AGREE-DISAGREE QUESTIONS: PROBLEMS AND SOME SOLUTIONS

Allyson L. Holbrook

Associate Professor of Public Administration and Psychology at the Survey Research Laboratory of the University of Illinois at Chicago
AGREE-DISAGREE QUESTIONS

- Simplest form – ask respondents whether they agree or disagree with a statement

- Do you agree or disagree with the following statement?

- “My health is good.”
Agree-Disagree Questions

- Can include intensity (agree strongly or agree somewhat)

- Do you strongly agree, somewhat agree, somewhat disagree, or strongly disagree with the following statement?

- “My health is good.”
AGREE-DISAGREE QUESTIONS

- Can include a midpoint (neither agree nor disagree)

- Do you strongly agree, somewhat agree, neither agree nor disagree, somewhat disagree, or strongly disagree with the following statement?

- “My health is good.”
AGREE-DISAGREE QUESTIONS

- Sometimes a showcard is used in in-person interviews

- but A-D questions are also used in telephone and self-administered questionnaires

- Often part of a scale or a set of items that uses the same response options.
IF YOU HAVE USED AGREE-DISAGREE SURVEY QUESTIONS...

- You are not alone.

- These questions are ubiquitous in survey research today...
  - Post webinar questionnaire.
  - Other examples
Political efficacy

- 'Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.'

- Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?
Do you agree strongly, agree somewhat, neither agree nor disagree, disagree somewhat, or disagree strongly with the following statement (HAND CARD TO RESPONDENT):

“Irish, Italians, Jewish and many other minorities overcame prejudice and worked their way up. Blacks should do the same without special favors.”
National Health and Nutrition Examination Survey (In-Person)

I am going to read a statement and I want you to let me know if you strongly agree, agree, neither agree nor disagree, disagree or strongly disagree with the statement.

I enjoy participating in PE or gym class.
GALLUP POLL (TELEPHONE)

Do you agree or disagree with each of the following statements about Muslims? Most Muslims around the world

...believe that women and men should have equal rights
...are accepting of other religions
...want peace
...are accepting of people of races other than their own
...are accepting of homosexuals
2013 Quinnipiac University Poll (Web Survey)

"In general, do you agree or disagree with the 1973 Roe v. Wade Supreme Court decision that established a woman's right to an abortion?"
EXAMPLE SCALES: HANDBOOK OF MARKETING SCALES (2010)

- Ten-Item and Five-Item Personality Inventories (Gosling, Rentfrow, and Swann 2003)
- Long-Term Orientation: LTO (Bearden, Money, and Nevins 2006)
- Maximization (Schwartz et al. 2002; Nenkov et al. 2008)
- Need for Cognition: NFC (Cacioppo and Petty 1982)
- Need for Touch: NFT (Peck and Childers 2003)
- Consumer’s Need for Uniqueness: CNFU (Tian, Bearden, and Hunter 2001)
EXAMPLE SCALES (cont.):

- Preference for Consistency: PFC (Cialdini, Trost, and Newsom 1995)
- Independent and Interdependent Self- Construals (Singelis 1994)
- Self-Concept Clarity: SCC (Campbell et al. 1996)
- Compulsive Buying Index (CBI): An Expanded Measure (Ridgway, Kukar-Kinney, and Monroe 2008)
- Hyperopia (Aversion to Luxury) (Haws and Poynor 2008)
- Consumer Spending Self-Control: CSSC (Haws and Bearden 2010)
- Moral Identity (Aquino and Reed 2002)
PERSONAL ANECDOTAL EXPERIENCE:

- Questionnaire Review Committee

- Almost every questionnaire has at least one A-D question.

- It is very difficult to persuade clients to change A-D questions.
WHY ARE AGREE-DISAGREE QUESTIONS SO POPULAR?

- They’re easy to write.

- It’s easy to ask about several unrelated topics using the same response scale.

