Characteristics of Single-Item Measures in Likert Scale Format

Aliosha Alexandrov
University of Wisconsin, Oshkosh, Wisconsin, USA
alexanda@uwosh.edu

Abstract: The use of single-item measures has been encouraged by several authors asserting that single-item measures are appropriate and can substitute multiple-item measures in many cases. This study focuses on the characteristics of single-item measures in Likert scale format. There are two motives behind it: first, the Likert scale has been called problematic and its usage discouraged by the very proponents of single-item measures; and second, the reverse wording of Likert items has led to many problems with multiple-item measures. Because the Likert scale is one of the most used scales in marketing and management, and more researchers may decide to use single-item measures in Likert scale format, it becomes necessary to answer the question if it is usable or not. This research scrutinizes the characteristics of the Likert scale in a positive-negative continuum: from positive to negative with different levels of intensities. Based on collected sample data for three popular computer brands, the main conclusion is that only positively worded Likert items with a fairly high level of intensity should be used as single-item measures. The supporting empirical evidence includes: (1) positively and negatively worded items are not true opposites, (2) items with reversed scores inflate means, (3) items with neutral intensity have unique conceptual meaning, (4) dependent variables are predicted best by independent variables with similar intensity and (5) negatively worded items contain a method factor that limits their ability to capture the measured concept. The results also suggest that the effect of the method factor is expressed more when respondents are not familiar with the object of the measured concept. The findings in this study provide guidelines for the practical use of measures in Likert format. Scales in other formats should undergo similar scrutiny.

Keywords: single-item measures, Likert scale, negatively-worded items, reversed items, C-OAR-SE

1. Introduction

The dominant paradigm of multiple-item scale development in marketing, as advanced by Churchill Jr. (1979), has been challenged by several authors (Drolet and Morrison 2001; Rossiter 2002; Bergkvist and Rossiter 2007). The main argument is that multiple-item measures are not always necessary and can be substituted by single-item measures in many cases. The C-OAR-SE procedure suggested by Rossiter (2002) has become the focal point of the recent debate. C-OAR-SE is a comprehensive methodology for the development of multiple- and single-item measures claiming to surmount some of the current pitfalls in scale development in marketing. The focus in C-OAR-SE is content validity. Rossiter (2005; 2008) claims that this is the only validity needed in scale development, and the typical item purification through statistical procedures is unnecessary because it can change the meaning of the measured concepts. Therefore, according to Rossiter (2005; 2008), if a scale has a precise definition, there is no need to examine other types of validities. An important practical assertion advanced by C-OAR-SE is that most concepts (e.g., purchase intentions) are concrete and understood unequivocally by raters and there is no need to use multiple-item scales to measure them; a single-item measure is sufficient.

Although the C-OAR-SE procedure is a solid argument in defense of single-item measures, it opens the door for a possible misconception. Specifically, Rossiter (2002; 2008) argues that the only appropriate scale for single-item measures is the semantic-differential scale. The Likert scale was called problematic and its use was discouraged because of the lack of a neutral point. According to Rossiter (2008, p.383), the Likert scale "produces hopelessly fuzzy scores." The danger with this assertion is that it annihilates a whole category of scales from the marketing research. It seems that the new attempt to relax scale development was restricted again. Furthermore, Rossiter does not provide comprehensive empirical or theoretical justification for his recommendation, except the lack of "psychological zero" and conceptual meaning.

Addressing some of the raised concerns, the purpose of this study is to clarify the use of single-item measures in Likert scale format. The Likert scale is one of the most popular scales in marketing and its status is unlikely to change. Therefore, as more researchers may decide to use it as a single-item measure, it becomes necessary to examine the characteristics of the Likert scale more closely. The intention of this study is *not* to avow the superiority of the Likert scale to any other scales, but to

Reference this paper as:
Alexandrov, A. "Characteristics of Single-Item Measures in Likert Scale Format" *The Electronic Journal of Business Research Methods Volume 8 Issue 1 2010, (pp.1 - 12), available online at www.ejbrm.com*

describe its behavior and provide practical recommendations. The research question this study answers is: Are single-item measures in Likert format usable?

2. Review of positive-negative asymmetry

Nunnally (1978) suggested that positively-worded items in Likert scales can be transformed into negatively-worded items and their scores can be reversed symmetrically afterwards. This practice continues even today, although it has been known that negatively-worded items introduce problems in multiple-item scales. Negatively-worded items often form a separate factor, independent of the main factor, and change the dimensionality of the construct (Herche and Engelland 1996; Mook et al. 1991; Tomas and Oliver 1999). Factors based on negatively-worded items have strong method effects and exhibit longitudinal invariance (Motl and DiStefano 2002; Horan et al. 2003). Negatively-worded items tend to lower the reliability of multi-item scales as measured by Cronbach's alpha by as much as 20% (Schriesheim et al. 1991; Barnette 2000), and confound measures in cross-cultural research, hampering measurement invariance (Wong et al. 2003). All of the above contribute to the positive-negative asymmetry, which is reviewed in the following section.

Positivity and negativity are not symmetrical: negative information weights more than positive information (Anderson 1965; Rodin 1978), and positive and negative affective states have low correlation (Diener and Emmons 1984; Watson et al. 1988). Cacioppo and Berntson (1994) advanced the concept of bivariate evaluative space, where positivity and negativity are distinct entities that can coexist independently. Cacioppo et al. (1999) summarized that the underlying cognitive processes are bivariate (i.e., positivity and negativity are different concepts), but the limiting physical conditions make them appear as bipolar (i.e., positivity and negativity are true opposites).

