A Comment on Employee Surveys: Negativity Bias in Open-Ended Responses

Reanna M. Poncheri, Jennifer T. Lindberg, Lori Foster Thompson and Eric A. Surface

Organizational Research Methods 2008; 11; 614 originally published online Aug 8, 2007;
DOI: 10.1177/1094428106295504

Organizational Research Methods
Volume 11 Number 3
July 2008 614-630

10.1177/1094428106295504
http://orm.sagepub.com
hosted at
http://online.sagepub.com

A Comment on Employee Surveys

Negativity Bias in Open-Ended Responses

Reanna M. Poncheri
North Carolina State University
Surface, Ward, and Associates
Jennifer T. Lindberg
Lori Foster Thompson
North Carolina State University
Eric A. Surface
Surface, Ward, and Associates

Recent technologies have reduced some of the major barriers to capturing, coding, and analyzing qualitative data from survey respondents. This has prompted a renewed interest in including open-ended questions on employee surveys and a corresponding need to better understand the potential biases of personnel who choose to provide comments. The present study used data from a climate survey (N = 661) to empirically examine qualitative comments and their relationship with quantitative survey ratings. Results revealed that relatively dissatisfied employees were more likely to provide comments than their more satisfied counterparts. Moreover, open-ended responses were disproportionately negative in tone and tended to echo commenters' closed-ended satisfaction ratings. For most survey dimensions studied, the length of comments increased as they became more negative in tone. Finally, the data revealed very few demographic differences between respondents who provided comments and those who did not.

Keywords: *open-ended survey responses; comments; organizational surveys; online personnel surveys; negativity bias*

Understanding the characteristics of open-ended survey comments and the respondents who provide them is an important concern for practitioners working in the current organizational environment. Open-ended questions (i.e., questions with no predetermined response options) have often been a source of controversy in survey research, with some of the original debates dating back to the 1940s (Converse, 1987). In these discussions, researchers concluded that the costs of including open-ended questions in surveys outweighed the benefits and added little information above and beyond what could be obtained from closed-ended questions (i.e., questions with predetermined response options; Converse, 1987). As a result of these early debates, closed-ended items became the dominant question

Authors' Note: Portions of this article were presented at the 21st annual meeting of the Society for Industrial and Organizational Psychology, Dallas, Texas, May 2006. The views expressed in this article are those of the authors and do not necessarily reflect the positions or policies of the research sponsors or the organizations that conducted the research. The usage of the words *study* and *research* in this article is not meant to imply that the investigation drew on any particular source of funding within or outside of the organization.

format in survey research (Geer, 1991; Krosnick, 1999), with open-ended items being less common.

Recently, there has been a renewed interest in open-ended questions, which is largely because of advances in technology (Fenlason, 2005; Kraut, 2005). From the organization's perspective, the increasing use of online surveys and computer-aided text analyses makes collecting and analyzing qualitative comment data much less time-consuming and expensive than it used to be (Borg, 2005; Fenlason, 2005). From the employee's perspective, the transition from paper to Web-based surveys may ease the process of providing comments, simply because many people type faster than they write (Thompson, Surface, Martin, & Sanders, 2003). One salient impact of this technological shift is that Web-based surveys tend to yield longer write-in comments than paper surveys (Kraut, 2006; Kulesa & Bishop, 2006).

Although technological advances provide advantages to organizations that wish to gather qualitative feedback from employees, it is important for researchers and practitioners to understand the risks that could be involved in collecting and reporting this type of information (Mossholder, Settoon, Harris, & Armenakis, 1995). Not only is it necessary to invest research time and energy into developing new ways to code and analyze large volumes of open-ended comments, but it is also important to conduct research to determine who provides comments and why. Because technology has reduced the major barriers to using open-ended questions (i.e., the time and expense associated with capturing, coding, and analyzing this type of data), the use of these questions is likely to increase. This foreseeable increase prompts a growing need for research focusing on the characteristics of open-ended comments.

Despite the practical need to better understand open-ended survey comments, there is a dearth of research related to this topic. Little is known about the basic characteristics of open-ended comments, the relationship between information obtained from quantitative (i.e., closed-ended) and qualitative (i.e., open-ended) results, or the characteristics of those who avail themselves of survey "comment blocks." The current study fills the gaps associated with this line of research by investigating these issues in the context of an organizational climate survey.

Positive-Negative Asymmetry (PNA) and Open-Ended Comments

Although research related to open-ended comments is somewhat sparse, a few important conclusions have emerged related to the characteristics of such comments. Perhaps the most commonly cited sentiment in this line of research is that most open-ended comments are negative (Borg, 2005; Macey, 1996). This finding is consistent with a phenomenon known as the PNA effect, which refers to the difference in the way that individuals process and attend to positively and negatively valued information (Lewicka, Czapinski, & Peeters, 1992; Peeters, 1971).

Numerous researchers (e.g., Dasborough, 2006; Gardham & Brown, 2001; Peeters, 1971) have conducted studies that lend support to the existence of the PNA effect. Research pertaining to the PNA effect generally falls into two categories. The first addresses the positivity bias, which refers to individuals' inclination toward favorable perceptions of reality and affective states when evaluating relatively novel targets (or targets that do not directly influence the individual) in neutral situations (Lewicka et al., 1992). Because it focuses on

novel targets and neutral situations, the positivity bias is not particularly pertinent to the organizational survey comment domain.

Conversely, the second category of PNA research, referred to as the negativity effect (or negativity bias), is highly relevant to survey comments. The negativity effect helps explain why one should expect open-ended responses to be disproportionately negative in tone. "The negativity effect is a reaction to specific stimuli and means higher impact of negative than of positive stimuli of the same intensity on behaviour, affect and cognitive representations of evaluated objects" (Lewicka et al., 1992, p. 426). The negativity effect (rather than the positivity bias) is presumed to take hold when personnel avail themselves of survey comment blocks because open-ended questions on organizational surveys ask respondents to reflect on specific, familiar, and relevant stimuli (not novel, neutral ones).

