

# EVIDENCE-BASED SURVEY DESIGN: THE USE OF A MIDPOINT ON THE LIKERT SCALE

Seung Youn (Yonnie) Chyung | Katherine Roberts | Ieva Swanson |  
Andrea Hankinson

Likert-type scales are often used in survey instruments, and practitioners and researchers need to clearly understand the appropriate use of a midpoint in these scales. The authors of this article explore research studies from various disciplines to indicate that there are circumstances when a midpoint should be included and others where it should not. They provide tables, summarizing the benefits and problems in each case as well as evidence-based strategies to employ.

## INTRODUCTION

One of the most popular response scales used in survey design is the Likert scale. In the 1930s, Rensis Likert, an American social psychologist, first introduced a 5-point psychometric scale for measuring a series of attitude-related propositions (Likert, 1932). The wording used in the initial Likert scale was: *Strongly Approve*, *Approve*, *Undecided*, *Disapprove*, and *Strongly Disapprove*. Over time, the wording changed from *Approve* to *Agree*, which resulted in the Likert scale we know today: *Strongly Agree* and *Agree* on one side, and *Disagree* and *Strongly Disagree* on the other side, with a midpoint in the middle.

Due to its simplicity and popularity, the traditional Likert scale expanded into many variations of *Likert-type* scales. These are commonly used in survey instruments designed for: measuring employee performance in the workplace (Purdey, 2013), communications research (Ryan, 1980), marketing research (Garland, 1991), political-opinion research (Raaijmakers, van Hoof, 't Hart, Verbogt, & Vollebergh, 2000), and psychometric research (Kulas & Stachowski, 2013). Its popularity, however, does not necessarily equate to its validity as a data collection method (Johns, 2005). Debates over validity arise from the variations in Likert-type scales in the context of these applications, such as the impacts of using the following formats:

- Including or not including a midpoint in the scale

- Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree
- Strongly Disagree, Disagree, Agree, Strongly Agree
- Using descending order vs. ascending order of the scale options
  - Strongly Agree, Agree, Neutral, Disagree, Strongly Disagree
  - Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree
- Measuring positively and negatively stated survey items with the Likert scale
  - The objectives were clear  
Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree
  - The objectives were unclear  
Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree
  - The objectives were not clear  
Strongly Disagree, Disagree, Neutral, Agree, Strongly Agree
- Using Likert-type scales or slider scales
  - Completely Dissatisfied 0 1 2 3 4 5 6 7 8 9 10 Completely Satisfied
  - A slider scale as shown in Figure 1

A vast amount of research has been done on these topics, and each topic is worth a separate article. The



FIGURE 1. A SLIDER SCALE



FIGURE 2. THE FIVE-POINT LIKERT SCALE ASSUMED AS AN INTERVAL SCALE



FIGURE 3. THE FIVE-POINT LIKERT SCALE AS AN ORDINAL SCALE



FIGURE 4. THE FOUR-POINT ORDINAL LIKERT SCALE WITH UNDECIDED AS A SEPARATE OPTION

purpose of this article is to provide an overview of the Likert scale's characteristics as a measurement method as well as to present research-based evidence and recommendations regarding the use of a midpoint on Likert-type scales.

## THE LIKERT SCALE AS A MEASUREMENT METHOD

### Types of Measurement

To understand the characteristics of the Likert scale as a measurement method, we need to understand different types of measurement. In the 1940s, Stanley Smith Stevens, another American psychologist, formulated four types (levels) of measurement: nominal, ordinal, interval, and ratio scales (Stevens, 1946). Simply put, the nominal scale contains names of equal quality (e.g., the six cells in Thomas Gilbert's (1996) behavior engineering model—*Data, Instruments, Incentives, Knowledge, Capacity, and Motives*). The ordinal scale contains rank-ordered levels (e.g., performance ratings such as *Poor, Fair, Good*). The interval scale contains levels with equal intervals (e.g., a scale consisting of -2, -1, 0, 1, 2). Finally, the ratio scale contains levels with equal ratios and a meaningful zero value (e.g., test scores of zero to 100).

