boise State University
- <https://www.boisestate.edu/opwl>



Yonnie Chyung, Ed.D.
Professor, Chair

*Department of Organizational
Performance and Workplace Learning,
Boise State University, Boise Idaho*



Jennifer Shamsy, M.S.
Adjunct Faculty



Evidence-Based Survey Design Principles

1

Include a midpoint or not?



2

Put Strongly Agree on the left or right side of the scale?



3

Use radio buttons or sliders?



4

Use positive statements only or negative ones too?



5

How to handle ceiling effects?





Our Topics

1

Various
survey
response
scales

2

Include or
exclude a
midpoint in the
Likert scale?

3

Use ascending
or descending
order of Likert
scale response
options?

4

Extreme Makeover
- Survey Edition

Closed-ended survey items with appropriate response scales



Anatomy of Closed-Ended Survey Items



1. Question

Response scale

2. Question

Response scale

3. Question

Response scale

Survey items for:

- Capturing facts?
- Measuring perceived quality/value?





Survey items for:

- Capturing facts?
- Measuring perceived quality/value?

A

How many times have you used this device?

- Never 1-2 times 3-5 times More than 5 times

B

How useful was the job aid provided to you?

- Not at all A little bit Somewhat Quite a bit Very much



Survey items to capture facts



Think about
the census
(2020census.gov)

When did you complete the program?

- Fall 2018
- Spring 2019
- Summer 2019
- Fall 2019
- Spring 2020

Did you use the help desk to get assistance?

- Yes
- No



Survey items to **measure perceived quality/value**

How would rate the quality of the workshop?

- Excellent
- Very Good
- Good
- Fair
- Mediocre
- Poor

The content presentation was interesting.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree



Statements vs. Questions

S1. I have a good relationship with my supervisor.

- Strongly disagree
- Disagree
- Neutral
- Agree
- Strongly agree

Q1. Do you have a good relationship with your supervisor?

- Not at all
- A little bit
- Somewhat
- Pretty much
- Very much



Statements vs. Questions

S1. I have a good relationship with my supervisor.

Agreement

- Strongly disagree Disagree Neutral Agree Strongly agree

Q1. Do you have a good relationship with your supervisor?

Degrees

- Not at all A little bit Somewhat Pretty much Very much



Various rating scales

Discrete rating scales

Continuous rating scales

Numerical rating scales

Verbal descriptor scales



Discrete rating scales



Strongly disagree



Disagree



Neutral



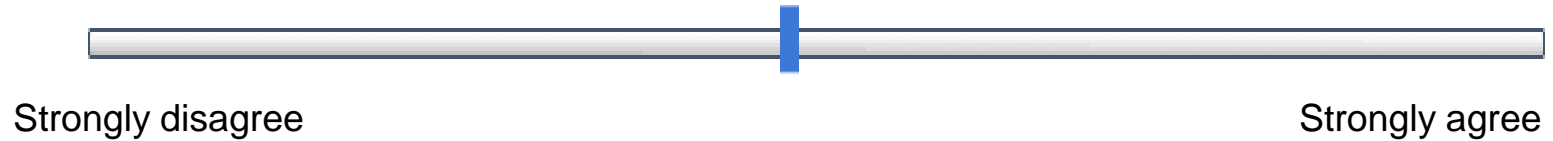
Agree



Strongly agree

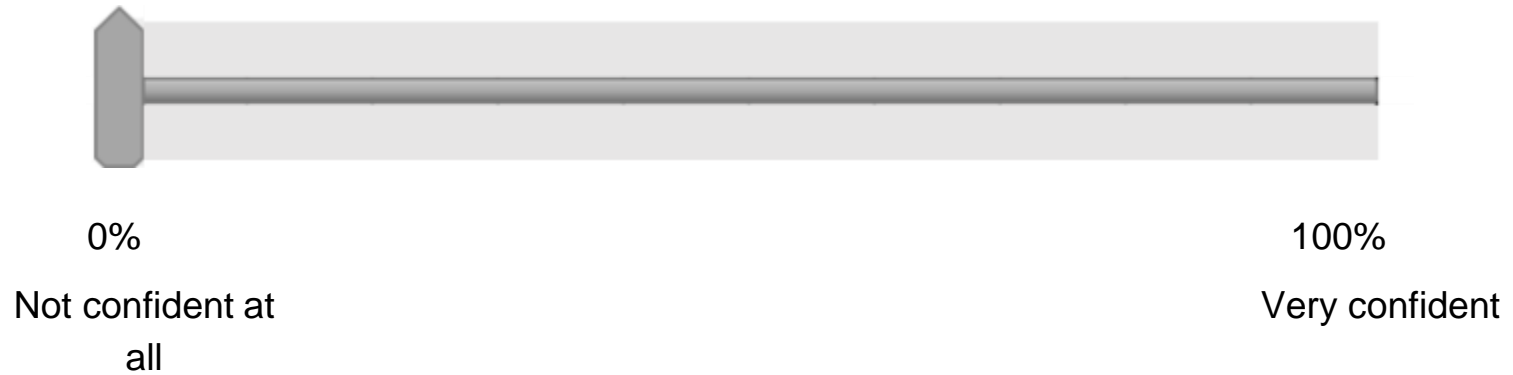


Continuous rating scales



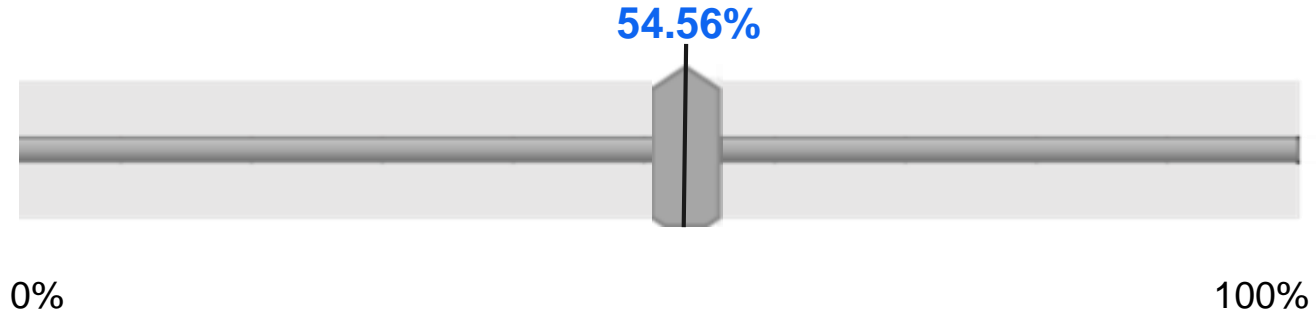


Web-based sliders



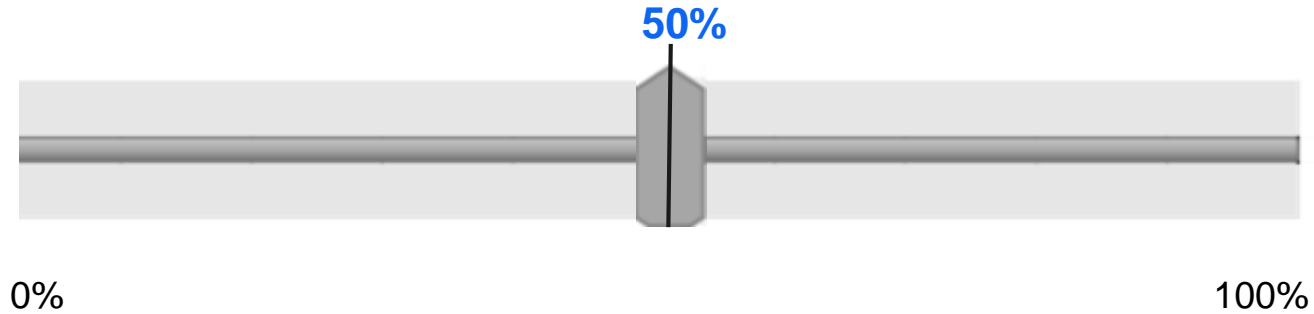


Web-based sliders - continuous scale





Web-based sliders - discrete scale



An 11-point scale?



Numerical rating scales

○

1

Very dissatisfied

○

2

○

3

○

4

○

5

Very satisfied



Verbal descriptor scales



Not at all



A little bit



Somewhat



Pretty much



Very much



Verbal descriptor scales



Strongly disagree



Disagree




Neutral



Agree



Strongly agree



If you need to present an average score of multiple survey items

- Use the same response scale
- Use a response scale that likely produces interval data
 - Numeric descriptor scale
 - Verbal descriptor like 5 pt Likert scale



Various rating scales

Discrete rating scales

Continuous rating scales

Numerical rating scales

Verbal descriptor scales

**Include or exclude a midpoint in
the Likert scale?**

What is a midpoint supposed to mean?