The negativity effect occurs for two reasons. First, negative stimuli have a higher affective impact than equally intense positive stimuli when individuals form overall evaluations of a stimulus. Second, individuals are more curious about negative than positive stimuli and therefore engage in greater cognitive elaboration of negative compared to positive events (Lewicka et al., 1992). In short, PNA theory suggests that personnel who are asked about workplace issues on a survey should be more mobilized by negative than positive events. Negative events may also lead to more fine-grained information processing. As a result, people who experience negative events may be stimulated to respond more strongly than those who experience positive events.

The negativity bias has been explored and supported in the literature (e.g., Rozin & Royzman, 2001). For example, Dasborough (2006) found support for the negativity bias in a study conducted in an organizational setting that explored the interactions between employees and their leaders. Specifically, Dasborough conducted a qualitative study that asked participants to recall workplace interactions that resulted in either positive or negative emotional reactions. She found that employees recalled significantly more negative interactions than positive interactions with their leaders. Employees' recollection of negative incidents involving their supervisors may generalize to recollection of other events or information available in the organizational environment as well. For example, if employees are asked to provide details about their level of satisfaction with the organization on a climate survey, they may be more likely to recall negative experiences than positive experiences.

Thus, negativity bias is one potential explanation for the finding that most open-ended comments tend to be negative. If individuals have a tendency to focus on negative information or events in general, employees may be more likely to recall negative events and information, as opposed to positive events and information, when responding to open-ended questions on organizational surveys. The tendency to focus on negative information or problems in the organization is further supported by the nature of many organizational surveys, which are often focused on organizational change initiatives. Employees completing these surveys may concentrate on negative information to bring attention to areas where they perceive improvement is needed.

In sum, initial empirical work (e.g., Borg, 2005; Macey, 1996) and theoretical justification (i.e., PNA) support the prevalence of negativity in open-ended comments. It is important to further investigate the tone of open-ended comments to determine if the findings from earlier research are replicable in a different organizational environment. The following hypothesis will therefore be tested:

Hypothesis 1: Open-ended comments produced in response to an employee survey will be disproportionately negative in tone.

Taylor (1991) argues that "other things being equal, negative events appear to elicit more physiological, affective, cognitive, and behavioral activity and prompt more cognitive analysis than neutral or positive events" (p. 67). In the study described earlier, Dasborough (2006) investigated the intensity associated with the recall of both positive and negative events and found that negative incidents were associated with a greater level of emotional arousal than positive incidents. As such, it can be argued that the increase in emotional and cognitive activity associated with negative information leads individuals to attend more to this type of information (Rozin & Royzman, 2001). This increased attention devoted to negative information could result in lengthier comments when negative rather than positive sentiments are expressed. Moreover, negative comments may be particularly long as respondents seek to clearly convey points of dissatisfaction.

An unpublished study by Borg (2005) supports the notion that negative comments tend to be lengthy. Borg found that in an online survey, negative comments were wordier than positive ones. Along these lines, the present study will investigate the relationship between comment tone and length:

Hypothesis 2: The length of comments will increase as they become more negative in tone.

The Relationship Between Closed-Ended and Open-Ended Responses

Not only is it important to understand the characteristics related to open-ended comments as described in the previous section, but it is equally important to examine the relationship between information obtained from closed-ended and open-ended questions. Both types of questions are often included in survey research to utilize multiple methods for gathering information from respondents (i.e., triangulation; Todd, Nerlich, McKeown, & Clarke, 2004). Although triangulation is often perceived as an approach for obtaining convergence (i.e., agreement from different methods), there is a possibility that information obtained using different methods will produce conflicting or inconsistent results (Mathison, 1988). It is important to explore whether or not information obtained from closed-ended and open-ended responses converges so that organizational decision makers can use information gleaned from survey results effectively.

The importance of this line of inquiry is further amplified by decision makers' tendency to favor certain types of feedback. Evidence suggests that managers receiving feedback from organizational surveys are highly interested in employee comments and may favor this type of feedback in comparison to quantitative results (Patton, 2002). Such a preference can pose problems, especially if the information provided by the quantitative results conflicts with information provided by the qualitative results. Past research suggests that verbal vividness affects recall (Schiefer, 1986). When salient, information is more recallable. Increased recallability can lead people to overestimate the frequency of an event (Lewandowsky & Smith, 1983). This overestimation of an event's frequency is referred to as an availability heuristic, whereby individuals make use of the ease with which an instance comes to mind to judge its frequency or commonness.

The availability heuristic has implications for organizational survey initiatives, thereby supporting the need for research investigating the relationship between responses to closed-ended and open-ended questions. Consider a manager receiving a survey feedback report containing both quantitative summary data and qualitative comments. If the survey comments are more memorable than the means and percentages reported, then an availability heuristic may cause the manager to later judge the opinions and satisfaction levels expressed in the open-ended comments as common and representative of the survey findings. To the extent that the open-ended comments do not represent the attitudes of the larger survey sample, this heuristic will lead to faulty conclusions. For example, a study conducted in a Michigan school system in the 1970s showed that when presented with the results from the study, school board members discredited the quantitative results in favor of open-ended comments (Patton, 2002). Therefore, it is important to investigate the extent to which open-ended comments are representative of the satisfaction levels conveyed in response to closed-ended questions.