The Likert scale measures attitudes using five points (anchors) where the third point in the middle represents neutrality. Whether the Likert scale is an ordinal scale or an interval scale has been an ongoing debate among researchers. For the Likert scale to be an interval scale, the distances between consecutive points on the scale must be the same, as shown in Figure 2. That is, the distance between *Strongly Disagree* and *Disagree* (the length of "a" in Figure 2) is the same as the distance between *Disagree* and

*Neutral* ("b"), which is the same as the distance between *Neutral* and *Agree* ("c"), as well as the distance between *Agree* and *Strongly Agree* ("d").

On the other hand, the distances between two consecutive points on the Likert scale ("a," "b," "c," and "d") may be different, as shown in Figure 3. This disqualifies the Likert scale from serving as an interval scale but characterizes it as an ordinal scale.

In addition, it is important to word a midpoint appropriately. For instance, *Neutral* (or *Neither Agree nor Disagree*) as a midpoint represents a neutral level of opinion. If, however, *Undecided* is used for the midpoint, it is questionable whether it is truly a midpoint of opinion between disagreement and agreement or whether it should be treated as an absence of opinion. In the latter case, the Likert scale is more like a 4-point ordinal scale with *Undecided* as a separate option off the scale, as illustrated in Figure 4. To collect data that is more likely to be interval than ordinal in nature, other opinions such as *I don't know* or *It depends* should not label a midpoint but could be presented as response options off the scale.

### What If the Likert Scale Is an Interval or Ordinal Scale?

Why does it matter whether the Likert scale is an interval scale or an ordinal scale? A more appropriate question to ask is, "When does it matter?" It matters when we intend to calculate average scores and perform certain statistical analyses on the data collected from the Likert scale. Depending on whether the Likert scale is an ordinal or interval scale, you are either limited or open to employing certain types of data analysis methods.

A 4-point Likert scale is an ordinal scale (as shown in Figure 4). Unlike an interval scale, an average score of "2.5" obtained from a 4-point Likert scale has as little meaning as a score of "fair and a half" from a scale labeled *poor, fair, good, and excellent* (Kuzon, Urbanchek, & McCabe, 1996, p. 266). With ordinal data, one should use median or mode rather than the mean as the measure of central tendency. Furthermore, you should describe a summary of ordinal data with frequencies or percentages of responses in each category (Jamieson, 2004). In addition, certain statistical analyses (such as parametric tests, like an independent samples *t* test or Pearson correlation *r*) require using normally distributed data likely obtained from interval or ratio scales. When data are obtained from ordinal scales such as a 4-point Likert scale

(Figure 4) or do not meet the normality assumption, you would use non-parametric tests such as Chi-Square (cross-tabulation), Mann-Whitney *U*, or Spearman rho.

The 5-point Likert scale can be an ordinal scale, as illustrated in Figure 3; however, it is often treated as an interval scale as shown in Figure 2. Proponents of treating the 5-point Likert scale as an interval scale argue that in practice, it is common to use a group of multiple items measured with the Likert scale. Thus, it is “perfectly appropriate to summarise the ratings generated from Likert scales using means and standard deviations, and it is perfectly appropriate to use parametric techniques like Analysis of Variance to analyse Likert scales” (Carifio & Perla, 2008, p. 1151).

The difference between a 4-point Likert scale and a 5-point Likert scale lies in whether or not a midpoint is included. The presence of a midpoint makes a difference in treating the scale as an interval scale only if survey respondents use the Likert scale’s midpoint for a true neutral meaning, as intended. In turn, the assumption of an interval scale influences the decision whether to use appropriate descriptive statistics among mean, median, mode, standard deviation, frequency, and percentage, and whether to employ appropriate types of inferential statistical analyses. In the following sections, we discuss research findings associated with respondents’ use of a midpoint in the Likert scale and present evidence-based recommendations for designing survey items using the Likert scale.

## RESEARCH FINDINGS ON THE USE OF A MIDPOINT ON LIKERT-TYPE SCALES

### How Do Respondents Interpret and Use a Midpoint?

In structured survey questionnaires, response scales such as the Likert scale are used “to allow respondents to express both the direction and strength of their opinion about a topic” (Garland, 1991, p. 66). You insert a midpoint on the Likert scale to allow respondents to express a neutral opinion between disagreement on one side and agreement on the other. When you present an equal number of options on both the disagreement and agreement sides, including or excluding a midpoint makes the scale an odd-numbered or even-numbered scale (e.g., a 5-point or 6-point scale). Likert-type scales without a midpoint are also characterized as *forced-choice scales*, as respondents are forced to choose either a disagreement or agreement option. For a midpoint of neutrality, *neutral* or *neither agree nor disagree* are often used.

