Likert Scale: Misuse of Mid-Point Anchor

tony lam[1]

[1]U. of Toronto

February 25, 2019

Abstract

The paper is a review paper based on lit search and authors' research experiences.

Likert Scale: Misuse of Mid-Point Anchor

Abstract

Validity of responses to surveys is affected by how respondents interpret the items and the anchors used in the rating scale. We show that under some circumstances, the mid-point anchor of the traditional 5-point Likert scale could be interpreted as a non-substantive response reflecting a lack of knowledge answer. An effective way to encourage respondents to read the mid-point anchor as a point representing their attitude in the middle of the continuum–is to provide a "Don't Know" option at the end of the rating scale with a gap to indicate that it is a non-substantive response option.

Key Words: Likert Scale, Survey, Scale Construction

Likert Scale: Misuse of Mid-Point Anchor

Surveys have long been recognized as an efficient method to gather opinion or factual information and to assess attitudes for large numbers of respondents (Tourangeau, Rips, & Rasinski, 2000). Decades ago, different scaling methods, such as Likert, Thurstone, Guttman, and Semantic Differential, were developed to measure attitude. According to Carifio and Perla (2008):"Likert methodology is one of the most commonly used methodologies in all fields of research, but particularly so in allied health, medicine and medical education" (p.1151).

In Likert's (1932) original formulation, the rating scale consisted of five points with "agree" and "disagree" labels on which the respondent indicated direction and intensity of agreement with declarative statements. In 2004, Jamieson echoed the first of the "seven deadly sins of statistical analysis" presented by Kuzon et al. (1996) and argued that use of a Likert scale yields data that are ordinal and hence parametric statistics should not be used to analyze these data. Since then, strong reactions from researchers have refuted her "ordinalist" view on Likert scale data analysis. (Pell, 2005; Carifio & Perla, 2007, 2008; Norman, 2010). [Her response to critics was published in 2005 in the same journal.]

The counter-argument is, even though the Likert scale format yields ordinal *item* responses, the derived *scale* scores are interval, and, even if ordinal, as long as there are 4 to 8 items per scale or subscale and 5 to 7 anchors points are used (Carifio & Perla, 2007), parametric statistics can still be applied because parametric tests are robust. The debate appears to have been settled in favor of the "intervalist" position. Norman (2010) declared: "Parametric statistics can be used with Likert data, with small sample sizes, with unequal variances, and with non-normal distributions, with no fear of 'coming to the wrong conclusion.' These findings are consistent with empirical literature dating back nearly 80 years. The controversy can

cease (but likely won't)." No response from the "ordinalists" has been heard since. In this paper, we put forth another often ignored problem in the use of the Likert scale.

Despite the controversies about the Likert scale, two issues all researchers have agreed on are (1) anchors must be chosen so they are roughly equidistant on the continuum underlying attitude being measured, and (2) the selection of an anchor must be a substantive response, a unit of information reflecting a point on the continuum.

Theoretically, every abstract construct exists on an underlying continuum. Discrete anchor points along the continuum are often provided for respondents to choose from. Choice of anchors in an attempt to create an equal interval rating scale has been researched for over three decades (e.g., Lam & Klockers, 1982). In addition to subjective judgment of the intensity of various anchors, through a rank ordering procedure, "agreement," "evaluation," and "frequency" anchors have been scaled. The derived numerical indices show the positions of the anchors on the underlying continuum to assist researchers in "choosing equally spaced response categories for summated rating scales" (Spector, 1976, p.374).

Likert used the five anchors: "strongly agree," "agree," "neither agree nor disagree," "disagree," and "strongly disagree." Since then, the middle anchor has been variously labeled as "Neutral" (Armstrong, 1987), "Undecided," (Armstrong, 1987; Lock, 2001), "Uncertain" or "not sure," (Klopfer & Madden, 1980) and "?" (DuBois & Burns, 1975; Klopfer & Madden,1987).11In a dissertation read by the first author, "don't know" was used as the mid-point anchor. In all of these cases, the distance between the mid-point anchor and the two adjacent anchors, "agree" and "disagree," seems the same. However, it is not clear if all these middle anchors are interpreted by respondents as an attitude with intensity at the middle of the agreement-disagreement bipolar continuum and *not* used as "don't know" (DK). Under some circumstances, the mid-point response to an item could be a non-substantive response reflecting a lack of knowledge answer. When such an item response is scored and included in the computation of the scale or sub-scale score, validity of the aggregated score suffers. We believe that this error has occurred but is typically undetected in survey research. The focus of this paper was to review the potential misuse of the middle anchor of the Likert scale and offer a strategy to minimize this potential misuse.