Discussions concerning the expected relationship between ratings and comments have centered on two different possibilities: (a) comments as providing alternative or additional information and (b) comments as providing supplemental information. Some researchers argue that open-ended comments provide alternative or additional information compared to that obtained from closed-ended questions (Mossholder et al., 1995). This view suggests that information provided in open-ended comments differs from information provided in closed-ended questions. Another perspective suggests that open-ended comments are best used to supplement information obtained from the quantitative data gathered from closed-ended questions (Jackson & Trochim, 2002). This implies that information obtained from open-ended comments is similar to that obtained from closed-ended questions.

In support of the latter viewpoint, Borg's (2005) unpublished study found that for 60% to 95% of open-ended comments provided in a survey, raters (i.e., an independent group of individuals who evaluated the comments) agreed that comments added little information when compared to closed-ended responses. Furthermore, Mossholder et al. (1995) found that there was a positive relationship between quantitative and qualitative measures in their study, lending support to the idea that information obtained from these methods should be related. In light of these trends, the following relationship between closed- and open-ended responses is expected:

Hypothesis 3: The satisfaction level (i.e., positive, negative, or neutral) individuals express in their open-ended comments will be positively related to the satisfaction level they express in response to closed-ended questions.

Investigating the Representativeness of Those Who Comment

In the previous section, we emphasized the importance of examining the relationship between closed- and open-ended responses for individuals who choose to provide open-ended comments. In this section, we argue that it is also important to gather information about individuals who do not provide open-ended comments to address a unique type of nonresponse error.

Nonresponse error is a commonly cited problem in survey research that occurs when respondents differ from nonrespondents in ways that could affect the study results (Dillman, 2000).

Nonrespondents are elusive, and gathering information about them is a challenging, somewhat paradoxical task (Rogelberg et al., 2003). As such, it is difficult and sometimes impossible to determine the extent of nonresponse error during a given study or survey initiative.

We maintain that the failure to submit surveys on request is not the only source of nonresponse error. Nonresponse error can also be viewed as meaningful differences between survey respondents who do and do not answer open-ended questions. Research has shown that a majority of survey respondents do not provide comments to open-ended questions (Borg, 2005; Siem, 2005). For example, Borg (2005) indicated that only 37% of respondents provided comments, whereas Siem (2005) reported that 40% of respondents submitted comments. Because of the fact that only a subset of the survey respondents provide open-ended comments, it is important to investigate any meaningful differences between those who comment and those who do not.

One way to accomplish this goal is to compare responses to closed-ended questions for employees who do and do not provide open-ended comments. In essence, this comparison addresses whether those who choose to respond to open-ended questions are disproportionately satisfied or dissatisfied when compared to their colleagues who did not offer comments. This has implications for the representativeness of the comments provided and the accuracy of the conclusions drawn if managers disproportionately weigh input provided by those who choose to offer open-ended responses.

As noted in Hypothesis 1, it is believed that most survey comments are negative (Borg, 2005; Macey, 1996). Coupled with the expectation that comments reflect the satisfaction level of the respondent (Hypothesis 3), it seems reasonable to suggest that those who provide comments may be less satisfied than those who do not. In fact, an unpublished study by Borg (2005) found that respondents who were less satisfied and had lower organizational commitment provided more open-ended comments than their colleagues. We anticipate a similar pattern and suggest the following:

Hypothesis 4: The closed-ended ratings produced by respondents who make comments will be more negative than the ratings produced by respondents who do not make comments.

Method

Participants

The participants in this study were members of a large military organization. More than 23 different work areas in the participating organization were represented in the sample. A total of 39.6% of the participants indicated that they were military personnel, whereas the remaining individuals were classified as either a civilian government employee (49.5%) or a civilian contract employee (5.7%). Of the participants who reported their gender, 69.6% were men. With regard to race or national origin, 0.4% were American Indian or Alaska Native, 2.0% were Asian or Pacific Islander, 14.8% were Black (not of Hispanic origin), 5.2% were Hispanic, 62.3% were White (not of Hispanic origin), and 6.3% indicated Other.

Survey Instrument

The entire workforce was asked to anonymously complete a recurrent Web-based climate survey. The survey was available online for a total of 19 days, during which 661 employees (77% of the workforce) completed it. The survey asked questions related to satisfaction with seven dimensions of organizational climate. Each dimension included between 5 and 20 closed-ended questions and 1 open-ended question. For the closed-ended questions, respondents were asked to indicate their level of agreement with various items on a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Higher ratings reflected greater levels of satisfaction, and ratings were averaged to produce a dimension score for each of the seven areas: Overall Satisfaction (5-item scale; $\alpha = .78$), Immediate Supervision (9-item scale; $\alpha = .96$), Senior Leadership, (7-item scale; $\alpha = .95$), Training and Development (6-item scale; $\alpha = .81$), Personnel Management (20-item scale; $\alpha = .94$), Team Cohesion (7-item scale; $\alpha = .89$), and Communications (8-item scale; $\alpha = .84$). The open-ended questions for each dimension were phrased as follows: "Use the space below for additional comments related to [dimension name]." These questions were presented at the end of each survey dimension through the survey.

Comment Coding

To test the hypotheses proposed in this study, it was necessary to code several elements of the open-ended comments. For each of the seven survey dimensions, a variable was created to indicate whether a comment was made (coded as 1) or not made (coded as 0). In addition, another variable was created to indicate the length of each comment (i.e., word count). When creating the variables for comment length, each comment was selected, and the word count was calculated objectively for each comment using the word count feature in Microsoft Word. Finally, two independent raters coded each open-ended response to indicate whether the overall tone of the comment was positive (coded as +1), negative (coded as –1), or neutral (coded as 0). If comments were mixed (i.e., positive, negative, and neutral), we counted the number of discrete negative statements, the number of discrete positive statements, and the number of discrete neutral statements. The tone that was most prevalent was used as the comment tone code. For example, if there were two negative statements and one positive statement, the comment was coded as negative. If there were two negative statements and two positive statements, the comment was coded as neutral. We scored overall comment tone, as opposed to the tone of discrete subcomments, to allow for comparisons with composite closed-ended responses (e.g., those conducted when testing Hypothesis 3).