Research has shown, however, that respondents do not always interpret and use a midpoint in the way that scale developers intended. Respondents might select a mid-

*Respondents may use a midpoint as a dumping ground when they are responding to survey items that are unfamiliar to them or to items that are ambiguous or socially undesirable.*

point even if their true opinion is not neutral. In Kulas and Stachowski’s (2009) research, the researchers asked 82 college students to complete survey items with a 5-point Likert scale (with *Neither agree nor disagree* as a midpoint) and also with four options: *It depends*, *Uncertain*, *Average*, and *Not applicable*. Of the four options, respondents who did select the midpoint most commonly chose *It depends* (a 50% probability).

Another concern is that respondents may use a midpoint as a *dumping ground* when they are responding to survey items that are unfamiliar to them, or items that are ambiguous or socially undesirable. In Nadler, Weston, and Voyles’s (2015) study, 635 college students described how they would interpret the midpoint *Neither agree nor disagree* on a 5-point Likert scale. Researchers organized their interpretations into 16 different themes. Some of the themes are:

- Unsure (13%), Undecided (8%), Confused (2%), or Need more information (2%)
- Neutral (10%), Neither (9%), or Middle point (4%)
- No opinion (15%)
- Don’t care (14%)
- Both agree and disagree (10%)
- Not applicable (3%)

Age also seems to influence the use of a midpoint. Raaijmakers, et al. (2000) conducted political opinion research with younger respondents: 3,220 Dutch youths aged 12–24 years in 1991 and 1,887 youths in 1994. The respondents completed survey items with a 5-point Likert scale accompanied by a *Don’t know* option. The researchers found that younger adolescents used the *Don’t know* option or did not respond considerably more often than older respondents. As age increased, *Don’t know* responses and non-responses decreased while midpoint responses increased. The researchers attributed this to *social desirability bias*, which explains that respondents may select an

option that they perceive to be more socially accepted or desirable. That is, selecting a midpoint could be viewed as more socially desirable than selecting a *Don't know* option or leaving it blank.

Garland (1991) made a similar observation about social desirability bias in marketing research he conducted with public grocery shoppers (age 15 or older), using a 4-point scale ( $n = 223$ ) and a 5-point scale ( $n = 225$ ). He found that 14% of the 5-point scale respondents who selected the midpoint (neither/nor) chose a negative option when the midpoint was absent from the scale. In other words, a considerable number of respondents selected a midpoint to avoid what they perceived as a socially undesirable behavior by selecting a negative option.

With the popularity of web-based surveys increasing, respondent expectations that the survey will require a response to each item may create *response tendencies* to complete all items. This creates situations where respondents feel compelled to answer questions about which they have no opinion and where they use the midpoint as an out. In Kulas, Stachowski, and Haynes' (2008) study, 118 college students completed an online survey of a personality assessment with a 5-point Likert scale (the traditional form) and completed it again with a 5-point Likert scale with an additional N/A option (the N/A form). Researchers found an overwhelming tendency for respondents to choose the midpoint on the traditional form if they had chosen N/A on the same item in the N/A form.

## Should You Use a Midpoint or Not?

These research problems associated with the use of a midpoint in response scales bring up the question of whether we should use a midpoint in survey instruments (i.e., whether to use odd-numbered or even-numbered scales). Simply omitting a midpoint from the scales, however, is not the best practice. The more important question that practitioners and researchers should seek to answer is not whether or not to include a midpoint, but rather *when* to omit or present a midpoint in a Likert-type scale.

For instance, research conducted by Matell and Jacoby (1972) with 360 college students revealed that more respondents selected the midpoint (*Uncertain*) on a 5-point scale than on 17 other Likert-type scales varying in the number of anchors (from 2 to 19 anchors), whereas the average completion time was shorter for the 5-point scale than for other scales (excluding the 2-point scale). The fact that many respondents quickly selected the midpoint on the 5-point scale can be explained as *satisficing behavior*: choosing a minimally acceptable response as soon as it is found, instead of putting effort to find an optimal response. Based on their findings, the researchers recommended that when survey designers need to minimize the

*The potential misuse of a midpoint as an “N/A proxy” or “dumping ground” can be reduced by first improving the clarity of survey items and presenting other options such as Not applicable, I don't know, or It depends.*

misuse of a midpoint, it is better to omit the midpoint on coarse scales (scales with few anchors) or increase scale sensitivity (by using many anchors) when including a midpoint.