- It seems efficient
  - Space in self-administered questionnaires
  - Time in both self-administered and interviewer-administered surveys

- It’s easy to find A-D items from past surveys to recycle.
PROBLEMS WITH INDIVIDUAL A-D ITEMS

1. A-D Questions are more cognitively difficult.

2. Acquiescence response bias
   - Systematic bias in prevalence estimates
   - Cultural (e.g., country or racial/ethnic) group differences

3. Difficulty in interpreting the meaning of a “disagree” response

4. Respondents often have to think through double negatives
PROBLEMS WITH INDIVIDUAL A-D ITEMS

1. **A-D Questions are more cognitively difficult.**

2. Acquiescence response bias
   - Systematic bias in prevalence estimates
   - Cultural (e.g., country or racial/ethnic) group differences

3. Difficulty in interpreting the meaning of a “disagree” response

4. Respondents often have to think through double negatives
1. A-D ITEMS ARE MORE COGNITIVELY DIFFICULT.

- Cognitive necessary to answer a question carefully and completely (Tourangeau, Rips, and Rasinski, 2000)
  - Comprehend question
  - Retrieve relevant information from memory
  - Integrate the information into a judgment
  - Map that judgment onto the response format being requested
1. **A-D Items are more cognitively difficult.**

- A respondent answering an A-D question must
  - Comprehend the A-D question as written
  - Determine what the underlying question is
  - Retrieve relevant information from memory
  - Integrate that information into a response to the underlying question
  - Map that response onto the A-D response scale
    - AD dimension different from that of the underlying question
    - Mapping more difficult.
FOR EXAMPLE:

- I am going to read a statement and I want you to let me know if you strongly agree, agree, neither agree nor disagree, disagree or strongly disagree with the statement. I enjoy participating in PE or gym class.
A respondent answering this question has to:

- Comprehend the question (words, meaning, etc.)
- Determine the underlying dimension:
  - the extent to which the R enjoys participating in PE or gym class
- Retrieve relevant information
- Form a judgment – I like PE class.
- Map that judgment into the A-D format requires
  - Take evaluation of PE class
  - Compare to statement
  - Link evaluation to correct A-D response to statement
ONE GOAL IN WRITING SURVEY QUESTIONS:

- To make the respondent’s cognitive task as simple as possible (all other things being equal)

- If the task of answering a survey question is difficult, some respondents will look for ways to make the task of question answering easier
  - When respondents go through the cognitive steps necessary to answer a survey question either incompletely or not at all → survey satisficing (Krosnick, 1991)
  - One easy way to do that is to agree with statements made by the interviewer
PROBLEMS WITH INDIVIDUAL A-D ITEMS

1. A-D Questions require extra cognitive steps

2. Acquiescence response bias
   • Systematic bias in prevalence estimates
   • Cultural (e.g., country or racial/ethnic) group differences

3. Difficulty in interpreting the meaning of a “disagree” response

4. Respondents often have to think through double negatives
2. **Acquiescence Response Bias (ARB)**

- **Definition:** Agreeing with assertions regardless of content.

- As a result, individual questions may be biased toward showing preferences in the agreement direction.

- ARB is more likely when respondent ability and motivation are low and when task difficulty is high
  - Ability: e.g., Cognitive skills: Education
  - Motivation: e.g., Interest in the survey or question topic
  - Location of the question in the survey
Acquiescence Response Bias Example:

- Often can’t identify individual agree responses as acquiescence

Acquiescence Response Bias Experiment:

- Half of respondents randomly assigned to be asked:
  - Do you agree or disagree with the following statement. Individuals are more to blame than social conditions for crime and lawlessness in this country.

