Wetzelhütter, Daniela. 2020. "Scale-Sensitive Response Behavior!? Consequences of Offering versus Omitting a 'Don't Know' Option and/or a Middle Category." *Survey Practice* 13 (1). https://doi.org/10.29115/SP-2020-0012.

ARTICLES

Scale-sensitive response behavior!? Consequences of offering versus omitting a "don't know" option and/or a middle category

Daniela Wetzelhütter[1][a]

[1] Healthcare-, Social- and Public Management, University of Applied Sciences upper Austria

Keywords: scale-sensitive, response behaviour, don't know option, middle category

https://doi.org/10.29115/SP-2020-0012

Survey Practice

Vol. 13, Issue 1, 2020

The effect of a "don't know" option (DK-O) and of a "middle category" (MC) on the survey response has been frequently examined. While researchers agree on some points, they disagree on how these options should be interpreted. For instance, a DK-O or MC may not necessarily be chosen due to a lack of a neutral/moderate opinion. This article addresses the consequences of offering versus omitting a MC or/and a DK-O. Data were collected by means of an online survey dealing with student participation (n=1,282). The results showed respondents who choose a DK-O may not have a neutral opinion. Conversely, respondents who choose a MC may not necessarily be undecided.

Introduction

This article begins from the intensively studied question of whether the presence or absence of answer options for expressing a lack of opinion, such as a "don't know" option (DK-O; e.g., Alwin and Krosnick 1991; Andrews 1984; de Leeuw, Hox, and Boevé 2016; O'Muircheartaigh and Helic 2000) or a neutral/moderate opinion, such as a midpoint (e.g., Baka, Figgou, and Triga 2012; Garland 1991; Hurley 1998; Moors 2008; Sturgis, Roberts, and Smith 2014), affects the quality of data on attitude. Researchers agree on some points but follow different convictions or have different preferences in other aspects. Reasons for respondents choosing a DK-O or middle category (MC) may be manifold and not necessarily due to a lack of a neutral/moderate opinion.

Krosnick and Fabrigar (1997) already summarized five explanations for choosing a DK-O, apart from a lack of knowledge, namely ambivalent attitudes when a MC is missing, neutral views when a neutral response is missing, inability to understand the question, uncertainty about the meaning of the scale points, and satisficing. Based on contradictory results of previous studies (conducted by different researchers, mainly in the 1980s and 1990s), Krosnick and Fabrigar (1997) conclude that omitting no-opinion options for attitude questions would lead to more reported "real" attitudes rather than

a E-mail: daniela.wetzelhuetter@fh-linz.at, Phone number: 0043 50804-52430

Figure 1: Literature recommendation for omitting vs. providing a DK-O or MC (if certain conditions apply)

Response category	Recommendation	Condition
Don't know	Omit	when respondents are familiar with the topic (Menold and Bogner 2016); otherwise, it reduces the proportion of "real" attitudes (Krosnick and Fabrigar 1997).
	Reverse (provide)	when respondents are unfamiliar with the topic.
Midpoint	Provide	when conceptually required (Krosnick and Fabrigar 1997), as it prevents response bias (Menold and Bogner 2016); otherwise, it decreases measurement quality.
	Reverse (omit)	when conceptually not required.

apparently false substantive answers by opinionless respondents. Nevertheless, contradictory findings (DeCastellarnau 2018) indicating that a DK-O influences or does not influence data quality have continued to be made—while some researchers still argue that a DK-O leads to incomplete, less valid and less meaningful data, others find no support for these effects.

In contrast, the reason for choosing a midpoint might be a neutral position (bipolar scale) or a moderate opinion (unipolar scale) (e.g., Krosnick and Fabrigar 1997), but also minimizing effort by satisficing (Krosnick 1991). Therefore, regarding the inclusion of a MC, Krosnick and Fabrigar (1997) point out, "Evidence on validity is mixed, and the theory of satisficing suggests that including midpoints may decrease measurement quality" (p. 148). Menold and Bogner (2016) agree with this, based on several findings from past and current studies. However, they point out that "most researchers recommend that a middle alternative should be offered in order to prevent respondents who have a moderate or neutral opinion from having to use an alternative category, thereby systematically distorting the data" (p. 4). Nonetheless, Streiner, Norman, and Cairney (2015) argue that the decision between an even or odd number of categories (the latter implies a sort of midpoint) within a unipolar scale is of little impact.