The biases associated with the processing of positive information were called the *positivity bias* (Markus and Zajonc 1985), and the biases associated with the processing of negative information were called the *negativity bias* (Kanouse and Hanson Jr. 1987). The *positivity bias* is a cognitive process referring to humans' readiness to generate positive content (Peeters and Czapinski 1990). One aspect of the bias is its linguistic expression. There are more positive than negative words in vocabularies, people ascribe more positive descriptions to a target, and it is common to have un-favorable terms defined as opposites of favorable terms (Adams-Webber 1997; Benjafield 1985; Matlin and Stang 1978; van Dijk et al. 2003). The *positivity bias* is an *a priori* hypothesis about reality; people approach or search for events expecting to find positivity (Peeters 1971; Markus and Zajonc 1985). This internal drive is called the unconditional optimism (Czapinski 1985). A unique aspect of the positivity bias is that it has a strong *subjective* component, which if reduced, eliminates the bias (Aderman 1969).

The *negativity bias* can be summarized in four ways: (1) negative evaluations are stronger than equivalent positive evaluations, (2) negative intensity increases faster than positive intensity when approaching corresponding events, (3) the combination of positive and negative stimuli results in a more negative result than their algebraic sum, and (4) negative events lead to more complex cognitive processes (Rozin and Royzman 2001).

Some unique findings in positive-negative asymmetry are worth mentioning. Positive events are with higher frequency, but less urgent (Rozin and Royzman 2001). The processing of information under negative mood is more systematic and accurate than under positive mood. Negative events provoke more causal attribution than positive events (Bohner et al. 1988). Processing positive information is more *subjective,* while processing negative information is more *objective*, analytical, and complex (Peeters and Czapinski 1990).

Considering the positive-negative asymmetry, it is obvious that the Likert and semantic-differential scales are different representations of the measured concepts. The Likert scale captures the presence or absence of a concept, and it does not necessarily assume that every concept has a corresponding opposite. On the other hand, the semantic-differential scale assumes that a concept always is restricted by two symmetrical opposite characteristics.

3. Theoretical advancement

The basic premise I step on distinguishes between the positivity bias under positive wording and the negativity bias under negative wording in Likert scales. I take a gradual approach and examine the properties of the Likert scale in a positive-negative continuum: from positive to negative with different

levels of intensities. Intensity herein refers to the extremeness of the argument with which a respondent needs to agree/disagree.

Positive and negative affective states can coexist and do not form a single dimension (Schimmack 2005). It is possible to elicit mixed feelings by positive and negative stimuli (Schimmack and Colcombe 2007). This asymmetry applies to many concepts, including liking and disliking (Herr and Page 2004), attitudes in general (Cacioppo and Berntson 1994), and others. Replicating previous studies, the following hypothesis is advanced:

H1: Positively and negatively worded items are not true opposites.

Neutrality, psychological zero, is a key attribute of semantic-differential scales, but it does not have the same significance in Likert scales. Because it is technically possible to have a Likert scale with a neutral wording, this category is examined as well. According to Edwards (1946), neutral items are non-differentiating because neutrality may be interpreted by raters as: ambiguity, ambivalence, irrelevance, or indifference. In addition, the logic dictates that a neutral Likert item should not be correlated with any other item. The existence of a non-zero correlation with an item, even in the slightest direction, suggests that a 'neutral' item is not really neutral.

H2: Items with neutral wording have a distinct conceptual meaning compared to items with non-neutral wording.

Often, an item is worded negatively, and its score is reversed to transform it into a positive item. This is a common practice based on the assumption that positivity and negativity are symmetrical. However, people are more loss aversive than gain oriented, and losses "loom" larger than gains (Kahneman and Knetsch 1991; Tversky and Kahneman 1991). Because negative information weights more than positive information, items with negative wording will be disagreed with more than items with positive wording. Therefore, a reversed negative item will have a higher mean than the corresponding positive item, and a reversed positive item will have a higher mean than the corresponding negative item.

H3a: Reversed items have higher means.

People have a natural tendency to prefer to remain in the status quo because the disadvantages of moving away may be larger than the advantages. This is called the status quo bias (Samuelson and Zeckhauser 1988). Similarly, when customers make a choice, they prefer a product with moderate attributes than with extreme attributes (Simonson and Tversky 1992). This implies that respondents would be more uncomfortable with extreme intensity items than with moderate intensity items because extreme statements are further away from the status quo. As a result, there will be a tendency to agree/disagree less with extreme items. Partial support for this contention comes from the negativity bias literature, where the negativity effect is pronounced more when intensity of the negative stimulus is higher (Czapinski 1986; Peeters and Czapinski 1990). Combined with Hypothesis 3a, this means that differences between the means of reversed items increase incrementally as intensity increases.

H3b: The difference between the means of reversed and non-reversed items with the same intensity increases as intensity increases.

The variability produced by a Likert scale can be decomposed into variability due to the main concept and a method factor, which I call a positive-negative method factor (PNMF). PNMF is the result of the cognitive biases due to the positive-negative processing asymmetry. Each Likert item will contain variance due to the measured concept and variance due to PNMF. For example, the differences in satisfaction scores obtained by positively or negatively worded items will be due to the actual differences of satisfaction among participants and the positive-negative biases elicited by the scale based on its wording. The presence of PNMF will attenuate the correlation of any two items if their intensities are different, but will increase the correlation if their intensities are similar. Therefore, PNMF can affect the predictive ability of scales.

H4: All else being equal, a dependent variable will be predicted best by independent variables with similar wording and intensity.

People tend to agree more on the negative (Bosson et al. 2006). "If positive evaluations reflect subjective preferences, which may vary according to the subjects' tastes while negative evaluations are more controlled by objective cues, then subjects may be expected to agree more on negative than on positive evaluations" (Peeters and Czapinski 1990, p.49). That is, more objective processing of items with negatively worded items can lead to increased agreement; therefore, the effect of PNMF will be expressed more uniformly on them.

H5: The effect of the positive-negative method factor is stronger on negatively-worded items than on positively-worded items.

4. Methodology

Data was collected from 153 undergraduate students at a mid-western university. The number of males and females was 76 and 77, respectively. The theoretical framework used to test the hypotheses was the theory of reasoned action (Ajzen and Fishbein 1980). It is a well established theory, which states that attitudes affect purchase intentions, which in turn affect purchase behavior. The collected data includes attitudes, purchase intentions, and simulated behavioral choice for three popular computer brands: Dell, HP, and Gateway. The reason for selecting three brands instead of one brand was to assure that the obtained results are not artifacts of a single brand.