- Half of respondents randomly assigned to be asked:
  - Do you agree or disagree with the following statement. Social conditions are more to blame than individuals for crime and lawlessness in this country.
Acquiescence Response Bias Example:

- Between subjects (Schuman and Presser, 1981)
  - Individuals are more to blame: 57% agree
  - Social conditions are more to blame: 59% agree

- Within subjects:
  - the same respondents were asked two opposing questions at different points in time (order rotated)
  - Agree with both = 25.6%
  - Disagree with both = 9.9%
CULTURAL DIFFERENCES IN ARB

- race/ethnicity

- cultural value indicators at both the country and individual levels

- deference norms → greater ARB
  - Cultures high in power distance
  - Cultures low in individualism.

- Evidence that ARB is greater among
  - Latino respondents (relative to non-Hispanic Whites)
  - Respondents from collectivistic countries (e.g., east asian countries) than those from individualistic countries (e.g., western countries)
  - Individuals who score lower on measures of individualism and higher on measures of collectivism
EXAMPLE:

- Members of four racial/ethnic groups came to the lab:
  - Non-Hispanic White (N=250)
  - Non-Hispanic Black (N=251)
  - Mexican-American (N=250)
  - Korean American (N=249)
- Self-administered questionnaire.
EXAMPLE:

- Rated agreement with the following two statements:

  - I tend to put off making important decisions until the last possible moment.

  - I usually make important decisions quickly and confidently.
PERCENT WHO AGREED WITH BOTH STATEMENTS:

- Overall: 26%

Across racial and ethnic groups:
- Non-Hispanic White = 15%
- Non-Hispanic Black = 26%
- Mexican-American = 31%
- Korean American = 31%

- Racial and ethnic differences in individualism
POSSIBLE RESULTS OF CULTURAL DIFFERENCES IN ARB

• The spurious appearance of systematic group differences in dimensions (actually due to ARB)

• Failure to find actual group differences in dimensions (hidden by differences in ARB).
PROBLEMS WITH INDIVIDUAL A-D ITEMS

1. A-D Questions require extra cognitive steps

2. Acquiescence response bias
   • Systematic bias in prevalence estimates
   • Cultural (e.g., country or racial/ethnic) group differences

3. Difficulty in interpreting the meaning of a “disagree” response

4. Respondents often have to think through double negatives
3. DIFFICULTY IN INTERPRETING THE MEANING OF A “DISAGREE” RESPONSE

'Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?
Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

Unclear whether a disagree response means:

- Respondent thinks politics and government always seem complicated
- Respondent thinks politics and government never seem complicated
3. Difficulty in Interpreting the Meaning of a “Disagree” Response

- Researchers often choose statements such as this that are moderate
  - Use of words like “some” “mostly” “often” “usually”
  - Item not helpful if everyone says “agree” or “disagree”
  - “Do you agree or disagree with the statement: My health is good?”
    - “I disagree, my health is excellent.”
Do you agree or disagree with each of the following statements about Muslims? Most Muslims around the world believe that women and men should have equal rights. They are accepting of other religions. They want peace. They are accepting of people of races other than their own. They are accepting of homosexuals.

Does disagree mean a R believes all Muslims hold these views or than no Muslims do?
Problems with individual A-D items

1. A-D Questions require extra cognitive steps

2. Acquiescence response bias
   - Systematic bias in prevalence estimates
   - Cultural (e.g., country or racial/ethnic) group differences

3. Difficulty in interpreting the meaning of a “disagree” response

4. Respondents often have to think through double negatives
4. RESPONDENTS OFTEN HAVE TO THINK THROUGH DOUBLE NEGATIVES.

'Public officials don't care much what people like me think.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?
4. **Respondents Often have to think through double negatives.**

- Disagree = Public officials DO care
- Requires thinking through a double negative
- Extra cognitive step
- Double negatives are understood and used in other languages differently – language comparability
- Avoid negatively worded statements, but...
USING MULTIPLE A-D ITEMS: A SOLUTION???