Briefly, arguments for omitting or providing a MC or/and a DK-O are inconsistent.

Ultimately, the question remains unanswered as to whether a MC and/or a DK-O should be offered or omitted when measuring attitudes. Figure 1 illustrates plausible recommendations which may guide researchers' decisions.

Both justifications presuppose that the researcher is able to assess the respective target group accordingly, although this is not always the case. This article aims to test an alternative approach. It empirically investigates the response behavior in terms of preferences for a MC or a DK-O when one or both of these options are present; the decision for or against a MC and/or a DK-O may be based on the target groups' reaction to various response scales but, is nevertheless, most probably a compromise.

Figure 2: Typification of answer strategies based on available scale format

Knowledge	Scale format (example)			
	Even (4c)	Uneven (4c+mc)	Even + DK (4c+dk)	Uneven + DK (4c+mc+dk)
Experienced	Regular answer			
Unexperienced	SSA		DK	DK
Moderate	SSA	MC	SSA	MC

Note: 4c= four categories, DK= don't know, MC=middle category, SSA= somewhat sufficient answer

Response Behavior: Methodological Considerations

Cognitive response theory (Tourangeau 1984) forms the starting point for the following considerations. Answering a question takes four steps: (1) comprehension of the question, (2) recall and retrieval of relevant information, (3) judgment and assessment based on the retrieved information, and (4) reporting an answer by linking it to the response category provided.

With reference to the response theory, respondents' knowledge about the topic in question can be assumed to be relevant to their ability to form an opinion and answer a question on attitudes properly. For instance, Lemert (1992) states that "popular knowledge can be so low about so many things that measuring preferences without first measuring knowledge can produce essentially meaningless percentages" (p. 41). Therefore, respondents might face difficulties when they are inexperienced. They may not have enough knowledge (not be able to recall/retrieve relevant information) that is helpful in order to form an opinion. Furthermore, respondents who have a moderate opinion may not be able to take a clearly positive or critical position. They must find an alternative solution when the required response category is not present. Accordingly, O'Muircheartaigh and Helic (2000) suggest that a midpoint should be provided in order to reduce measurement error, since people in this category may make random choices if there is none. In both cases (when respondents lack experience or have moderate opinions), if the ideal answer category (a DK-O or MC) is missing, a somewhat sufficient solution will instead be found—not least because, as experience shows, respondents usually try to fit into the given answer categories and avoid item nonresponse. In this context, "somewhat sufficient" means the occurrence of certain response sets as a consequence of the scale, for example, positive answers instead of the "right ones" when a DK-O or MC are missing.

Figure 2 shows respondents' different response strategies—derived from previously mentioned preliminary considerations —which vary according to the given scale format.

At first glance, an uneven scale including a DK-O promises to be a suitable version on most occasions. However, the impact of most respondents' knowledge on the topic is also assumed to be relevant for the outcome.

For instance, if the respondents generally have experience in the topic, the presence of a DK-O makes less sense. Conversely, there might be a higher risk of response distortion, which was previously mentioned, if respondents have little or no experience of the survey topic, but there is no DK-O. In that case (see Krosnick and Fabrigar 1997), systematic response patterns that change the frequency distribution of the result are more likely to occur than random and thus equally distributed responses. A tendency towards positive (agreeing) answers would be a plausible reaction. If this is true, the following Hypothesis One is supported:

> H1: The absence of a DK-O increases the share of positive response categories.

Furthermore, if respondents are able to clearly agree or disagree with specific statements, the presence of a MC might be pointless. Conversely, respondents with a moderate attitude might have trouble coping with the task of taking a side. In that case, as previously argued, systematic response patterns that change the frequency distribution of the result are more likely than random and equally distributed responses. If this is true, the following Hypothesis Two is supported:

> H2: The absence of a MC increases the share of positive response categories.

The verification or refutation of the previously mentioned hypotheses might serve as guiding factors for the decision for or against a response scale, as shown later.

Data and Methods

Data

Data were collected by means of an online survey (optimized for smartphones) dealing with student participation at a university in Austria. It can be reasonably assumed that this survey/topic is appropriate for the intended analyses, since the students' engagement and therefore insights into university policies differ. For example, in 2017, and the years before that[1], less than 25% of the students took part in the election for the students' union in Austria. Therefore, it is safe to assume that the sample consists of students with more experience (voters) and others with less experience (nonvoters).

