



Survey Methodology

When creating an employee survey, most people think about the crafting of the questions. There are also a lot of other decisions needed to make the survey as valuable as possible for unlocking the value of employee feedback.

This whitepaper provides details on the decisions behind our efforts to provide maximum value, while reducing time and effort to implement and analyze.

Kinsey Smith, Analyst, Insights & Analytics Team



Table of Contents

Confidentiality Thresholds	3
Background	3
Our thresholds	4
Comments	5
Background	5
Our approach	5
Response scale	6
7-point Likert-type scale	6
Neutral midpoint	6
Phrasing of items	7
Testing	7
Declarative sentences	7
Reversal	7
Insufficient effort responding	7
Reliability and validity	9
Background	9
Reading Level	9
Reliability	9
Validity	9
Time	10
References	11

Confidentiality Thresholds

Background

When weighing whether or not to participate in a survey, responders consider many factors, including: confidentiality or privacy, the likelihood of their responses being identified, and the potential ramifications if they were to be identified (Couper, Singer, Conrad, & Groves, 2008). When considering the potential for harm, responders consider their trust in the organization, topic sensitivity, and attitudes toward privacy and surveys in general. Responder willingness to participate is influenced by perceived risk-benefit ratio, personal benefit, and social benefit (Couper et al., 2008).

Survey invitees' opinions regarding privacy and confidentiality influence survey participation, with individuals more likely to respond to surveys asking about sensitive topics (such as attitudes about the workplace) when perceived confidentiality is high (Mathews, 2013; Singer, Van Hoewyk, & Neugebauer, 2003; Singer, von Thurn, & Miller, 1995; Thompson & Surface, 2007; Tourangeau & Yan, 2007). Specifically, individuals worried about the possibility of confidentiality breaches and the misuse of survey results are less likely to participate in a survey (Singer et al., 2003). Perceived confidentiality is particularly important as it relates to potentially identifying information, including demographics and departments.

Employees may feel that confidentiality is threatened based on their responses, particularly to demographic questions. Responders may realize that their pattern of responses to demographic questions would be fairly unique within the organization (Lusty, 2009). For example, an individual in the Finance department, recently hired, at the management level, may be the only employee who fits such parameters. While our organization only provides aggregated survey data that meet our confidentiality thresholds, employees may lack awareness of or confidence in this part of the process.

When the survey sponsor (driving force behind a survey) is the employer, it can further complicate survey participation and perceptions of confidentiality. Employees already have a relationship with their organization, along with a history of how previous surveys have been executed, data shared, and results acted on (Rogelberg, Spitzmüller, Little, & Reeve, 2006). When they believe the survey sponsor will take the results seriously, individuals are more likely to participate in a survey (Haunberger, 2011). Employees may be less concerned with confidentiality if previous surveys at the organization have helped established trust (Saari & Scherbaum, 2011). On the other hand, they may perceive more risk of punishment, in which case they may be less likely to participate (Rogelberg et al., 2006; Tourangeau & Yan, 2007).

Organizations should be concerned about the possibility of employees choosing not to respond to the survey, as their data can be incredibly valuable. Employees who deliberately refuse to participate (rather than not responding due to other factors) may have more negative attitudes about the organization and its use of survey data, along with potentially having higher intentions of leaving the organization (Rogelberg,

Conway, Sederburg, Spitzmüller, Aziz, & Knight, 2003; Rogelberg, Luong, Sederburg, & Cristol, 2000). By not responding, their feedback will obviously be absent, and the overall results for the organization may appear to be more positive than they truly are (Rogelberg et al., 2000),

Confidentiality remains a concern even when individuals decide to participate in a survey. Employees want to know who will be able to access results and how those results will be used, especially in the case of sensitive information (Dillman, Smyth, & Christian, 2014). When responders perceive a lower degree of confidentiality or anonymity, their scores decrease in quality and become more socially desirable for sensitive topics (Durant, Carey, & Schroder, 2002; Futrell, Stem, & Fortune, 1978; Heerwegh, Vanhove, Matthijs, & Loosveldt, 2005; Rosenfeld, Booth-Kewley, Edwards, & Thomas, 1996). One study illustrated that employees with lower anonymity can provide higher ratings for their managers (Kantor, 1991). This is not altogether surprising, as employees who worry that confidentiality will be breached or that their responses will lead to reprisals may therefore choose to be less candid, both in the personal information and opinions they share (Simsek & Veiga, 2001).