- A-D Items are often part of multi-item scales
  - Easy to generate multiple items that use A-D response options

- Using multiple A-D items – some positively worded/coded and some negatively worded/coded
  - Strategy for addressing concerns with ARB
  - Acquiescent responses cancel each other out
PROBLEMS WITH USING MULTIPLE A-D ITEMS:

- Inefficient use of respondent time
- Increases respondent fatigue $\rightarrow$ increases the probability of ARB
- Low negative correlations between A-D reports of opposite statements
  - I tend to put off making important decisions until the last possible moment.
  - I usually make important decisions quickly and confidently.
  - $r = -.13$, $p<.01$; $N=992$
- Respondents who acquiesce end up in the middle of the scale when responses are combined into an index
PROBLEMS WITH USING MULTIPLE A-D ITEMS:

- May artificially reduce variance (dispersion) of index
  - May do so differently across cultural or racial/ethnic groups
  - May do so differently depending upon question difficulty or context (e.g., location in the questionnaire)

- Using multiple items with the same response scale opens the door for other response effects like nondifferentiation
  - Using a limited range of responses
  - Straightlining
**So What To Do?**

- Revise A-D items to use item-specific response formats

- Ascertain (or decide) what the underlying dimension of interest is

- Write a question using item-specific responses to measure that underlying dimension

- Questions with item-specific responses provide higher quality data than A-D items
  - (Saris, Revilla, Krosnick, & Shaeffer, 2010 Survey Research Methods)

- Smaller number of higher quality items
Example: (2012 ANES Experiment)

- Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

Underlying dimension:

- How often politics and government are too complicated to understand
- How complicated politics and government are
- How well the R feels s/he understands politics and government
EXAMPLE: 2012 ANES

- 'Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

- How often do politics and government seem so complicated that you can't really understand what's going on? ALWAYS, MOST OF THE TIME, ABOUT HALF THE TIME, SOME OF THE TIME, or NEVER? (rotated)
NHANES Question:

- I am going to read a statement and I want you to let me know if you strongly agree, agree, neither agree nor disagree, disagree or strongly disagree with the statement. I enjoy participating in PE or gym class.

- Underlying dimension is
  - how much the respondent liked PE or gym class
Do you like or dislike participating in PE or gym class?

[IF LIKE]: Do you like it a lot or somewhat?

[IF DISLIKE]: Do you dislike it a lot or somewhat?

OR

How much of the time do you enjoy participating in PE or gym class? Always, most of the time, about half of the time, some of the time, or never?
Do you agree or disagree with each of the following statements about Muslims? Most Muslims around the world

...believe that women and men should have equal rights
...are accepting of other religions
...want peace
...are accepting of people of races other than their own
...are accepting of homosexuals

Underlying dimension is how many Muslims hold particular views.
How many Muslims around the world do you think believe that women and men should have equal rights? All Muslims, most Muslims, some Muslims, a few Muslims, or no Muslims?

How many Muslims around the world do you think are accepting of other religions? All Muslims, most Muslims, some Muslims, a few Muslims, or no Muslims?

How many Muslims around the world do you think want peace? (All Muslims, most Muslims, some Muslims, a few Muslims, or no Muslims?)

How many Muslims around the world do you think are accepting of people of races other than their own? (All Muslims, most Muslims, some Muslims, a few Muslims, or no Muslims?)

How many Muslims around the world do you think are accepting of homosexuals? (All Muslims, most Muslims, some Muslims, a few Muslims, or no Muslims?)
"In general, do you agree or disagree with the 1973 Roe v. Wade Supreme Court decision that established a woman's right to an abortion?"

- Underlying dimension is presumably opinions about legalized abortion, although the question gets at that very indirectly.
  - Support for Supreme Court
  - Trust in Government?
Alternative item-specific question:

- Which of the following best represents your opinion about abortion: 1) abortion should not be legal under any circumstances, 2) abortion should be legal only in case of rape, incest, or when the woman’s life is in danger, 3) abortion should be legal for reasons other than rape, incest, or danger to the woman or 4) abortion should be legal under all circumstances.
**Does it matter?**

- Do A-D and item-specific questions obtain different responses?

