# The Impact of Negation in Survey Research 

Marci Morrow Enos

Negation may win elections, but it creates misunderstandings in survey research. Accusations of "negative campaigning" and "negative advertising" abound in political races. The implica tion is that candidates who use negativity take unfair advan tage, since it grabs the public's attention. Negativity made headlines in the Republican presidential primary in South Carolina when Senator John McCain blamed his loss on Governor George W. Bush's "negative message of fear" (Berke, 2000, February 20).

My story is about negation's effects - not in politics, but in survey research. Long ago I was involved in the development of questionnaires to elicit students' attitudes toward school. The questionnaire items were thoughtfully chosen and closely targeted but, when the results were analyzed by Rasch methodology (Rasch, 1993/1960), a disturbing pattern emerged. The response format used four categories. Positive and negative items were included. Everything was done according to standard research methods. Negative items, which asked about the "bad" aspects of the attitudes examined, were reversed coded so that the respondent's reactions would be "in line" with their responses to the positive items. The problem emerged when the scales were analyzed with Rasch methodology. Many of the negative items misfit and were found to be measuring a variable different from the positive items.

This experience stuck with me and has led me to investigate this phenomenon. Social scientists should try to be as smart as politicians. Politicians understand the unique power of negation. Social scientists seem to think it is just affirmation flipped over!

## Abuse of the Positive and <br> Misuse of the Negative

The once popular concept of self-esteem has taken a beating in recent years. A New York Times article (Johnson, 1998, May 5) criticized a self-esteem survey instrument (Rosenberg, 1979) used in a study of educational change in the California school system. Educators and researchers expressed disappointment in the project. The results did not yield the expected correlations with aptitude and achievement and, therefore, could not predict the direction of academic progress.


Marci Morrow Enos
Marci Enos has had a varied career in the worlds of education and psychology. Currently, she is part of a private practice offering psychological services in Glenview and Chicago, Illinois while also pursuing research interests with Ben Wright at the University of Chicago. She has been an elementary school teacher, a counselor and teacher for severely disturbed adolescents, an assistant professor of education at Roosevelt University of Chicago, a test author with a publishing company and, at Michael Reese Hospital of Chicago, she was a clinical therapist and the director of a research program designed to help hearing handicapped infants and their parents. She is married and the mother of two wonderful daughters.

The whole idea of "self-esteem" was called into question.
While conceding that the self-esteem studies may have suffered "distortions in how self-esteem statistics had been gathered," the Times article cites several prominent educators who bash self-esteem as a construct:

- Research [indicates] that esteem is not in and of itself a strong predictor of success. The idea that high self-esteem is the exclusive province of those with admirable achievements has been rejected.
- Questions have been raised about the size of [self-esteem] effects and the direction of effects and whether in fact it's a mixed blessing to even have high self-esteem.
- Criminals and juvenile delinquents . . often have high self-esteem.
- Self-esteem . . . mutated instead into a kind of crutch that explains ... low achievement.

The baby was being pitched out with the bath water. The belief that the constellation of ideas and opinions we have about ourselves shapes how we behave makes sense. These ideas, under a variety of names - self-image, self-esteem, identity, ego, self awareness or self-concept - have long been used by human behavior researchers such as Bloom (1976), Brookover (1964), Coopersmith (1967), Epps (1969), Purkey (1970), and Rosenberg (1965). What was wrong? I reexamined the Rosenberg Scale to find out why this instrument did not lead to useful results.

Raw scores were used in the computation of esteem scores. But raw scores are not linear (Wright \& Stone, 1979), and perhaps that was the problem. The inches on a yardstick are useful only because each inch is the same as the one before it and the one after. One yardstick is like another. My height is the same using my yardstick and the one at my doctor's office. Because of this uniformity, my height is predictable.

Perhaps the Rosenberg Scale (Table 1) is too abbreviated. It has only ten questions. Five of them are worded positively. This may be too few to delineate such a complex variable.

