

EFFECTS OF A NEUTRAL ANSWER CHOICE ON THE RELIABILITY AND VALIDITY OF ATTITUDE AND OPINION ITEMS

Clément Dassa
Jean Lambert
Département de médecine sociale et préventive
Université de Montréal

Régis Blais
Département d'administration de la santé
Université de Montréal

Diane Potvin
Natalie Gauthier
Département de médecine sociale et préventive
Université de Montréal

Abstract: The objective of this study is to determine whether a middle alternative in the response choices to a questionnaire influences the reliability and validity of survey responses. A 32-item questionnaire regarding maternity care was administered to a sample of 1,390 persons consisting of 597 physicians, 723 nurses, and 70 midwives. The sample was randomly divided into two groups, one receiving a questionnaire with five answer choices, including a middle option, and the other receiving a similar questionnaire but without the middle option. Results showed that including a neutral option had little effect on overall reliability and validity; however, it did allow better psychometric coherence when the items were considered globally. The effect of the middle option also depends on the opinions the surveyed persons hold about the study issue.

Résumé: L'objectif de cette étude était de déterminer si un point milieu dans les choix de réponse à un questionnaire influence la fiabilité et la validité des réponses à une enquête. Un questionnaire de 32 items concernant la périnatalité a été administré à un échantillon de 1 390 personnes composé de 597 médecins, 723 infirmières et 70 sages-femmes du Québec. L'échantillon a été réparti au hasard en deux groupes, l'un recevant le questionnaire à cinq choix de réponse, incluant un point milieu, et l'autre recevant le même questionnaire sans point milieu. Les résultats ont montré qu'en général la présence d'un point milieu avait

peu d'effet sur la fidélité et la validité. Cependant, elle permet une meilleure cohérence psychométrique quand les items sont analysés ensemble. L'effet d'un point milieu dépend aussi de l'opinion des personnes interrogées à l'égard de l'objet d'étude.

Since the early 1940s, surveys have increasingly become one of the main sources of data in the social sciences and program evaluation (Mason, 1996). Many studies have been conducted to determine whether survey responses were influenced by the wording of questions and the questionnaire format. A number of factors — question sequence, length, syntax, and presence of given answer choices rather than open-ended questions, among others — may influence the results and thereby the conclusions drawn from them. Excluding (omitted form) or including (offered form) a middle option in a multiple-choice question, indicating a neutral opinion, has received special attention. Some authors favor the inclusion of such an option, as its absence, they argue, demands that a respondent with a neutral opinion give a different answer, thus biasing results. Others, however, recommend that a middle option not be included in order to force the respondent to take sides. In this debate, Payne (1951) suggested taking into account the purpose of the question. If the purpose is to learn the leanings (predispositions, propensities) of the respondent, it is preferable to omit the middle option; if it is to assess the convictions (certainties, firm opinions) of the respondent, it is better to offer a middle answer option. It is important to note that the expression middle “option” or “choice” indicates a neutral opinion, which is different from a “no opinion” or “don’t know” answer.

Over the past 20 years, many studies have focused on the middle option. It bears noting that all these studies allowed the respondent to choose the middle option even when it was not explicitly offered (omitted form).

In a study on the validity and components of measurement error, Andrews (1984) examined three indicators of the quality of a questionnaire item: construct validity, reliability, and variance due to method. His results showed that omitting or including a middle option had no significant effect on any of these three quality indicators. Nor did the mode of data collection (phone, face-to-face, or written questionnaire) affect these measures. However, the number of answer choices and offering of a “don’t know” option were found to be two important characteristics in influencing indicators. Alwin and Krosnick (1991) used a similar approach, estimating an indica-

tor of item reliability. They attempted to explain the variation in their indicator using different variables, including the presence or absence of a middle option. These authors indicated that items with three options (where a middle alternative is offered) were more reliable than those with two or four options (without a middle alternative). However, the reliability of items using a five-point scale (offered form) was found to be similar to that of items using four points.

One issue often examined is what the respondents choosing the middle option would have done had this option not been offered to them. The hypothesis generally tested is that this category of respondents behaves like other respondents. In other words, the patterns of answers in the omitted form and the offered form, without the middle answers, should not differ significantly. Several authors (Ayidiya & McClendon, 1990; Kalton, Collins, & Brook, 1978; Presser & Schuman, 1980; Schuman & Presser, 1977, 1981) have obtained results confirming this hypothesis. On the other hand, Kalton, Roberts, and Holt (1980) found that the increase in the number of respondents choosing the middle option, when offered, came disproportionately from other answer choices. They failed to observe a general tendency in the way the distributions varied. A more recent study could neither confirm nor disprove this hypothesis (Bishop, 1987).

Another important question studied is the relationship between respondents' characteristics and their choosing the middle option, and whether this relationship varies with the form used (offered or omitted form). Presser and Schuman (1980) and Schuman and Presser (1981) observed that the association between, on the one hand, education, gender, age, and race and, on the other hand, the answer to a particular question was independent of the form used. However, for many items a relationship was found between these variables and the choice of the middle option; but the relationship varied with the item.