Interrater reliability statistics were calculated for the comments overall and for each dimension to ensure that comment tone could be reliably scored. The statistics calculated to assess interrater reliability were the Pearson product-moment correlation (r) and Cohen's kappa (κ). For both of these statistics, higher values indicate higher levels of reliability, and values closer to 1.00 are desirable. Pearson's r is calculated based on the ratio of true score variance in ratings divided by the total variance in ratings (Gregory, 2000). Cohen's κ accounts for the fact that raters will agree by chance a certain percentage of the time and

Table 1
Number and Percentage of Positive, Neutral, and Negative Comments Per Dimension

	Comment Tone								
	Positive		Neutral		Negative		Total		
Dimension	*n*	%	*n*	%	*n*	%	*n*	χ^2 (*df* = 2)	*p*
1. Overall satisfaction	14	11.3	28	22.6	82	66.1	124	62.39	< .001
2. Immediate supervision	23	31.5	21	28.8	29	39.7	73	1.43	.491
3. Senior leadership	21	31.3	11	16.4	35	52.2	67	13.02	.001
4. Training and development	13	18.6	18	25.7	39	55.7	70	16.31	< .001
5. Personnel management	8	13.3	10	16.7	42	70.0	60	36.40	< .001
6. Team cohesion	3	6.3	16	33.3	29	60.4	48	21.13	< .001
7. Communications	12	16.0	30	40.0	33	44.0	75	10.32	.006
Total	94	18.2	134	25.9	289	55.9			

therefore is a more conservative estimate of agreement than Pearson's *r*. For Cohen's κ, values between .40 and .75 are considered good, and values above .75 are considered to be excellent levels of agreement (von Eye & Mun, 2005). In addition, the percentage of absolute agreement (P) was calculated and represents the number of ratings that were identical between raters divided by the number of rating opportunities. These analyses indicated an acceptable level of reliability overall (r = .87, κ = .78, P = 87.04) and for all seven dimensions: Satisfaction (r = .86, κ = .75, P = 87.10), Immediate Supervision (r = .94, κ = .88, P = 91.78), Senior Leadership (r = .85, κ = .75, P = 85.07), Training and Development (r = .90, κ = .78, P = 87.14), Personnel Management (r = .82, κ = .70, P = 85.00), Team Cohesion (r = .78, κ = .65, P = 81.25), and Communications (r = .86, κ = .83, P = 89.33). After the comments were independently coded, the two raters met to resolve any disagreements. The resolved codes were used in all subsequent analyses.

Results

Overall, 225 (34%) of the respondents provided at least one comment and generated a total of 517 open-ended responses across the seven comment blocks. To address the first hypothesis and determine if open-ended comments were disproportionately negative in tone, a series of chi-square goodness-of-fit tests were conducted, which examined the proportion of the three comment types (positive, negative, and neutral) within each survey dimension. As shown in Table 1, the data revealed that employees offered a fairly even number of positive, negative, and neutral comments in response to the Immediate Supervision comment block. However, the remaining dimensions produced significant chi-square results, indicating that an imbalance of positive, negative, and neutral attitudes characterized remarks about Satisfaction, Senior Leadership, Training and Development, Personnel Management, Team Cohesion, and Communications. Descriptive statistics demonstrated a larger percentage of negative comments than positive and neutral ones for

Table 2
Relationship Between Comment Tone and Comment Length, Per Dimension

	Comment Tone							
	Positive (+1)		Neutral (0)		Negative (–1)		Relationship Between Comment Tone and Length	
Dimension	*M*	*SD*	*M*	*SD*	*M*	*SD*	*r*	*p*
1. Overall satisfaction	25.64	19.05	30.82	30.95	37.26	37.40	–.120	.185
2. Immediate supervision	18.83	11.24	16.24	10.02	63.76	63.71	–.424	< .001
3. Senior leadership	28.05	16.87	33.55	34.56	59.86	76.82	–.247	.044
4. Training and development	19.46	15.51	24.83	11.75	44.97	36.77	–.354	.003
5. Personnel management	32.00	21.76	28.10	22.08	43.40	35.37	–.167	.202
6. Team cohesion	18.00	7.21	26.06	15.52	52.52	34.97	–.424	.003
7. Communications	44.92	36.43	19.90	35.19	46.64	35.69	–.125	.287

Note: The descriptive statistics in this table represent the comment word count means and standard deviations within each dimension.

each of the survey dimensions named above. To determine whether the negative comments significantly exceeded the neutral and positive ones, we conducted dimension-level chi-square analyses that first compared the percentage of negative and positive responses and then compared the percentage of negative and neutral responses. Results revealed that negative comments significantly ($p < .01$) exceeded positive ones within a majority of the seven dimensions: Satisfaction, Training and Development, Personnel Management, Team Cohesion, and Communications. Similarly, negative comments significantly ($p < .01$) exceeded neutral ones within a majority of the dimensions: Satisfaction, Senior Leadership, Training and Development, and Personnel Management.

Finally, the percentage of positive, negative, and neutral comments was calculated for the sample overall (see Table 1). Aggregated across all survey dimensions, 56% of the comments were negative, 26% were neutral, and 18% were positive. In combination with the chi-square analyses reported above, this pattern of results indicated support for Hypothesis 1.

Hypothesis 2 predicted that the length of comments would increase as they became more negative in tone. Comments ranged in length from 2 to 438 words, and the average comment consisted of 38 words ($SD = 39.87$). Table 2 shows the average positive, neutral, and negative comment length per dimension.