**Strongly
Disagree**

Disagree



Agree

**Strongly
Agree**

Research by Kulas & Stachowski (2009)

- 82 adults (college students)
- Midpoint = It depends? Uncertain? Average? Not applicable?

**Strongly
Disagree**

Disagree



Agree

**Strongly
Agree**

Research by Kulas & Stachowski (2009)

- 82 adults (college students)
- Midpoint = It depends? Uncertain? Average? Not applicable?

**Strongly
Disagree**

Disagree



Agree

**Strongly
Agree**

It depends



Research by Nadler, Weston, & Voyles (2015)

686 adults (college students)

*When asked a question on the scale below,
how do you interpret **the middle response**?*



Strongly
disagree



Disagree



Neither

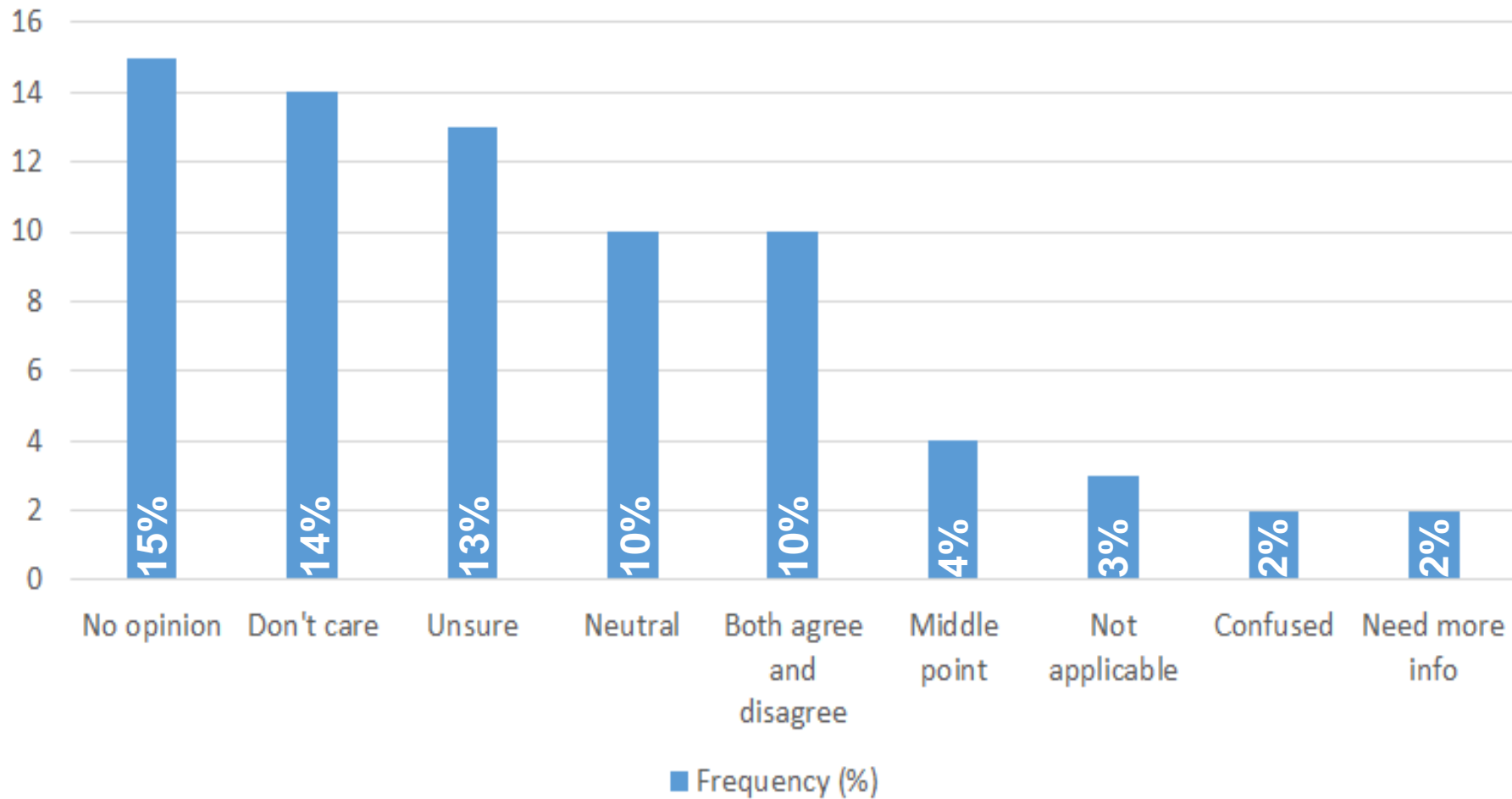


Agree



Strongly agree

Answer _____





Midpoint as a dumping ground

- So, should you exclude a midpoint?



Research by Johns (2005)

3,881 respondents (UK)

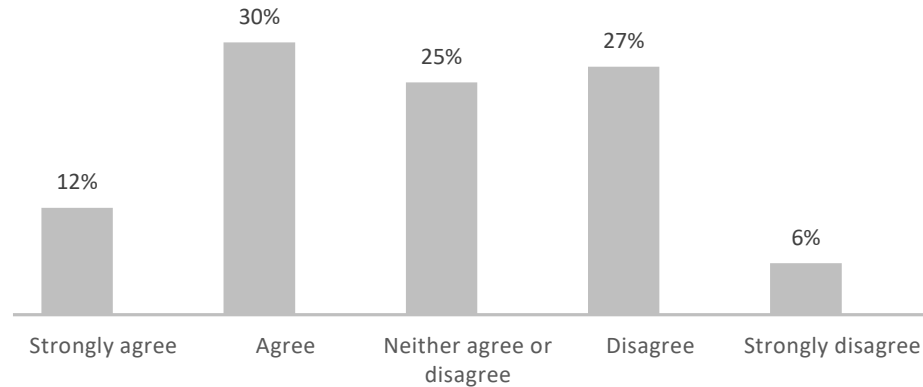
Likert items measured economic left-right attitudes

Four versions (modes) using 4-point and 5-point Likert scales with and without a “No opinion” option

Mode A (n=960)	Mode B (n=1005)	Mode C (n=923)	Mode D (n=993)
1. Strongly Agree	1. Strongly Agree	1. Strongly Agree	1. Strongly Agree
2. Agree	2. Agree	2. Agree	2. Agree
3. Neither agree nor disagree		3. Neither agree nor disagree	
4. Disagree	3. Disagree	4. Disagree	3. Disagree
5. Strongly Disagree	4. Strongly Disagree	5. Strongly Disagree	4. Strongly Disagree
No Opinion	No Opinion		