In total, 17,491 students were contacted via email. Almost every student was contactable by email. Of them, 2,483 (14.2%) responded to the invitation, and 1,923 (11%) answered at least one question. In total, 7.3% (n=1,282) of the invited students finished the survey by reaching the final page. Previous

1 https://www.oeh.univie.ac.at/wahlkommission

Table 1: Group affiliation based on the availability of a MC and/or a DK-O (n=1,282)

		DK-O	
		Yes	No
MC	Yes	4c+mc+dk (22.9%)	4c+mc (26.4%)
	No	4c+dk (24.9%)	4c (25.8%)

Note: 4c= 4 categories, mc= middle category, dk= "don't know" option

surveys have shown that this is a typical response rate for the target group. Nonetheless, the students who dropped out are excluded in order to avoid biased results based on item nonresponses caused by the time at which they dropped out. Consequently, the following analyses are based on 1,282 cases.

Experimental Design

In order to explore the students' preferences depending on the availability of a MC and/or DK-O, an experimental design was used based on a unipolar scale. This is because a unipolar scale has the advantage of higher reliability (Alwin 2007).

Nonetheless, the researcher has to decide whether a moderate position (middle category) and/or a "don't know" option is needed. Therefore, all invited students were randomly assigned to one of four conditions. The answer format of the rating scales provided to each group, concerning three sets of items, differed due to the absence of a DK-O and/or a MC. In order not to overtax the respondents, a moderate number of response categories was chosen: four categories excluding a "don't know" option (4c/4c+mc) and five including it (4c+dk/4c+mc+dk). This restriction was selected because Alwin (2007) reports "the superiority of four- and five-category scales for unipolar concepts" (p. 196).

Table 1 shows that the survey population (after data cleansing) was divided into four comparably sized groups of about 23% to 26%. Accordingly, the dropout rates did not statistically significantly differ (chi^2=4.412; p=0.220) based on group assignment.

Extracts of the questions and the given answer categories are included in the appendix. At this point, it should be mentioned that only the endpoints were clearly labelled as "very satisfied" and "not satisfied at all." To avoid "face-saving don't knowers," the MC was tagged with a face emoji (see figure A-1 in the appendix) in the case of two small sets of items and not tagged at all in the case of one large set of items.

Results

Descriptive results based on the group affiliation previously mentioned (Table 1) are shown in Figure 3. More precisely, the average proportion of extreme responses (ERs; e.g., "very satisfied" and "not satisfied at all"), of chosen middle categories (MCs), of mild response styles (MLRSs, response categories