- Does the quality of data obtained from A-D and item-specific questions differ?


- Analysis of items from 2012 ANES.
2012 ANES Experiments:

- Survey
  - U.S. eligible voters
  - Face-to-face sample: N=2,054
  - Internet sample (GfK internet panel): N=3,860
  - Pre-election – two months before election
  - Post-election – day after election into January, 2013
2012 ANES Experiments:

- Experiments:
  - Half of respondents randomly assigned to be asked 4 political efficacy questions using the A-D format in pre-election interview
  - Other half of respondents randomly assigned to be asked parallel item-specific questions in pre-election interview
  - All respondents asked other set of items in post-election interview
    - Half of respondents:
      - Pre-election – A-D questions
      - Post-election – item specific questions
    - Half of respondents:
      - Pre-election – item specific questions
      - Post-election – A-D questions
EXAMPLE: (2012 ANES EXPERIMENTS)

- 'Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

- How often do politics and government seem so complicated that you can't really understand what's going on? ALWAYS, MOST OF THE TIME, ABOUT HALF THE TIME, SOME OF THE TIME, or NEVER? (rotated)
'I feel that I have a pretty good understanding of the important political issues facing our country.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

How well do you understand the important political issues facing our country? EXTREMELY WELL, VERY WELL, MODERATELY WELL, SLIGHTLY WELL, or NOT WELL AT ALL? (rotated)
'Public officials don't care much what people like me think.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

How much do public officials care what people like you think? A GREAT DEAL, A LOT, A MODERATE AMOUNT, A LITTLE, or NOT AT ALL? (rotated)
'People like me don't have any say about what the government does.' Do you AGREE STRONGLY, AGREE SOMEWHAT, NEITHER AGREE NOR DISAGREE, DISAGREE SOMEWHAT, or DISAGREE STRONGLY with this statement?

How much can people like you affect what the government does? A GREAT DEAL, A LOT, A MODERATE AMOUNT, A LITTLE, or NOT AT ALL? (rotated)
ANES Experiments: Examples

Questions:

- How does the distribution of responses vary for questions with agree-disagree and item-specific response options? (pre-election)

- What is the association between the standard and revised items?

- Do the standard and revised items differ in their predictive validity?
ANES Experiments: Examples

Questions:

- How does the distribution of responses vary for questions with agree-disagree and item-specific response options? (pre-election)

- What is the association between the standard and revised items?

- Do the standard and revised items differ in their predictive validity?
'Sometimes, politics and government seem so complicated that a person like me can't really understand what's going on.

How often do politics and government seem so complicated that you can't really understand what's going on?
'I feel that I have a pretty good understanding of the important political issues facing our country.'

How well do you understand the important political issues facing our country?
'Public officials don't care much what people like me think.'

How much do public officials care what people like you think?
People like me don't have any say about what the government does. How much can people like you affect what the government does?
ANES Experiments: Examples

Questions:

• How does the distribution of responses vary for questions with agree-disagree and item-specific response options? (pre-election)

• What is the association between the standard and revised items?

• Do the standard and revised items differ in their predictive validity?
Correlations among items across time: (Pre-Post)

Politics and government complicated:
- A-D / Item-specific: $r=.04$, $p=.05$, $N=2965$
- Item-specific / A-D: $r=.02$, $p=.41$, $N=2920$

Understanding of important political issues:
- A-D / Item-specific: $r=.01$, $p=.74$, $N=2960$
- Item-specific / A-D: $r=.03$, $p=.12$, $N=2924$

Public official care what people like me think:
- A-D / Item-specific: $r=.03$, $p=.12$, $N=2956$
- Item-specific / A-D: $r=.11$, $p<.01$, $N=2918$

People like me affect government
- A-D / Item-specific: $r=.04$, $p=.05$, $N=2957$
- Item-specific / A-D: $r=.10$, $p<.01$, $N=2921$
What do these correlations mean?