When we intend to develop a linear variable, it is important to use a range of items. The scale should include some easy items, some a bit more challenging, and some that are hard. It is unrealistic to expect only five posi-
tive items to carry the weight of self-esteem on their backs.
The Rosenberg directions say to score the negative items in the opposite direction from the positive and add them to the positive scale. Social science research has long utilized this positive plus reversed negative strategy to combat a "mind set" in the respondents. Wright and Masters (1982), citing Angell (1907), discuss this practice of constructing attitude measures from equal numbers of positive and negative state-ments-done with the hope of "balancing out" the effects of individual response styles. Wright and Masters show us that this strategy does not correct the problem. It is more important is to discover whether all items "provide consistent information about a person's attitude before combining them to obtain a single attitude for that person" (p. 135).

## Why Isn't Negative the Opposite of Positive?

In De Anima, Aristotle wrote that "knowledge of the soul admittedly contributes greatly to the advance of truth in general and, above all, to our understanding of Nature" and noted further that "to attain any assured knowledge about the soul is one of the most difficult things in the world" (McKeon, 1973, p. 155).

We test designers, attempting to understand our "souls," face this difficult task when we develop survey instruments. We devise affirmative statements, targeted on our variable, which we expect respondents high in the trait will affirm. Our dream is that our respondents will treat the negative statements in a manner consistent with the way they affirm positive statements. If they "mildly agree" with a positive statement, they will "mildly disagree" with its opposite. Were this to happen, a smooth, linear variable would emerge when positives and reversed negatives are added together. Rasch analysis shows this does not happen.

This analysis reveals that to say " No " to a statement is not the equivalent of saying "Yes" to its opposite. IfI should strongly reject the statement, "I hate you," it does not follow that I would strongly endorse the statement, "Ilove you," or even, "I like you." A negative statement is not the opposite of a positive one.
There is No "Just" in "Just Say No!"
"No" is a big deal-an
important thing to say. Ask the mother of any two-year-old. Of the many ways we try to control of our lives, an important one is our ability to refuse, to abstain, to object, to fight back, to resist, to say "No!" We don't "just" say it randomly, without some preparation, some adjustment of our mental state. Biological, developmental, linguistic, and psychological necessities are the antecedents of this behavior.

In "On Negation" (1925), Freud understands negation of a thought as a way of denying that we could have ever had that thought, thereby allowing repressed ideation to enter our consciousness. By negation, we can think about forbidden ideas.

What others might think keeps us from confessing ideas we fear would cause us shame or disapproval. We can think about forbidden ideas by denying them or by joking about them. "Thou shalt not kill," presumes our capacity for such behavior. We fear death. Yet jokingly we say, "Oh, you'll die when you hear this!" or, "I almost died when he said that!" or, "It scared me to death!"

## When Less Is More: Separating Analyses

To understand negative vs. positive, I developed a longer self-esteem test from the Rosenberg items. The new test, "Thinking About Myself" (Table 2), has twenty items, ten negative and ten positive. The response format has four categories: "Strongly Disagree," "Disagree," "Agree," and "Strongly Agree."

Three forms were composed. In Form M (Table 2), the twenty negative and positive items were intermixed. In Form P, the ten positive items were given first, followed by the ten negative items. In Form N , the ten negative items were given first, followed by the ten positive items.

The forms were administered to graduate students. Some students took Form N, while others took Form P. All students took Form M, with items intermixed.

Responses from all three versions (Forms N , P , and M) were combined into one analysis. Responses were analyzed three ways: (1) responses to the 20 negative and positive items of all three forms together; (2) responses to the 10 negative items across all three forms; and (3) responses to the 10 positive items across all three forms. Because the category "Strongly Disagree" was hardly chosen for the positive items, "Strongly Disagree" and "Disagree" were combined. Responses to "Strongly Agree" and "Agree" were combined for the reversed-coded negative items.