It has also been suggested that the strength of the respondent's opinion on a given item may be related to the effects of including or omitting a middle option. Studies have shown that less intensely opinionated respondents tended to be attracted to the middle option, whether this option is offered or not (Bishop, Hippler, Schwarz, & Strack, 1988; Krosnick & Schuman, 1988; Presser & Schuman, 1980; Schuman & Presser, 1981). It should also be noted that the position of the item in the questionnaire and how the respondent is asked about the intensity of his/her opinion may influence the relation between intensity of opinion and the form used.

Studies on the significance of the middle option showed that, for some items, the selection of this option varied according to whether the wording used referred to “ambivalence,” “neutrality,” or “uncertainty” (Goldberg, 1981; Klopfer & Madden, 1980). However, using a five-point Likert scale, Armstrong (1987) showed that including a middle option indicating either “neutrality” or “uncertainty” made little difference for the 11 items studied. Armstrong recommends, from a semantic point of view, the utilization of a middle option indicating “neutrality,” which is less ambiguous.

A number of issues raised in this literature are relevant to the present study. First, although the effect of including or omitting a middle option on the reliability of answers to various items has been studied (Alwin & Krosnick, 1991; Andrews, 1984), the relative contribution of the middle option to the measure of reliability has yet to be assessed. Second, regarding the finding that construct validity does not seem to be affected by the presence or absence of a middle alternative (Andrews, 1984), it should be noted that this result is based on a study that simultaneously examined subjective items (attitudes or opinions) and items pertaining to objective facts. It is thus difficult to draw specific conclusions regarding the validity of attitude or opinion questions. Thus, analysis of construct validity would be needed for this particular type of item to assess the effect of the middle option. Third, although many authors have tested whether the inclusion of a middle option influences the distribution of respondents for other answer choices, the results were mixed. Finally, little is known about the relationship between the respondent’s characteristics and the middle option. It would be especially interesting to document the overall profile of respondents who choose this option in a questionnaire measuring attitudes or opinions.

The present study attempts to fill these gaps by measuring the effect of a middle option in a questionnaire assessing the opinions of physicians, nurses, and midwives on various maternity-care issues and the role of midwives. The specific objectives were:

1. to estimate the contribution of the middle option to the Kappa statistic in a test-retest reliability analysis;
2. using log-linear models, to assess the influence of the type of scale (with or without the middle option) on the distribution of respondents in the answer choices for each separate item, and to complete this evaluation with dual analysis for all items where a homogeneity test is significant, in order to take into account the interrelationships among items;

3. to assess the effect of the middle option on the overall reliability and the factorial structure of the scales;
4. to assess the differential effect of the middle option on the factorial structure of the scales for each of the groups studied; and
5. to describe the effect of the middle option on the identification of the variables characterizing the respondents and contributing to discriminant validity.

METHOD

In spring 1991 we conducted a survey by mail of three groups of maternity-care professionals in Quebec: physicians practicing obstetrics (family physicians and obstetrician-gynecologists); maternity-care nurses (working in hospitals or community health centers); and midwives who were members of the two provincial midwife associations (Blais, Lambert, Maheux, Loisel, Gauthier, & Framarin, 1994; Blais, Maheux, Lambert, Loisel, Gauthier, & Framarin, 1994). The survey was done to assess opinions about selected maternity-care issues and the introduction of midwives into the Quebec health-care system. Physicians were identified through a mail and telephone survey of all Quebec hospitals where deliveries were performed. Nurses (female only) were identified through the *Ordre des infirmières et des infirmiers du Québec* (the Quebec college of nurses). Midwives were identified through their respective associations. Systematic random samples were drawn for male family physicians (80%) and nurses (50%). Given their relatively small numbers, all midwives, obstetrician-gynecologists, and female family physicians were included in the study. The final sample included 1,744 persons, comprised of 844 physicians, 808 nurses, and 92 midwives.

The self-administered questionnaire had four sections. The first two sections consisted of 16 items each that assessed the respondents' attitudes and opinions on various maternity-care issues and the need for midwives, respectively. The remaining two sections collected data on the forms of midwifery practice preferred by respondents, and on their demographic and professional background.

As a preliminary study, a test-retest analysis was conducted on the first two sections after a one-month interval with 78 persons from the three study groups. The questionnaire used included 31 items to be responded to using a five-point scale, including a middle option (1 = strongly disagree, 2 = somewhat disagree, 3 = neither agree nor

disagree, 4 = somewhat agree, 5 = strongly agree). One of the items on the final questionnaire was not included in this preliminary study.

The main analysis of this study, aimed at assessing the effect of the middle option, was performed using the 32 items included in the first two sections of the questionnaire (see Appendix). The respondents of each group were randomly divided into two subgroups. The first received a questionnaire with five-point scales, including a middle option representing a neutral opinion ("offered form"). The second received a similar questionnaire with four-point scales, that is, sans a middle option ("omitted form"). The wordings of the four answer choices other than the middle option were identical in the two forms. Unlike in most earlier studies, the middle option was not accepted when it was not explicitly offered.