Prior to examining Hypothesis 2, Levene statistics were computed, testing the homogeneity of variance assumption per survey dimension to determine whether the variability of comment length was equivalent across the three tone groupings. Results revealed significant heterogeneity of variance within three of the seven dimensions: Immediate Supervision, Training and Development, and Team Cohesion. For these three dimensions, a Brown-Forsythe F* test was conducted in lieu of the standard one-way ANOVA, and Dunnett's T3 post hoc comparisons, which do not assume homogeneity of variance, were used in lieu of traditional post hoc tests. For the remaining four dimensions, (i.e., Overall Satisfaction,

Senior Leadership, Personnel Management, and Communications), the three comment tone groups were compared via standard one-way ANOVAs, and Bonferroni post hoc tests were used for subsequent pairwise comparisons.

Results comparing the length of positive, negative, and neutral comments revealed significant ($p < .01$) differences within 4 of the 7 dimensions: Immediate Supervision, Training and Development, Team Cohesion, and Communications. Post hoc analyses indicated that negative comments were significantly ($p < .05$) longer than neutral ones within all four of these dimensions. Moreover, negative comments were significantly ($p < .05$) longer than positive ones within 3 of the 4 dimensions: Immediate Supervision, Training and Development, and Team Cohesion. The fourth dimension, Communications, diverged slightly from this trend. As noted, negative Communications comments were longer than neutral ones; however, they did not differ from positive ones in length. Neutral and positive comments did not differ significantly in terms of length.

To further examine Hypothesis 2, a series of bivariate correlations were computed to determine the association between comment tone and length within each of the seven survey dimensions. As shown in Table 2, all correlations were in the expected direction. Comment length significantly increased as the tone became more negative within the following dimensions: Immediate Supervision, Senior Leadership, Training and Development, and Team Cohesion.[1] In short, ANOVA and correlational analyses supported Hypothesis 2 within most but not all of the survey dimensions.

Next, we examined the relationships between the quantitative and qualitative results from the climate survey. Hypothesis 3 predicted that the satisfaction level (i.e., positive, negative, or neutral) of open-ended comments would significantly relate to the satisfaction level expressed in response to closed-ended questions. As demonstrated by the descriptive statistics presented in Table 3, the closed-ended satisfaction ratings provided by those who made negative comments appeared to be lower than the ratings provided by those who made neutral comments, which appeared lower than the ratings of those who made positive comments. Significance tests were conducted to examine the reliability of these differences.

First, Levene statistics were computed, testing the homogeneity of variance assumption within each survey dimension to determine whether the variability of closed-ended survey ratings was equivalent across the three tone groupings. Results revealed significant heterogeneity of variance within 3 of the 7 dimensions: Immediate Supervision, Senior Leadership, and Training and Development. Brown-Forsythe F* tests and Dunnett's T3 post hoc comparisons, which do not assume homogeneity of variance, were used in lieu of traditional post hoc tests.

Omnibus tests revealed significant ($p < .001$) group differences for all of the dimensions except Team Cohesion. Follow-up post hoc tests showed that the closed-ended dimension ratings provided by those who made negative comments were significantly ($p < .05$) lower than the ratings provided by those who made positive comments for all six of the significant dimensions. Moreover, the negative comment providers' ratings were significantly ($p < .05$) lower than the ratings provided by those who made neutral comments for all of the six significant dimensions except Senior Leadership. The ratings generated by neutral and positive comment providers only differed within the Training and Development dimensions, where positive comment providers offered significantly ($p < .05$) higher ratings than neutral comment providers.

Table 3
Relationship Between Comment Tone and Satisfaction Level Rated in Response to Closed-Ended Questions, Per Dimension

	Comment Tone							
	Positive (+1)		Neutral (0)		Negative (–1)		Relationship Between Comment Tone and Satisfaction Rating	
Dimension	*M*	*SD*	*M*	*SD*	*M*	*SD*	*r*	*p*
1. Overall satisfaction	4.64	0.65	4.11	0.64	3.69	0.75	.408	< .001
2. Immediate supervision	4.68	0.39	4.34	0.81	2.82	1.22	.649	< .001
3. Senior leadership	4.58	0.46	3.92	1.35	2.97	1.16	.582	< .001
4. Training and development	4.51	0.45	3.97	0.55	3.50	0.78	.508	< .001
5. Personnel management	4.26	0.81	4.11	0.46	3.25	0.86	.451	< .001
6. Team cohesion	4.00	0.43	3.56	0.75	3.16	1.09	.260	.075
7. Communications	4.00	0.38	4.08	0.62	3.25	0.68	.467	< .001

Note: The descriptive statistics in this table represent the closed-ended satisfaction rating means and standard deviations per dimension.

Next, correlations between comment tone and the ratings assigned to each of the seven survey dimensions were calculated. The results were consistent with the pattern of significance demonstrated in the omnibus analyses described above. As shown in Table 3, positive correlations were found for 6 of the 7 survey dimensions.[1] Thus, Hypothesis 3 was largely supported.

Hypothesis 4 predicted that the closed-ended ratings produced by respondents who made comments would be more negative than the ratings produced by respondents who did not make comments. Independent samples *t* tests were conducted for each of the seven dimensions to examine this prediction. The results, shown in Table 4, revealed significant differences for all seven dimensions. Individuals who provided comments had more negative quantitative ratings than those who did not provide comments. Hypothesis 4 was therefore supported.