It may also be desirable to omit a midpoint and offer instead an *I don't know* option when respondents are comfortable with the survey topic (they have formed their opinions), when they are under strong social desirability pressures (John, 2005), or when it is expected that some respondents have little or no involvement in the survey topic (Weems & Onwuegbuzie, 2001).

On the other hand, including a midpoint gives respondents the opportunity to express a neutral opinion, especially on obscure topics (Johns, 2005). A midpoint can also improve psychometric properties such as instrument reliability (Adelson & McCoach, 2010). You can reduce the potential misuse of a midpoint as an “N/A proxy” or a “dumping ground” by first improving the clarity of survey items (Kulas & Stachowski, 2013) and presenting other options off the scale such as *Not applicable*, *I don't know*, or *It depends* (Kulas et al., 2008; Raaijmakers et al., 2000). Table 1 presents a summary of the evidence-based recommendations based on the literature we reviewed.

Another reason for including a midpoint in the Likert scale is to use it as an interval scale and thus apply appropriate statistical analyses, as discussed previously. Including a midpoint in the Likert scale, however, has raised discussions on how it affects the validity and reliability of the survey instrument.

## Does a Midpoint Affect the Validity and Reliability of the Scales?

To generate accurate and reliable data, survey designers must be sure that their survey items actually measure what the designers intended to measure (validity). To test for consistency (reliability), designers typically use a set of many related survey items. As discussed previously, the

TABLE 1

EVIDENCE-BASED RECOMMENDATIONS ON THE USE OF A MIDPOINT ON THE LIKERT SCALE

	INCLUDE A MIDPOINT (STRONGLY DISAGREE, DISAGREE, NEUTRAL, AGREE, STRONGLY AGREE)	OMIT A MIDPOINT (STRONGLY DISAGREE, DISAGREE, AGREE, STRONGLY AGREE)
Benefit	It allows respondents to express their true neutral or indifferent opinion; respondents are not forced to agree or disagree.	Eliminates the possibility that respondents will misuse the midpoint.
Problem	Respondents may use the midpoint as a <i>dumping ground</i> when: <ul style="list-style-type: none"> <li>• They don't know enough about the content asked in the survey.</li> <li>• They are ambivalent about the topic.</li> <li>• They don't care about the topic.</li> <li>• They think their answer depends on other factors.</li> <li>• They want to provide the responses that are more socially acceptable.</li> </ul>	Respondents are not provided with an opportunity to express their neutral opinion. Forcing respondents to take a side may produce biased data.
When	Respondents are familiar with the topic and should be allowed to express a neutral opinion. It is important to use the scale as an interval scale for the statistical analysis purpose.	Respondents are: <ul style="list-style-type: none"> <li>• Unfamiliar or uncomfortable with the survey topic.</li> <li>• Not expected to have formed their opinion about the topic.</li> <li>• Under strong social desirability pressures.</li> <li>• Likely to show satisficing behavior.</li> </ul>
Strategies to Use	Improve clarity of survey items as respondents tend to select a midpoint when they are uncertain about the meaning of the items. Present additional options such as <i>Not applicable</i> , <i>I don't know</i> , or <i>It depends</i> .	Offer an <i>I don't know</i> or <i>Not applicable</i> option instead of forcing respondents to choose an option.

presence of a midpoint can promote satisficing behavior and social desirability bias. Respondents may interpret and use a midpoint as a dumping ground or an easy way out (Kulas & Stachowski, 2009) or select a midpoint rather than selecting socially undesirable options (Garland, 1991; Raaijmakers et al., 2000). Another example is Guy and Norvell's (1977) study that involved 200 college students using both a 5-point Likert scale with a midpoint and a 4-point Likert scale without it. They found that when a midpoint was absent, not only were the midpoint responses distributed to the neighboring options but there were also increased numbers of non-responses. One of the findings was that:

- When a midpoint was present, the distribution of responses showed: SA (205), A (422), N (291), D (402), SD (178), No response (2).
- When a midpoint was omitted, the distribution of responses showed: SA (127), A (568), D (573), SD (202), No response (30).