All of this information lends itself toward our principle of a minimum response threshold for reporting summarized data. Using a confidentiality threshold when reporting results helps reassure employees that their responses will remain confidential, which increases overall survey participation, along with responses to more sensitive items (Biga, McCance, & Massman, 2011; Singer et al., 1995). Beyond increasing employee confidence in our survey process, our confidentiality threshold also allows us to help organizations avoid accidentally identifying responders, which would damage both employee trust and organization improvement efforts.

Our thresholds

Our objective is to balance two competing interests. On the one hand, we want employees to feel completely safe being completely candid. On the other hand, we want to be able to share results with as many managers as possible so that the feedback can be used to improve conditions for the employees providing the feedback.

With those two interests in mind, we have chosen these confidentiality thresholds:

- Five for summarized quantitative results
- Ten for comments
- Thirty for response count reporting (such as histograms and heatmaps)

We have found these values work for most organizations, though if you are at a low-trust organization, we do provide the option of choosing higher custom thresholds.

Comments

Background

We believe that qualitative information (i.e. comments) contributes significantly to the value of our survey, well beyond the value of quantitative scores alone. However, open-ended comments can often seem overwhelmingly negative. Various studies suggest that 56-67% of comments are negative in tone (Borg & Zuell, 2012; Poncheri, Lindberg, Thompson, & Surface, 2008).

While the tone of comments generally reflect the responder's score on that topic, responders with lower scores are more likely to comment (Borg & Zuell, 2012; Poncheri et al., 2008). Further, negative comments are generally longer (i.e., word count) than positive ones (Borg & Zuell, 2012; Poncheri et al., 2008). Thus, even if the scores are fairly evenly split between positive and negative, the negative comments are likely to exceed positive ones in both number and word count.

Our approach

Our goals with respect to comments are to:

1. Be able to connect comments to quantitative data, which can provide context on whether a comment is an isolated complaint or a hint about a bigger issue.
2. Ask the right questions of the right people. For this, we use the same principles that make for good conversation:
 - Ask questions about the topics the employee is passionate about
 - Ask questions where the employee has a unique perspective.
3. Where possible, balance positive and negative sentiment. If things are bad, the best way to improve it is to build on the bright spots. If things are good, it's easier to keep them that way by staying vigilant.

To achieve the first goal, we have two conditional questions associated with each standard Likert item, one for generating positive feedback, and the other for generating improvement feedback. We choose which one is appropriate for a responder based on their Likert response.

To achieve the second and third goals, responders themselves receive four conditional questions based on their responses to the standard survey statements. The first question asks about the topic the responder was most positive about. Similarly, the second question asks about the topic the responder was most negative about. The third and fourth questions still reflect the employee's sentiment on various topics, but also target questions that have limited feedback on the survey at that point.

This algorithm allows us to ask employees for their feedback on topics they are passionate about, while providing more actionable feedback to the organization through a healthy balance of positive and constructive comments.

Response Scale

7-point Likert-type scale

An important consideration when creating a survey is the response scale. Some aspects of response scale creation are fairly simple: it is generally preferable to use a balanced scale, with an equal number of positive and negative options. This allows us to collect feedback from both ends of the spectrum, without biasing the scale in one direction or the other. Employee preferences are also fairly straightforward: they prefer a scale that fully captures their opinions while remaining simple to use, only including as many options as is necessary (McDonald, 2004).

Researchers, however, debate the best number of response options. Generally, they strive for a response scale that sufficiently captures responder sentiment, with a distribution of responses across all response options, but no one response option over-represented (Schall, 2003). Including more response options allows researchers to employ more sophisticated analysis techniques, and glean more insight from the results (Foddy, 1994; Pearse, 2011), particularly when looking for differences between groups.

Employee surveys results generally fall toward the positive end of the scale, so using more response options (e.g. a 7-point scale, rather than 5-point) allows us to collect richer data. By capturing more granularity, we can better differentiate between degrees of sentiment. This is especially relevant to our survey; the richer data supports the high accuracy of our sophisticated probability-based individualized benchmarking that we use for identifying focus areas and celebrations.

Neutral midpoint

Another topic of debate in the field is whether to include a neutral option in the response scale. There is a fear that responders will over-use the midpoint of a scale. However, when a scale has more response options, fewer responders choose the neutral or uncertain response option (Foddy, 1994). This may suggest that individuals were previously using the neutral/uncertain option because they did not see another option that reflected their opinions (Foddy, 1994). By including a neutral option, we hope to avoid people artificially placing themselves into a negative or positive response category, which could lead to misrepresentation of that responder's attitudes (Fink, 1971).