Statistical analyses performed were as follows. (1) A reliability analysis was conducted using the 78 respondents from the preliminary study. The Kappa statistic was computed, then broken down for each answer choice on the scale in order to determine the relative contribution of the middle option to this statistic. (2) A univariate analysis of the answer patterns was performed for each item of the four- and five-point scales using log-linear models to study the distribution of respondents according to the presence or absence of the middle option. (3) An analysis of construct validity for each of the first two sections of the questionnaire was performed using factor analysis separately for the four- and five-point-scale respondents to check whether the covariance structure among items was similar. The internal consistency of the constructs was also examined. (4) A series of multidimensional analyses (factorial and correspondence) was conducted on the groups (physicians, nurses, midwives) to assess the differential effect of the two types of scales for each profession. (5) Finally, using multiple linear regression, the influence of potential explanatory factors, such as age and gender (controlling for each profession) on the choice of the middle option was studied for the first two questionnaire sections.

RESULTS

Respondent characteristics

The overall response rate was 80%, for a total of 1,390 respondents: 709 respondents for the omitted form and 681 respondents for the offered form. The response rate varied with the profession: 71% for

physicians ($n = 597$), 89% for nurses ($n = 723$), and 76% for midwives ($n = 70$). All nurses and midwives as well as 24.8% of physicians were females. The subsamples for the physicians, nurses, and midwives had, respectively, the following characteristics: (a) mean age (standard deviation): 42.89 (10.47), 47.86 (6.96), 41.86 (8.07); (b) percentage of respondents who studied in Canada: 89.6%, 99.4%, 41.2%. Most physicians were generalists (66.4%), the rest were specialists in obstetrics and gynecology, whereas nurses were divided between community nurses (41.6%) and hospital nurses (58.4%).

For the offered form, the percentages of respondents choosing the middle option indicate a difference among the three groups. In the maternity care section of the questionnaire, the physicians chose the neutral position more often (18.36%) than the nurses (13.99%) or the midwives (6.86%). The order of choice changed for the midwife section of the questionnaire: 22.62% for nurses, 18.36% for physicians, and 8.52% for midwives. Moreover, the individuals who more often chose the middle option in the section on maternity care also tended to express a neutral opinion in the section on midwives (the Pearson correlation coefficient between the number of times respondents chose the middle option in the maternity care section and the midwife section was 0.44). However, the midwife section had a larger mean number of neutral answers than did the maternity care section ($p = 0.0001$ for the t -test).

Test-retest reliability

The Kappa statistic and the relative contribution of each answer choice to this value, calculated in the preliminary study, are provided in Table 1. The middle option contributed least to the Kappa statistic with means of 7.80% and 13.61% for sections A (maternity care) and B (midwives), respectively. The number of times the middle option ranked most influential to least influential on the Kappa statistic was computed and can be summarized as follows. In 14 of 31 cases, the middle option contributed least to the Kappa statistic. Moreover, its relative contribution was found to rank in the last and second-last positions 25 times out of 31; this result coincides with the weighted mean of the middle option, which is clearly inferior to other means.

These findings suggest that the middle answer choice is the least reliable (i.e., a respondent opting for the middle alternative might easily answer the question differently if it were posed a second time).

Table 1
Kappa statistic and relative contribution of each answer choice to the Kappa statistic in the preliminary study

Section	Kappa statistics	Strongly disagree (%)	Somewhat disagree (%)	Neither agree nor disagree (%)	Somewhat agree (%)	Strongly agree (%)
Section A:	0.45	17.91	20.02	7.80	25.37	28.91
Maternity care	0.15	-0.19	0.00	-2.85	7.25	-0.85
(15 items)	0.63	47.59	35.86	19.60	44.22	59.96
Section B:	0.52	21.49	19.66	13.61	21.81	23.43
Midwives	0.41	5.24	11.36	1.22	6.62	7.96
(16 items)	0.65	38.24	29.95	25.87	36.13	49.93
Total	0.49	19.76	19.83	10.80	23.53	26.08
(31 items)	0.15	-0.19	0.00	-2.85	6.62	-0.85
	0.65	4.759	35.86	25.87	44.22	59.96

Note. The results are presented in the following order in each cell: mean, minimum, and maximum.

This phenomenon did not seem to be as strongly evident for the remaining answer choices. Section B, concerning midwives, appeared to be more reliable (Kappa = 0.52) than section A on maternity care (0.45). The relative contribution of the middle option to the Kappa statistic was also noted to be higher for section B than section A. The practice of midwifery seems to be a subject on which respondents have a more definite opinion than on maternity care.

Modeling of answer patterns

Log-linear models revealed that for eight items (six in section A of the questionnaire and two in section B) the type of scale used significantly ($p < 0.04$) influenced the distribution of respondents in their answer choices (homogeneity test). In other words, the percentage of respondents choosing the middle option when it was offered came disproportionately from other answer choices. For these eight items, the significance of the homogeneity test and differences (percentages) between the observed and expected frequencies under the independence hypothesis are presented in Table 2.