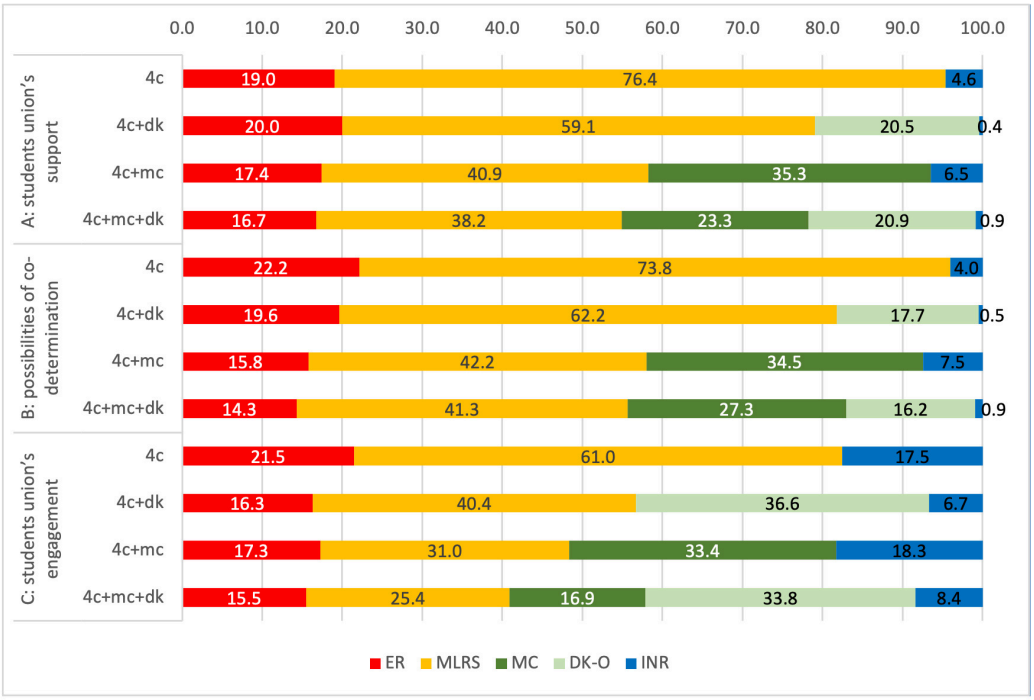

Figure 3: Frequency distribution, summarized per set of items (A, B, C), differentiated by group affiliation (in percent)

Note: The first set (A) includes four items, the second (B) includes three, and the third (C) includes 13—per scale, one (A, B) or three (C) missing data items per respondent were tolerated to form the mean value index.

Note: n: 4c=331, 4c+dk=319, 4c+mc=339, 4c+mc+dk=293;

Note: Only the endpoints of the answer categories were labeled "very satisfied" and "not satisfied at all." The answer categories of the first two sets of items were tagged with face emojis instead of points (see figures in the appendix).

Note: ER=extreme response (e.g. "very satisfied" and "not satisfied at all"), MC=middle category, MLRS=mild response style (=response categories next to the extreme response categories), DK-O="don't know" option; INR=item nonresponse.

next to the ER categories), of chosen "don't know" options (DK-Os), and of item nonresponses (INRs), are presented per set of items. It is clear that offering a MC in addition to a DK-O (4c+mc+dk vs. 4c+dk) did not essentially affect the proportion of participants that chose a DK-O, showing a decrease of about one to three percentage points. Conversely, offering a DK-O in addition to a MC (4c+mc+dk vs. 4c+mc) decreased the proportion of participants that chose a MC by about 7 to 17 percentage points, and reduced the proportion of participants that chose a INR by about six to ten percentage points. Moreover, the proportion of participants that chose a MLRS decreased more and more if a DK-O, a MC or a combination of the two was available. This was similarly true for ERs, but in a limited way.

The calculation of the average number of MCs and DK-Os chosen (see Table 2) for all three sets of items confirms the outcome presented previously. The risk or chance of choosing a MC was almost twice as high if no DK-O was available compared to one being available. On average, seven (4c+mc) or four (4c+mc+dk) out of 20 possible options were chosen. The difference is statistically significant. On the other hand, the level of risk or chance of choosing a DK-O did not depend on whether a MC was available. On average, about six MCs out of 20 possibilities were selected.

Table 2: Average number of MCs and DK-Os chosen considering the scale format (n=1,282)

	MC		DK-O	
Group	4c+mc	4c+mc+dk	4c+dk	4c+mc+dk
Mean (0–20)	7.09	3.95	6.42	6.11
T-Test	t=9.408 (p<.000)		t=.649 (p=.517)	