Further, we find that the neutral option is generally not overused on our survey. This is due, in part, to the fact that we test our survey items prior to including them on the standard survey, discarding those with a high degree of neutrality. Moreover, our scale also has more response options, decreasing the likelihood of responders choosing the neutral option. Finally, we see more responses on the positive end of the scale, given the topic of our survey and our responder population. Most importantly, we are proponents of the neutral category precisely because it captures apathy, fear, or a lack of knowledge, information, or opinion on the part of responders - all of which we feel is valuable information.

Phrasing of Items

Testing

Our survey statements were each chosen after thorough testing of multiple variations with thousands of responders, and we found that these phrasings resulted in the least amount of confusion, highest response rates, fullest distribution of responses, and most candor.

Declarative sentences

We include a number of statements that are purposefully phrased negatively or positively to generate the best possible responses, a common survey technique (Roberts, Laughlin, & Wedell, 1999). By providing a declarative sentence, we can gauge responders' opinions through the level of agreement they indicate with their response (DeVellis, 2003). Statements that are neutral or moderate in tone may not provide much useful information, as responders will be too inclined to agree (DeVellis, 2003).

Reversal

Reversal (the practice of including a mix of both "positive" and "negative" statements) is a common approach in designing surveys. Many survey practitioners believe that reversal decreases inattention bias, which is the tendency of people to stop thinking carefully about survey statements and answer using a more narrow range of responses. By including statements where positive responses would use both ends of the response scale, responders may be more careful in responding, and we can weed out individuals who did not respond to the survey thoughtfully.

As with anything, there is also a downside, which is that this approach can cause confusion where people don't understand that reversal has happened. However, while we can correct for confusion by using tools such as benchmarking, we cannot correct inattention.

It should be noted that responses to a negatively-phrased statement do not necessarily reflect the direct opposite of a similar positively-phrased statement. In general, responders are more willing to provide a strongly positive response than a strongly negative one (Foddy, 1994). To simplify our reporting, we want higher scores to indicate more positive results - as such, we reverse the scoring on these statements, and adjust the phrasing to reflect that in the survey results.

Insufficient effort responding

When collecting and interpreting survey data, it is important to eliminate discrepancies between what we are trying to measure and how employees respond to the survey, also known as response bias.

One particular issue is that response data may suffer from insufficient effort responding (IER), where responders intentionally randomly select their responses, or do not devote the proper amount of attention to items or instructions. IER is particularly concerning when the average score differs significantly from the scale midpoint, as it can introduce more variance in the survey results, inflating or even creating correlations between variables (Huang, Liu, & Bowling, 2015). As we often see average scores well above the midpoint of our scale, IER is a concern. To counteract some of the variance IER may introduce, our algorithm considers response patterns of individuals, noting when they do not respond as expected to negatively-phrased items based on their other responses (using reversal), or display other patterns that may indicate inattention while responding.

Insufficient effort responding is less of a concern when the topic of the survey is salient to the responder. IER is also less likely to be an issue when participation in the survey is optional (Meade & Craig, 2012). By providing a survey whose topic is relevant to employee concerns, with optional participation, we hope to further limit IER.

Reliability and Validity

Background

A good survey has a reading level appropriate to the intended population, is reliable, has acceptable validity, and doesn't take too long to complete. So, let's unpack each of these.

Reading Level

All respondents should interpret a survey item in a similar way, and that starts with easy-to-understand language. During our testing of potential survey items, we use a proprietary measure to determine if respondents do not understand an item. This "confusion" measure is one of the many hurdle an item must pass as part of our process of choosing the best items for our survey.

Reliability

Reliability is the extent to which a test or measuring procedure yields the same score for the same person from one administration to the next.

However, because our survey measures employee perceptions at a given moment, we expect their responses to change from one survey to the next. Our survey results can then provide actionable data for organizations to improve based on employee perceptions at that time and also highlight changes in perception over time.

Another form of reliability determines how similar the survey items are to one another. In research and academia, tests often include numerous items that measure the same variable. By using more than one item, the cumulative score reflects a more accurate respondent score through triangulation.

However, this follows the law of diminishing returns - as you use more items to measure a single variable, each item adds less and less value. Further, respondents face survey fatigue, becoming less likely to complete a test the longer it goes on. To maximize response rate, surveys should minimize the number of items and time required as much as possible.

These two competing issues mean that we need to strike a balance between having enough items to accurately measure a variable and few enough that we don't exhaust our busy respondents. For this reason, we assess each variable with a single item that has been tested over many years and millions of respondents. Additionally, we follow up each item with open-ended questions to gather even more insight.