First, we note that the effect of the form (offered or omitted) was more often present in section A, which consisted of general questions on maternity care. Second, when contrast analysis is used, several interesting patterns emerged. For all section A questions except A4, the

Table 2**Results of the log-linear modeling of answer patterns for the items presenting a significant interaction**

Item (Homo- geneity test)	Form	Answer Choices ²				
		1 Strongly disagree	2 Somewhat disagree	3 Neither agree nor disagree	4 Somewhat agree	5 Strongly agree
A3 ($p=0.014$)	Omitted	+0.1	+3.0		-3.6	+0.5
	Offered	-0.1	-3.1	—	+3.7	-0.5
A4 ($p=0.003$)	Omitted	-1.4	-1.1		+4.3	-1.8
	Offered	+1.6	+1.1	—	-4.5	+1.9
A6 ($p=0.002$)	Omitted	+0.4	+2.0		+0.4	-2.9
	Offered	-0.4	-2.1	—	-0.4	+3.0
A11 ($p=0.001$)	Omitted	-0.9	+1.9		-4.0	+3.0
	Offered	+1.0	-2.0	—	+4.1	-3.1
A13 ($p=0.037$)	Omitted	-0.3	+1.6		-1.4	0.0
	Offered	+0.3	-1.7	—	+1.5	-0.1
A15 ($p=0.004$)	Omitted	+0.5	+2.7		-1.2	-2.1
	Offered	-0.5	-2.9	—	+1.3	+2.1
B4 ($p=0.020$)	Omitted	+1.7	-1.8		+2.1	-2.0
	Offered	-1.8	+1.8	—	-2.2	+2.1
B5 ($p=0.007$)	Omitted	-0.6	-1.2		+4.0	-2.2
	Offered	+0.6	+1.2	—	-4.2	+2.3

² Differences (%) between observed and expected frequencies under the independence hypothesis.

absence of a neutral point led respondents to opt more frequently for answer 2 (“somewhat disagree”) than answer 4 (“somewhat agree”) than did respondents offered a neutral point. As for item A4 and the two section B items concerning midwives, the reverse phenomenon was apparent. For items A6 and A15, the tendency extended to the extreme choices: the absence of a neutral point prompted respondents to opt for answer 2 (“somewhat disagree”) more often than choices 4 and 5 (“somewhat agree” and “strongly agree”).

Construct Validity and Internal Consistency

One aspect of construct validity, factorial validity, was assessed in this study. Classical exploratory factor analyses using the method of maximum likelihood estimation were performed separately for the respondents of the offered and omitted forms and for each of sections A (maternity care) and B (midwives). Results pertaining to the factorial structure, communalities, and the proportion of ex-

plained variance are presented in Table 3. Eight items were removed from section A and three from section B because they did not contribute to any identified factor. The items retained from section A allow identification of two factors. The first factor pertains to the *humanization* of maternity care, and explained 48.7% of the total variance of the offered form. The second refers to the tendency not to use obstetrical interventions, or *non-interventionism*, and explained 50.5% of variance. The two factors are correlated (correlations of 0.62 for the offered form and 0.67 for the omitted form were obtained with oblimin rotation). A single factor, representing *openness to midwives*, was found in the reduced section B; this factor accounts for 61.8% of the variance of the offered form and 59.9% of the variance of the omitted form.

The reliability of the scores for both sections, as measured by Cronbach's alpha, was high: the coefficient was around 0.85 for section A and 0.95 for section B. Moreover, omitting the middle option did not affect the overall reliability of the two sections. The internal consistency coefficients (Cronbach's alpha) of the reduced sections (to 8 and 13 items for sections A and B, respectively) were equal or slightly higher than the coefficients of the initial unreduced sections (0.839 and 0.939 for the offered form of section A and B, and 0.850 and 0.938 for the omitted form). As the value of the Cronbach's alpha coefficient tends to decrease when the number of items diminishes, these results suggest that the reduced sections allowed a good measure of the identified constructs.

Further factor analyses were applied for the physician and nurse groups separately (the sample of midwives, $N = 36$, was too small for this analysis). Results showed that the factorial structures previously found for the total population remain essentially unchanged for nurses and physicians: two relatively stable and correlated factors emerged from section A, and a single stable factor emerged from section B. However, with the nurse group, for section A, the scale without the middle option yielded somewhat more variance than did the scale with the middle option: 45.7% and 39.7%, respectively. This difference of 6% is larger than the one observed for the total sample (1.8%). The middle option seems to have reduced the construct validity of the maternity care scale for nurses. The situation was reversed for physicians (35.5% for the offered form and 32.4% for the omitted form). In section B, the effect of the middle option was similar to that in section A. For the physician group, the presence of the middle alternative yielded 8.3% more variance than its absence, thus contributing to improve the factorial solution retained.

Table 3
Factor analyses results and reliability coefficients

Section A (maternity care) <i>N</i> =655 (omitted) <i>N</i> =627 (offered)									
Factors	Items	Loadings		Communalities		Explained variance (%)		Cronbach's alpha	
		offered	omitted	offered	omitted	offered	omitted	offered	omitted
1	A14	0.745	0.701	0.595	0.568				
	A8	0.695	0.682	0.475	0.543				
	A1	0.694	0.787	0.463	0.589			0.804	0.837
	A2	0.614	0.690	0.359	0.431				
	A3	0.585	0.657	0.410	0.453				
2	A12	0.927	0.824	0.730	0.574				
	A10	0.532	0.553	0.505	0.494			0.730	0.723
	A4	0.470	0.536	0.356	0.389				
1 & 2					48.7	50.5	0.839	0.855	
Section B (midwives) <i>N</i> =638 (omitted) <i>N</i> =639 (offered)									
Factor	Items	Loadings		Communalities		Explained variance (%)		Cronbach's alpha	
		offered	omitted	offered	omitted	offered	omitted	offered	omitted
1	B16	0.894	0.896	0.800	0.802				
	B15	0.876	0.853	0.768	0.728				
	B8	0.870	0.838	0.757	0.703				
	B6	0.849	0.847	0.721	0.718				
	B3	0.794	0.779	0.630	0.607				
	B1	0.793	0.762	0.629	0.580				
	B12	0.791	0.789	0.626	0.622	61.8	59.9	0.954	0.951
	B2	0.789	0.754	0.623	0.568				
	B13	0.775	0.751	0.600	0.565				
	B4	0.714	0.727	0.511	0.529				
	B9	0.707	0.712	0.500	0.507				
	B7	0.697	0.694	0.486	0.482				
	B5	0.615	0.618	0.379	0.382				