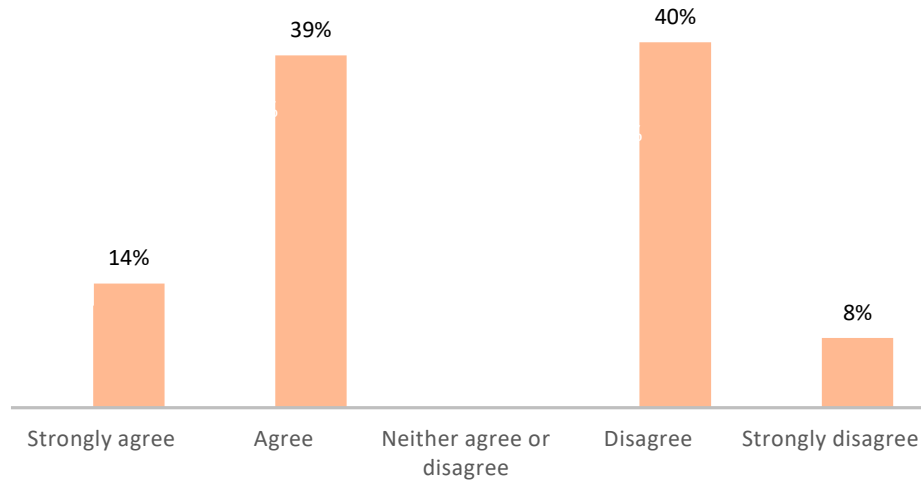


5-point Likert scale (with “Neither”) – Mode C



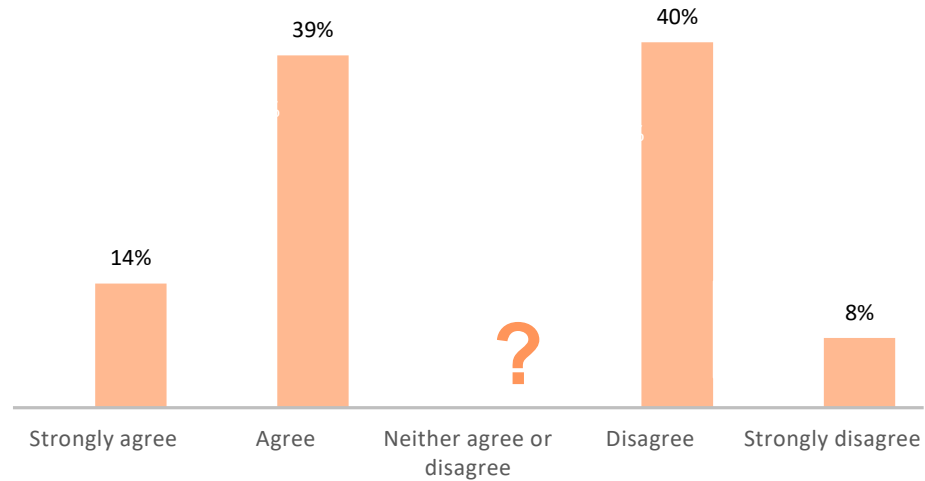


4-point Likert scale (without “Neither”) – Mode D



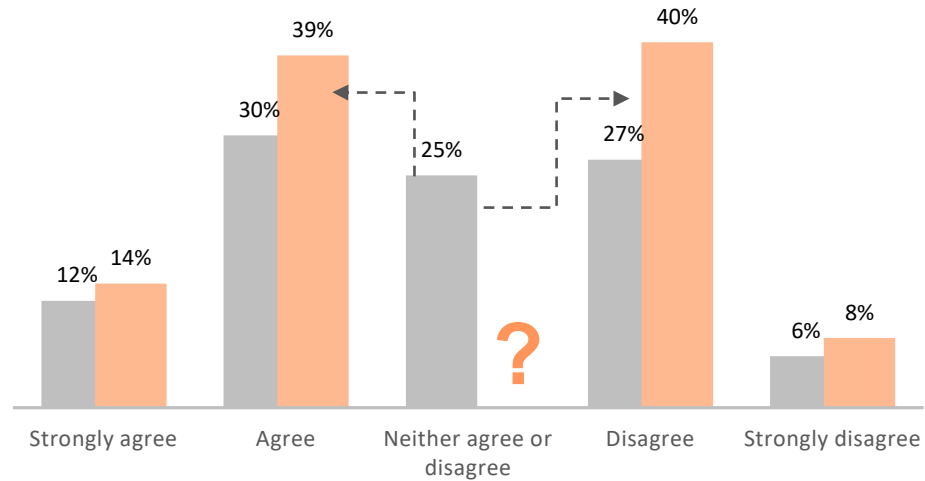


Where would “Neither” go?



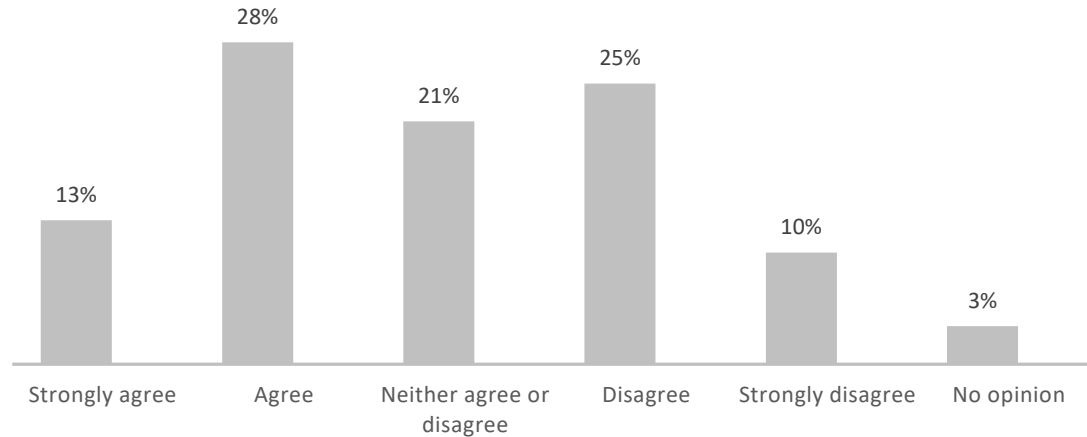


Where did “Neither” go?



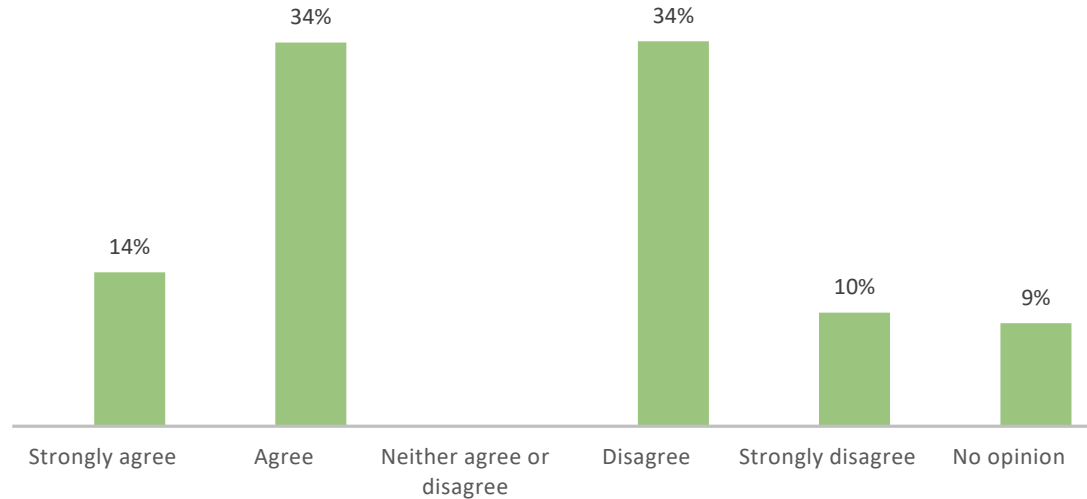


5-point Likert scale + “No opinion” – Mode A



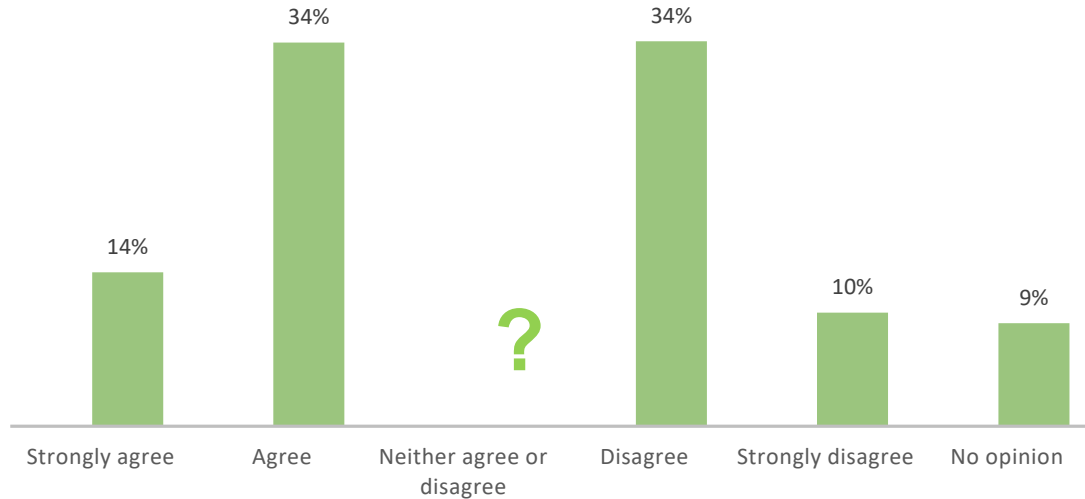


4-point Likert scale + “No opinion” – Mode B



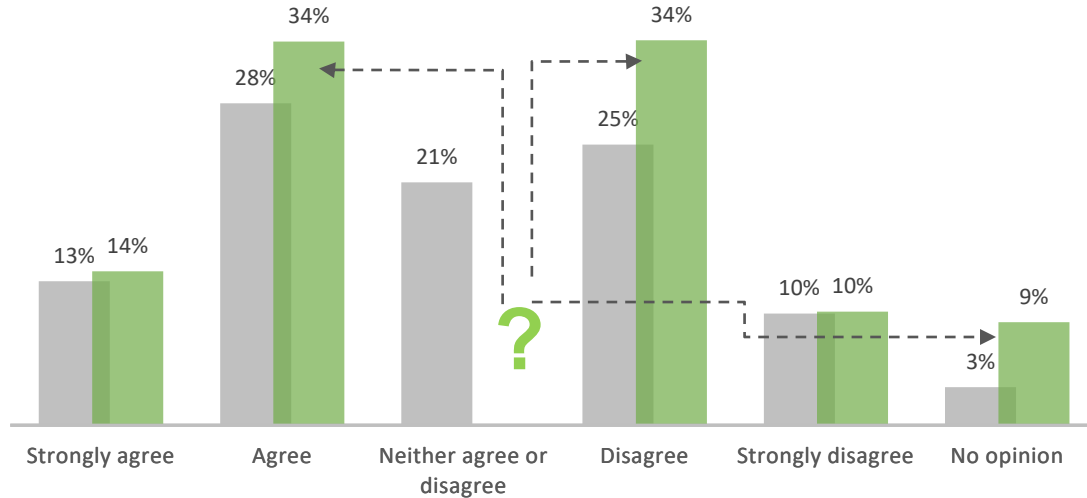


Where would “Neither” go?