Interpretation of the Middle Anchor of Likert Scale

There has been substantial debate about using an even or odd number of rating scale points in attitude scaling with surveys. When an odd-numbered rating scale is used, researchers are confronted with an issue: how to label the mid-point. Not surprisingly, research on interpretation of the middle option finds a lack of consistency regarding respondent interpretation of the midpoint (Baka, Figgou, & Triga, 2012; Kalton, Roberts & Holt, 1980; Nadler, Weston, & Voyles, 2015). When respondents were allowed to choose the label they associated with a rating scale mid-point, different labels were used for items assessing attitudes towards capital punishment and Sunday observance, suggesting different meanings for the middle response category as a function of item content. Nadler et al. (2015) found interpretation of the mid-point as "don't care," "no opinion," "unsure," "neutral," "equal/both," and "neither" was found with a sample of U.S. undergraduates asked about politics in the U.S. Baka et al. (2012) found the mid-point was interpreted as lacking in knowledge or indifferent, ambivalent, disputing aspects of the question, or had no information on the topic. When respondents choose the mid-point as their response, it is possible the respondents are telling us that they do not have a response because they either have no opinion, are unclear/uncertain what their opinions are, or don't care; that they do have a response and the response is that they have an opinion that is moderate in intensity, or they do not understand the question, or have no information about the topic. In sum, the question is: are the mid-point Likert scale responses substantive or non-substantive?

In attitude scaling, if respondents have access to the necessary information and neither agree nor disagree with the issues and they choose the mid-point anchor to indicate their state of mind, the response is a substantive response and it should be coded and add to the composite scale score. On the other hand, a non-substantive response is a response that does not contribute or add to the measurement of the target attitude. However, the meaning of a non-substantive response is unclear, as it is multifaceted and could indicate carelessness,

confusion, or intentional behavior (Francis & Busch, 1975). A non-substantive response can be a DK response implying that respondents cannot retrieve the relevant information from their long-term memory because they either do not have that information or cannot retrieve the information at the time to form an opinion. Alternatively, a non-substantive response is a random response if respondents guess because they do not understand the question or if they simply satisfice because they do not wish to spend the effort to provide a thoughtful response. It can also be a systematic but irrelevant response, due to deliberate efforts to present a certain impression or to some unconscious stable response style such as leniency, neutrality, or an acquiescent response tendency.

The Use of the DK Option

Inclusion of a DK option in attitude scales can be problematic when a Likert scale is used because its meaning overlaps with the middle option. As noted above, depending on the label used, respondents interpret the middle option differently, including as a response reflecting an absence of agreement or disagreement with a statement (substantive response) and as a lack of a response (non-substantive response) (Klopfer & Madden, 1980). The former interpretation is consistent with Likert's intention, a neutral attitude toward the object reflected by the item, and the latter justifies selection of a DK response choice.

Illustrative of this potential confusion, researchers found respondents with clearly defined attitudes who initially provided DK responses to questions related to those attitudes (Gilljam & Granberg, 1993; O'Muircheartaigh, Krosnick, & Helic, 2000). Gilljam and Granberg (1993) found Most (80%) of those subjects indicating DK to the first question took a position on the subsequent two questions, and the positions they took were predictive of subsequent behavior in responding to three questions related to a nuclear referendum. In contrast, Feick (1989) reported DK responses are provided by truly uninformed respondents as well as by respondents providing equivocal responses to attitudinal questions. Schuman and Presser (1996) reported that 30% of respondents would provide a substantive opinion on a law *they know nothing about* if *not* given an explicit DK option.

While Alwin and Krosnick (1991) found that allowing DK responses had no effect on the reliability of attitude survey questions, Schuman and Presser (1996) reported mixed findings with respect to the impact of inclusion of the DK option on the correlation between items. They reported that in addition to attitude intensity, education and knowledge levels also affected the use of the DK response, and those with less developed or weaker attitudes used the DK option more often. Francis and Busch (1975) also found increased use of DK with lower educational and SES levels.

These findings raise the question of what DK means. Does a DK response mean a lack of access to information required to formulate and provide a response, uncertainty or a neutral response, or unwillingness to provide a response? Is it interpreted differently by informed and uninformed respondents? Without disentangling the meaning of the DK and the middle-option response, the validity of the research findings is threatened.