For the nurse group, the middle option reduced the variance expressed from 56.0% to 52.5%.

Dual scaling analysis

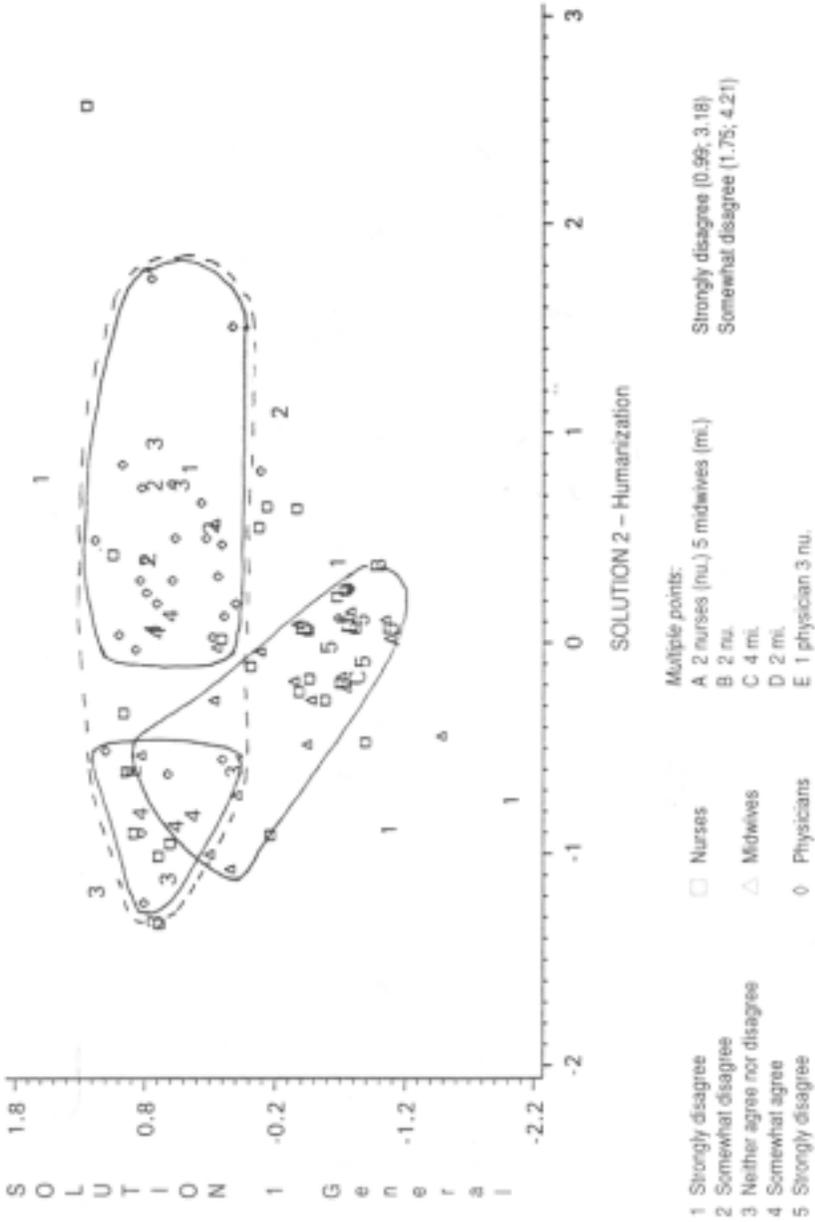
The following analysis considered only the six items of section A that each exhibited heterogeneity in answer patterns by type of scale, that is, with or without the middle option (Table 2). Two series of analyses were used to jointly study, in a multivariate approach, the differentiation effect of the middle option. The first type of analysis – Nishisato (1984) dual analysis for successive categories – involves

the sequence of the four or five options of the scales. The second type of dual analysis relaxes the sequence requirement for all groups, by considering the responses as multiple choice. It is important to note that these two types of analyses are based on mutually orthogonal components forming a dual space allowing simultaneous representation of answer choices and subjects. Moreover, dual analysis is feasible with relatively small samples, which allowed, in the present case, inclusion of the midwives ($N = 36$ valid cases); due to the limitations of the software used, randomly selected subsamples of physicians ($N = 150$) and nurses ($N = 150$) were used to round out the studied sample.