Finally, exploratory follow-up analyses were conducted to examine whether demographic subgroups in our sample differed in their tendency to provide comments. Chi-square tests of independence were computed, with respondents cross-classified according to demographic grouping and whether or not they provided open-ended comments. Racial group differences were found for 2 of the 7 dimensions. Specifically, a smaller percentage of Blacks (1.9%) made comments than Whites (12.4%) and other racial groups (11.3%) in the Training and Development dimension, $\chi^2(2, N = 634) = 9.78$, $p = .008$. There was also a significant ethnicity effect within the Team Cohesion dimension, $\chi^2(2, N = 634) = 6.27$, $p = .044$. Blacks (1.9%) were less likely to comment than Whites (9.0%). Meanwhile, those who self-identified as Other (6.2%) did not significantly differ from Blacks and Whites in their tendency to comment on Team Cohesion. Additional chi-square analyses revealed that comment tone did not differ significantly across racial groups. Moreover, no differences

Table 4
Mean Satisfaction Ratings Produced by Those Who Did and Did NOT Provide Comments

	No Comment			Comment				
Dimension	*n*	*M*	*SD*	*n*	*M*	*SD*	t^a	*p*
1. Overall satisfaction	537	4.22	0.67	124	3.89	0.78	4.29	< .001
2. Immediate supervision	588	4.23	0.82	73	3.84	1.24	2.60	< .05
3. Senior leadership	594	4.11	0.846	67	3.63	1.25	3.05	< .01
4. Training and development	591	4.18	0.63	70	3.81	0.78	3.89	< .001
5. Personnel management	601	4.01	0.70	60	3.52	0.92	3.97	< .001
6. Team cohesion	613	3.94	0.78	48	3.35	0.98	4.11	< .001
7. Communications	586	3.99	0.71	75	3.70	0.73	3.22	.001

Note: *t* tests were conducted in lieu of a MANOVA because different individuals commented on different survey dimensions. Whether or not a comment was made was coded separately for each of the seven survey dimensions; therefore, it was possible for a given individual to be assigned to the "comment" group for one dimension and the "no comment" group for another dimension.
a. The degrees of freedom were adjusted downward, and a separate variances test was employed when the two groups failed to demonstrate homogeneity of variance.

were observed between men and women or between military and civilian personnel in their tendency to provide open-ended comments.

Discussion

The findings from this study begin to shed light on the characteristics of open-ended comments and the attitudes of respondents who provide them. Recent technological developments have eased the process of providing, capturing, coding, and analyzing survey comments. These and other advancements will most likely lead to the increased appearance of open-ended questions on organizational surveys. Although the information obtained from such questions can be helpful to organizational decision makers, the findings from our study indicate that caution must be taken when interpreting and using survey comments.

Our results revealed that relatively dissatisfied employees were more likely to provide comments than their more satisfied counterparts (Hypothesis 4). In terms of comment tone, open-ended responses tended to echo the closed-ended satisfaction ratings of the commenter (Hypothesis 3), a finding that is consistent with Borg's (2005) and Mossholder et al.'s (1995) results. Taken together, these two trends help explain why open-ended comments were disproportionately negative in tone (Hypothesis 1) within most of the survey dimensions studied. This finding provides empirical support for the PNA effect in general and the negativity bias in particular. However, it is important to note that there were exceptions to the negativity rule. That is, negative comments outnumbered both positive and neutral ones in most but not all of the survey dimensions examined.

This research also demonstrated that for most of the survey dimensions studied, the length of comments increased as the comments became more negative in tone (Hypothesis 2). This

suggests that negative comments may be particularly lengthy as respondents attempt to express their dissatisfaction thoroughly. This could be dangerous for consumers of survey research if the availability heuristic is operating, especially if the consumer is unaware of the negative bias present in the comments. Finally, the findings from our exploratory follow-up analyses indicated that respondents who provided comments were demographically similar to those who did not provide comments, with the exception of racial group differences found within 2 of the 7 survey dimensions examined.

Practical Implications

The results of this study have important practical implications, particularly for consumers of survey research. The finding that open-ended comments are disproportionately negative lends support to a suggestion cited among survey practitioners: Recipients of survey feedback reports should be warned about the negative tone of open-ended comments so that they do not place undue emphasis on the content of these comments (Borg, 2005; Macey, 1996). As a practical suggestion, survey researchers might consider providing feedback recipients with reports that compare the tone of open-ended comments to the tone of quantitative data collected in the study.

It is important to note that automatically presenting managers with all of the open-ended comments provided by survey respondents may mask important group differences and trends. Therefore, individuals who conduct survey research may wish to examine and report any potential subgroup differences. The results of our research suggest that respondents who provided comments were demographically similar to those who did not; however, this finding could be specific to the population we examined and should perhaps be investigated on a case-by-case basis. It is possible that other organizational settings are characterized by other relevant differences, which may affect the representativeness of the comments provided. In one of our follow-up analyses, we explored differences between two personnel systems (i.e., military and civilian personnel). In other arenas, it may be appropriate to examine alternative personnel system differences (i.e., union vs. nonunion employees, blue-collar vs. white-collar employees, etc.) or differences in responses between organizational departments. If analyses reveal group differences between those who do and do not respond to open-ended questions, leaders should be informed of this disproportionality and instructed to interpret the information accordingly, particularly when attitudinal differences exist between the groups in question. Differential commenting trends between units or departments are important to explore not only because they can help ensure appropriate information processing but also because they may alert organizational decision makers to potential problems in a particular department. If the majority of open-ended comments are negative and are provided by one department, action can be taken to correct the problem at its source.

Limitations and Research Implications

Although this study addresses a noteworthy void in the literature, it should be viewed in the context of several limitations. Comments were coded using three tone categories (i.e., positive, negative, and neutral). Though consistent with other research in this area (e.g., Borg, 2005),

this coding scheme has limitations. In particular, these codes are fairly simplistic considering the depth of some of the comments provided by survey respondents. The decision was made to score a respondent's overall comment tone as opposed to providing subscores for each discrete component of a given comment to facilitate statistical comparisons with quantitative ratings. Nevertheless, it is important to note that the coding scheme may have masked important complexities.