Based on the fact that the number of no responses increased from two to 30, it is plausible that some of the

people who selected the midpoint when it was present did not use it for a true neutral meaning. These are potential threats to an instrument's validity and reliability.

On the other hand, other studies did not find that the presence and absence of a midpoint affected instrument validity and reliability differently. In Adelson and McCoach's (2010) study with 606 children (grades three to six), researchers tested the validity of a 2-factor model of the instrument by using a 5-point Likert scale with a midpoint and a 4-point Likert scale without it. Regardless of whether a midpoint was present or not, they found a good model fit for the instrument; however, the use of a 5-point Likert scale resulted in a significantly higher reliability estimate for the instrument. Alwin and Krosnick (1991) also observed a pattern of increased reliability with an increasing number of anchors (from three to four, to seven, and to nine) in response scales, although the reliabilities of the 4-point and 5-point scales were not significantly different from each other.

In Leung's (2011) study with 1,217 secondary-school students (13 to 18 years of age) in Macau, the presence of a midpoint in the 5- and 11-point scales compared to the



4- and 6-point scales without a midpoint did not make any differences in factor loadings (a measure of construct validity) and Cronbach's alpha (a measure of reliability). The 6- and 11-point scales, however, produced data that met the normality assumption whereas the data obtained from the 4- and 5-point scales did not.

In Kulas et al.'s (2008) study with college students discussed previously, researchers found no significant difference in variance and reliability estimates of data obtained from the traditional form with a 5-point Likert scale and the N/A form with a 5-point Likert scale and an additional N/A option. However, the researchers warned that the lack of effect on reliability and validity estimates could have been due to the low frequency of the overall N/A responses.

As shown previously, there has been inconsistent evidence regarding the impacts of inclusion of a midpoint on validity and reliability. Krosnick and Fabrigar (1997) also indicate that experimental studies have failed to show a clear pattern in the effects of a midpoint on scale reliability. The research findings, however, seem to point to another issue to consider: the number of anchors used in response scales.

## How Many Anchors Should You Use?

When determining the number of anchors to use in response scales, the survey designer should consider various factors, including strategies to minimize potential biases, as illustrated previously. Research has shown that, as the number of anchors increased, the use of a midpoint decreased; thus, it is better to use many anchors in scales when you need to discourage respondents' use of a midpoint (Matell & Jacoby, 1972). Respondents may also prefer having a sufficient number of options from which to choose. With 149 adults, Preston and Colman (2000) studied respondents' preferences for response scales based on three criteria: (1) was easy to use, (2) was quick to use, and (3) allowed adequate expression of feelings. The response scales differed in the number of options included (2, 3, 4, 5, 6, 7, 8, 9, 10, 11, and 101) and were end-anchored (e.g., *Very poor* 1 2 3 4 5 *Very good*). The data showed that 5-, 7-, and 10-point scales were the easiest to use, while (unsurprisingly), 2-, 3-, and 4-point scales were the quickest to use. On the other hand, respondents felt that the scales with more options (9-, 10-, 11-, and 101-point scales) allowed them greater expression of feelings. The researchers concluded that overall, the 10-point scale scored best, followed by the 7-point and 9-point scales.

In contrast, Chang (1994) points to a problem in using many anchors in scales. By comparing the reliability of two even-numbered Likert scales—a 4-point scale (*Disagree, Somewhat disagree, Somewhat agree, Agree*) and a

6-point scale (*Strongly disagree, Disagree, Somewhat disagree, Somewhat agree, Agree, Strongly agree*)—the researcher found a tendency of respondents to skip categories in the 6-point scale and an inability to differentiate between two similar categories such as *Strongly disagree* and *Disagree*, leading to a possibility that respondents may use the categories interchangeably.

This leads to the importance of considering information-processing issues in survey design. Chen, Yu, and Yu (2015) conducted an eye-tracking experiment with a sample of 184 Chinese people to measure respondents' reaction time in answering a question as evidence of their cognitive effort. They found that a 5-point scale required the shortest reaction time, compared to 4-, 6-, 7-, 8-, and 9-point scales. Also when considering the data showing potential *acquiescence bias* (respondents' tendency to choose positive options) and *extreme response bias* (respondents' tendency to disproportionately select extreme responses), the researchers determined five as the optimal number of anchors and emphasized keeping an optimal balance in the scale design in order to lower the respondents' cognitive effort while maximizing the communication of information within the scale.