Dual analysis for successive categories was performed for all the subsamples combined to serve as a basis of reference. The results indicate that the items retained their relative positions on the continuum of the main dimension (the first component). Omitting the middle option caused a shift toward the positive side of the continuum; A4 shifted from the middle option to the adjacent option "somewhat agree," and A6 shifted from "somewhat agree" to "strongly agree." The movements of the items on the main continuum from separate analyses for each profession can be summarized as follows. The absence of the middle option prompted physicians to express their disagreement with A3: "Much progress is still to be made in humanizing maternity care" (A3). In contrast, midwives reinforced their agreement regarding the reduction of obstetrical intervention to make maternity care more human (A4). Nurses, on the other hand, failed to change their opinion in the absence of a middle option. These results complete those obtained with log-linear models. One must remember that this latter analysis, which was univariate, did not take into account the interrelationships among items, whereas dual analysis, which is multivariate in nature, did.

Dual analysis for multiple-choice data allows the simultaneous representation of the respondents and the different options of each item in spaces defined by orthogonal dimensions. This method also permits calculation of the optimal Cronbach's alpha coefficient for each dimension. Figure 1 jointly represents physicians, nurses, midwives, and the options of the six items including the middle option. Because dual analysis axes are arbitrarily oriented, it must be noted that in the present case, and for both dimensions, the negative scores indicate agreement with the trait defined by the dimension and positive scores express disagreement. The first dimension (Solution 1) was defined by all items and was hence named "general dimension"; it explained 13.51% of the variance. The second dimension (Solu-

Figure 1
Distribution of subjects and options in two-dimensional space for the five-point scale. (To avoid overloading the figure, a random subsample [90 persons, 30 for each profession] is represented.)



tion 2), defined by items A3, A4, and A6, had a correlation of 0.623 with item A3 and thus was named the “humanization dimension”; it explained 9.27% of the variance.

Two distinct groups of respondents emerged: physicians and midwives. Midwives had opinions clearly differing from those of physicians; nurses, for their part, took a median position. Midwives were homogeneously grouped around all the option 5s (“strongly agree”), whereas physicians exhibited a position that differed with the dimension. On the first dimension, physicians positioned themselves around option 3 (middle) and 4 (“somewhat agree”). On the second dimension, they divided into two subgroups: one revolving around option 2s (“somewhat disagree”) of items A3, A4, and A6, and the other around option 4s (“somewhat agree”) of these same items. In other words, the first subgroup of physicians disagreed somewhat with items stating that much effort remained to be made to humanize maternity care, that reducing obstetrical intervention makes maternity care more human, and that birthing rooms in hospitals should be used more often, and the second subgroup of physicians agreed somewhat with these statements. Overall, in the first dimension, results for option 3s (middle) were much closer to those for physicians than for midwives or nurses.

The two subgroups of physicians identified for the second dimension (“humanization”) were compared for relevant exogenous variables, using either a *t*-test or chi-square, depending on the type of data. Compared with physicians who agreed more on the “humanization” dimension, physicians who disagreed more strongly (first subgroup) were older ($p < 0.021$), included proportionately more men ($p < 0.009$), and practiced alone more often ($p < 0.024$), less in hospitals ($p < 0.001$), and perhaps less in community health centers ($p < 0.08$).

Finally, the same analysis performed on the omitted form of the questionnaire confirmed the general configuration of the relative positions for the three professions, but showed differences when the form offering a middle option was used. The first dimension of the omitted form was less well defined than that of the offered form, and the second dimension was restricted to the aspect “reduction of intervention” of the “humanization” dimension found with the offered form. This conceptual reduction was associated with a decrease in Cronbach’s alpha coefficients of the underlying dimensions (from 0.83 for the offered form to 0.73 for the omitted form for dimension 1, and from 0.66 to 0.53 for dimension 2). Hence, when the three

professions were examined together, the presence of the middle option produced more coherent latent dimensions (i.e., the dual scores defined by the underlying dimensions were more reliable). Finally, no relevant partitioning of the physician group could be detected with the omitted form, suggesting that the questionnaire version with the middle option contributed more to the discriminant validity of the physician's dual scores.

DISCUSSION

Overall, reliability, as measured by Cronbach's alpha, was not influenced by the presence or absence of the middle option. However, examination of the stability of each single item showed that the middle option was the answer choice that contributed least to the Kappa statistic. Thus, excluding the middle option would probably increase item stability. Results pertaining to overall reliability coincide with those obtained by Alwin & Krosnick (1991) for four- and five-point scales and by Andrews (1984). However, these authors did not attempt to assess the isolated effect of the middle option on their reliability indicators for the same scale. It should be noted that the theme of the items may be a factor that affects reliability. In the present case, the "midwife" issue seemed more reliable than the "maternity care" issue, probably because respondents had a more definite opinion on the former.

Results regarding construct validity showed that using a scale with four points (omitted form) versus five points (offered form) did not affect the underlying structure of the questionnaire. These results confirm the more general results obtained by Andrews (1984) on construct validity of questionnaires measuring attitudes, opinions, and facts.

The effect of including or excluding a middle option was examined from two perspectives: univariate, focused on each separate item, and multivariate, taking the whole instrument into consideration. Modeling of the patterns of answers pertained to the first perspective, and multidimensional analysis referred to the second. It must be stressed that these two approaches, while complementary, are based on distinct analysis strategies. The first details the distribution of answers for each item without considering the interrelationships among items, and the second approach aims to summarize the information revealed by the patterns of answers in terms of underlying constructs.