Another limitation of this research is related to the sample. It is uncertain whether our results generalize to other kinds of employees. Replication in private industry should examine the degree to which the trends uncovered in the current study extend to nonmilitary settings. Additional external validity issues involve whether these findings generalize to other surveys, such as those that are new to an organization, those that include more and fewer open-ended questions, and those introducing comment blocks with other types of wording.

A related concern is generalizability to different types of organizational surveys. This study was based on findings from a climate survey that included closed-ended and open-ended items. These items elicited feedback related to individuals' satisfaction with various aspects of their organizational climate. As such, respondents may have been more likely to negatively evaluate the organizational climate to encourage organizational change. Although climate surveys are commonly used in organizational research, there are many other types of surveys that are used in organizations. Other survey types may not yield the same pattern of results when examined from this perspective. For example, open-ended questions included on 360-degree feedback surveys may not be predominantly negative but instead may be primarily positive as a result of leniency error. The generalizability of our findings to open-ended questions presented on instructor evaluations and post-training evaluation surveys (i.e., reaction measures) also warrants investigation.

In terms of future work, studies investigating the way in which organizational decision makers process quantitative and qualitative results would be extremely valuable. Such research could help determine the extent to which the availability heuristic operates in the organizational survey process. This agenda should include controlled studies examining the degree to which qualitative data overshadow contradictory quantitative data when individuals recall the results of a survey feedback report. Outside of the survey domain, work by Shedler and Manis (1986) has demonstrated that showing participants male and female photographs alongside a proportionally dissimilar list of male and female names biased later judgments of the proportion of men and women whose names appeared on the list. The authors of this study suggested that vividness affected the cognitive availability of the information presented to their participants, which in turn affected later judgments regarding the nature of the data.

Are managers and organizational decision makers susceptible to these types of biases when reflecting on survey feedback reports? The present study demonstrated that those who chose to respond to open-ended questions were relatively dissatisfied and therefore unrepresentative of the sample as a whole. Moreover, dissatisfied employees tended to provide lengthy comments, perhaps to paint a detailed picture of their frustrations. If such detail increases the vividness of comments, recall of results may be skewed. Clearly, research is needed to determine whether managers and organizational decision makers place undue weight on these lengthy comments when later estimating the predominant attitudes within the workforce.

Research is also needed to develop our understanding of respondents who provide open-ended comments. Specifically, future work should seek to identify individual differences (e.g., personality, motivation, and previous experience with surveys) that affect both the tendency to provide comments and the tone of those comments. For example, it would be interesting to explore how negative affectivity or agreeableness is related to a tendency to provide open-ended comments. If results from such a study showed that less agreeable individuals, for example, provide more comments than more agreeable individuals, it would contribute to our understanding of the sources of negativity in open-ended comments. A targeted program of research investigating individual differences in comment tendencies would shed light on whether the inclination to comment is situational, dispositional, or some combination of both.

Furthermore, although the present study focused on group-level trends in providing comments, future work could also examine respondent profiles through the use of latent profile analysis. This may lead to the identification of certain types of commenters (e.g., individuals who rarely comment, individuals who usually comment) based on their frequency and length of response to open-ended questions. In addition, the potential for group differences (structural or demographic) in comment tone and the tendency to provide comments could be explored using multilevel modeling. In general, using hierarchical linear modeling techniques could reveal the factors influencing survey responses (quantitative and qualitative) across multiple levels of analysis (individual, supervisory, workgroup, division, organization), allowing practitioners to more accurately target the source of perceptions concerning a particular survey topic.

Finally, more research is needed to examine both question wording and the placement of open-ended comments on organizational surveys. In our survey, respondents were asked to provide dimension-level comments after quantitative items were presented for each survey dimension. However, responses to these questions could have changed if the question were worded differently or placed at a different location on the survey. In terms of question wording, Kulesa and Bishop (2006) described four types of open-ended questions that may affect the types of responses provided: general descriptive, general prescriptive, specific descriptive, and specific prescriptive. Additional research is needed to examine how question wording influences responses. Also, open-ended questions tend to be placed at the end of organizational surveys and are usually not required, whereas closed-ended questions are more often required (i.e., programmed to prevent respondents from progressing through the survey before rating all items presented). Both the placement of open-ended questions and the fact that they are not required may contribute to the perception that they are unimportant. More research is needed to determine how question placement and response requirements affect the frequency and nature of comments.

Although a great deal of future work is needed to fully understand open-ended survey responses, the present study begins to address an area of empirical and practical concern. Contemporary technologies have reduced barriers that previously discouraged the collection of qualitative survey data. As a result, the use of open-ended questions is apt to increase, and the comments they encourage are likely to play a more prominent role in organizational surveys than ever before. Researchers should continue to explore the nature of open-ended responses so that practitioners can use the feedback provided by survey comments in ways that are informative, yet appropriate.

Note

1. Unequal variances generally do not affect correlation results, except to reduce the power of the significance tests (Bobko, 1995). Significant correlations were found within each of the dimensions that failed to demonstrate homogeneity of variance. Thus, although the violation of this assumption was not ideal, it did not appear to affect the conclusions of the correlational analyses.

References

Bobko, P. (1995). *Correlation and regression: Principles and applications for industrial/organizational psychology and management*. New York: McGraw Hill.

Borg, I. (2005, April). *Who writes what kinds of comments? Some new findings*. In A. I. Kraut (Chair), Grappling with write-in comments in a web-enabled survey world. Symposium conducted at the 20th annual conference of the Society for Industrial and Organizational Psychology, Los Angeles.

Converse, J. M. (1987). *Survey research in the United States: Roots and emergence 1890-1960*. Los Angeles: University California Press.

Dasborough, M. T. (2006). Cognitive asymmetry in employee emotional reactions to leadership behaviors. *The Leadership Quarterly*, *17*, 163-178.

Dillman, D. A. (2000). *Mail and Internet surveys: The tailored design method* (2nd ed.). New York: John Wiley.