Similarly, in a study with 1,207 male Dutch-speaking internet users (aged from 15 to 65 with a median age of 49) comparing 4-, 5-, 6-, and 7-point Likert-type scales, Weijters, Cabooter, and Schillewaert (2010) concluded that including a midpoint resulted in lower levels of extreme responses and that using fully labeled scales with a midpoint produced lower levels of mis-responses when reversed items were used. They recommended using fully labeled 5-point scales for the general population and fully labeled 7-point scales for populations with high levels of verbal skills and experience with using survey questionnaires.

Based on these research findings, it seems that the decision as to the number of anchors to use should depend largely on what the survey designer deems important in terms of ways to lower the respondents' cognitive effort and boost the ease of taking the survey, while allowing respondents to express their opinions sufficiently, especially when respondents are under time pressure. For these situations, the use of 5- or 7-point scales is commonly suggested.

## SUMMARY

During the past several decades, researchers have studied the effects of the presence or absence of a midpoint in response scales on respondents' behavior and have discussed both benefits and challenges, as well as suggesting ways to mitigate respondents' potential misuse of a midpoint. The decision to enhance instrument reliability and validity by

TABLE 2

## RESEARCH EVIDENCE FOR THE USE OF A MIDPOINT ON THE LIKERT SCALE

FOCUS	AUTHORS (YEAR)	RECOMMENDATIONS BASED ON RESEARCH FINDINGS
Inclusion vs. exclusion of a midpoint	Garland (1991)	Offering a midpoint is largely based on researchers' preference. Social desirability bias can be minimized by omitting a midpoint.
	Johns (2005)	Omit a midpoint when respondents are comfortable with the subject matter and under social desirability pressures. Vary formats within a questionnaire to increase respondents' awareness.
	Kulas et al. (2008)	Include an <i>N/A</i> option in addition to a midpoint, especially in online survey questionnaires.
	Kulas and Stachowski (2009)	Improve clarity of survey items. Offer an <i>It depends</i> option.
	Kulas and Stachowski (2013)	Use a midpoint cautiously; improve clarity of survey items. Offer alternative <i>I don't know</i> or <i>It depends</i> options.
	Nadler et al. (2015)	Use a midpoint with caution; clearly define the meaning of a midpoint for respondents. Offer a <i>No opinion</i> option and treat <i>No opinion</i> as missing data.
	Raaijmakers et al. (2000)	Do not remove a midpoint simply based on the problem associated with its dual meaning. Provide an <i>I don't know</i> option.
	Ryan (1980)	Differentiate between <i>I don't know</i> , <i>Neutral</i> , and <i>No opinion</i> . <i>Neutral</i> is a definite opinion, while <i>No opinion</i> is not.
Validity and reliability of the instrument and the number of anchors	Weems and Onwuegbuzie (2001)	Omit a midpoint when respondents are less involved in the content. Provide other response categories (e.g., <i>I don't know</i> ).
	Adelson and McCoach (2010)	Can use a five-point scale including a midpoint when surveying with children in grades three to six.
	Alwin and Krosnick (1991)	Be aware that response scales with more categories are more reliable; fully labeled 7-point scales are more reliable than those not so labeled; reliability does not improve by offering <i>I don't know</i> .
	Chang (1994)	When using many anchors in scales, be aware of respondents' inability to differentiate between two similar categories such as <i>Strongly disagree</i> and <i>Disagree</i> .
	Chen et al. (2015)	Keep an optimal balance in the scale design in order to lower the respondents' cognitive effort while maximizing the communication of information within the scale. Use a 5-point scale, as it is the best option from an information-processing perspective.
	Guy and Norvell (1977)	Be aware that if respondents are familiar with using a 5-point scale, the use of a 4-point scale may produce somewhat distorted data.
	Leung (2011)	Use an 11-point scale for improved psychometric performance (meeting the normality assumption).