Results of log-linear analysis showed that, for 25% of the items, the patterns of answers were distributed differently according to the form used (omitted or offered). It is difficult to identify a general trend in the way these distributions varied. Such results support those obtained by Kalton, Roberts and Holt (1980). However, a closer look at the eight items in question yields some interesting observations. The items for which respondents moved toward "disagreement" in the absence of the middle option (A3, A6, A11, A13, and A15) were all items that directly criticized current obstetrical practice. With the exception of A2, all the other items of section A were either general principles or positive comments about current obstetrical practice. Results showed that those who had more neutral opinions were physicians and nurses. In the absence of a neutral option, these two groups of professionals, who are in fact those mainly responsible for maternity care at present, will probably tend to protect their interests (i.e., their practice, if it is criticized).

In section A, only item A4 leaned toward "agreement." Contrary to the other five items, this one was not a direct criticism of the health-care system, but rather a principle or a belief. Those physicians and nurses who expressed a neutral opinion on this subject would probably tend to agree more with a laudatory principle than with a criticism of the system for which they are more or less responsible. In section B, items B4 and B5 are two arguments commonly used by individuals and organizations who are against the practice of midwifery. Again, physicians and nurses, who hold more neutral opinions than midwives, would tend to adopt a position "protective" of their interests in the absence of a middle option. As these results show, the effect of the form (omitted or offered) on the distribution of the answer choices depends largely on the nature of the subject studied and the degree of the respondent's implication with this subject (Bishop, 1990; Krosnick & Schuman, 1988; Presser & Schuman, 1980; Schuman & Presser, 1981).

When these items were examined as a set using dual analysis for successive categories, it was found that the absence of the middle option failed to alter the relative order of the items on the dominant dimension. This confirmed the stability of the underlying dimensions of the instrument. Also, the redefinition of the categories of the scale, due to the exclusion of the middle alternative, shifted item A4 from the neutral category toward "agreement," in keeping with the result of the univariate analysis. A more detailed examination of each group identified the midwives as the source of this positive shift (toward "agreement"). Item A3 did not change category but

negatively regressed (toward “disagreement”) compared to other items. In this case, it was the physicians who were responsible for the shift along the scale. Item A6 moved slightly in the positive direction across the boundary of the two positive categories. Therefore, these multidimensional analyses reveal that while the effect of the middle option did not change the underlying constructs, it nonetheless varied according to the group studied.

Dual analysis for multiple-choice data revealed that the presence of a middle option helped define more reliable latent dimensions. Furthermore, only the middle option made possible the definition of measures (dual scores based on latent dimensions) that allowed the identification of two distinct physician subgroups. Using the middle option, it was possible to show that the physicians whose answer profiles were closer to those of midwives were younger, included more women, and practiced more in hospitals.

In summary, this study examined the influence of a neutral choice on the reliability and validity of answers in a questionnaire measuring attitudes and opinions. Results showed that including a neutral option had little effect on overall reliability and validity. However, if the middle option seemed to decrease the stability of the items taken separately, it allowed better psychometric coherence (better reliability and construct validity) when the items were considered globally through latent dimensions and when the make-up of the groups was taken into consideration.

Finally, not offering a middle option forced respondents to make a choice in one direction or the other. For some items, this factor failed to alter the general answer patterns, but patterns did change for still others. In the present case and for certain items, the absence of the middle option revealed the tendency of physicians and nurses to protect current obstetrical practice.

In general, evaluation studies, whether normative or criterion referenced, aim at formulating judgements based on the rigorous assessment of differences among entities (subjects or objects). Therefore, including a middle point in the measurement scale should be considered whenever the results are used to compare different categories of respondents on underlying parsimonious constructs. It seems that, as the psychometric qualities fluctuate with the make-up of the sample, the presence of a middle option appears to improve the measurement of opinions and attitudes among groups whose positions are more moderate. In situations where we suspect polarization of opinions or attitudes, the middle option may be less necessary.

ACKNOWLEDGEMENTS

This study was supported by grants from the Social Sciences and Humanities Research Council of Canada, the National Health Research and Development Program (Health and Welfare Canada), and a National Health Research Scholar Award to Dr. Blais. We thank the physicians, nurses, and midwives who generously contributed their time to this study.

REFERENCES

- Alwin, D.F., & Krosnick, J.A. (1991). The reliability of survey attitude measurement. *Sociological Methods and Research*, 20, 139–181.
- Andrews, F. (1984). Construct validity and error components of survey measures: A structural modeling approach. *Public Opinion Quarterly*, 48, 409–442.
- Armstrong, R.L. (1987). The midpoint on a five-point Likert-type scale. *Perceptual and Motor Skills*, 64, 359–362.
- Ayidiya, S.A., & McClendon, M.J. (1990). Response effects in mail surveys. *Public Opinion Quarterly*, 54, 229–247.
- Bishop, G.F. (1987). Experiments with the middle response alternative in survey questions. *Public Opinion Quarterly*, 51, 220–232.
- Bishop, G.F. (1990). Issue involvement and response effects in public opinion surveys. *Public Opinion Quarterly*, 54, 209–218.
- Bishop, G.F., Hippler, H.J., Schwarz, N., & Strack, F. (1988). A comparison of response effects in self administered and telephone surveys. In R. Groves (Ed.), *Telephone survey methodology* (pp. 321–34). New York: Wiley.
- Blais, R., Lambert, J., Maheux, B., Loiselle, J., Gauthier, N., & Framarin, A. (1994). Controversies in maternity care: Where do physicians, nurses and midwives stand? *BIRTH*, 21, 63–70.
- Blais, R., Maheux, B., Lambert, J., Loiselle, J., Gauthier, N., & Framarin, A. (1994). Midwifery defined by physicians, nurses and midwives: The birth of a consensus? *Canadian Medical Association Journal*, 150, 691–697.
- Goldberg, L.R. (1981). Unconfounding situational attributions from uncertain, neutral, and ambiguous ones: A psychometric analysis of

- descriptions of oneself and various types of others. *Journal of Personality and Social Psychology*, 41, 517–552.
- Kalton, G., Collins, M., & Brook, L. (1978). Experiments in wording opinion questions. *Applied Statistics*, 27, 149–161.
- Kalton, G., Roberts, J., & Holt, D. (1980). The effects of offering a middle response option with opinion questions. *The Statistician*, 29, 65–78.
- Klopfers, F.J., & Madden, T.M. (1980). The middlemost choice on attitude items: Ambivalence, neutrality, or uncertainty? *Personality and Social Psychology Bulletin*, 6(7), 97–101.
- Krosnick, J.A., & Schuman, H. (1988). Attitude intensity, importance, and certainty and susceptibility to response effects. *Journal of Personality and Social Psychology*, 54, 940–952.
- Mason, G. (1996). Recent advances in questionnaire design for program evaluation. *Canadian Journal of Program Evaluation*, 11(1), 73–94.
- Nishisato, S. (1984). *Dual scaling*. Toronto: MicroStats.
- Payne, S.L. (1951). *The art of asking questions*. Princeton, NJ: Princeton University Press.
- Presser, S., & Schuman, H. (1980). The measurement of a middle position in attitude surveys. *Public Opinion Quarterly*, 44, 70–85.
- Schuman, H., & Presser, S. (1977). Question wording as an independent variable in survey analysis. *Sociological Methods and Research*, 6, 151–170.
- Schuman, H., & Presser, S. (1981). *Questions and answers in attitude surveys*. New York: Academic Press.

Appendix

List of 32 items studied

Section A (Maternity care)

- A1. During prenatal followup done by physicians, women receive sufficient information about nutrition and physical condition.*
- A2. Psychological aspects of pregnancy are often neglected during physicians' prenatal followup.
- A3. Much progress is still to be made in humanizing maternity care.
- A4. Reducing obstetrical intervention (e.g. cesarean sections, episiotomies) humanizes maternity care.

- A5. All women should have two routine ultrasounds during pregnancy.*
- A6. More use should be made of hospital birthing rooms.
- A7. All women in labour should receive intravenous drip.*
- A8. Current maternity care practice respects the wishes of each pregnant woman.*
- A9. Continuous electronic fetal monitoring should be limited to high risk pregnancies.
- A10. The current caesarean section rate in Quebec is medically justified.*
- A11. Routine use of episiotomies is unacceptable nowadays.
- A12. Reducing obstetrical intervention (e.g. cesarean sections, episiotomies) would increase perinatal and maternal mortality and morbidity.*
- A13. Vaginal births after previous cesarean could be much more common.
- A14. At present, women are permitted to participate fully in decisions concerning the management of labor and birth.*
- A15. It is important that the professional providing prenatal care be the one to deliver the baby.
- A16. Current rates of obstetric intervention reflect women's requests.*

Section B (Midwives)

- B1. Maternity care can be humanized without introducing midwifery.*
- B2. Midwives can restore women's confidence in their capacity for natural childbirth.
- B3. Because of their approach, midwives can more easily detect certain risks, such as low birth weight and premature babies.
- B4. Giving birth without the aid of a physician is dangerous for the health of the mother and the baby.*
- B5. The increasing number of female physicians makes it unnecessary to legally recognize midwifery in Quebec.*
- B6. Midwives are needed because modern medicine has an overly pathological view of pregnancy and birth.
- B7. A negligible minority of women will want to be taken care of by midwives.*
- B8. Midwives will better meet the needs of pregnant women than the current professionals providing maternity care.
- B9. Introducing midwifery will reduce intervention rates (e.g. cesarean sections, episiotomies).
- B10. In the long run, midwives will take the place of general practitioners who practise obstetrics.
- B11. Midwifery constitutes a threat to the nurse's role.*
- B12. Legalization of midwifery will create healthy competition, improving the quality of care.
- B13. Recognizing midwifery will allow obstetrician/gynecologists to focus their efforts on high risk pregnancies.
- B14. Recognizing midwifery will decrease obstetrician/gynecologists' clientele.
- B15. Rather than in midwife services, we should invest elsewhere in maternity care.*
- B16. The recognition of midwifery in Quebec today is a must.

* Scores on these items were inverted using the following formulas.

Offered form: $6 - \text{original score} = \text{new score}$

Omitted form: $5 - \text{original score} = \text{new score}$