According to C-OAR-SE, when a concept is concrete, it can be measured with a single-item scale. The concreteness criterion means that a concept should not be measured as being manifested or formed by other indirect concepts because all raters understand it unequivocally. Therefore, to be measured, a concrete concept should be included directly in the scale. For example, measuring attitude by a Likert single-item scale can take the form of "My attitude with BrandX is positive." If a concept is not included in the wording of the scale, it would mean that there is a description that conveys a better meaning than the concept itself, which automatically raises the question if there are other suitable or even better descriptions, countering the concreteness argument.

Because the focus of the study was on the use of single-item measures, the participants were asked straight questions about their attitudes and purchase intentions. Attitudes and purchase intentions were measured using Likert scales by asking the same question with five different levels of intensity from positive to negative. One semantic-differential question was included as a control variable for one of the tests. Purchase intention was related to the confidence in the brand of the rater's next computer. The simulated behavioral choice was measured by asking respondents to pick a brand. Four choices were given: the three aforementioned brands and 'Other.' The three brand choices contained the brand logos and a typical desktop image from the corresponding brand so that more brand related cues were evoked. The simulated behavior was coded as a binary variable: one if the brand was selected and zero if it was not selected. All measures are listed in the appendix. The descriptive statistics of the measured variables are presented in Table 1.

5. Results

Hypothesis 1 was tested by calculating the linear correlations among all attitudes and purchase intentions variables for the three brands (Table 1). The 95% confidence intervals of the correlations of all opposite pairs of items, based on the Fisher's z' score transformation, were considerably above -1.00. For example, the correlation of the most positive and negative attitude items for HP is -.51, which at a sample size of 153, leads to a 95% confidence interval ranging from -.61 to -.38. The average correlation of all pairs with most opposite intensities was -.54, and the average correlation of all opposite pairs with medium intensity was -.40. Correlations of such magnitude, although negative, are not evidence that positively and negatively worded items are true opposites. The averaged correlation of -.54 corresponds to a shared variance (squared correlation) of .29, which is not sufficient to suggest equivalence. These results are interpreted as support for Hypothesis 1.

To identify the unique nature of items with neutral intensity, Hypothesis 2 was tested using exploratory factor analysis with maximum likelihood extraction and direct oblimin rotation. The use of factor analysis on items measuring one concept but with different intensity is appropriate. If such items do not load on a single factor, it could be interpreted as either (1) they do not have much in common and are essentially different, or (2) the effect of other factors is so predominant that it overtakes the main common variance of the measured concept.

Table 1: Descriptive statistics of all variables

Dell	M	SD	A_5	A_4	A_3	A_2	A_1	PI_5	PI_4	PI_3	PI_2	A_1
A_5	3.31	1.17	-									
A_4	3.34	1.03	0.50	-								
A_3	3.13	1.21	-0.33	-0.04 ^{ns}	-							
A_2	2.26	1.08	-0.59	-0.35	0.21	-						
A_1	1.77	1.10	-0.52	-0.41	0.14 ^{ns}	0.82	-					
PI_5	2.54	1.32	0.65	0.26	-0.30	-0.45	-0.37	-				
PI_4	2.57	1.16	0.54	0.43	-0.17	-0.43	-0.38	0.82	-			
PI_3	2.99	1.27	0.06 ^{ns}	0.19	0.23	-0.22	-0.21	0.14 ^{ns}	0.35	-		
PI_2	2.93	1.31	-0.46	-0.26	0.31	0.59	0.50	-0.67	-0.63	-0.27	-	
PI_1	2.74	1.53	-0.52	-0.27	0.29	0.62	0.60	-0.74	-0.74	-0.37	0.87	-
A_{SD}	3.54	0.96	0.71	0.35	-0.28	-0.53	-0.45	0.65	0.58	0.16	-0.46	-0.53
HP	M	SD	A_5	A_4	A_3	A_2	A_1	PI_5	PI_4	PI_3	PI_2	A_1
A_5	3.34	0.97	-									
A_4	3.30	0.87	0.59	-								
A_3	2.89	0.98	-0.25	0.06 ^{ns}	-							
A_2	2.19	0.99	-0.47	-0.30	0.19	-						
A_1	1.81	1.06	-0.50	-0.42	0.07 ^{ns}	0.69	-					
PI_5	2.65	1.13	0.56	0.39	-0.13 ^{ns}	-0.35	-0.36	-				
PI_4	2.75	1.05	0.52	0.50	-0.07 ^{ns}	-0.29	-0.33	0.88	-			
PI_3	3.15	1.18	0.10	0.26	0.31	-0.06	-0.20	0.26	0.39	-		
PI_2	2.67	1.16	-0.39	-0.28	0.29 ^{ns}	0.51	0.41	-0.49	-0.49	-0.14 ^{ns}	-	
PI_1	2.48	1.38	-0.42	-0.37	0.16 ^{ns}	0.54	0.58	-0.58	-0.59	-0.30	0.83	-
A_{SD}	3.46	0.75	0.47	0.43	-0.16 ^{ns}	-0.28	-0.31	0.58	0.54	0.18	-0.31	-0.40
Gateway	M	SD	A_5	A_4	A_3	A_2	A_1	PI_5	PI_4	PI_3	PI_2	A_1
A_5	2.67	1.00	-									
A_4	2.79	0.93	0.79	-								
A_3	3.13	1.20	0.16	0.30	-							
A_2	2.58	1.11	-0.41	-0.30	-0.11 ^{ns}	-						
A_1	2.40	1.25	-0.43	-0.42	-0.21	0.79						
PI_5	2.07	1.00	0.67	0.52	0.13 ^{ns}	-0.21	-0.19	-				
PI_4	2.21	0.95	0.56	0.57	0.25	-0.13 ^{ns}	-0.14 ^{ns}	0.89	-			
PI_3	2.99	1.29	0.22	0.33	0.51	-0.03 ^{ns}	-0.11 ^{ns}	0.41	0.56	-		
PI_2	3.27	1.23	-0.31	-0.20	-0.07 ^{ns}	0.58	0.48	-0.43	-0.38	-0.18	-	
PI_1	3.17	1.40	-0.42	-0.35	-0.18	0.57	0.60	-0.49	-0.47	-0.31	0.85	-
A_{SD}	2.88	0.80	0.52	0.47	0.13	-0.45	-0.45	0.37	0.34	0.23	-0.37	-0.39

^{ns} Not Significant at *p-value < 0.05*
M – mean, SD – standard deviation
A – attitude, PI – purchase intention
Index 5 refers to the highest positive intensity, index 1 refers to the highest negative intensity, and index 3 refers to a neutral intensity. Index $_{SD}$ indicates overall attitude measured by a semantic-differential scale.