- Relatively low associations between revised and standard items.

- Possibility that efficacy is just very unstable?
  - These associations examined over just a few months
  - Evidence from prior studies suggests that the standard items have over time reliabilities of .30-.45 across a 4 year time span (e.g., McPherson, Welch, and Clark, 1977)
ANES Experiments: Examples

Questions:

• How does the distribution of responses vary for questions with agree-disagree and item-specific response options? (pre-election)

• What is the association between the standard and revised items?

• Do the standard and revised items differ in their predictive validity?
**Predictive validity:**

- Political efficacy is positively associated with activism (voting, campaign contributions, etc.)

- **Internal efficacy**
  - Index of “politics and government complicated” and “Understanding of important political issues” items
  - Index calculated separately for respondents who were asked A-D questions and those who were asked item-specific questions

- **External efficacy**
  - Index of “public official care what people like me think” and “people like me affect government” items
  - Index calculated separately for respondents who were asked A-D questions and those who were asked item-specific questions
**PREDICTIVE VALIDITY:**

- Index of political activism during 2012 campaign and election.

- During post election, respondents were asked whether they had:
  - Voted
  - Attended a rally for a party or candidate
  - Wore a button or posted a sign for a candidate or party
  - Did any other work for a party or candidate
  - Contribute to a candidate
  - Contribute to a political party
  - Contribute to any other organization working for or against a candidate

- Campaign activism was the total number of these behaviors each respondent reported performing during the campaign/election.
CORRELATION BETWEEN EFFICACY AND ACTIVISM:

Graph showing the correlation between internal efficacy index and external efficacy index, with two categories: A-D Questions and Item-specific Questions.
ANES EXPERIMENTS:

- Distribution of responses to A-D and item-specific questions differed

- Association between responses to A-D and item-specific questions was very low

- Evidence that responses to item-specific questions have greater predictive validity than do responses to A-D questions.
But, But, But what about…

- Comparing results to previous studies

- If data comparability is absolutely a must, use old items, but understand and try to estimate and control for potential error.
  - Include ARB experiments or
  - Include items to measure ARB to use as a covariate
  - *NOT* ideal solution
  - Doesn’t account for many of the problems with A-D questions (double negatives; uncertainty in interpreting disagree responses)
But, But, But what about…

- Longitudinal data collection

- Include new and old questions as a means to transition.
  - 2012 ANES good example
But what about…

- Length of time (or amount of space) for revised items

- Anand, Parsons, and Owens (2010) AAPOR presentation
  - Respondents randomly assigned to receive either a series of questions using item-specific response options or parallel questions with yes/no response options
  - Did not significantly affect interview length
  - Some evidence that interviewers perceived respondents as less cooperative and more likely to want to end the interview in the item-specific condition

- Using item-specific items does not appear to substantially increase administration time
  - Trade-off with data quality
  - Minimizes needed sample size and items
**Final Conclusions:**

- **Agree-disagree questions:**
  - Not good!!!!!
  - Many approaches (e.g., controlling for ARB, using lots of items some of which are reverse coded) not effective at dealing with all the problems with A-D questions.
  - Recommended alternative: questions with item-specific response options
FINAL CONCLUSIONS:

- Particularly problematic:
  - Culturally diverse populations
  - Group differences
  - Circumstances or individuals where ARB is most likely
  - Negative statements
  - Statements that use moderate qualifiers (e.g., most, usually, sometimes)
FINAL CONCLUSIONS:

- Very limited set of circumstances where you may want to use A-D items:
  - Direct comparison with previous data in which they were used
  - Longitudinal data collections
  - Even in these cases, there may be ways to move away from A-D items.
THANK YOU!

Questions?

I can be reached at allyson@uic.edu.
REFERENCES:


