



Effect of Survey Response Format on Participant Response

Lewis Wittry, Kristina Bell, Lisa Colclasure, Chris Taylor, Sharon Sears, Ph.D.
Psychology Department, Fort Lewis College, Durango, CO 81301



Abstract

Self-report scales are used in abundance across several different professional areas. It is important to understand how wording, anchors, and formats can affect final results. In addition, short forms of scales are important for use in busy clinical and medical settings. Our experiment tested two different short anxiety scale formats in order to determine if responses would vary, and whether they would correlate with longer anxiety scale versions. For this study, we chose a Likert 10-point scale and a Visual Analogue scale represented by a line with beginning and end anchors. Our results did not support our hypothesis that we would find a difference in anxiety scores between the two scale formats. However, we did find order interaction effects such that when people responded to the number, they scored higher if they saw that version first before the line version. Finally, both short scales correlated with a longer version. Although both scale types can be used in gathering data, we did find that the Likert scale was easier to score and left less possibility of researcher error.

Background Literature

- 436 college level economics and education students
 - Results demonstrated that when presented with response scales that started with positive categories, there were higher rates of positive response. Subsequently, when presented with negative categories, participant's responses were neither high nor low (Krebs & Hoffmeyer-Zlotnik, 2010).
- 1737 students
 - Study aimed to find out if response format would have an effect on student's ranking of life satisfaction/dissatisfaction. Results demonstrated that when given a response format ranging from -5 to +5, participants ranked life dissatisfaction as being less than when given a response format of 0 to +10. This study demonstrates how response format can influence survey results (Mazaheer & Theuns, 2008).
- 192 Critically ill patients
 - Study aimed to develop a shorter version of the STAI to use with critically ill hospital patients. This study developed a survey that was high in validity when compared to the STAI and BSI, and found that the new survey was less cumbersome for hospital patients (Abed, Hall & Moser, 2011).

Today's Date: _____

"Right now, I am feeling..."

Please indicate your response on each line with an X

Not at All Very Much So

1. Comfortable _____

2. Worried _____

3. Pleasant _____

4. Nervous _____

5. At Ease _____

6. Frightened _____

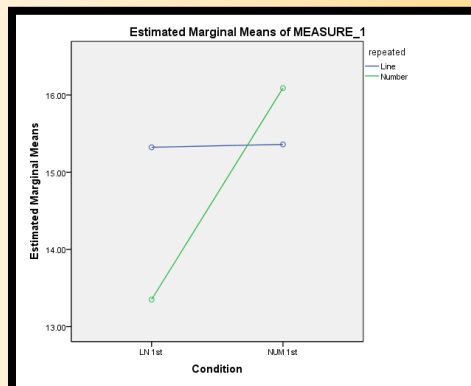
7. Pain _____

Today's Date: _____

"Right now, I am feeling..."

Not At All Very Much So

1. Comfortable	0	1	2	3	4	5	6	7	8	9	10
2. Worried	0	1	2	3	4	5	6	7	8	9	10
3. Pleasant	0	1	2	3	4	5	6	7	8	9	10
4. Nervous	0	1	2	3	4	5	6	7	8	9	10
5. At Ease	0	1	2	3	4	5	6	7	8	9	10
6. Frightened	0	1	2	3	4	5	6	7	8	9	10
7. Pain	0	1	2	3	4	5	6	7	8	9	10



Methods

- Participants** -Convenience sample of 44 Intro to Psychology students at Fort Lewis College
- Design** -Mixed model (within and between groups)
-paired sample t-test and 2x2 ANOVA
- Materials** -Likert scale survey
-Visual Analogue survey (measured in millimeters)
-State Anxiety Inventory survey (long version - 20 items)

Results

- No difference in overall anxiety response based on survey format (LN $M = 15.34$, $SD = 10.08$; NUM $M = 14.78$, $SD = 9.62$; $t = .865$, $p = .392$)
- Found positive correlation with a previously validated Anxiety Inventory (STAI-NUM $r = .864$, $p = .000$; STAI-LN $r = .894$, $p = .000$)
- Scores were higher on the NUM survey when presented with the NUM survey first (see graph for means; repeated*condition $F = 4.80$, $p = .034$)
- Likert scale is much easier to use as a research tool

Discussion

This research deconstructed suspected differences between the Likert scale and the Visual Analogue Scale finding that those differences did not really exist. By using a short, easily understandable scale, we were able to remove any wording or width effects from our results. The original hypothesis pertaining to differences in scores between the NUM and LN conditions was not supported by our findings. Instead, we found that the replies to the NUM scale and the replies to the LN scale were highly correlated. These findings are similar to the outcomes of Hasson & Arnetz, 2005 who found no differences between a Likert scale, (our NUM) and a Visual Analogue scale, (our LN). The present study did produce some puzzling results. The crossover effect in the NUM category suggests that the NUM scale produced statistically higher results when viewed as the first scale. Previous research has found conflicting results as to the best possible response format of a scale (Hasson & Arnetz, 2005). Although this study did not support our hypothesis, it did serve to add another small piece to the puzzle of what makes the best scale. Our findings do support existing research pointing to minimal differences in scoring between Likert scales and Visual Analogue Scales, as we did not find significant differences between the two scale types.