Note: t=t-value of the independent samples test, p=significance

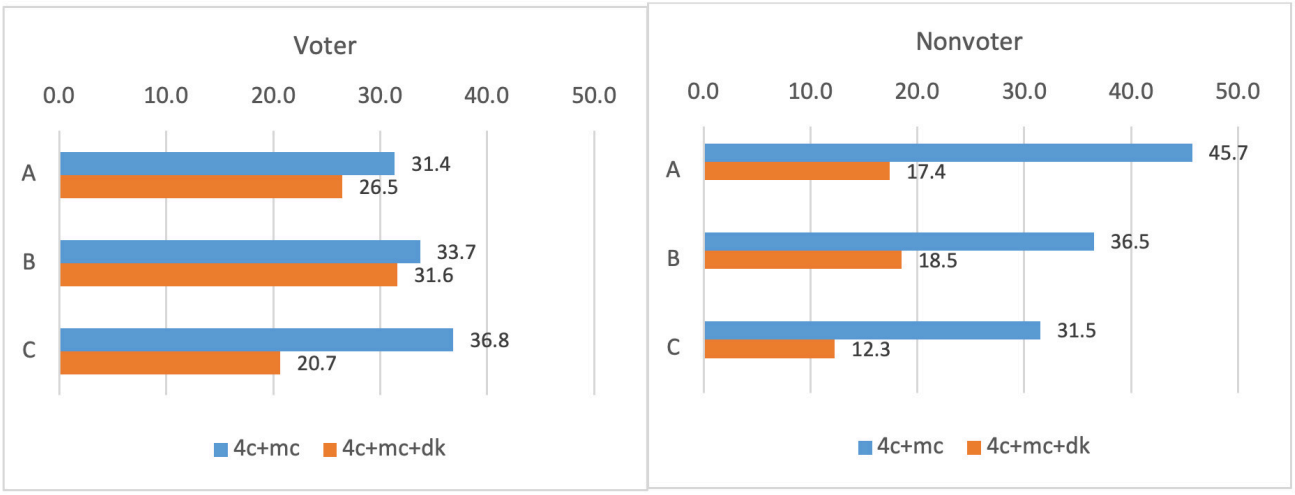

Figure 4: Proportion of MCs chosen, per set of items (A, B, C), depending on the availability of a DK-O, differentiated by voters vs. nonvoters (in percent)

Note: Differences between the proportion of participants choosing a MC was statistically significant with regard to the item set "C" of voters and all three item sets of nonvoters (p<.05)

Figure 4 illustrates that choosing a MC depended on the respondents' experience with the subject of the survey, "student participation", which was measured by participation in the elections of the student representatives (voter vs. nonvoter). Although the proportion of voters and nonvoters that chose a MC was reduced when a DK-O was available, the extent to which this occurred varied. The difference was much smaller among voters (with just one statistically significant result) than among nonvoters (with three statistically significant results). In contrast, the proportion of voters and nonvoters that chose a DK-O was independent of whether a MC was present—only small, statistically insignificant differences were observed (see Figure 5).

But what does this mean for the outcome of the survey? Table 3 should give an answer to this question, as it presents the frequency distribution of the three sets of items (A, B, C). In order to detect differences in the distribution of the responses, the "don't know" and/or "moderate" responses were treated as missing.

Based on item sets A and B, the presence of a DK-O (comparing 4c vs. 4c+dk and 4c+mc vs. 4c+mc+dk) did not statistically significantly change the respective share of responses to other categories, when the DK-O responses

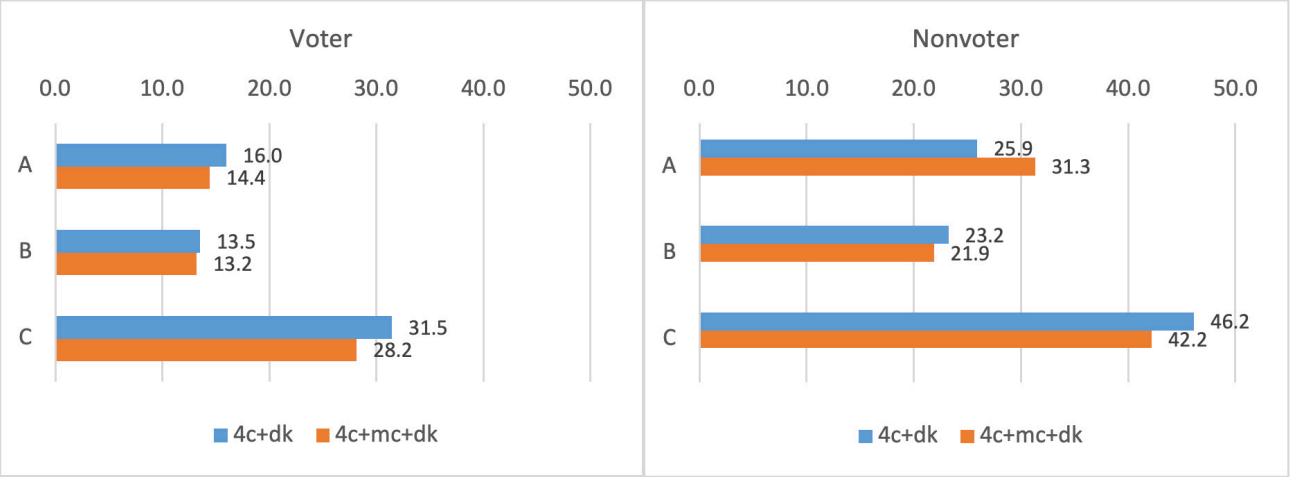

Figure 5: Proportion of DK-Os chosen, per set of items (A, B, C), depending on the availability of a MC, differentiated by voters vs. nonvoters (in percent)