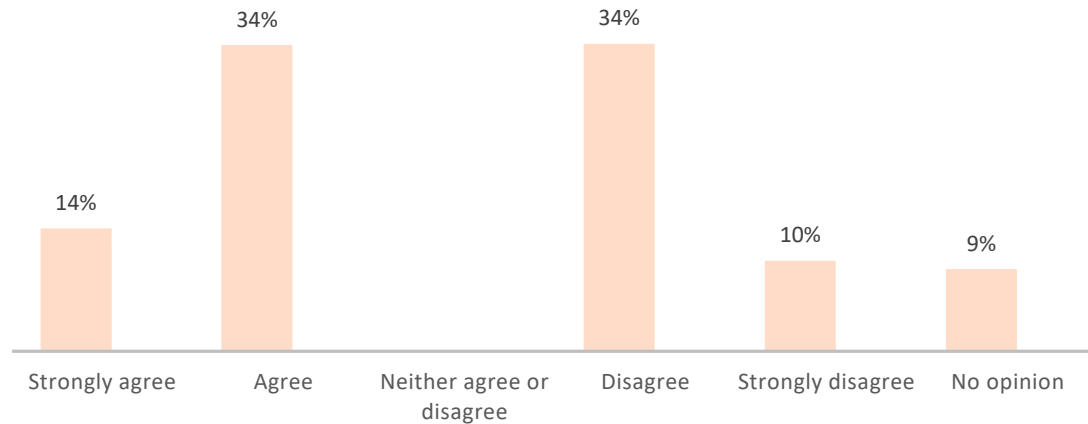


Where would “Neither” go?



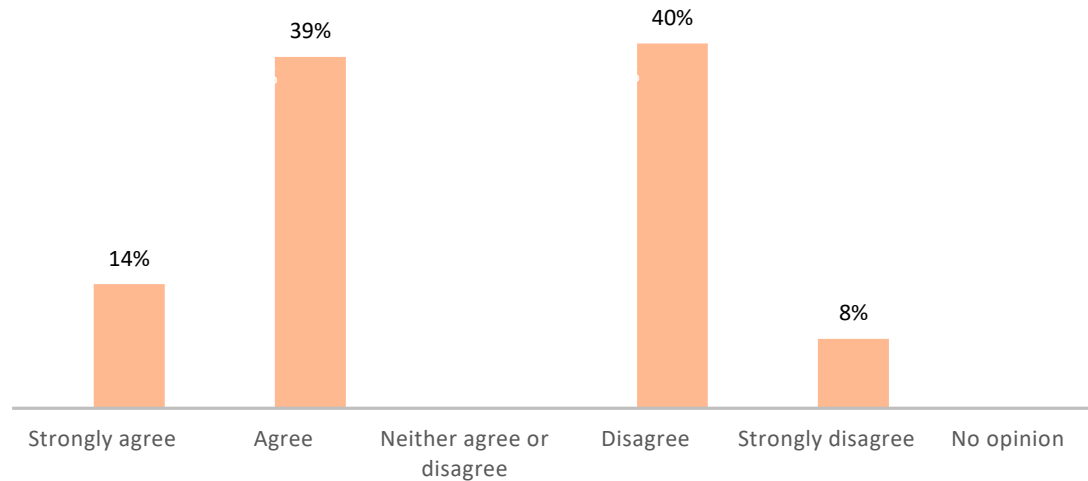


4-point Likert scale (with “No opinion”) – Mode B



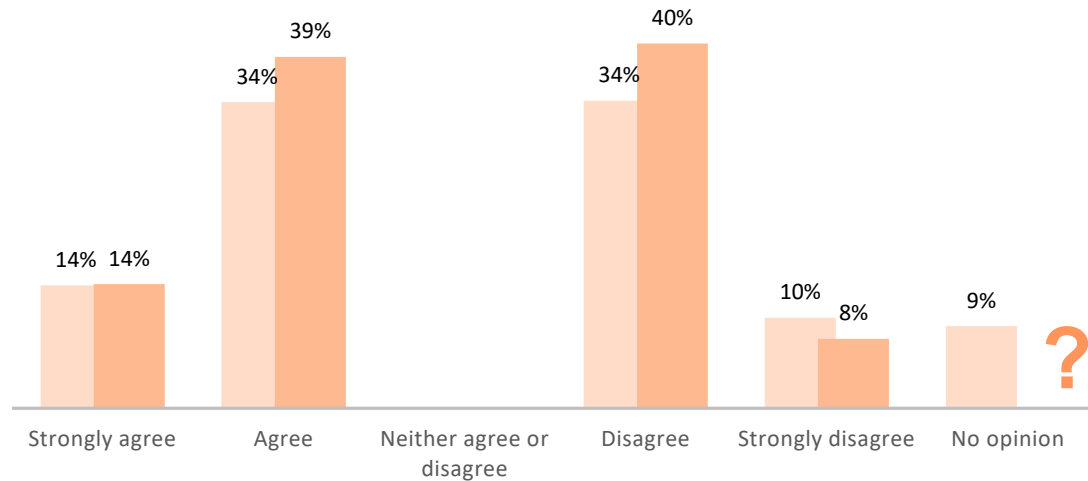


4-point Likert scale (without “No opinion”) – Mode D

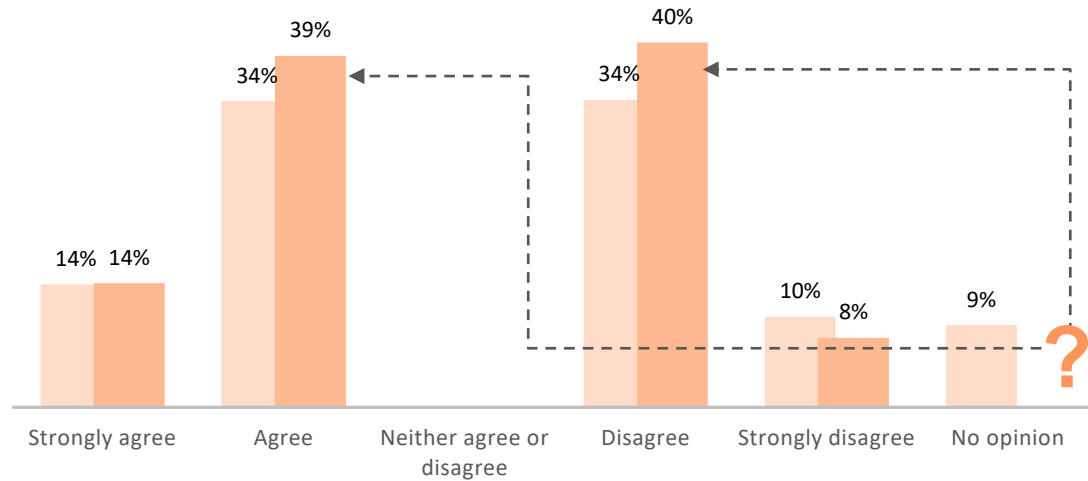




Where would “No opinion” go?

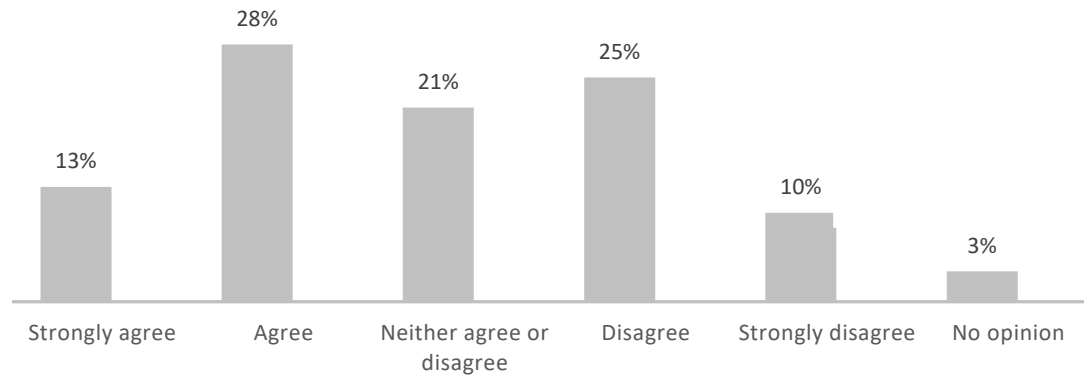


Where would “No opinion” go?