In a study with University students (Lam, Allen, & Green, 2011), the authors examined how respondents with (informed) and without (uninformed) the necessary knowledge to form an opinion utilized Neutral and DK response options when presented alone or together and when placement of the option varied—placed as a mid-point or at the end of the scale. Five findings are notable. First, both informed and uninformed respondents used the DK and Neutral anchors interchangeably, implying a blurred semantic distinction between the two terms as suggested by the literature.

Second, when *both the DK and Neutral options were provided* , uninformed respondents chose the DK more often than the Neutral, and, the proportion of DK response for the uninformed group was much higher than for the informed group. This finding suggests higher validity of responses when both DK and Neutral options are provided than when only one of the two options is provided.

Third, when *both the DK and Neutral options were provided* , both informed and uninformed respondents selected the DK option more than the Neutral option when either one of these two options was located in the middle of the rating scale. These results are consistent with earlier findings of Krosnick (1999) and

Schuman and Presser (1996) who observed greater use of the nonresponse option when located earlier in the response scale; i.e., a primacy effect for options presented in a written format. The meaning of the DK response option is apparent when it is placed physically as the mid-point, and it is chosen as the appropriate response. This finding suggests a placement effect.

Fourth, when offered *either a DK or Neutral option* but not both, when presented in the middle of the rating scale, both informed and uninformed participants chose that option more often than when presented at the end of the scale. This finding further supports a placement effect.

Fifth, if *either a DK or Neutral option was provided,* for the uninformed respondents, when the Neutral option was placed at the end of the response scale (with an extra space separating it from the other rating options), they were more likely to interpret it as a non-substantive response option than when the option was placed in the middle of the scale. However, the informed respondents tended to choose either the neutral or DK option more often when placed in the middle rather than at the end of the response scale. This observation implies that informed respondents may rely on placement as a satisficing strategy when they do have the necessary knowledge to respond to an item, but perhaps lack sufficient attitude strength to reply meaningfully (Schuman & Presser, 1996). The uninformed respondents lacked the knowledge to provide substantive responses, meaning they did not need a satisficing strategy. Instead, they simply looked for a way to indicate an absence of attitude or evaluative judgment or opinion by selecting a no response option.

Conclusion

Our review of the literature and our own empirical study lead us to conclude that the DK label and labels used for the middle position of the Likert scale are interpreted similarly; this interpretation is affected by whether or not both the DK and mid-point anchor are used in the rating scale and the positions of these two anchors either alone or together.

If respondents lack knowledge or interest about attitudes assessed in a survey or are incapable of understanding or responding to a survey item, forcing a response increases the likelihood of random error and consequently decreases reliability and validity of the survey results. It may be easier for respondents to be aware of a lack of necessary knowledge or information to respond to a factual question than to an attitude question.[26] Interpretation of a DK response to factual questions is straightforward and providing the DK response option to factual questions should enhance validity of the survey results. In factual surveys, higher frequency of DK responses reflects less familiarity with the topic area, and less ability to respond. Low involvement in the target issues (Durand & Lambert, 1988), and less exposure to the target issues and low degrees of topical knowledge (Falkenberry & Mason, 1978) also affect use of DK. Consequently, inclusion of the DK option in surveys of factual knowledge should improve validity of responses. Our review suggests that this recommendation is generalizable to attitude items.

Those without knowledge about an attitudinal item or issue prefer to acknowledge their lack of evaluative response when allowed to do so. Therefore, in order to ensure that uninformed respondents provide the most accurate response possible, we recommend that a DK option be provided to all Likert rating scale formats, especially if a 5-point or any other uneven-numbered format is used. Ideally, this should be placed at the end of the rating scale with formatting (like extra spaces separating the substantive scale from DK) to indicate to the respondents that the DK response option is not part of the rating scale.

We also advise that a middle option should be provided to respondents, and it should be labeled as Neutral or neutral-sounding (such as "Neither Satisfied nor Dissatisfied"). First, this allows for a substantive response that can be interpreted as a middle state of endorsement, intensity, or simply a logical middle ground between two polar anchors (such as "Moderate" between "Liberal" and "Conservative"). As we have noted, Likert's intention was that this middle point be third in a five-point scale; but can be generalized to any odd-numbered response scale. Also, when provided with a separate DK or other non-substantive response option (such as "Not Applicable"), it ensures that apparently substantive responses are just that, that the Neutral response option is available for those whose "true" attitudinal position is a middle one, and that the middle point does not have to serve multiple functions. This will be especially true when scale points

are clearly labeled as they are intended to be used. We hope we have contributed to the list of guidelines for avoiding misuse of the Likert scale. In closing, we want to suggest that other rating scale anchors such as frequency (e.g., often, seldom) and evaluative (e.g., poor, good) labels should be explored as they might function more effectively than Likert scale format for certain content to elicit reliable and valid responses.

