2006

The User-Friendliness of Alternative Answer Formats

Sara Dolnicar
University of Wollongong, s.dolnicar@uq.edu.au

Bettina Grun
Vienna University of Technology, Austria, bettina@uow.edu.au

Publication Details

This paper was originally published as: Dolnicar, S & Grun, B, The User Friendliness of Alternative Answer Formats, ANZMAC 2006 CD Proceedings (Australia and New Zealand Marketing Academy Conference, Brisbane, Queensland, 4-6 December 2006.

The User-Friendliness of Alternative Answer Formats

Abstract
Despite the increasing resistance of consumers to participate in market research and the vast amount of literature on the methodological superiority of certain answer formats over others, the issue of user-friendliness of different answer formats has not been investigated extensively in the past. We contribute to this area of research by investigating respondents' preferences for one of five answer formats. The preference is not measured hypothetically, respondents are invited to choose their preferred format and complete the questionnaire in the respective version. Results indicate that ordinal (polytomous and dichotomous) scales are the respondents' favourite choices. These favourite answer formats are – after having experienced them in the questionnaire – perceived as equally pleasant and equally capable of capturing the respondents' opinions.

Disciplines
Business | Social and Behavioral Sciences

Publication Details
This paper was originally published as: Dolnicar, S & Grun, B, The User Friendliness of Alternative Answer Formats, ANZMAC 2006 CD Proceedings (Australia and New Zealand Marketing Academy Conference, Brisbane, Queensland, 4-6 December 2006.

The User-Friendliness of Alternative Answer Formats

Sara Dolnicar, University of Wollongong
Bettina Grün, Vienna University of Technology

*Authors listed in alphabetical order.

Abstract

Despite the increasing resistance of consumers to participate in market research and the vast amount of literature on the methodological superiority of certain answer formats over others, the issue of user-friendliness of different answer formats has not been investigated extensively in the past. We contribute to this area of research by investigating respondents' preferences for one of five answer formats. The preference is not measured hypothetically, respondents are invited to choose their preferred format and complete the questionnaire in the respective version. Results indicate that ordinal (polytomous and dichotomous) scales are the respondents' favourite choices. These favourite answer formats are – after having experienced them in the questionnaire – perceived as equally pleasant and equally capable of capturing the respondents' opinions.

Introduction

One of the main measurement instruments in the social sciences is the questionnaire. Consequently, a lot of research work has been undertaken to assess how a questionnaire should be designed to measure what it is supposed to measure in the most valid and practically feasible way. The present paper focuses on one particular aspect of questionnaire design: the answer format. More precisely: the number of answer alternatives that are provided to respondents.

The most comprehensive review of research investigating the effects of the number of answer options conducted in the field of marketing (Cox, 1980) concludes that two and three-point scales are generally inadequate and recommends the use of a 7-point scale. This recommendation is reflected in current marketing research practice which is dominated by the use of 5 or 7-point ordinal answer formats (Van der Eijk, 2001). Most of the prior work comparing answer formats, however, has led to contradictory conclusions with respect to which answer format is optimal, typically using criteria such as reliability, validity or structural equivalence to assess optimality. One aspect that has so far been neglected as a criterion for the comparison of answer formats is user friendliness. User friendliness is central to the validity of survey findings because lack of user-friendliness is known (1) to lead to low-quality data (Herzog & Bachman, 1981; Johnson, Lehmann, & Horne, 1990), and (2) to decrease the willingness of respondents to participate in survey studies (Hardie and Kosomitis, 2005; Bednell & Shaw, 2005).

Only four studies have so far investigated this aspect: Jones (1968), Preston and Colman (2000), Dolnicar (2003) and Dolnicar, Grün and Leisch (2004). These studies resulted in different conclusions: Jones and Preston and Colman find binary scales to be perceived as quicker, but less attractive to respondents (less reliable, less interesting, less accurate, less suitable for expressing feelings, and more ambiguous). Dolnicar (2003) and Dolnicar, Grün and Leisch (2004) find that polytomous scales are perceived as more difficult while a dichotomous format is quicker and perceived as quicker.

The aim of this study is to extend research in the area of answer format comparisons by (1) investigating respondents' preferences for certain answer formats, and (2) doing this in a

choice-based rather than an evaluative manner whereby the respondents actually decide which response format they want to use for a specific survey. We hypothesise that:

H1: There are clear preferences for certain answer formats by respondents.
H2: There is an effect of gender on answer format preference.
H3: The chosen answer format will have an effect on the time needed to complete the questionnaire.
H4: The chosen answer format will have an effect on the perceived pleasantness of the questionnaire.
H5: The chosen answer format will have an effect on the perceived ability to express one's personal view when completing the questionnaire.