Fenlason, K. J. (2005, April). *Comments: Where have we been? Where are we going?* Symposium conducted at the 20th annual conference of the Society for Industrial and Organizational Psychology, Los Angeles.

Gardham, K., & Brown, R. (2001). Two forms of intergroup discrimination with positive and negative outcomes: Explaining the positive-negative asymmetry effect. *British Journal of Social Psychology*, *40*, 23-34.

Geer, J. G. (1991). Do open-ended questions measure "salient" issues? *Public Opinion Quarterly*, *55*, 360-370.

Gregory, R. J. (2000). *Psychological testing: History, principles, and applications* (3rd ed.). Boston: Allyn & Bacon.

Jackson, K. M., & Trochim, W. M. K. (2002). Concept mapping as an alternative approach for the analysis of open-ended survey responses. *Organizational Research Methods*, *5*, 307-336.

Kraut, A. I. (2005, April). *Grappling with write-in comments in a web-enabled survey world*. Symposium conducted at the 20th annual conference of the Society for Industrial and Organizational Psychology, Los Angeles.

Kraut, A. I. (2006). Moving the needle: Getting action after a survey. In A. I. Kraut (Ed.), *Getting action from organizational surveys: New concepts, technologies, and applications* (pp. 1-30). San Francisco: Jossey-Bass.

Krosnick, J. A. (1999). Survey research. *Annual Review of Psychology*, *50*, 537-567.

Kulesa, P., & Bishop, R. J. (2006). What did they really mean? New and emerging methods for analyzing themes in open-ended comments. In A. I. Kraut (Ed.), *Getting action from organizational surveys: New concepts, technologies, and applications* (pp. 238-263). San Francisco: Jossey-Bass.

Lewandowsky, S., & Smith, P. W. (1983). The effect of increasing the memorability of category instances on estimates of category size. *Memory and Cognition*, *11*, 347-350.

Lewicka, M., Czapinski, J., & Peeters, G. (1992). Positive-negative asymmetry or "when the heart needs a reason." *European Journal of Social Psychology*, *22*, 425-434.

Macey, W. H. (1996). Dealing with the data: Collecting, processing, and analysis. In A. I. Kraut (Ed.), *Organizational surveys* (pp. 204-232). San Francisco: Jossey-Bass.

Mathison, S. (1988). Why triangulate? *Educational Researchers*, *17*(2), 13-17.

Mossholder, K. W., Settoon, R. P., Harris, S. G., & Armenakis, A. A. (1995). Measuring emotion in open-ended survey responses: An application of textual data analysis. *Journal of Management*, *21*, 335-355.

Patton, M. Q. (2002). *The nature of qualitative inquiry. Qualitative research and evaluation methods* (3rd ed.). Thousand Oaks, CA: Sage.

Peeters, G. (1971). The positive-negative asymmetry: On cognitive consistency and positivity bias. *European Journal of Social Psychology*, *1*, 455-474.

Rogelberg, S. G., Conway, J. M., Sederburg, M. E., Spitzmüller, C., Aziz, S., & Knight, W. E. (2003). Profiling active and passive nonrespondents to an organizational survey. *Journal of Applied Psychology*, *88*, 1104-1114.

Rozin, P., & Royzman, E. B. (2001). Negativity bias, negativity dominance, and contagion. *Personality and Social Psychology*, *5*(4), 296-320.

Schiefer, H. J. (1986). Effect of verbal and nonverbal vividness on students' information processing. *Perceptual and Motor Skills*, *63*, 1106.

Shedler, J., & Manis, M. (1986). Can the availability heuristic explain vividness effects? *Journal of Personality and Social Psychology*, *51*, 26-36.

Siem, F. (2005, April). *History of survey comments at the Boeing Company*. In K. J. Fenlason (Chair), Comments: Where have we been? Where are we going? Symposium conducted at the 20th annual conference of the Society for Industrial and Organizational Psychology, Los Angeles.

Taylor, S. E. (1991). Asymmetrical effects of positive and negative events: The mobilization-minimization hypothesis. *Psychological Bulletin*, *110*(1), 67-85.

Thompson, L. F., Surface, E. A., Martin, D. L., & Sanders, M. G. (2003). From paper to pixels: Moving personnel surveys to the Web. *Personnel Psychology*, *56*, 197-227.

Todd, Z., Nerlich, B., McKeown, S., & Clarke, D. D. (Eds.). (2004). *Mixing methods in psychology: The integration of qualitative and quantitative methods in theory and practice*. New York: Psychology Press.

von Eye, A., & Mun, E. Y. (2005). *Analyzing rater agreement: Manifest variable methods*. Mahwah, NJ: Lawrence Erlbaum.

Reanna M. Poncheri is an industrial/organizational psychology doctoral student at North Carolina State University and an associate and consultant at Surface, Ward, and Associates in Raleigh, North Carolina. She is involved in research related to survey methodology, training effectiveness, and situational predictors of job performance.

Jennifer T. Lindberg is an industrial/organizational psychology doctoral student at North Carolina State University and a research and development assistant at Kaplan DeVries Inc. in Greensboro, North Carolina. Her research focuses on the dark side of leader behavior, with specific emphasis on maladaptive leader personality and the qualitative study of leadership.

Lori Foster Thompson is an associate professor of psychology at North Carolina State University and a senior research fellow with the Army Research Institute. Her research focuses on employee surveys and technology-driven workplace changes. Her editorial board memberships include *TIP*, *Journal of Organizational Behavior*, and *Ergometrika*, where she is associate editor.

Eric A. Surface is a principal at Surface, Ward, and Associates (SWA), an organizational consulting and personnel research firm located in Raleigh, North Carolina. He works on projects with military, nonprofit, and corporate clients. His research focuses on training effectiveness, workplace survey methodology, and performance and proficiency measurement.