(Continued)

TABLE 2

Continued

FOCUS	AUTHORS (YEAR)	RECOMMENDATIONS BASED ON RESEARCH FINDINGS
	Matell and Jacoby (1972)	Use many anchors in scales when needing to minimize respondents' (mis)use of a midpoint; as the number of anchors decreased, the use of a midpoint increased.
	Preston and Colman (2000)	Be aware of respondents' preference toward having a sufficient number of options from which to choose (e.g., a 10-point scale, followed by 7-point and 9-point scales). Trade off reliability and validity if needed to improve factors such as respondent preferences.
	Weijters et al. (2010)	Use fully labeled 5-point scales for the general population and fully labeled 7-point scales for populations with high cognitive skills.

choosing to include or exclude a midpoint on the response scale cannot be easily isolated from other factors, such as scale sensitivity, age and education level of respondents, absence or presence of a midpoint label, and choice of a midpoint label. When designing a scale, we suggest employing the strategies discussed previously to reduce response bias and the possibility of respondents misusing the midpoint as a catch-all. For reference, Table 2 includes a summary of the research evidence used in generating our recommendations. 🌞

## References

- Adelson, J.L., & McCoach, D.B. (2010). Measuring the mathematical attitudes of elementary students: The effects of a 4-point or 5-point Likert-type scale. *Educational and Psychological Measurement*, 70(5), 796–807. <https://doi.org/10.1177/0013164410366694>
- Alwin, D.F., & Krosnick, J.A. (1991). The reliability of survey attitude measurement. In D.F. Alwin (Ed.), *Sociological methods & research* (Volume 20, Issue 2, pp. 139–181). Thousand Oaks, CA: Sage.
- Carifio, J., & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education*, 42(12), 1150–1152. <https://doi.org/10.1111/j.1365-2923.2008.03172.x>
- Chang, L. (1994). A psychometric evaluation of 4-point and 6-point Likert-type scales in relation to reliability and validity. *Applied Psychological Measurement*, 18(3), 205–215.
- Chen, X., Yu, H., & Yu, F. (2015). What is the optimal number of response alternatives for rating scales? From an information processing perspective. *Journal of Marketing Analytics*, 3(2), 69–78. <https://doi.org/10.1057/jma.2015.4>
- Garland, R. (1991). The mid-point on a rating scale: Is it desirable? *Marketing Bulletin*, 2, 66–70. Available at: [http://marketing-bulletin.massey.ac.nz/V2/MB\\_V2\\_N3\\_Garland.pdf](http://marketing-bulletin.massey.ac.nz/V2/MB_V2_N3_Garland.pdf)
- Gilbert, T.F. (1996). *Human competence: Engineering worthy performance* (Tribute ed.). Washington, D.C: The International Society for Performance Improvement.
- Guy, R.F., & Norvell, M. (1977). The neutral point on a Likert scale. *The Journal of Psychology*, 95, 199–204.
- Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education*, 38(12), 1217–1218. <https://doi.org/10.1111/j.1365-2929.2004.02012.x>
- Johns, R. (2005) One size doesn't fit all: Selecting response scales for attitude items. *Journal of Elections, Public Opinion and Parties*, 15(2), 237–264. <https://doi.org/10.1080/136898805001788>
- Krosnick, J.A., & Fabrigar, L.R. (1997). Designing rating scales for effective measurement in surveys. In L. Lyberg, P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz, & D. Trewin (Eds.), *Survey measurement and process quality* (pp. 141–164). New York, NY: John Wiley & Sons, Inc. <https://doi.org/10.1002/9781118490013.ch6>
- Kulas, J.T., & Stachowski, A.A. (2009). Middle category endorsement in odd-numbered Likert response scales: Associated item characteristics, cognitive demands, and preferred meanings. *Journal of Research in Personality*, 43(3), 489–493. <https://doi.org/10.1016/j.jrp.2008.12.005>
- Kulas, J.T., & Stachowski, A.A. (2013). Respondent rationale for neither agreeing nor disagreeing: Person and item contributors to middle category endorsement intent on Likert personality indicators. *Journal of Research in Personality*, 47(4), 254–262. <https://doi.org/10.1016/j.jrp.2013.01.014>