Hypothesis H1 is based on the assumption that respondents feel most comfortable with familiar formats and given that the ordinal format dominates marketing research, it is likely that respondents will be most attracted to ordinal answer formats. If it is distinguished between dichotomous and polytomous ordinal answer formats, it is assumed that answer formats with more answer categories are preferred according to Cox (1980, p. 420) who postulates that "scales with two or three alternatives […] tend to frustrate and stifle respondents." Hypothesis H2 is interesting to investigate because the response style literature provides a vast amount of empirical evidence that respondents who differ in socio-demographic characteristics show different usage patterns of answer formats. It seems plausible that differences in using answer formats will be reflected in differences in answer format preferences. Hypotheses H3-H5 are used to investigate and complement previous findings.

Data

The data was collected at the University of Wollongong among students attending a first year lecture in marketing. The students were asked to complete a survey on water recycling. Two different questionnaires were handed out: one included behavioural intentions with respect to recycled water and the other brand image evaluation. Both surveys also included a shortened version of the scale known as the New Environmental Paradigm (Dunlap, Van Liere, Mertig and Jones, 2000), i.e. the measured construct are attitudes. The questionnaire type (behavioural intentions or brand image evaluation) was randomly assigned to students. The students were informed that they could choose their favourite answer format out of five different formats. The five answer formats were binary (yes – no, dichotomous), 3-point scale, 7-point scale, continuous-also known as visual analogue scale-(e.g. Wewers and Lowe 1990; respondents were asked to mark a point on a line) and percentage (respondents were asked to indicate a percentage). The answer formats were presented on a separate sheet together with an exemplary question. The respondents were asked to (1) study the options, (2) indicate their favourite answer format and (3) complete the corresponding questionnaire. In addition, the following information was collected from students: the actual beginning and end time of completing the questionnaire, language spoken with their parents, perceived simplicity ("Simple"), perceived pleasantness ("Pleasant"), perceived speed ("Quick") and perceived ability to express their feelings ("Express feelings"). The responses on the perception of scale were recorded in the same way for all questionnaire versions, namely using a five-point bipolar ordinal scale. The survey resulted in 263 completed questionnaires. Of these 263, 136 (52%) students had completed the brand image questionnaire and 127 the version including behavioural intentions (48%). 155 respondents (59%) were female and 107 (41%) male.

Results

All computations and graphics for the empirical analysis have been done using the R statistical software package (R Development Core Team, 2006). Table 1 gives the absolute and relative frequencies of choices for the five answer formats. The formats are sorted with respect to their popularity, i.e. the order is 7-point scale, binary, 3-point scale, continuous and percentage.

Table 1: Absolute (Rel.) Frequency of Answer Format Choice

		Gender		Language	
Format	Total	Female	Male	English	Chinese
7-point scale	91 (0.35)	58 (0.37)	33 (0.31)	74 (0.39)	9 (0.23)
Binary	84 (0.32)	48 (0.31)	36 (0.34)	58 (0.31)	12 (0.30)
3-point scale	60 (0.23)	37 (0.24)	23 (0.21)	36 (0.19)	16 (0.40)
Continuous	19 (0.07)	9 (0.06)	10 (0.09)	14 (0.07)	2 (0.05)
Percentage	9 (0.03)	3 (0.02)	5 (0.05)	6 (0.03)	1 (0.03)
Total	263 (1.00)	155 (1.00)	107 (1.00)	188 (1.00)	40 (1.00)

Favourite Answer Format

To investigate if certain answer formats are preferred, it is first checked if a single favourite answer format exists. Pairwise comparisons of the proportions are undertaken using Pearson's Chi-Squared test with Yates' continuity correction and Holm's method to correct for multiple testing (Holm, 1979). The results are given in Table 2. The results indicate that no single most popular answer format exists. However, the ordinal multi-category answer formats (binary, 3-point, 7-point) are generally preferred to formats where the answer is recorded on a nearly continuous scale, such as the continuous and percentage scale. While the 7-point scale is the most popular, the binary scale is chosen almost as frequently. The 3-point scale is significantly less popular than the 7-point scale with respect to a significance level of 0.05, while the null hypothesis of equal proportions can not be rejected for a comparison of binary with 3-point scale. Percentage is the least popular scale, but the difference to the continuous scale is not significant.

To avoid confounding effects due to language differences, we have recomputed the test separately for the English-speaking and the Chinese-speaking sub-samples. The results for the English-speakers are the same as for the total sample, minor deviations from the total sample results occur, however, for the Chinese-speaking subsample: the difference between 3-point scale and binary or 7-point scale and continuous and binary or 7-point is not significant.