Validity

Validity is the extent to which an item or test measures what it is intended to measure.

However, validity requires an accumulation of supporting evidence. It is also influenced by context, purpose, and population.

Face validity is the extent to which a survey appears to measure what it intends to measure. So, if our survey is intended to measure OrgHealth™ and Engagement, do the survey items and resulting response data seem to actually measure those two concepts? This measure is one of the first things people notice. Through continued conversations with our survey customers and analysis of employee comments in relation to statement scores, we track and refine our statements to drive face validity.

Content validity is the extent to which a survey statement adequately represents the variable it measures. In other words, does the survey statement adequately assess what it's meant to assess? Through rigorous testing of survey item and open-ended question verbiage, we have refined each statement to be concise yet representative.

Criterion-related validity considers the more practical issue of how a test will relate to other variables. One form, predictive validity, focuses on how well a test or measure can predict future variables. For instance, our Top Workplaces Fund has demonstrated that organizations that score well on our survey outperform their peers in the Russell 2500 Index over one, three, and five year periods. We would also encourage you to speak with your account manager about connecting our survey to the performance metric most important to your organization.

Time

We want to ensure that our survey is robust enough to yield actionable data, yet concise enough that it can be completed quickly and easily. By using one survey item for each variable we assess and leveraging our dynamic survey technology to solicit qualitative comments from the most passionate respondents, our survey provides you with invaluable employee input while protecting your organization's time.

References

- Biga, A., McCance, A. S., & Massman, A. J. (2011). Identified employee surveys: Lessons learned. *Industrial and Organizational Psychology: Perspectives on Science and Practice*, 4(4), 449-451. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1111/j.1754-9434.2011.01370.x>
- Borg, I., & Zuell, C. (2012). Write-in comments in employee surveys. *International Journal of Manpower*, 33(2), 206-220. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1108/01437721211225453>
- Couper, M. P., Singer, E., Conrad, F. G., & Groves, R. M. (2008). Risk of disclosure, perceptions of risk, and concerns about privacy and confidentiality as factors in survey participation. *Journal of Official Statistics*, 24(2), 255. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/1266793740?accountid=12259>
- DeVellis, R. F. (2003). *Scale development: Theory and applications (2nd ed.)*. Thousand Oaks, CA: Sage Publications, Inc.
- Dillman, D. A., Smyth, J. D., & Christian, L. M. (2014). *Internet, phone, mail, and mixed mode surveys: The tailored design method (4th ed. ed.)* John Wiley & Sons Inc, Hoboken, NJ. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/1604739658?accountid=12259>
- Durant, L. E., Carey, M. P., & Schroder, K. E. E. (2002). Effects of anonymity, gender, and erotophilia on the quality of data obtained from self-reports of socially sensitive behaviors. *Journal of Behavioral Medicine*, 25(5), 439-67. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1023/A:1020419023766>
- Fink, H. C. (1971). Fictitious groups and the generality of prejudice: An artifact of scales without neutral categories. *Psychological Reports*. 29, 359-365.
- Foddy, W. (1994) *Constructing questions for interviews and questionnaires: Theory and practice in social research*, Cambridge University Press, Cambridge.
- Futrell, C. M., Stem, D. E., Jr., & Fortune, B. D. (1978). Effects of signed versus unsigned internally administered questionnaires for managers. *Journal of Business Research*, 6(2), 91. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/196320937?accountid=12259>
- Haunberger, S. (2011). To participate or not to participate: Decision processes related to survey non-response. *BMS, Bulletin De Methodologie Sociologique*, 109(1), 39-55. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1177/0759106310387721>

Heerwegh, D., Vanhove, T., Matthijs, K., & Loosveldt, G. (2005). The effect of personalization on response rates and data quality in web surveys. *International Journal of Social Research Methodology*, 8(2), 85-99. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1080/1364557042000203107>

Huang, J. L., Liu, M., & Bowling, N. A. (2015). Insufficient effort responding: Examining an insidious confound in survey data. *Journal of Applied Psychology*, 100(3), 828-845. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1037/a0038510>

Kantor, J. (1991). The Effects of Computer Administration and Identification on the Job Descriptive Index (JDI). *Journal of Business and Psychology*, 5(3), 309-323. Retrieved from <http://www.jstor.org/stable/25092290>

Lusty, D. (2009). Find out what your people really think. *Human Resource Management International Digest*, 17(4), 32-36. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1108/09670730910963325>