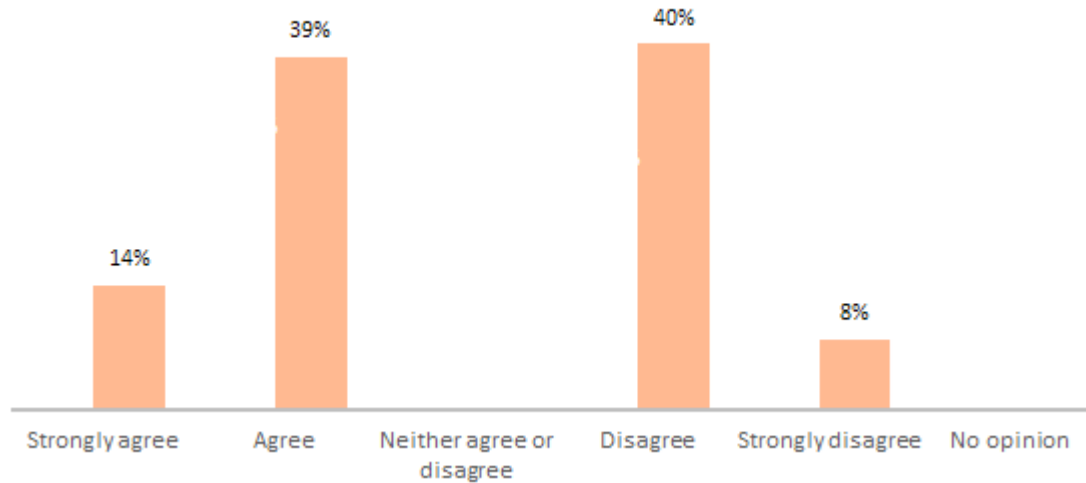


5-point Likert scale + “No opinion” – Mode A



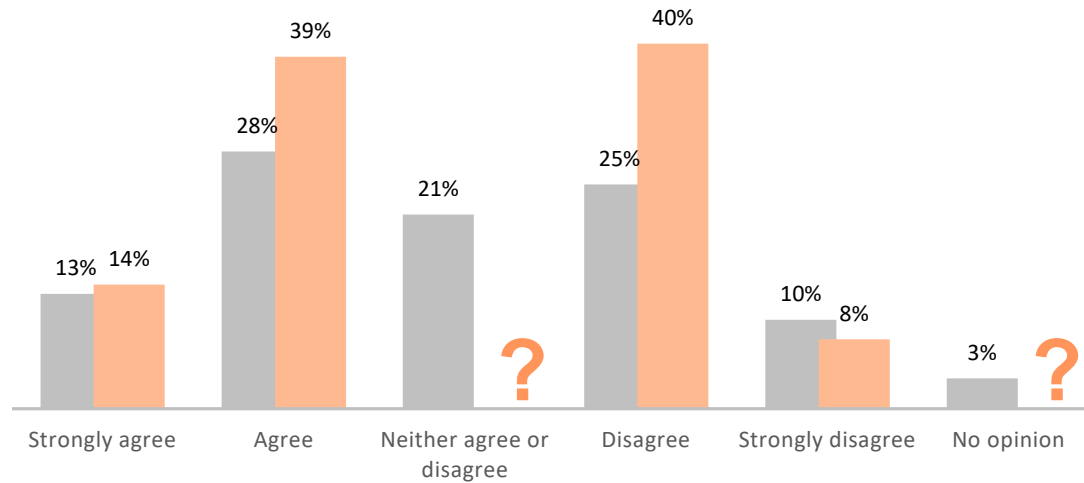


4-point Likert scale (without “No opinion”) – Mode D

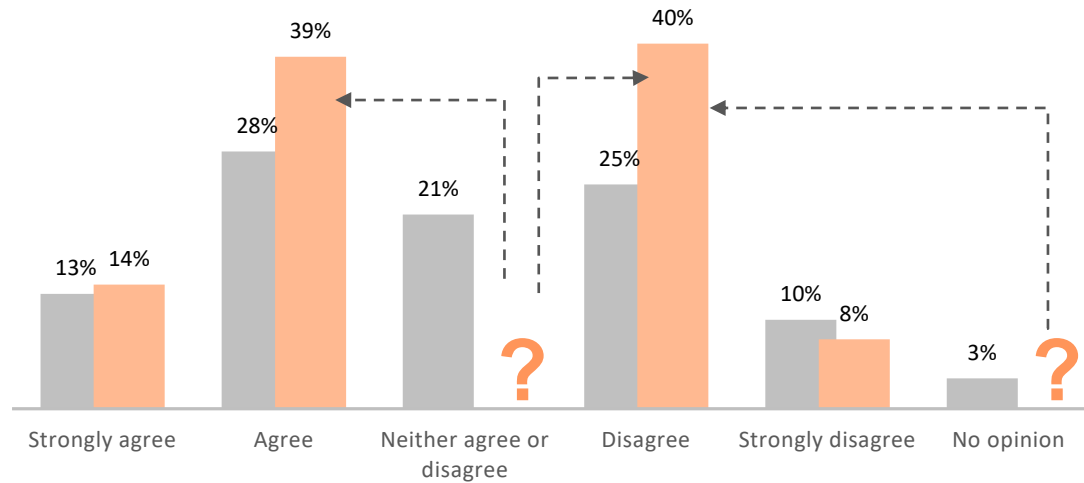


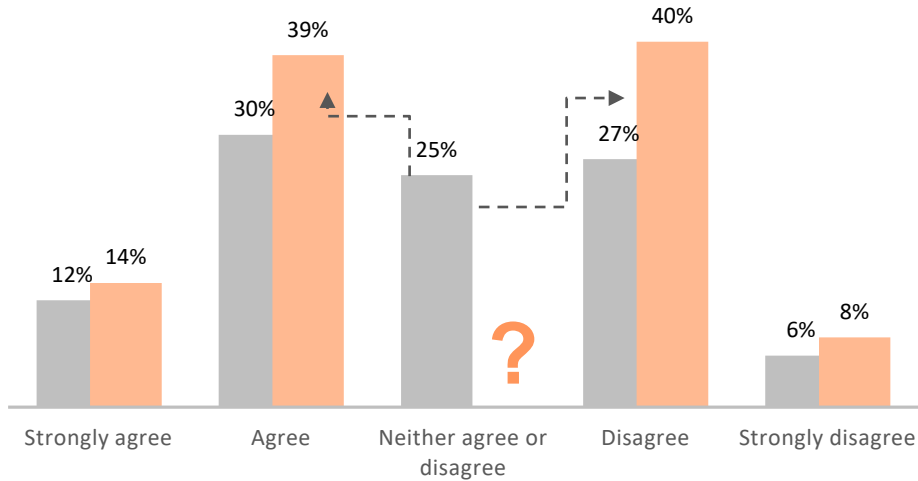


Where did “Neither” and “No opinion” go?



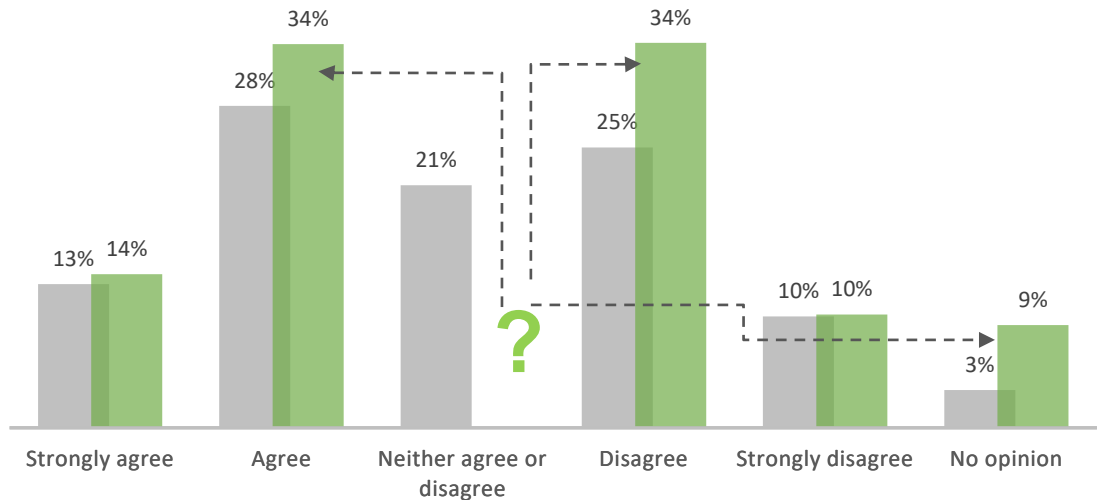
Where did “Neither” and “No opinion” go?