- Kulas, J., Stachowski, A., & Haynes, B. (2008). Middle response functioning in Likert-responses to personality items. *Journal of Business & Psychology*, 22(3), 251–259. <https://doi.org/10.1007/s10869-008-9064-2>
- Kuzon, W.M., Urbanchek, M.G., & McCabe, S. (1996). The seven deadly sins of statistical analysis. *Annals of Plastic Surgery*, 37(3), 265–272.
- Leung, S-O. (2011) A comparison of psychometric properties and normality in 4-, 5-, 6-, and 11-point Likert scales. *Journal of Social Service Research*, 37(4), 412–421. <https://doi.org/10.1080/01488376.2011.580697>
- Likert, R. (1932). A technique for the measurement of attitudes. In R.S. Woodworth (Ed.), *Archives of Psychology* (Vol. 22, No. 140, pp. 5–55). New York, NY: The Science Press.
- Matell, M.S., & Jacoby, J. (1972). Is there an optimal number of alternatives for Likert-scale items? Effects of testing time and scale properties. *Journal of Applied Psychology*, 56(6), 506–509.
- Nadler, J.T., Weston, R., & Voyles, E.C. (2015). Stuck in the middle: The use and interpretation of mid-points in items on questionnaires. *The Journal of General Psychology*, 142(2), 71–89. <https://doi.org/10.1080/00221309.2014.994590>
- Preston, C.C., & Colman, A.M. (2000). Optimal number of response categories in rating scales: Reliability, validity, discriminating power, and respondent preferences. *Acta Psychologica*, 104(2000), 1–15.
- Purdey, B. (2013). Occupant stimulus response workplace productivity and the vexed question of a measurement. *Workplace Productivity*, 31(11), 505–520. <https://doi.org/10.1108/F-03-2012-0021>
- Raaijmakers, Q.A.W., van Hoof, A., 't Hart, H., Verbogt, T.F.M.A., & Vollebergh, W.A.M. (2000). Adolescents' midpoint responses on Likert-type scale items: Neutral or missing values? *International Journal of Public Opinion Research*, 12(2), 208–216. <https://doi.org/10.1093/ijpor/12.2.209>
- Ryan, M. (1980). The Likert scale's midpoint in communication research. *Journalism Quarterly*, 57(2), 305–313.
- Stevens, S.S. (1946). On the theory of scales of measurement. *Science*, 103(2684), 677–680. <https://doi.org/10.1126/science.103.2684.677>
- Weems, G.H., & Onwuegbuzie, A.J. (2001). The impact of midpoint responses and reverse coding on survey data. *Measurement and Evaluation in Counseling and Development*, 34(3), 166–176.
- Weijters, B., Cabooter, E., & Schillewaert, N. (2010). The effect of rating scale format on response styles: The number of response categories and response category labels. *International Journal of Research in Marketing*, 27(2010), 236–247.

SEUNG YOUN (YONNIE) CHYUNG, Ed.D., is a professor in the Department of Organizational Performance and Workplace Learning in the College of Engineering at Boise State University (<http://opwl.boisestate.edu/faculty-staff/faculty/yonnie-chyung/>). She teaches graduate courses on Program Evaluation and Quantitative Research in Organizations. She runs a workplace-oriented research central (WORC) lab where teams of practitioners and researchers conduct research on various topics in the HPI context. She may be reached at [yhyung@boisestate.edu](mailto:yhyung@boisestate.edu)

KATHERINE ROBERTS is a senior instructional designer with WellMed under UnitedHealth Group's umbrella in San Antonio, Texas. She is a graduate assistant for Dr. Seung Youn (Yonnie) Chyung and attends Boise State University as a graduate student in the Organizational Performance and Workplace Learning program. She expects to complete her Master's Degree in May of 2017. She may be reached at [katherin.roberts@gmail.com](mailto:katherin.roberts@gmail.com)

IEVA SWANSON, MS, is an instructional designer for Scottrade, Inc. in St. Louis, Missouri. Since 2007, her professional and volunteer experience has included facilitation, instruction, learning management system administration, and multi-media course design. She earned her MS in Organizational Performance and Workplace Learning from Boise State University in 2016. She may be reached at [ievaswanson@u.boisestate.edu](mailto:ievaswanson@u.boisestate.edu)

ANDREA HANKINSON, MA, MS, is the performance excellence consultant at the Public Service Commission, a department of the Government of Alberta, in Edmonton, Alberta, Canada. She holds an MS in Organizational Performance and Workplace Learning from Boise State University and an MA in English from Simon Fraser University. She has worked extensively in online and distance learning and is a freelance writer and instructor of writing. She may be reached at [andrea.hankinson@gmail.com](mailto:andrea.hankinson@gmail.com)

Copyright of Performance Improvement is the property of John Wiley & Sons, Inc. and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.