Using the WINSTEPS computer program (Linacre, 2000) employing Rasch analysis, linear measures (logits) were constructed from raw scores. This made it possible to compare item calibrations across questionnaire versions. The analysis of combined positive and negative items yielded the "map" of items shown in Table 3.

spondents found it easier to reject negative items than to affirm positive items. The very hardest item was also a negative one. It was very hard to reject feeling "Worthless," although it
was easy to affirm being "Worthwhile." Being worthwhile was not seen as the opposite of being worthless. Only two negative items were successfully seen as the obverse of their positives: "Satisfied - Dissatisfied" and "Very proud - Not proud" which

When asked later, they said the questionnaire made them feel uncomfortable by confronting them immediately with a string of negative ideas about themselves. This was an unexpected, serendipitous observation, yet in line with what we

| Table 4Positive Items Only (Measure Order) |  | Table 5Negative Items Only (Measure Order) |  |
| :---: | :---: | :---: | :---: |
| Measure | Esteem Idea | Measure | Esteem Idea |
| 71.7 | Useful. . . . . . . . Hardest to Affirm | 81.7 | -Worthless. . . . . . .Hardest to Reject |
| 59.2 | Positive Attitude | 66.7 | -Do Poorly |
| 56.9 | Very Proud | 53.5 | -Not Proud |
| 51.8 | Accomplish Things | 52.3 | -Bad Qualities |
| 48.1 | Good Person | 50.9 | -Not Positive |
| 45.7 | Self Respect | 50.9 | -Dissatisfied |
| 44.4. | Satisfied | 41.1 | - Oseless |
| 44.4 | Successes | 40.0 | -No Self Respect |
| 39.3 | Worthwhile | 40.0 | -A Failure |
| 38.3 | Good Qualities . . . . Easlest to Affirm | 27.1 | -No Good. . . . . . . . . Easlest to Reject |

are close on the variable line. The inclusion of negative and positive items muddles our ability to interpret this analysis.

The story improves when we look at the positive and negative data side-by-side in measure order (Tables 4 \& 5).

By separating them, we can discuss more lucidly what the easy and hard items are on each subscale and better understand the story the respondents are telling us. When we draw arrows between the positive items and their negative counterparts, we see differences in location on the measure line. Most egregious are "Useful Useless," "Worthwhile - Worthless," and "Good qualities - Bad qualities." These so-called reversals evoked different reactions between positive and negative.

The Principal Components (Standardized Residual) Factor Plot and related analysis of the combined positive and negative items shows in another way how respondents reacted to the questionnaire. These two tables (Tables 6 \& 7) show us how the negative items drop like stones to the bottom of the analysis. The standardized residuals of the negative items, except for "No Good," are all in the bottom half of the factor loadings, indicating once again that respondents treated negative items differently from positive.

Some students were observed to be in distress while taking Form N of the questionnaire. They complained and squirmed in their chairs.
observed to be the impact of negative stimuli. Although (Tables $8 \& 9$ ) border on the astonishing, they are more understandable in light of that revelation from the students.




The results yielded by the principal components analysis of the students' responses to the positive items were very interesting. Both the pictorial representation of the plot (Table 8) and the table of standardized residual correlations (Table 9) are shown. For the positive items, all except one of the students who took Form N are located in the upper (positive) region of the factor loadings (in bold, with asterisks). Note that a large portion of the variance ( 17.06 units) is explained by this factor.

The principal components analysis for the negative items looks very different (Table 10). On that one, the Form N students are scattered among positive and negative loadings in the expected, random way. The dramatic reaction of Form N students to the negative item bombardment was manifested mainly when they responded to the positive items. We could not have learned this if we had not analyzed the positive and negative items separately.

These analyses demonstrate the difference between

| FACTOR 1 FROM PRINCIPAL COMPONENT ANALYSIS OF STANDARDIZED RESIDUAL CORRELATIONS (SORTED BY LOADING) -FACTOR 1 EXPLAINS 10.75 OF 41 VARIANCE UNITS |  |  |  |  |  |  |
| :---: | :---: | :---: | :---: | :---: | :---: | :---: |
|  |  | \|LO | ORE | $\begin{aligned} & \text { NEIT } \\ & \text { MNS } \end{aligned}$ | $\begin{array}{ll\|l\|} \text { STI } \\ \text { SQ } \end{array}$ | $\begin{aligned} & \text { Y } \\ & 3 R \end{aligned}$ |
|  | 1 | . 93 | 44.9 | . 21 | 17 |  |
|  | 1 | . 93 | 44.9 | . 21 | . 17 \|B | $17 \mathrm{nF2}$ * |
|  | 1 | . 93 | 44.9 | . 21 | . 17 \|c | $24 \mathrm{nma}+1$ |
|  | 1 | . 93 | 44.9 | . 21 | . 17 ID | $26 \mathrm{mF1}$ । |
|  | 1 | . 931 | 44.9 | . 21 | . 17 IE | 27 mF2 |
|  | 1 | . 93 | 44.9 | . 21 | . 17 IF | 37 mM 2 । |
|  | 1 | . 88 I | 105.7 | . 40 | . 14 I6 | 20 nrl * |
|  |  | . 88 | 105.7 | 40 | 14 \| ${ }^{\text {H }}$ | $42 \mathrm{mF1}$ । |
|  | 1 | . 69 | 56.0 | 1.04 | 1.03 II | $39 \mathrm{mF1}$ |
|  | 1 | . 42 | 75.3 | 1.11 | 1.14 J | 38 mm 4 |
|  |  | . 35 | 65.8 | . 97 | . 92 lk | $29 \mathrm{mM1}$ |
|  | 1 | . 35 | 65.8 | . 97 | . 92 [L | $30 \mathrm{mM1}$ |
|  |  | . 30 | 61.0 | 1.00 | 1.00 IM | 36 mm 1 |
|  |  | . 25 | 75.3 | . 95 | . 95 in | 40 mF 2 |
|  | 1 | . 23 | 61.0 | 74 | . 6410 | 6 pM1 |
|  | 1 | . 221 | 75.3 | 1.10 | 1.12 IP | $21 \mathrm{nmC} \times 1$ |
|  |  | . 20 | 94.3 |  | . 7718 | $9 \mathrm{pF1}$ |
|  |  | . 03 | 80.5 | . 00 | . 93 IR | 3 pM1 |
|  |  | -. 83 | 105.7 | 1.40 | 53 | 41 mM 1 |
|  | 1 | -. 72 । | 33.5 | 1.14 | 1.53 lb | 28 mM2 |
|  |  | -. 62 | 28.2 | 3.34 | 8.33 lc | $16 \mathrm{nma*}$ |
|  | 1 | -. 50 | 56.0 | 1.17 | 1.31 ld | 10 pM4 |
|  | 1 | -. 40 । | 43.8 | . 73 | . 58 Ie | $49 \mathrm{mM1}$ । |
|  |  | -. 39 | 61.0 | 2.29 | 2.20 If | 48 mF 4 |
|  |  | -. 39 | 33.5 | 1.14 |  |  |
|  | 1 | -. 36 | 56.0 | . 35 | . 29 lh | $23 \mathrm{nr2*}$ |
|  |  | -. 32 | 80.5 |  | . 4511 | 46 mM1 |
|  | 1 | -. 31 | 65.8 | 1.36 | 1.54 ij | 31 mF 2 |
|  | 1 | -. 301 | 56.0 | 1.35 | 1.35 lk | 32 mm 4 |
|  | 1 | -. 15 | 70.5 |  | . 6711 | 34 mF 2 |
|  | 1 | -. 14 । | 56.0 | 1.55 | 1.74 lm | $2 \mathrm{pM4}$ |
|  | 1 | -. 12 | 39.1 | 2.76 | 2.85 ln | $25 \mathrm{mM4}$ |
|  | 1 | -. 12 | 80.5 | . 88 | . 8510 | 45 mM 1 |
|  | 1 | -. 10 | 51.5 | . 45 | . 35 Ip | 11 nir2* |
|  | 1 | -. 10 | 80.5 | . 84 | . 8319 | $47 \mathrm{mF1}$ |
|  | 1 | -. 10 | 50.7 | . 42 | . 33 is | 1 pF 1 |
|  | 1 | -. 10 | 50.7 | . 42 | . 33 Ir | $35 \mathrm{mF3}$ |
|  | 1 | -. 10 | 65.8 |  | . 71 lt | 22 nma* |
|  | 1 | -. 08 | 70.5 | 1.79 | 2.0010 | 15 niz* 1 |
|  | 1 | -. 04 | 65.8 | . 80 | . 8217 | $33 \mathrm{mF2}$ |
|  | 1 | -. 04 | 75.3 | 1.03 | . 94 Is | $14 \mathrm{nr3}+1$ |

how we react to positive and negative statements. Remember, these are just sentences on a piece of paper, no curse words, no punches were thrown, no mud was slung - or so we thought.

Although a small study, this analysis revealed a surprising trend. The map of mixed items along the logit "ruler" is muddled by the inclusion of both negative and positive items. The "story" about what is easier and harder to believe about ourselves is immediately clearer when negative and positive items are analyzed separately. When we look at the principal components analysis of the mixed items, we see the negative items showing they are a separate factor.

The principal components analyses give us evidence
that the students who took Form N , with all the negative items first, experienced a common reaction. What was it? Anger? Anxiety? Depression? Whatever it was altered their behavior when they took the positive items both at the end of their Form N questionnaires and also mixed on the Form M questionnaire. For a few moments, Form N students were more like one another than they were before they undertook this task. This hints at the impact of other kinds of more active, traumatic negative experience, particularly educational evaluations.

The "Moral of the Story" is that negation is not the opposite of affirmation. Negativity has a powerful effect. Negative and positive items are not additive. This study brings out the inherent and previously unacknowledged confusion that occurs when we use an arithmetical maneuver to solve what is actually a profound psychological misunderstanding.

The usefulness of the idea of self-esteem (and many other "self" ideas examined over the years, such as motivation, aspiration, and sense of control) might not be at an end after all. What needs to be abandoned is the way survey instruments which attempt to target these variables are analyzed. Thoughtful analysis using Rasch methodology to construct useful measures from responses will make it possible to construct stable inferences from these old friends.

## References

Angell, F. (1907). On judgments of "like" in discrimination experiments. American Journal of Psychology, 18, 253-260.

Berke, R. L. (2000, February 20). Bush halts McCain in South Carolina by drawing a huge republican vote. The New York Times, p. 1.

Bloom, B. S. (1976). Human characteristics and school leaming. New York: McGraw-Hill.

Brookover, W. B., \& Thomas, S. (1964). Self-concept of ability and academic achievement. Sociology of Education, 37, 271-278.

Coopersmith, S. (1967). The antecedents of self-esteem. San Francisco: W. L. Freeman.

Epps, E. (1969). Correlates of achievement among northern and southern urban Negro students. Journal of Social Issues, 25, 55-71.

Freud, S. (1959). Negation. In L. Strachey (Ed. and Trans.) Sigmund Freud: Collected Papers. Vol. 5, (pp. 181-185). New York: Basic Books. (Original work published 1925)

Johnson, K. (1998, May 5). Self-image is suffering from lack of esteem. The New York Times, p. B12.

Linacre, J. M. (2000). WINSTEPS (Version 2.98) [Computer software]. Chicago: MESA Press.

Linacre, J. M., \& Wright, B. D. (1998). A user's guide to BIGSTEPS WINSTEPS: Rasch-model computer programs. Chicago: MESA Press.

McKeon, R. P. (1973). Introduction to Aristotle (2nd ed.) (pp. 153. 245). Chicago: University of Chicago Press.

Purkey, W. W. (1970). The self and academic achievement. Englewood Cliffs, NJ: Prentice-Hall.

Rasch, G. (1993). Probabilistic models for some intelligence and attainment tests. Chicago: MESA Press. (Original work published 1960) Rosenberg, M. (1965). Society and the adolescent self-image. Princeton, NJ : Princeton University Press.

Rosenberg, M. (1979). Conceiving the self. New York: Basic Books. Wright, B. D., \& Masters, G. N. (1982). Rating scale analysis: Rasch measurement. Chicago: MESA Press.

Wright, B. D., \& Stone, M. H. (1979). Best test design: Rasch measurement. Chicago: MESA Press.