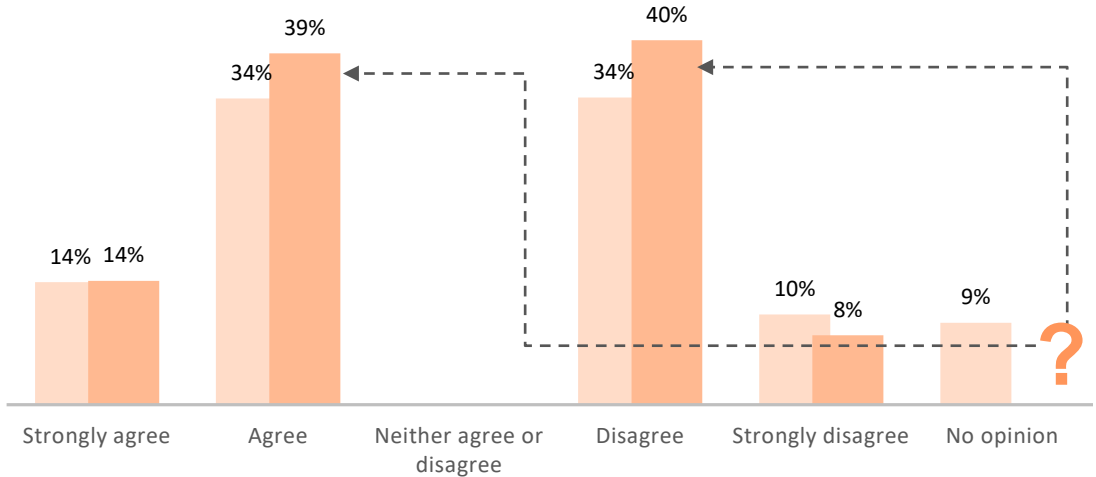




4-point Likert scale
= Forced-choice scale

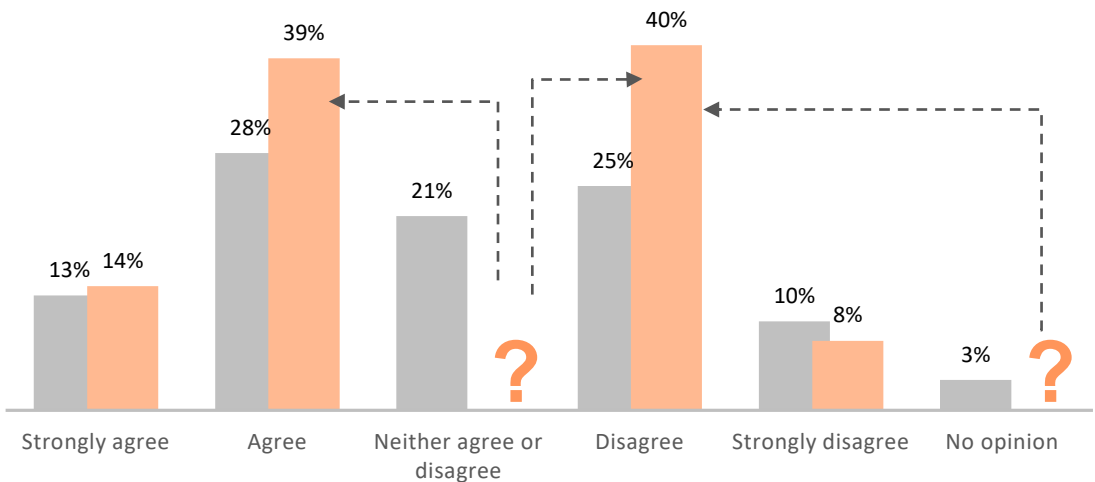
Allow to differentiate
“Neutral” opinion and
“no opinion” or “I don’t
know”





4-point Likert scale
=
Forced-choice scale

Allow to differentiate
“Neutral” opinion and
“no opinion” or “I don’t
know”



Research-based suggestions

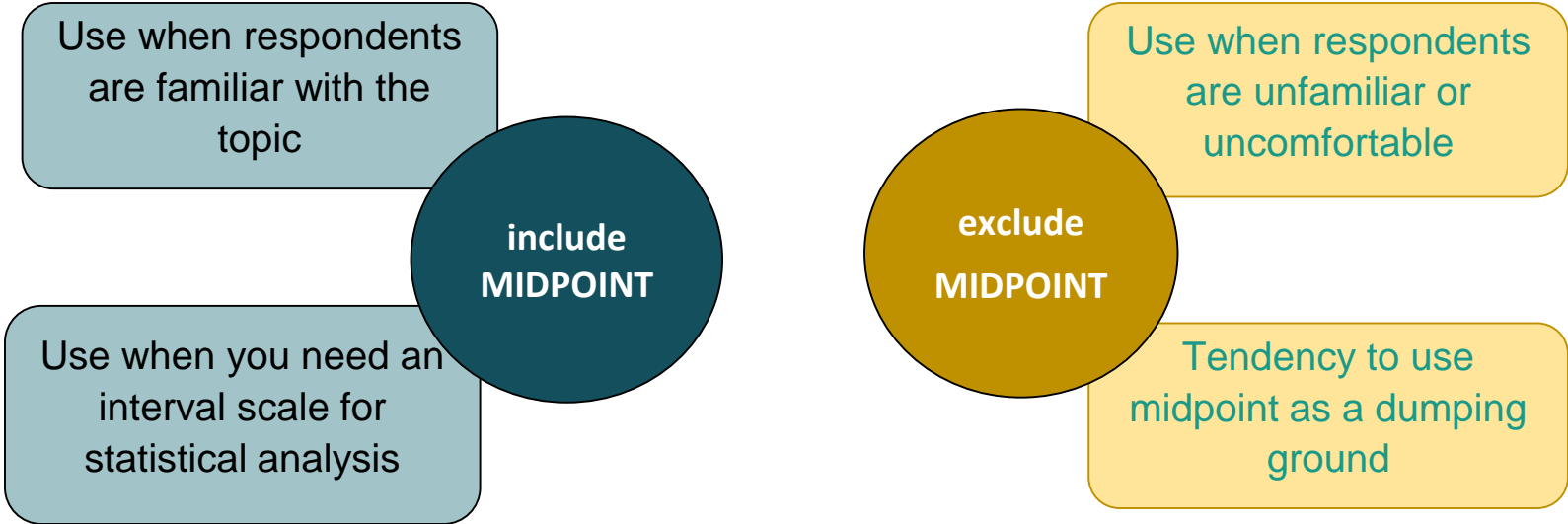


Use when respondents
are familiar with the
topic

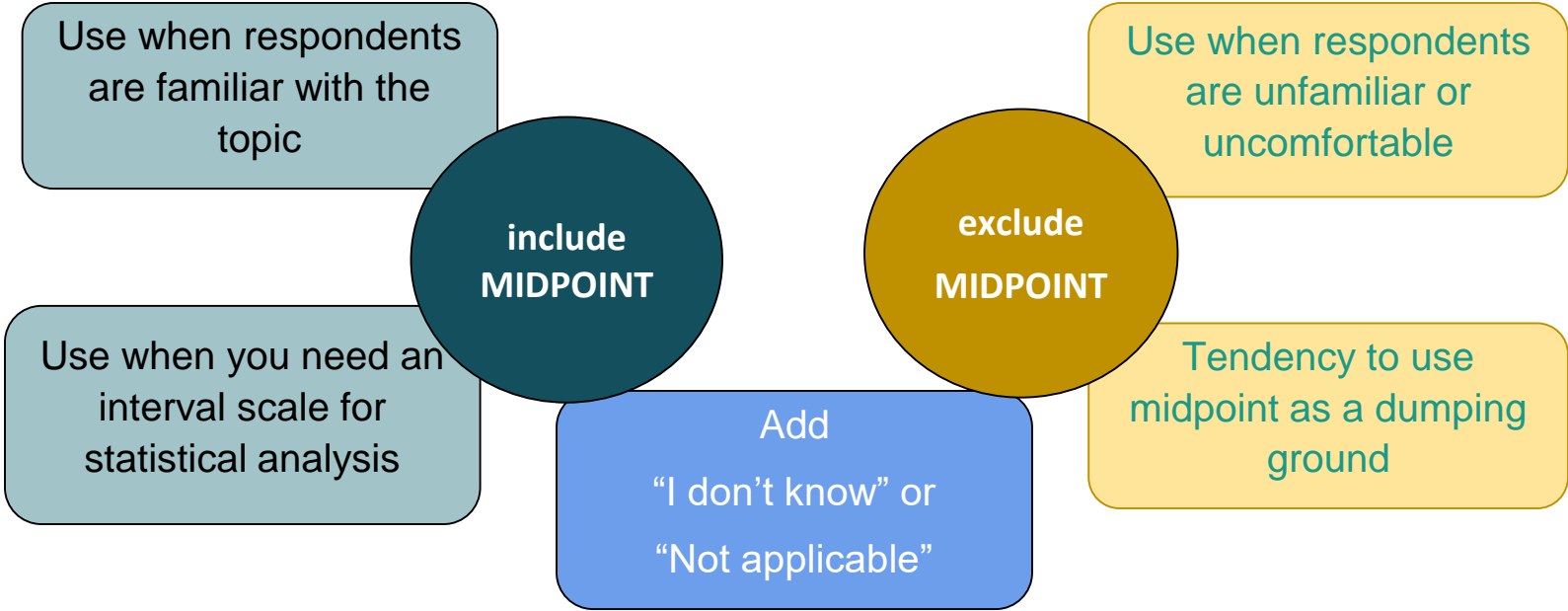
Use when you need an
interval scale for
statistical analysis