Mathews, K. R. (2013). *Understanding faculty survey nonrespondents: Their characteristics, organizational citizenship behaviors, workplace attitudes, and reasons for nonparticipation* (Order No. 3592335). Available from ProQuest Dissertations & Theses Global. (1436228545). Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/1436228545?accountid=12259>

McDonald, J. L. (2004). *The optimal number of categories for numerical rating scales* (Order No. 3134422). Available from ProQuest Dissertations & Theses Global. (305208667). Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/305208667?accountid=12259>

Meade, A. W., & Craig, S. B. (2012). Identifying careless responses in survey data. *Psychological Methods*, 17, 437–455. <http://dx.doi.org/10.1037/a0028085>

Pearse, N. (2011). Deciding on the scale granularity of response categories of likert type scales: The case of a 21-point scale. *Electronic Journal of Business Research Methods*, 9(2), 159-171. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/897652808?accountid=12259>

Poncheri, R. M., Lindberg, J. T., Thompson, L. F., & Surface, E. A. (2008). A comment on employee surveys. *Organizational Research Methods*, 11(3), 614-630. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1177/1094428106295504>

Roberts, J. S., Laughlin, J. E. & Wedell, D. H. (1999). Validity Issues in the Likert and Thurstone Approaches to Attitude Measurement. *Emotional and Psychological Measurement*, 59(2), pp 211-233.

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*(6), 1104-1114. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1037/0021-9010.88.6.1104>

Rogelberg, S. G., Luong, A., Sederburg, M. E., & Cristol, D. S. (2000). Employee attitude surveys: Examining the attitudes of noncompliant employees. *Journal of Applied Psychology, 85*(2), 284-293. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1037/0021-9010.85.2.284>

Rogelberg, S. C., Spitzmüller, C., Little, I., & Reeve, C. L. (2006). UNDERSTANDING RESPONSE BEHAVIOR TO AN ONLINE SPECIAL TOPICS ORGANIZATIONAL SATISFACTION SURVEY. *Personnel Psychology, 59*(4), 903-923. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/220137782?accountid=12259>

Rosenfeld, P., Booth-Kewley, S., Edwards, J. E., & Thomas, M. D. (1996). Responses on computer surveys: Impression management, social desirability, and the big brother syndrome. *Computers in Human Behavior, 12*(2), 263-274. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/61501312?accountid=12259>

Saari, L., & Scherbaum, C. (2011). Identified Employee Surveys: Potential Promise, Perils, and Professional Practice Guidelines. *Industrial and Organizational Psychology, 4*(4), 435-448. doi:10.1111/j.1754-9434.2011.01369.x

Schall, M. (2003). Best practices in the assessment of hotel-guest attitudes. *Cornell Hotel and Restaurant Administration Quarterly, 44*(2), 51-65. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/209712203?accountid=12259>

Simsek, Z., & Veiga, J. F. (2001). A primer on internet organizational surveys. *Organizational Research Methods, 4*(3), 218. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/195077582?accountid=12259>

Singer, E., Van Hoewyk, J., and Neugebauer, R. (2003). Attitudes and Behavior: The Impact of Privacy and Confidentiality Concerns on Participation in the 2000 Census. *Public Opinion Quarterly, 65*, 368–384.

Singer, E., von Thurn, D. R., & Miller, E. R. (1995). Confidentiality assurances and response: A quantitative review of the experimental literature. *Public Opinion Quarterly, 59*(1), 66-77. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1086/269458>

Thompson, L. F., & Surface, E. A. (2007). Employee surveys administered online: Attitudes toward the medium, nonresponse, and data representativeness. *Organizational Research Methods, 10*(2), 241-261. Retrieved from <http://ezproxy.mnsu.edu/login?url=https://search-proquest-com.ezproxy.mnsu.edu/docview/195088839?accountid=12259>

Tourangeau, R., & Yan, T. (2007). Sensitive questions in surveys. *Psychological Bulletin, 133*(5), 859-883. doi:<http://dx.doi.org.ezproxy.mnsu.edu/10.1037/0033-2909.133.5.859>



Contact us at info@energage.com
397 Eagleview Blvd. Ste 200
Exton, PA 19341
(484) 323 6300
<http://energage.com>

The Energygage platform helps realize your workplace's full potential by combining cutting-edge HR technology, Top Workplaces insights, and personalized guidance. We combine neuroscience, organizational development, and over 14 million survey responses into clear next steps for you to develop an employee-centric approach to success.



Certified B Corporation

We are certified by the nonprofit B Lab to meet rigorous standards of social and environmental performance, accountability, and transparency.