References

Alwin, D., & Krosnick, J. (1991). The reliability of survey attitude measurement: The influence of question and respondent attributes.*Sociological Methods and Research, 20* , 139-181.

Armstrong, R. L. (1987). The mid-point on a Likert-type scale.*Perceptual and Motor Skills, 64,* 359-362.

Baka, A., Figgou, L., & Triga, V. (2012). "Neither agree, nor disagree": A critical analysis of the middle answer category in Voting Advice Applications. *International Journal of Electronic Governance* , *5* (3), 244–263.

Bauer, M. (1996). Socio-demographic correlates of DK-responses in knowledge surveys: Self-attributed ignorance of science. *Social Science Information, 35,* 39-68.

Carifio, J. & Perla, R. (2008). Resolving the 50-year debate around using and misusing Likert scales. *Medical Education, 42,*1150-1152.

Carifio, J. & Perla, R. (2007). Ten common misunderstanding, misconceptions, persist myths and urban legends about Likert scales and Likert response formats and their antidotes. *Journal of Social Sciences, 3* (3), 106-116.

DuBois, B., & Burns, J. A. (1975). An analysis of the meaning of the question mark response category in attitude scales. *Educational and Psychological Measurement, 35,* 869-884.

Durand, R. M., & Lambert, Z. L. (1988). Don't know responses in surveys: Analyses and interpersonal consequences. *Journal of Business Research, 16,* 169-188.

Falkenberry, G. D., & Mason, R. (1978). Characteristics of nonopinion and no opinion response groups. *Public Opinion Quarterly, 42,*533-545.

Feick, L. M. (1989). Latent class analysis of survey questions that include "don't know" responses. *Public Opinion Quarterly, 53,*525-547.

Francis, J. D., & Busch, L. (1975). What we don't know about "I don't knows". *Public Opinion Quarterly, 34* , 207-218.

Gilljam, M., & Granberg, D. (1993). Should we take "don't know" for an answer? *Public Opinion Quarterly, 57,* 348-357.

Jamieson, S. (2004). Likert scales: How to (ab)use them. *Medical Education, 38,* 1212- 1218.

Jamieson, S. (2005). Uses and misuses of Likert scales/Author's reply.*Medical Education,* 39-97.

Kalton, G., Roberts, J., & Holt, D. (1980). The effects of offering a middle response option with opinion questions. *The Statistician, 29,* 11-24.

Klopfer, F. J., & Madden, T. M. (1980). The middlemost choice on attitude items: Ambivalence, Neutrality or Uncertainty?*Personality and Social Psychology Bulletin, 6,* 97-101.

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

Kuzon, W.M., Urbanchek, M.G., & MaCabe, S. (1996). The seven deadly sins of statistical analysis. *Annals of Plastic Surgery, 37,*265-272.

Lam, T.C.M., & Klockars, A.J. (1982, Winter). Effect of choice of anchor points on the equivalence of questionnaire items. *Journal of Educational Measurement, 19(* 4), 317-322.

Lam, T.C.M., Allen, G. & Green, K.E. (2010. April). *Is "Neutral" on a Likert Scale The Same As "Don't Know" for Informed and Uninformed Respondents? Effects of Serial Position and Labeling on Selection of Response Options.*Paper presented at the annual meeting of the National Council of Measurement in Education.

Likert, R. (1932). A technique for the measurement of attitudes.*Archives of Psychology, 140,* 55.

Lock, C.L. (2001). *The influence of a large-scale assessment program on classroom practices* . Unpublished Doctoral thesis, Queen's University, Kingston, Ontario, Canada.

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,* 71-89.

Norman, G. (2010). Likert scales, levels of measurement and the "laws" of statistics. *Advances in Health Science Education, 15,* 625-632.

O'Muircheartaigh, C., Krosnick, J. A., & Helic, A. (2000). Middle alternatives, acquiescence, and the quality of questionnaire data. Unpublished paper.

Pell, G. (2005) Uses and misuses of Likert scales. *Medical Education,* 39-37.

Schuman, H., & Presser, S. (1996). *Questions and answers in attitude surveys: Experiments on question form, wording and context.*Thousand Oaks, CA: Sage.

Spector, P.E. (1976). Choosing response categories for summated rating scales. *Journal of Applied Psychology, 61* (3), 374-375.

Tourangeau, R., Rips, L. J., & Rasinski, K. (2000). *The psychology of survey response* . Cambridge, UK: Cambridge University Press.