**include
MIDPOINT**

Research-based suggestions



Research-based suggestions



**Ascending order or
descending order of
response options**



Response order

Ascending Order

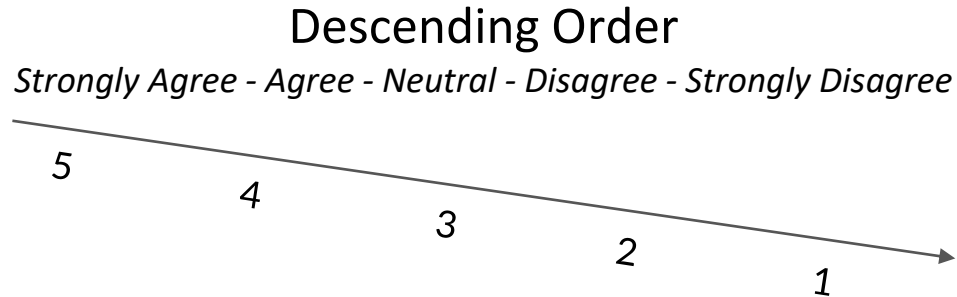
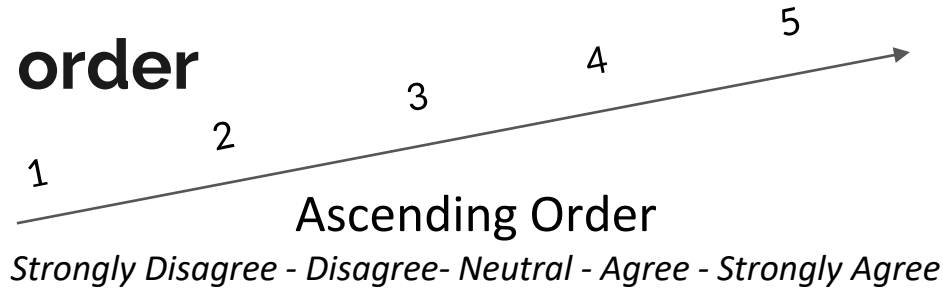
Strongly Disagree - Disagree - Neutral - Agree - Strongly Agree

Descending Order

Strongly Agree - Agree - Neutral - Disagree - Strongly Disagree

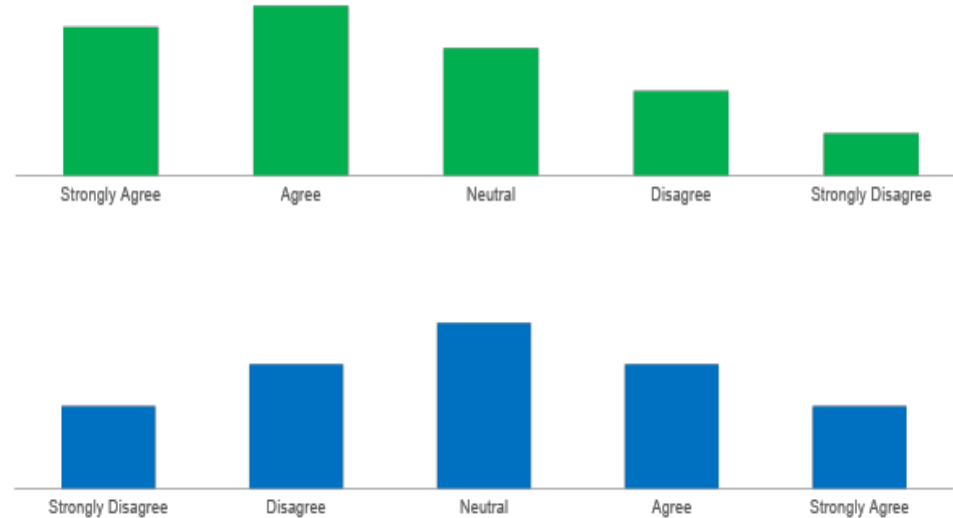


Response order



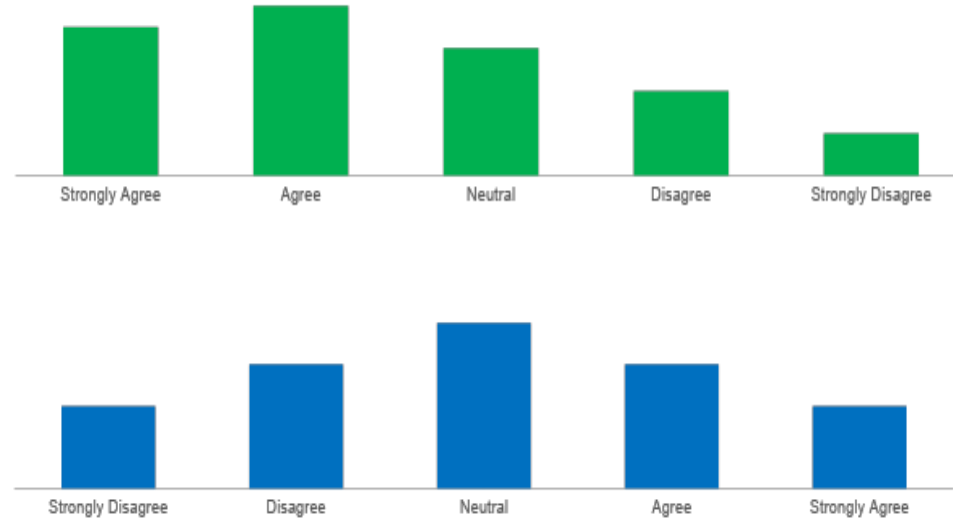
Research shows

- *Friedman et al. (1994)*
- Asked 208 college students to rate their attitudes towards college
- Randomly assigned to one of the two questionnaires (w/different response orders)
- 10 item paper survey using a 5 point Likert-scale
- Students' average rating from the descending order scale was significantly more positive than the one from the ascending order scale.



Research shows

- Maeda (2015)
- Web Survey of 1,653 US Adults
- Randomly assigned to one of the questionnaires with different response orders
- 15 item Power of Foods Scale
- There was a significant difference between the ascending and descending groups
- The descending ordered group's mean shifted 3.3% higher





Research shows

- Hartley and Betts (2010)
- 465 adults in the UK (academic writers, reviewers)
- Randomly assigned to one of the four conditions (different response orders)
- Asked to rate an abstract using one of the four 11-point scales

1.	Clear	10 9 8 7 6 5 4 3 2 1 0	Unclear
2.	Clear	0 1 2 3 4 5 6 7 8 9 10	Unclear
3.	Unclear	10 9 8 7 6 5 4 3 2 1 0	Clear
4.	Unclear	0 1 2 3 4 5 6 7 8 9 10	Clear

Which one generated the highest average score?

Research shows

- Hartley and Betts (2010)
- 465 adults in the UK (academic writers, reviewers)
- Randomly assigned to one of the four conditions (different response orders)
- Asked to rate an abstract using one of the four 11-point scales
- #1 scale generated the most positive score

1.	Clear	10 9 8 7 6 5 4 3 2 1 0	Unclear
2.	Clear	0 1 2 3 4 5 6 7 8 9 10	Unclear
3.	Unclear	10 9 8 7 6 5 4 3 2 1 0	Clear
4.	Unclear	0 1 2 3 4 5 6 7 8 9 10	Clear

Which one generated the highest average score?