Note: Differences based on voters and nonvoters were statistically insignificant (p>.05)

Table 3: Frequency distribution, per set of items (A, B, C) (excluding "don't know" and "moderate" responses)

Set of items	Answer scale	Very satisfied			Not satisfied at all	n
A: student union support	4c	13.4	55.8	**24.3**	6.4	317
	4c+dk	17.0	50.5	**24.1**	8.4	250
	4c+mc	**26.1**	53.5	13.0	7.3	177
	4c+mc+dk	**25.0**	52.8	11.9	10.3	137
Tukey's HSD, p-level: .05		i vs. iii i vs. iv ii vs. iii ii vs. iv	-	i vs. iii i vs. iv ii vs. iii ii vs. iv	-	-
B: possibilities of co-determination	4c	10.4	43.6	33.6	12.4	321
	4c+dk	11.0	38.7	**36.7**	13.7	263
	4c+mc	15.3	45.1	27.8	11.8	206
	4c+mc+dk	12.4	43.4	29.9	14.3	166
Tukey's HSD, p-level: .05		-	-	ii vs. iii	-	-
C: student union engagement	4c	17.4	**51.1**	**23.1**	8.4	264
	4c+dk	18.5	40.6	**30.8**	10.1	133
	4c+mc	**28.6**	38.9	17.6	**15.0**	88
	4c+mc+dk	**39.0**	40.2	9.0	11.8	53
Tukey's HSD, p-level: .05		i vs. iii i vs. iv ii vs. iii ii vs. iv	i vs. ii i vs. iii	i vs. ii i vs. iv ii vs. iii ii vs. iv	i vs. iii	-

Note: The first set (A) includes four items, the second (B) includes three, and the third (C) includes 13—per scale, one (A, B) or three (C) missing data items were tolerated to form the mean value index

Note. i=4c, ii=4c+dk, iii=4c+mc, iv=4c+mc+dk, p-level=significance level

Note. HSD=Honestly Significant Difference; Values that differ significantly are highlighted in gray - the higher value is highlighted in bold

(as well as the MC responses) were set as missing. In these cases, a tendency for positive responses could not be detected. This means the students might have enough knowledge to form an opinion regarding the question if they are forced to. However, item set C reveals two statistically significant differences (C: i vs. ii). The respective questions probably require greater insight. In this case, the slightly positive answers "somewhat satisfied" decreased while the slightly negative answers "somewhat dissatisfied" increased. However, H1 is refuted. The presence of a MC (comparing 4c vs. 4c+mc and 4c+dk and 4c+mc+dk) did change the share of several answer categories if the MC (as well as the DK-O) responses were set as missing. For two item sets (sets A and C), this procedure mainly increased the proportion of "very satisfied" respondents and mainly reduced the proportion of "somewhat dissatisfied" respondents, but also statistically significantly (but only partly) reduced the proportion of "not at all satisfied" respondents. Item set B was only marginally affected. This means that the students showed systematic response patterns (negative responses were reduced) depending on the availability of a MC. H2 is verified.

However, excluding respondents who showed systematic responses might not be an option—as this would generally distort the data collected. Accordingly, alternative ways are discussed below.

Conclusions and Recommendations

Scale-sensitive response behavior was empirically tested in terms of preferences for a MC or DK-O when answering questions of attitude. The results suggest that respondents who choose a DK-O most probably do not have a moderate opinion: the presence of a MC in addition to a DK-O does not change the proportion of such respondents. However, they might have enough knowledge to form a positive or critical opinion regarding the question if they are forced to, since the respective share of other response categories does not statistically significantly change when a DK-O is present compared with when it is not—provided that those who choose the MC are not included in the analysis.

Conversely, respondents choosing a MC do not necessarily have a moderate view: the presence of a DK-O in addition to a MC does reduce the proportion of the latter. Moreover, the absence of a MC does change the share of other regular answer categories—then respondents would opt for a different regular answer category. Therefore, a systematic response pattern occurs, depending on the availability of a MC.