Table 2: Pairwise Comparison of Proportions

	7-point scale	Binary	3-point scale	Continuous
Binary	0.579	-	-	-
3-point scale	0.015	0.074	-	-
Continuous	< 0.001	< 0.001	< 0.001	-
Percentage	< 0.001	< 0.001	< 0.001	0.161

If the ordinal scales are combined and compared to the continuous scales this shows that 89.3% of the respondents chose an ordinal scale and that this proportion is significantly higher than 50% ($\chi^2 = 161.4$, df = 1, $p < 0.001$). H1 (There are clear preferences for certain answer formats by respondents) is confirmed. While no single most preferred answer format can be determined, the ordinal (polytomous and dichotomous) formats are clearly perceived as preferable to the continuous versions.

Gender Effect on Answer Format Choice

The absolute and relative frequencies of the answer format choices cross-tabulated with gender and language are given in Table 1. Fisher's exact test is used to investigate if there is a significant association between choice of answer format and one of the concomitant variables. The null hypothesis of independence can not be rejected for gender ($p = 0.44$). A separate analysis for the English and the Chinese respondents give similar results (English: $p = 0.52$; Chinese: $p = 0.48$). H2 (There is an effect of gender on answer format preference) is not confirmed based on the available data.

Duration and Perception of Alternative Answer Formats

The duration and respondents' perception of alternative answer formats is analysed using ANOVA with linear models. In the analysis of duration only those respondents (97%) were included where end minus begin time gave a positive number assuming that the respondents needed at least one minute to complete the questionnaire. For the analysis of respondents' perception of alternative answer formats equal distances between the five answer categories of the perception question are assumed. Analyses of variance are conducted, with duration and perception as dependent variables and different answer formats and questionnaire type as independent variables. Before analysing the duration the logarithm is taken because the variable is distributed skewed to the right otherwise. Results are provided in Table 3.

Table 3: ANOVA for Duration and Perception of Scale

	log(Duration)			Simple			Pleasant		
	Df	F	p	df	F	p	df	F	p
Type	1	23.16	< 0.001	1	4.62	0.033	1	7.47	0.007
Format	4	2.62	0.036	4	5.12	< 0.001	4	0.74	0.566
Type:Format	4	1.31	0.266	4	1.34	0.257	4	0.15	0.961
Residuals	226			244			244		

	Quick			Express feelings		
	Df	F	p	df	F	p
Type	1	8.09	0.005	1	6.63	0.011
Format	4	3.31	0.012	4	0.20	0.938
Type:Format	4	1.04	0.385	4	0.96	0.432
Residuals	244			244		

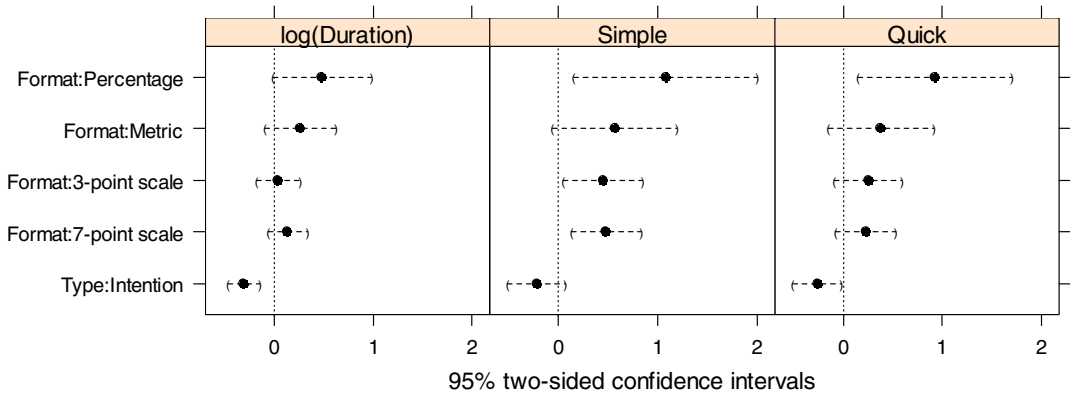
The ANOVA for duration indicates that there is a significant influence of questionnaire type and answer format, but no interaction effect. With respect to respondents' perceptions of alternative answer formats the interaction effect is not significant for any of the variables. This means that the influence of the answer formats on the evaluation of the questionnaires does not depend on the questionnaire type. This can be seen as a reassurement that the conclusions drawn with respect to answer format can be generalised to other questionnaires.

The influence of the questionnaire type is always significant, while the answer format leads to a significant difference only for the items "Simple" and "Quick". It can consequently be concluded that all answer formats are perceived as equally "Pleasant" and suitable to "Express feelings". Simultaneous confidence intervals are derive for the coefficients of the relevant linear models without interaction effect using Dunnett's method (Dunnett, 1955) and are given in Figure 1. The intervals are plotted as horizontal lines where the limits of the intervals are given by round brackets and the estimates by a point. The binary scale is used as the base category for the answer format and the behavioural intentions for the questionnaire type. It can be seen that with respect to duration or perception, the questionnaire including behavioural intentions always has lower or equal scores. With respect to simplicity the

percentage, 3-point and 7-point scale are perceived as significantly less simple than the binary scale. With respect to quickness a significant difference to the binary scale can only be determined for the percentage scale.

H3 (The chosen answer format will have an effect on the time needed to complete the questionnaire) is confirmed as the influence of answer format on duration is significant and the continuous answer formats took longer to complete. H4 (The chosen answer format will have an effect on the perceived pleasantness of the questionnaire) and H5 (The chosen answer format will have an effect on the perceived ability to express one's personal view when completing the questionnaire) are not confirmed, as no significant differences with respect to either pleasantness or the ability to express one's feelings could be determined.

Figure 1: Simultaneous Confidence Intervals for Duration and Perception of Scale



Conclusions and Future Work

Respondents have different preferences with respect to answer formats. No single answer format is preferred by the majority of respondents, but a clear preference of polytomous and dichotomous answer formats was determined as opposed to continuous formats. The answer format preference does not depend on gender. In terms of user-friendliness, the continuous answer formats were found to take longer to complete, all formats are perceived as equally pleasant and as providing the same opportunity to express feelings. With respect to simplicity and quickness the respondents choosing the binary scale gave the most favourable evaluation. These evaluations could potentially be blurred by the fact that the answer format was not imposed on the respondents, but that they actually chose it themselves and thus might be evaluating the answer format they chose as more favourable than they would have evaluated the same scale if imposed on them.

The implications for marketing researchers from this study are that there are answer formats which are preferred by respondents and can thus be used to make the market research experience as pleasant as possible and, as a consequence, increase the willingness of respondents to participate in survey research. Interestingly, binary format is among the most preferred, which opens up the opportunity of not only offering respondents a pleasant scale but also benefiting from the quickness of this format and saving fieldwork cost. Optimally, the suitability and preference of answer formats should be pre-tested for the construct under study before the actual fieldwork phase is undertaken. In addition, other measurement issues also have to be taken into account, for example accuracy, reliability and validity.

Acknowledgements: This research was supported by the Australian Research Council (through grants DP0557257 and LX0559628), the University of Wollongong through internal research grant schemes and the Austrian Academy of Sciences (ÖAW) through a DOC-FFORTE scholarship for Bettina Grün.

References

Bednell, D. H. B., and Shaw, M., 2003. Changing response rated in Australian market research, Australasian Journal of Market Research, 11(1), 31-41.

Cox, E. P., 1980. The optimal number of response alternatives for a scale: A review. Journal of Marketing Research 17 (4), 407-422.

Dolnicar, S., 2003. Simplifying three-way questionnaires - Do the advantages of binary answer categories compensate for the loss of information? ANZMAC CD Proceedings.

Dolnicar, S., Grün, B., and Leisch, F., 2004. Time efficient brand image measurement - Is binary format sufficient to gain the market insight required? CD Proceedings of the 33rd EMAC conference.

Dunnett, C. W., 1955. A multiple comparison procedure for comparing several treatments with a control. Journal of the American Statistical Association 50(272), 1096-1121.

Dunlap, R. E., Van Liere, K. D., Mertig A. G., and Jones R. E., 2000. Measuring Endorsement of the New Ecological Paradigm: A Revised NEP Scale. Journal of Social Issues 56(3), 425-442.

Hardy, T., and Kosomitis, N., 2005. The F word in market research: high impact public relations in the 21st century. CD Proceedings of the Australian Market & Social Research Society.

Herzog, A.R., and Bachman, J.G. (1981). Effects of questionnaire length on response quality. Public Opinion Quarterly, 45(4), 549-559.

Holm, S., 1979. A simple sequentially rejective multiple test procedure. Scandinavian Journal of Statistics 6, 65-70.

Johnson, M. D., Lehmann, D. R., and Horne, D. R., 1990. The effects of fatigue on judgments of interproduct similarity, International Journal of Research in Marketing, 7, 35-43.

Jones, R. R., 1968. Differences in Response Consistency and Subjects' Preferences for Three Personality Inventory Response Formats. Proceedings of the 76th Annual Convention of the American Psychological Association, 247-248.

Preston, C. C., and Colman, A. M., 2000. Optimal number of response categories in rating scales: reliability, validity, discriminating power, and respondent preferences. Acta Psychologica 104, 1-15.

R Development Core Team, 2006. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Van der Eijk, C., 2001. Measuring agreement in ordered rating scales. Quality & Quantity 35, 325-341.

Wewers, M. E., and Lowe, N. K., 1990. A critical review of visual analogue scales in the measurement of clinical phenomena. Research in Nursing and Health 13(4), 227-236.