Response bias associated with descending order

- Primacy effect
- Left-side selection bias
- Acquiescence bias (yea-saying bias)
- Social desirability bias
- Satisficing



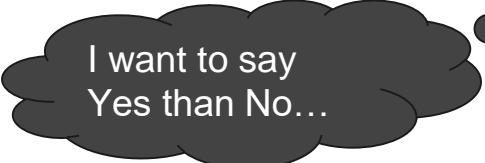
I see this first...



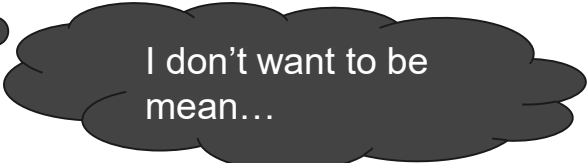
Sure, that sounds good...

Descending Order

Strongly Agree - Agree - Neutral - Disagree - Strongly Disagree



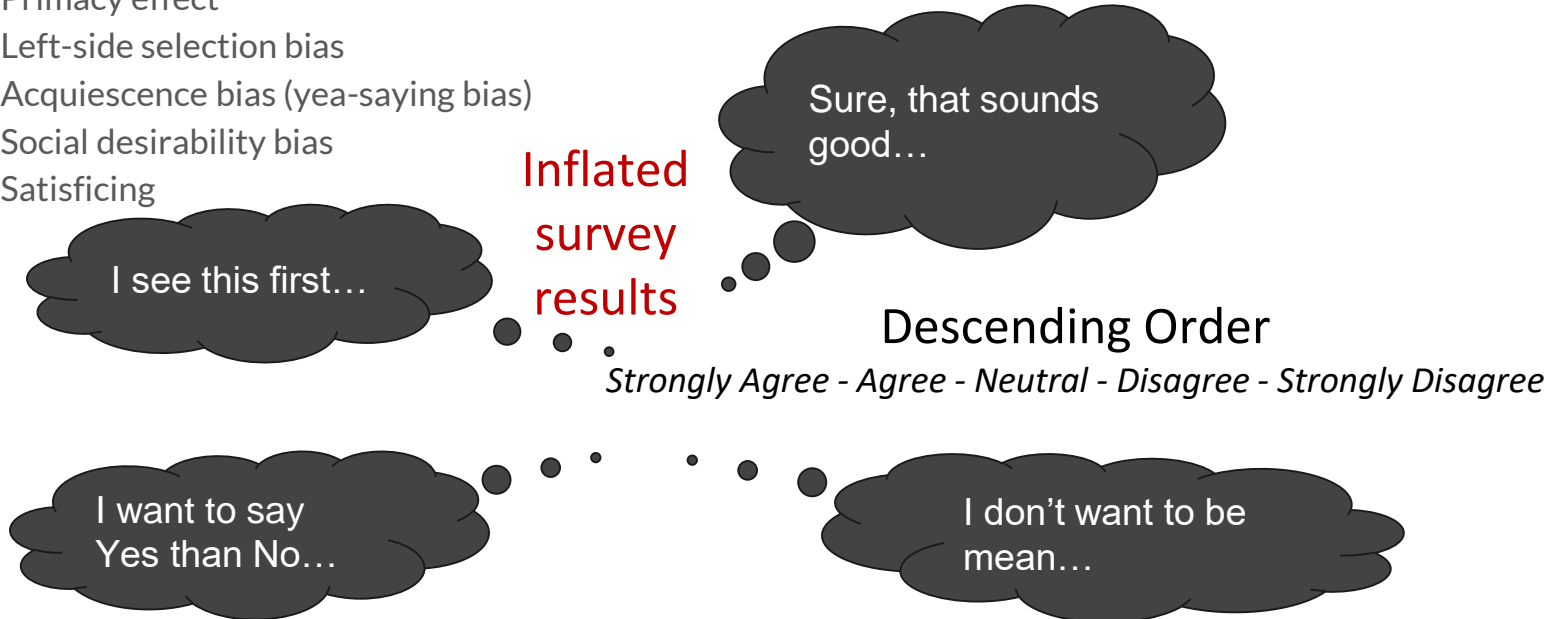
I want to say Yes than No...



I don't want to be mean...

Response bias associated with descending order

- Primacy effect
- Left-side selection bias
- Acquiescence bias (yea-saying bias)
- Social desirability bias
- Satisficing





Research-based suggestions?

- Use ascending-ordered scales to avoid inflated survey results

Strongly Disagree - Disagree - Neutral - Agree - Strongly Agree

- Use descending-ordered scales (and interpret the data) with caution

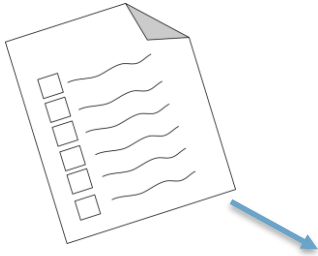
Strongly Agree - Agree - Neutral - Disagree - Strongly Disagree

Research-based suggestions?

Nicholls, Orr, Okubo, and Loftus (2006)

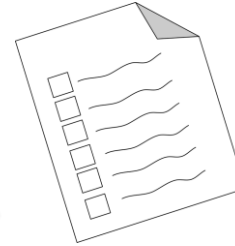
Descending response order

Strongly Disagree - Disagree - Neutral - Agree - Strongly Agree



Ascending response order

Strongly Agree - Agree - Neutral - Disagree - Strongly Disagree



average



To generate an average score?

A

Q1. I am fairly compensated for the quality of my work.

SD - D - N - A - SA

Q2. I have a good relationship with my coworkers.

SD - D - N - A - SA

Q3. I have resources to do my job well .

SD - D - N - A - SA

B

Q1. Do you think you are fairly compensated for the quality of your work?

Yes - Probably yes - Probably no - No

Q2. I have a good relationship with my coworkers.

SD - D - N - A - SA

Q3. Do you have resources to do your job well?

Not enough - Maybe - Enough



To generate an average score?

A

Q1. I am fairly compensated for the quality of my work.

SD - D - N - A - SA

Q2. I have a good relationship with my coworkers.

SD - D - N - A - SA

Q3. I have resources to do my job well .

SD - D - N - A - SA

C

Q1. I am fairly compensated for the quality of my work.

Not at all - A little bit - Somewhat - Quite a bit - Very much

Q2. I have a good relationship with my coworkers.

Not at all - A little bit - Somewhat - Quite a bit - Very much

Q3. I have resources to do my job well .

Not at all - A little bit - Somewhat - Quite a bit - Very much

Extreme Makeover – Survey Edition



Resources (you can download the files)

References:

<https://drive.google.com/file/d/1mnj7lpUC8Qg1z4LFpHBqCFk3Uv9aG4kt/view?usp=sharing>

Extreme Makeover - Survey Edition (Questionnaire):

<https://drive.google.com/file/d/1wu22vnPzA1YUVkl9lipfpKtUbdTarsGH/view?usp=sharing>

Extreme Makeover - Survey Edition (Answer):

<https://drive.google.com/file/d/1WjheKV-2vqIPkyAAIBk4EUFj2Ka0mail/view?usp=sharing>

Evidence-Based Survey Design Principles

1

Include a midpoint or not?



2

Put Strongly Agree on the left or right side of the scale?



3

Use radio buttons or sliders?



4

Use positive statements only or negative ones too?



5

How to handle ceiling effects?



Chyung et al.'s evidence-based survey design articles

- Chyung, S. Y., Roberts, K., Swanson, I., & Hankinson, A. (2017). Evidence-based survey design: The use of a midpoint on the Likert scale. *Performance Improvement Journal*, 56(10), 15-23. <https://doi.org/10.1002/pfi.21727>
- Chyung, S. Y., Kennedy, M., & Campbell, I. (2018). Evidence-based survey design: The use of ascending and descending order of Likert-type scale options. *Performance Improvement Journal*, 57(9), 9-16. <https://doi.org/10.1002/pfi.21800>
- Chyung, S. Y., Barkin, J., & Shamsy, J. (2018). Evidence-based survey design: The use of negatively-worded items in surveys. *Performance Improvement Journal*, 57(3), 16-25. <https://doi.org/10.1002/pfi.21749>
- Chyung, S. Y., Swanson, I., Roberts, K., & Hankinson, A. (2018). Evidence-based survey design: The use of continuous rating scales in surveys. *Performance Improvement Journal*, 57(5), 38-48. <https://doi.org/10.1002/pfi.21763>
- Chyung, S. Y., Hutchinson, D., & Shamsy, J. A. (2020). Evidence-based survey design: Ceiling effects associated with response scales. *Performance Improvement Journal*, 59(6), 6-13. <https://doi.org/10.1002/pfi.21920>