Based on this outcome, two consecutive steps for future projects can be advised. (1) Regarding the questionnaire design: Questions that trigger a lack of knowledge should be avoided — this is a known basic rule. Nonetheless, a DK-O should not be omitted in general, presuming that the proportion of inexperienced respondents is relatively high. Including a filter question before

a question block dealing with a topic that requires previous experience could help avoid overtaxing these respondents. In contrast, offering or omitting a MC is a more difficult decision. If the group of respondents with a moderate opinion is very small, a MC can be used to identify respondents with a systematic response behavior. Either way, pretests should be conducted to determine the respondents' reasons for choosing a MC. (2) Regarding the main survey: The sample composition should be determined as regard the participants' knowledge on the topic (experienced, inexperienced, and moderate). The proportions of different types of survey participants should be specified and considered in the data analyses.

Finally, these results are an important step for the further investigation of scale-based responses. Subsequently, further investigation is required as to what extent random choices vs. systematic choices are made depending on the scale format and the respondents' knowledge on the topic.

Limitations have to be considered. Data were collected by means of an online survey dealing with student participation at a university in Austria. Therefore, the results are not generalizable but form a starting point for further investigations. However, the results might be different, when another number of answer categories (e.g. six instead of four) is chosen. Moreover, the concept of the measuring scales seems to be optimizable. This will be followed up in a further research project.

Corresponding Author:

Daniela Wetzelhütter, University of Applied Sciences Upper Austria, Garnisonstraße 21, 4020 Linz, Austria

E-mail: daniela.wetzelhuetter@fh-linz.at, Phone number: 0043 50804-52430

Submitted: June 22, 2020 EDT, Accepted: September 02, 2020 EDT

REFERENCES

Alwin, Duane F. 2007. *Margins of Error*. Hoboken, NJ: John Wiley & Sons, Inc. https://doi.org/10.1002/9780470146316.

Alwin, Duane F., and Jon A. Krosnick. 1991. "The Reliability of Survey Attitude Measurement." *Sociological Methods & Research* 20 (1): 139–81. https://doi.org/10.1177/0049124191020001005.

Andrews, Frank M. 1984. "Construct Validity and Error Components of Survey Measures: A Structural Modeling Approach." *Public Opinion Quarterly* 48 (2): 409. https://doi.org/10.1086/268840.

Baka, Aphrodite, Lia Figgou, and Vasiliki Triga. 2012. "'Neither Agree, nor Disagree': A Critical Analysis of the Middle Answer Category in Voting Advice Applications." *International Journal of Electronic Governance* 5 (3/4): 244. https://doi.org/10.1504/ijeg.2012.051306.

DeCastellarnau, Anna. 2018. "A Classification of Response Scale Characteristics That Affect Data Quality: A Literature Review." *Quality & Quantity* 52 (4): 1523–59. https://doi.org/10.1007/s11135-017-0533-4.

Garland, Ron. 1991. "The Mid-Point on a Rating Scale: Is It Desirable? Research Note 3." *Marketing Bulletin* 2):66–70. http://marketing-bulletin.massey.ac.nz/V2/MB_V2_N3_Garland.pdf.

Hurley, John R. 1998. "Timidity as a Response Style to Psychological Questionnaires." *The Journal of Psychology* 132 (2): 201–10. https://doi.org/10.1080/00223989809599159.

Krosnick, Jon A. 1991. "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys." *Applied Cognitive Psychology* 5 (3): 213–36. https://doi.org/10.1002/acp.2350050305.

Krosnick, Jon A., and Leandre R. Fabrigar. 1997. "Designing Rating Scales for Effective Measurement in Surveys." In *Survey Measurement and Process Quality*, edited by L. Lyberg, P. Biemer, M. Collins, E. De Leeuw, C. Dippo, N. Schwarz, and D. Trewin, 6:141–64. Hoboken, New York: John Wiley & Sons, Inc.

Leeuw, Edith D. de, Joop J. Hox, and Anja Boevé. 2016. "Handling Do-Not-Know Answers." *Social Science Computer Review* 34 (1): 116–32. https://doi.org/10.1177/0894439315573744.