Table 2: Exploratory factor analysis (structure matrix)

	Factors						
Dell	1	2	3	4	5	6	7
A_5	.60	**-.66**	-.01	-.03	-.23	.03	-.13
A_4	.28	**-.52**	-.15	-.06	.10	.16	-.15
A_3	-.31	.16	-.03	.12	**.41**	.30	-.05
A_2	-.46	**.86**	-.03	.26	.04	.22	-.04
A_1	-.39	**.87**	-.07	.23	-.02	.15	-.14
PI_5	**.92**	-.41	.00	-.10	-.11	-.09	-.20
PI_4	**.86**	-.44	-.10	-.10	.20	.01	-.29
PI_3	.25	-.21	-.20	-.13	**.69**	-.10	-.13
PI_2	**-.78**	.51	-.03	.24	-.09	.49	.06
PI_1	**-.85**	.57	-.01	.25	-.16	.38	.08
HP							
A_5	.02	-.11	**-.72**	-.15	-.05	-.06	-.07
A_4	-.10	-.06	**-.62**	-.11	.19	.05	-.05
A_3	.00	.18	.17	.12	**.31**	.21	-.12
A_2	-.05	.16	**.49**	.44	.12	.41	-.13
A_1	-.07	.16	**.53**	.48	-.07	.28	-.13
PI_5	.07	.08	**-.84**	-.19	.12	-.22	-.25
PI_4	.05	.13	**-.84**	-.16	.27	-.22	-.28
PI_3	.14	.07	-.27	-.04	**.68**	-.13	-.13
PI_2	-.11	.16	**.59**	.39	.04	.72	.05
PI_1	-.10	.12	**.69**	.49	-.16	.62	.05
Gateway							
A_5	.12	-.11	-.18	-.46	.06	.15	**-.76**
A_4	.03	-.09	-.11	-.41	.25	.25	**-.69**
A_3	-.14	-.02	.02	-.18	**.62**	.12	-.21
A_2	-.13	.22	.23	**.85**	.07	.29	.20
A_1	-.11	.23	.17	**.89**	-.12	.16	.21
PI_5	.28	.02	-.20	-.18	.19	-.15	**-.92**
PI_4	.24	.08	-.15	-.11	.39	-.14	**-.90**
PI_3	.08	-.01	-.11	-.06	**.70**	-.09	-.44
PI_2	-.34	.17	.09	**.61**	-.06	.68	.34
PI_1	-.35	.11	.16	**.69**	-.24	.55	.43

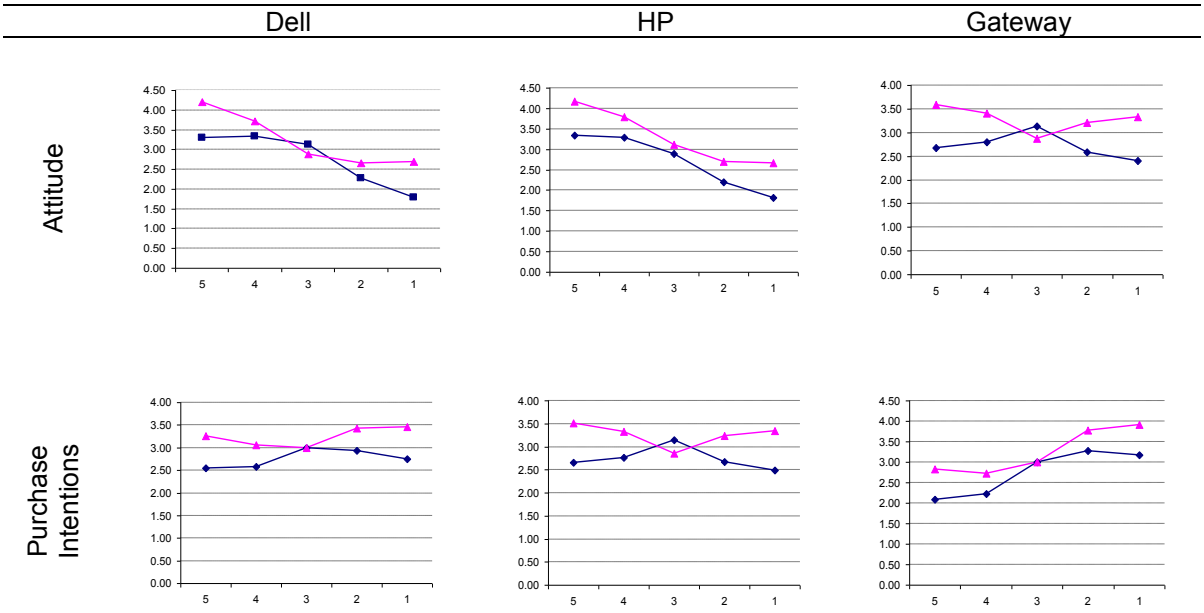
A – attitude, PI – purchase intention

The explained variance is 68%

All attitudes and purchase intentions variables for the three brands were included in the analysis (Table 2). Ideally, all attitudes and purchase intentions items for the three brands should load on distinct factors, and according to Hypothesis 2, the neutral items should not be associated with either of them. The results from the factor analysis revealed that all neutral items loaded on a single distinct factor. The interpretation of the other six is not straightforward, and more attention is devoted to this in the non-hypothesized results section. However, the solid neutral factor that emerged across the three brands is interpreted as support for Hypothesis 2. Hypothesis 3a states that all reverse means are higher than the means of the opposite items. Hypothesis 3b states that the difference between reversed and non-reversed means increases incrementally from the neutral point. The visual inspection of the plotted means of all measures (Figure 1) suggests that the two hypotheses have merit. The results of multiple paired-samples t-tests provided full support (not presented to save page space). First, I tested the significance of the difference between reversed and non-reversed items for each of the five levels of intensity. All differences were significant with the exception of the neutral point. Next, using paired-sample t tests, I tested if the differences of reversed and non-reversed means differed significantly across adjacent intensities. All of them were significant, meaning that the

mean differences between reversed and non-reversed items increase incrementally from the neutral point. Therefore, Hypotheses 3a and 3b are supported.

Figure 1: Comparison of means and reversed-means



- Means; - Reversed means;
Index 5 refers to the highest positive intensity, index 1 refers to the highest negative intensity, and index 3 refers to a neutral intensity.

Preliminary analysis of the correlations between attitudes and purchase intentions for each brand suggested the validity of Hypothesis 4. The highest correlations were at the diagonal, suggesting that the correlations between items with similar intensity were highest (Table 1). The analysis was continued using multiple linear regressions. Consistent with Ajzen and Fishbein (1980), I used attitudes as predictors of purchase intentions. Five regressions were estimated for each brand, corresponding to the five levels of intensity of purchase intentions. Each level of purchase intentions was regressed on all five attitude items with different intensities (i.e., $PI_{ji} = f(A_{j5}, A_{j4}, A_{j3}, A_{j2}, A_{j1})$, where i is intensity index and j denotes brands). As expected, the significant independent variables for each regression tended to be those with intensities similar to the intensity of the dependent variable. This pattern is visible in Table 3.

Table 3: Regression of purchase intentions on attitude

Dell	A_5	A_4	A_3	A_2	A_1	R^2
PI_5	0.68*	-0.10	-0.09	-0.13	0.03	0.43
PI_4	0.34*	0.24*	-0.03	-0.18	0.03	0.33
PI_3	-0.11	0.18	0.28*	-0.30	-0.02	0.14
PI_2	-0.12	-0.03	0.19*	0.52*	0.07	0.39
PI_1	-0.28	0.07	0.16	0.37*	0.38*	0.46
HP	A_5	A_4	A_3	A_2	A_1	R^2
PI_5	0.53*	0.10	0.00	-0.09	-0.04	0.32
PI_4	0.35*	0.35*	0.00	-0.03	-0.02	0.32
PI_3	0.02	0.24	0.37*	0.09	-0.21	0.18
PI_2	-0.07	-0.18	0.25	0.43*	0.06	0.33
PI_1	-0.04	-0.23	0.13	0.33*	0.43*	0.40
Gateway	A_5	A_4	A_3	A_2	A_1	R^2
PI_5	0.72*	-0.01	0.03	-0.02	0.11	0.47
PI_4	0.35*	0.32*	0.11	-0.02	0.15	0.40
PI_3	0.04	0.28	0.49*	0.08	0.03	0.30
PI_2	-0.19	0.12	-0.01	0.54*	0.06	0.34
PI_1	-0.29	0.05	-0.07	0.30*	0.36*	0.41

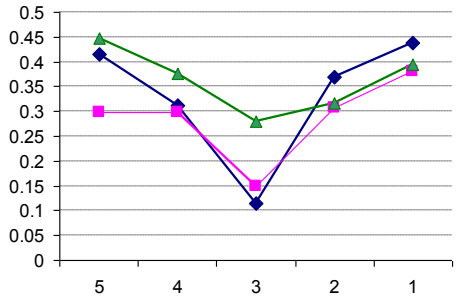
* Significant at *p-value* < *0.05*
A – attitude, PI – purchase intention

The fact that only a few independent variables were significant per regression was due to multicollinearity, which in this case was desirable. In regression with multicollinearity, the item that shares the highest level of variance with the dependent variable will be significant and will 'kick-out' other independent variables that share similar but less variance with the predicted variable. This happens because the variance of non-significant predictors already is accounted for by the items with highest covariance with the independent variable. The presence of the common component among items measuring the same concept can lead to multicollinearity, and any difference in their variances should be due to PNMF. Therefore, one can expect that it is this difference that will affect the inclusion or exclusion of items in the regression. The diagonal pattern observed in Table 3 demonstrates that PNMF affects the predictive ability of measures with similar intensity. Therefore, Hypothesis 4 is supported. One unexpected finding is the U shaped pattern of R^2 in the positive-negative continuum, which suggests not only that items with similar intensities are predicted better, but also that the explained variance increases as a function of intensity.

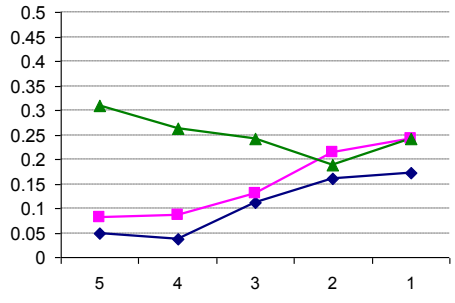
Hypothesis 5 states that the effect of the positive-negative method factor is more pronounced in negatively-worded items. To test this claim, it was necessary to separate the two sources of variability as captured by Likert items. An alternative measure not affected by the positive-negative method factor present in Likert scales in the same way was required. For this purpose, I measured the attitude toward each brand by a single-item semantic-differential scale in addition to the Likert scale. Semantic-differential measures also may be affected by a method factor specific to them (not discussed in this study), but a semantic-differential item measuring attitude also would capture variability due to the attitude it measures. Consequently, if the Likert item is regressed on the semantic-differential item, then the covariance due to the common concept will be accounted for by the semantic-differential scale and the residuals produced by the regression will contain mostly variability due to PNMF. If it is true that PNMF is expressed more in negatively worded items, then the explained variance in purchase intentions as a function of attitude residuals will be higher for negatively worded than for positively worded purchase intentions. The shared attitude component from all Likert attitude measures was removed by regressing them on the semantic-differential attitude measure, and the resulting residuals containing the variability of PNMF were saved. Then all fifteen regressions, as in Table 3, were repeated and purchase intentions of all intensities were regressed on the saved residuals (i.e., $PI_{ji} = f(ResA_{j5}, ResA_{j4}, ResA_{j3}, ResA_{j2}, ResA_{j1},$ where i is intensity index, j denotes brands, and $ResA$ is attitude residuals). In this case, any explained variance of the purchase intention items would be due to the remaining variance of PNMF in the residuals. The resulting R^2_{res} (variance explained by residuals) were compared to R^2_A (variance explained by the attitude Likert items) of the original measures. Attenuation coefficients were calculated as the ratio of R^2_A and R^2_{res} and their graphs are presented in Figure 2.

Figure 2: Explained variances in purchase intentions

a) Explained Variance by Attitudes (R^2_A)



b) Explained Variance by Attitude Residuals (R^2_{res})



Attenuation coefficient (R^2_A/ R^2_{res})

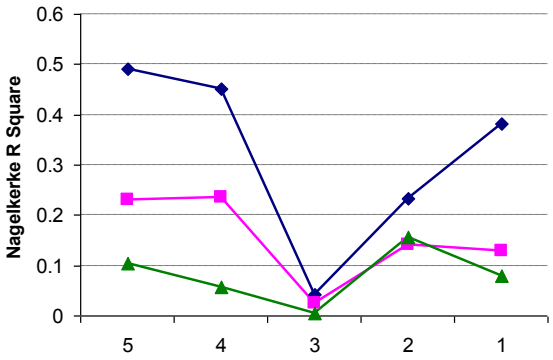
	5	4	3	2	1
Dell	8.49	8.43	1.02	2.30	2.54
HP	3.68	3.48	1.14	1.44	1.57
Gateway	1.45	1.44	1.15	1.69	1.63

Del, HP; Gateway;

Index 5 refers to the highest positive intensity, index 1 refers to the highest negative intensity, and index 3 refers to a neutral intensity.

The results demonstrated that the explained variance of the positive purchase intention items is reduced more than the explained variance of the negative purchase intention items. This result is consistent for Dell and HP. For example, the attenuation coefficient of Dell for the item with highest positive intensity tells that R^2_{res} is 8.48 times attenuated compared to its original value R^2_A. The effect for HP is less pronounced, with a coefficient of 3.68. In comparison, the attenuation coefficient for the item with highest negative intensity is only 2.54 for Dell and 1.57 for HP. The attitude residual explained 18% and 25% in negative purchase intention items and 5% and 7% in positive purchase intention items for Dell and HP, which makes a strong case that PNMF is contained more in negatively worded items than in positively worded items. The exception is Gateway, where the removal of the attitudinal component doesn't seem to affect much R^2_{res}. Therefore, Hypothesis 5 is partially supported. It should be noted that the only significant attitude residuals exhibited a diagonal pattern identical to the pattern in Table 3. This result provides additional support that PNMF plays a critical role in the relationships of Likert scale items.

Two interesting results obtained in a post-hoc analysis are worth mentioning. First, the predictive ability of purchase intentions items with different intensity on purchase behavior was examined. Using binary logistic regression, the purchase behavior (PB) for each brand was regressed on all purchase intentions with different intensities, one independent variable at a time (i.e., $PB_j = f(PI_{ji})$, where i is intensity index and j denotes brands). Fig. 3 shows the Nagelkerke R^2 of fifteen binary logistic regressions for the three brands. It should be noted that binary logistic regression does not have a real R^2 reflecting goodness-of-fit. The Nagelkerke R^2 is a pseudo-measure of R^2 representing strength of association. However, as it ranges from 0 to 1, it is a good indication about the overall quality of the regression. The items with higher intensity on Fig. 3 explained more variance than those with moderate intensity, and the effect is slightly more pronounced for positive items. Similar results were obtained during the test of Hypothesis 4, where explained variance was highest for regressions including items with higher intensity, no matter whether positively or negatively worded.



Del, HP; Gateway;
Index 5 refers to the highest positive intensity, index 1 refers to the highest negative intensity, and index 3 refers to a neutral intensity.

Figure 3: Explained variance in purchase behavior as a function of the intensity of purchase intention items

Second, when measuring the simulated purchase behavior, of the 153 respondents, 89 selected Dell, 39 selected HP, 9 selected Gateway, and 36 selected 'Other.' In addition, half of the respondents indicated that they were slightly to not at all familiar with Gateway. Considering these results, a logical question is: To what was the variability in the Gateway scales due, if respondents were not very familiar with this brand? It seems that when a person does not have an opinion, he/she will respond to a positive item in a way he/she usually responds to all positive items, and to a negative item in a way he/she usually responds to all negative items. As a result, the variance captured by measures will be dominated by PNMF, which can significantly affect dimensionality of constructs. Indeed, Table 1 reveals that the variables for Dell loaded as expected: attitudes and purchase intentions loaded on separate factors. For HP, the results were inconclusive; but for Gateway, both attitude and purchase intention items split and loaded on positive and negative factors. The results have a pattern corresponding to the level of brand familiarity of the respondents. For the familiar brand, the

dimensionality was as expected, but for the unfamiliar brand, the dimensionality seems to have been affected heavily by PNMF. In other words, when respondents do not have an opinion and construct it on the spot, a negative factor is likely to emerge.

6. Discussion

Several important results were found. First, it was replicated that opposite items are not true opposites. Second, the means of reversed items are higher, and the differences between them increase as a function of item intensity. This difference achieved almost 20% on the most positive side for Dell. Third, independent variables predict better dependent variables with similar intensity, and the stronger the intensity, the more the explained variance seems to be. Finally, PNMF can account for up to 30% of the variance in a dependant variable, even after the removal of the main component as an actual predictor. Obviously, the implications of such results are significant, and the question begging an answer is if Likert scales are usable. My stance is that despite the results, Likert scales have a place in single-item measures. Considering that (1) the natural tendency of humans toward the positive and their positivity bias as an *a priori* hypothesis about reality; (2) the subjective nature of the positivity bias compared to the objective nature of the negativity bias, leading to a higher level of the measured concept captured by positive items; and (3) the better predictive ability of positive items and their normal sound, then the logical conclusion is that a *Likert scale should be positively-worded with a fairly high level of intensity.* Finally, the reviewed literature and the presented empirical evidence demonstrate that the bipolar assumption of the semantic-differential scale is not necessarily true. The presentation of positivity and negativity as opposites may be restraining. Also, the unique nature of the neutral factor is a warning sign against focusing on neutral points as basis for justification. Therefore, semantic-differential and other scales, which use positive and negative wording, should undergo similar scrutiny. Different biases and method effects specific to them may be discovered.

7. Appendix 1: Survey instruments

Question	Abbreviations used in the data analysis [c]
Attitude (Likert)[a]	
My attitude toward Dell is very positive	A_5
My attitude toward Dell is somewhat positive	A_4
My attitude toward Dell is neutral	A_3
My attitude toward Dell is somewhat negative	A_2
My attitude toward Dell is very negative	A_1
Attitude (semantic-differential) [b]	
What is your overall attitude toward DELL	A_{SD}
Purchase Intentions (Likert) [a]	
I am very confident that my next purchase of a computer will be Dell	PI_5
I am somewhat confident that my next purchase of a computer will be Dell	PI_4
I am not sure if my next purchase of a computer will be Dell	PI_3
I am somewhat confident that my next purchase of a computer will NOT be Dell	PI_2
I am very confident that my next purchase of a computer will NOT be Dell	PI_1
Purchase Behavior	
If you needed a computer today which brand would you purchase? (click on a picture)	
	
[a] The answer format was from 'Strongly Disagree' to 'Strongly Agree'; [b] The answer format was from 'Very Negative' to 'Very Positive'; [c] Index 5 refers to the highest positive intensity, index 1 to the highest negative intensity, and index 3 refers to a neutral intensity. SD - semantic differential A – attitude, PI – purchase intention	

References

Adams-Webber, J. R. (1997) "Positive-Negative Asymmetry in the Evaluation of Familiar Versus Unfamiliar Persons and Objects", *Journal of Constructivist Psychology*, Vol 10, No. 2, pp139-152.

Aderman, D. (1969) "Effects of Anticipating Future Interaction on the Preference for Balanced States", *Journal of Personality and Social Psychology*, Vol 11, No. 3, pp214-219.

Ajzen, I. and Fishbein, M. (1980) *Understanding Attitudes and Predicting Social Behavior,* Englewood Cliffs, NJ: Prentice-Hall.

Anderson, N. H. (1965) "Averaging Versus Adding as a Stimulus-Combination Rule in Impression Formation", *Journal of Experimental Psychology*, Vol 70, No. 4, pp394-400.

Barnette, J. J. (2000) "Effects of Stem and Likert Response Option Reversals on Survey Internal Consistency: If You Feel the Need, There Is a Better Alternative to Using Those Negatively Worded Stems", *Educational and Psychological Measurement*, Vol 60, No. 3, pp361-370.

Benjafield, J. (1985) "A Review of Recent Research on the Golden Section", *Empirical Studies of the Arts*, Vol 3, No. 2, pp117–134.

Bergkvist, L. and Rossiter, J. R. (2007) "The Predictive Validity of Multiple-Item versus Single-Item Measures of the Same Constructs", *Journal of Marketing Research*, Vol 44, No. 2, pp175-184.

Bohner, G., Bless, H., Schwarz, N., and Strack, F. (1988) "What Triggers Causal Attributions? The Impact of Valence and Subjective Probability", *European Journal of Social Psychology*, Vol 18, No. 4, pp335-345.

Bosson, J. K., Johnson, A. B., Niederhoffer, K., and Swann, W. B. (2006) "Interpersonal Chemistry Through Negativity: Bonding by Sharing Negative Attitudes About Others", *Personal Relationships*, Vol 13, No. 2, pp135-150.

Cacioppo, J. T. and Berntson, G. G. (1994. "Relationship Between Attitudes and Evaluative Space: A Critical Review, with Emphasis on the Separability of Positive and Negative Substrates", *Psychological Bulletin,* Vol 115, No. 3, pp401-423.

Cacioppo, J. T., Gardner, W. L., and Berntson, G. G. (1999) "The Affect System Has Parallel and Integrative Processing Components: Form Follows Function", *Journal of Personality and Social Psychology*, Vol 76, No. 5, pp839-855.

Churchill Jr., G. A. (1979) "A Paradigm for Developing Better Measures of Marketing Constructs", *Journal of Marketing Research*, Vol 16, No. 1, pp64-73.

Czapinski, J. (1986) "Informativeness of Evaluations in Interpersonal Communication. Effects of Valence, Extremity of Evaluations and Ego-Involvement of Evaluator", *Polish Psychological Bulletin*, Vol 17, No.(3-4), pp155–164.

Czapinski, J. (1985) *Wartościowanie. Zjawisko Inklinacji Pozytywnej. O Naturze Optymizmu* (*Valuation - The Positivity Bias Phenomenon. One the Nature of Optimism),* Wrocław: Ossolineum.

Diener, E. and Emmons, R. A. (1984) "The Independence of Positive and Negative Affect", *Journal of Personality and Social Psychology*, Vol 47, No. 5, pp1105-1117.

Drolet, A. L. and Morrison, D. G. (2001) "Do We Really Need Multiple-Item Measures in Service Research?", *Journal of Service Research*, Vol 3, No. 3, pp196-204.

Edwards, A. L. (1946) "A Critique of 'Neutral' Items in Attitude Scales Constructed by the Method of Equal Appearing Intervals", *Psychological Review*, Vol 53, No. 3, pp159-169.

Herche, J. and Engelland, B. (1996) "Reversed-Polarity Items and Scale Unidimensionality", *Journal of the Academy of Marketing Science*, Vol 24, No. 4, pp366-374.

Herr, P. M. and Page, C. M. (2004) "Asymmetric Association of Liking and Disliking Judgments: So What's Not to Like?", *Journal of Consumer Research*, Vol 30, No. 4, pp588-601.

Horan, P. M., DiStefano, C., and Motl, R. W. (2003) "Wording Effects in Self-Esteem Scales: Methodological Artifact or Response Style?", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol 10, No. 3, pp435–455.

Kahneman, D. and Knetsch, J. L. (1991) "The Endowment Effect, Loss Aversion, and Status Quo Bias", *Journal of Economic Perspectives*, Vol 5, No. 1, pp193-206.

Kanouse, D. E and Hanson Jr., L. R. (1987) "Negativity in Evaluations", in Jones, E. E., Kanouse, D. E., Kelley, H. H., Nisbett, R. E., Valins, S., and Weiner, B. (Eds.), *Attribution: Perceiving the Causes of Behavior* (pp47-62.), Hillsdale, NJ: Lawrence Erlbaum Associates.

Markus, H. and Zajonc, R. B. (1985) "The Cognitive Perspective in Social Psychology", InLindzey, G. and Aronson, E. (Eds.), *Handbook of Social Psychology, 3rd Ed.* (pp137-229), New York: Random House.

Matlin, M. W. and Stang, D. J. (1978) *The Pollyanna Principle: Selectivity in Language, Memory, and Thought*, Cambridge, MA: Schenkman.

Mook, J., Kleijn, W. C., and Van der Ploeg, H. M. (1991) "Symptom-Positively and -Negatively Worded Items in Two Popular Self-Report Inventories of Anxiety and Depression", *Psychological Reports*, Vol 69, No. 2, pp551-560.

Motl, R. W. and DiStefano, C. (2002) "Longitudinal Invariance of Self-Esteem and Method Effects Associated with Negatively Worded Items", *Structural Equation Modeling: A Multidisciplinary Journal*, Vol 9, No. 4, pp562–578.

Nunnally, J.C. (1978) *Psychometric Theory*, 2nd Ed., New York: McGraw-Hill.

Peeters, G. (1971) "The Positive-Negative Asymmetry: On Cognitive Consistency and Positivity Bias", *European Journal of Social Psychology*, Vol 1, No. 4, pp455-474.

Peeters, G. and Czapinski, J. (1990) "Positive-Negative Asymmetry in Evaluations: The Distinction Between Affective and Informational Negativity Effects", *European Review of Social Psychology*, Vol 1, No. 1, pp33-60.

Rodin, M. J. (1978) "Liking and Disliking: Sketch of an Alternative View", *Personality and Social Psychology Bulletin*, Vol 4, No. 3, pp473-478.

Rossiter, J. R. (2008) "Content Validity of Measures of Abstract Constructs in Management and Organizational Research", *British Journal of Management*, Vol 19, No. 4, pp380-388.

--------- (2005) "Reminder: A Horse Is a Horse", *International Journal of Research in Marketing*, Vol 22, No. 1, pp23-25.

--------- (2002) "The C-OAR-SE Procedure for Scale Development in Marketing", *International Journal of Research in Marketing*, Vol 19, No. 4, pp305-335.

Rozin, P. and Royzman, E. B. (2001) "Negativity Bias, Negativity Dominance, and Contagion", *Personality and Social Psychology Review*, Vol 5, No. 4, pp296-320.

Samuelson, W. and Zeckhauser, R. (1988) "Status Quo Bias in Decision Making", *Journal of Risk and Uncertainty*, Vol 1, No. 1, pp7-59.

Schimmack, U. (2005) "Response Latencies of Pleasure and Displeasure Ratings: Further Evidence for Mixed Feelings", *Cognition & Emotion*, Vol 19, No. 5, pp671-691.

Schimmack, U. and Colcombe, S. (2007) "Eliciting Mixed Feelings with the Paired-Picture Paradigm: A Tribute to Kellogg (1915)", *Cognition & Emotion*, Vol 21, No. 1, pp1546-1553.

Schriesheim, C. A., Eisenbach, R. J., and Hill, K. D. (1991) "The Effect of Negation and Polar Opposite Item Reversals on Questionnaire Reliability and Validity: An Experimental Investigation", *Educational and Psychological Measurement*, Vol 51, No. 1, pp67-78.

Simonson, I. and Tversky, A. (1992) "Choice in Context: Tradeoff Contrast and Extremeness Aversion", *Journal of Marketing Research*, Vol 29, No. 3, pp281-295.

Tomas, J. M. and Oliver, A. (1999) "Rosenberg's Self-Esteem Scale: Two Factors or Method Effects", *Structural Equation Modeling*, Vol 6, No. 1, pp84-98.

Tversky, A. and Kahneman, D. (1991) "Loss Aversion in Riskless Choice: A Reference-Dependent Model", *Quarterly Journal of Economics*, Vol 106, No. 4, pp1039-1061.

Van Dijk, W. W., Danner, U., Nieweg, M., and Sumter, S. (2003) "Positive-Negative Asymmetry in the Evaluation of Trivial Stimuli", *Journal of Social Psychology*, Vol 143, No. 6, pp783-784.

Watson, D., Clark, L. A., and Tellegen, A. (1988) "Development and Validation of Brief Measures of Positive and Negative Affect: The PANAS Scales", *Journal of Personality and Social Psychology*, Vol 54, No. 6, pp1063-1070.

Wong, N., Rindfleisch, A., and Burroughs, J. E. (2003) "Do Reverse-Worded Items Confound Measures in Cross-Cultural Consumer Research? The Case of the Material Values Scale", *Journal of Consumer Research*, Vol 30, No. 1, pp72-91.