Lemert, James B. 1992. "Effective Public Opinion." In *Public Opinion, the Press, and Public Policy*, edited by J. D. Kennamer, 41–61. Westport, CT: Praeger Publisher.

Menold, Natalja, and Kathrin Bogner. 2016. "Design of Rating Scales in Questionnaires." *Mannheim, Germany: GESIS – Leibniz Institute for the Social Sciences*. https://doi.org/10.15465/GESIS-SG_EN_015.

Moors, Guy. 2008. "Exploring the Effect of a Middle Response Category on Response Style in Attitude Measurement." *Quality & Quantity* 42 (6): 779–94. https://doi.org/10.1007/s11135-006-9067-x.

O'Muircheartaigh, Jon A.Krosnick, and Armen Helic. 2000. "Middle Alternatives, Acquiescence, and the Quality of Questionnaire Data." Working Paper 0103.

Streiner, David L., Geoffrey R. Norman, and John Cairney. 2015. *Health Measurement Scales. A Practical Guide to Their Development and Use*. 5th ed. Oxford: Oxford University Press. https://doi.org/10.1093/med/9780199685219.001.0001.

Sturgis, Patrick, Caroline Roberts, and Patten Smith. 2014. "Middle Alternatives Revisited." *Sociological Methods & Research* 43 (1): 15–38. https://doi.org/10.1177/0049124112452527.

Tourangeau, R. 1984. "Cognitive Science and Survey Methods." In *Cognitive Aspects of Survey Methodology: Building a Bridge Between Disciplines*, edited by T. Jabine, M. StTraf, J. Tanur, and R. Tourangeau, 73–100. Washington, DC: National Academy Press.

Appendix

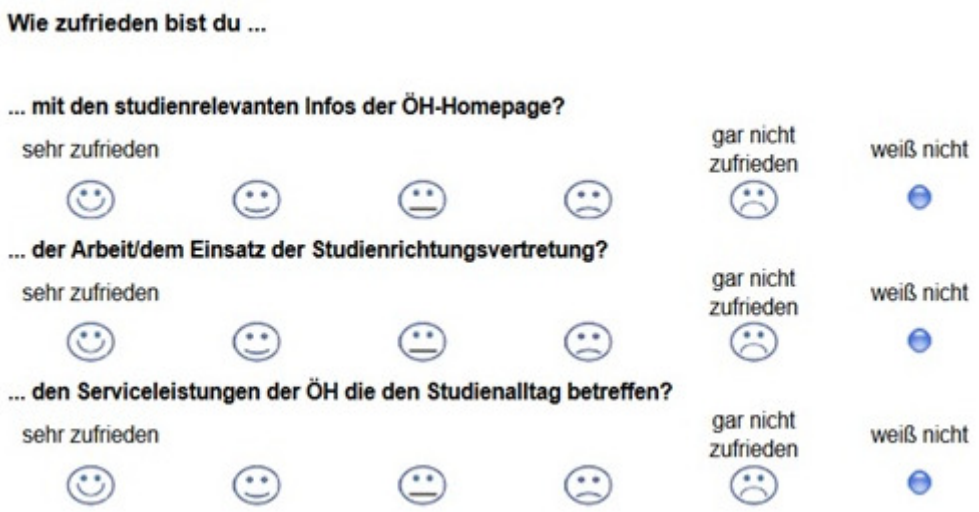

Figure A-1 Answer categories tagged with face emojis

Figure A-2: Answer categories tagged with radio buttons

Note: Translated, the question reads as follows: Figure A-1: *"How satisfied are you ...", "... with the study related information on the student unions (ÖH) homepage?", "...the work/the engagement of student representation of the study program?", "...the services of the student union (ÖH) concerning the everyday life of the students?"*; Figure A-2: *"How satisfied are you with how the student union (ÖH) has stood up for the following topics so far?"*. The listed topics are *"Remission of tuition fees," "Definition of admission criteria for studying," "Library: longer opening hours,"* and *"Library: book orders."* The end points of the regular answer categories of all three sets of items were labeled *"very satisfied"* and *"not satisfied at all."* The label of the sixth option was *"don't know."* The answer categories were tagged either with face emojis or with buttons. The DK-O was made noticeable as an additional, irregular category in order to ensure that the conceptual midpoint was visible, independently of the presence of a DK-O.

Extracts of questions and given answer categories