



“Translating” between survey answer formats[☆]

Sara Dolnicar^{a,*}, Bettina Grün^{b,1}

^a Institute for Innovation in Business and Social Research, University of Wollongong, Wollongong, NSW 2522, Australia

^b Department of Applied Statistics, Johannes Kepler Universität Linz, 4040 Linz, Austria

ARTICLE INFO

Article history:

Received 1 April 2011

Received in revised form 1 October 2011

Accepted 1 November 2011

Available online 2 March 2012

Keywords:

Questionnaire design

Survey research

Answer formats

Likert

ABSTRACT

Survey research remains the most popular source of market knowledge, yet researchers have not yet established one consistent technique for measuring responses. Some market research companies offer respondents two answer options; others five or seven. Some answer formats use middle points on the answer scales, others do not. Some formats verbalize all answer options, some only the endpoints. The wide variety of answer formats that market research companies and academic researchers use makes comparing results across studies virtually impossible. This study offers guidance for market researchers by presenting empirical translations for the answer formats they most commonly use, thus enabling easier comparisons of results.

© 2012 Elsevier Inc. All rights reserved.

1. Introduction

Organizations heavily use survey research to learn about consumer behavior, preferences, and perceptions. While repeat surveys by the same organization using the same market research company typically use the same answer format, this does not occur in studies that different organizations, market research companies, or academic researchers conduct, which makes comparing results across different studies virtually impossible.

A good example of this problem occurs in research into the stated acceptance for recycled water. Researchers first conducted studies in this area in the early 1970s, and continue to conduct them internationally. Two Australian examples illustrate the point well. They were both published in 2006 and refer to the same geographic region, yet report acceptance levels for drinking recycled water of 11% and 47% respectively; a difference that suggests that how the questions are asked, and what answer options are offered, significantly affect results. Hurlimann (2006), who reports the higher acceptance level, asked respondents how happy they would be using recycled water, and offered a ten-point scale ranging from *not at all happy to use recycled water to extremely happy to use recycled water*. The authors added responses with the value of six or more on the ten-point scale to determine the 47% acceptance level. Dolnicar and Schäfer (2006) report the lower acceptance level of 11%. They asked

respondents in that study a scenario question and offered five fully verbalized answer options; the 11% acceptance level represents the respondents who selected the *very likely* answer option.

The consequences of such measurement inconsistencies and the absence of guidance on how to compare results across studies are that recycled water usage studies have produced many heterogeneous and incompatible numbers, instead of making definitive contributions to the body of knowledge. Such dissimilar results appear in many contexts, because no strategies are available for comparing survey results that employ different answer formats. The lack of tools to compare results effectively weakens our ability to draw valid conclusions and develop a body of knowledge in certain research areas.

The present study addresses the problem of heterogeneous and incompatible survey results by offering empirical translations that support comparisons of results across studies, regardless of the answer formats employed. The tools that this study generates should be particularly useful to market researchers, academic researchers, and users of market research studies. Specifically, this study provides translations that allow practitioners to compare: the forced-choice full binary answer format against other answer formats in common use; answer formats with middle points against answer formats without middle points; Likert-type and bipolar answer formats; and answer formats with fully verbalized options against endpoint-labeled answer formats.

In offering empirical translations to compare results from different survey methodologies, this study contributes to the theoretical understanding of answer formats in survey research, and is of direct practical value to market researchers, academic researchers and users of market research results.

This study does not determine a single, most-valid answer format. Rather, it accepts that different studies use different answer formats, and the consequent virtual impossibility of comparing results across

[☆] This research was funded by the Australian Research Council under the Discovery and Linkage International Grant Schemes and by the Austrian Science Fund (FWF) under Elise-Richter grant V170-N18. Authors are listed in alphabetical order.

* Corresponding author. Tel.: +61 2 4221 3862; fax: +61 2 4221 4210.

E-mail addresses: sarad@uow.edu.au (S. Dolnicar), Bettina.Gruen@jku.at (B. Grün).

¹ Tel.: +43 732 2468 6829; fax: +43 732 2468 6800.

studies. This paper is the first to provide guidance for translating different answer formats onto one another. Such guidance is important when comparing findings across studies, or comparing results over time in longitudinal studies, because researchers often encounter dissimilar answer formats. In addition, researchers frequently binarize multi-categorical data using the middle point to split respondents. This study demonstrates that such binarization does not actually match the internal translation process of respondents, which leads to invalid data transformations before data analysis even starts. The presented translations address problems associated with changed or different answer formats, and validity in the binarization of multi-categorical data.

The context of brand image measurement limits empirical investigation in the present study; traditionally, the free-choice binary or pick any/n answer format dominates commercial research (such as in brand tracking studies). According to Rossiter (2011, p. 75), brand-attribute beliefs, which brand image studies measure, are the single most common construct measured in marketing research. Also, interactions often occur between the construct under study and the answer format; and therefore, results may deviate somewhat for other constructs under study (Dolnicar & Grün, 2007a, 2009).

1.1. Prior work

Prior work that relates to this study resides in two areas. First are studies that seek the best answer formats. Second are studies that attempt to translate between answer formats. The research debate over the best answer format is as old as survey research itself. Authors tend to (rather passionately) take one of two positions: either they propose that binary measures are sufficient (Bendig, 1954; Dolnicar & Grün, 2007a, 2007b; Dolnicar, Grün, & Leisch, 2011; Komorita & Graham, 1965; Martin, Fruchter, & Mathis, 1974; Matell & Jacoby, 1971a, 1971b; Schutz & Rucker, 1975), or they tend to reject absolutely binary measures and instead use multi-category answer formats. Within the latter group, views differ regarding the optimal number of answer options, with recommendations ranging from five (Boote, 1981; Jenkins & Taber, 1977; Lissitz & Green, 1975; Remmers & Ewart, 1941), to six (Finn, 1972; Green & Rao, 1970), to seven (Cicchetti, Showalter, & Tyrer, 1985; Miller, 1956; Oaster, 1989; Symonds, 1924) and nine (Hancock & Klockars, 1991), and 18 or more (Champney & Marshall, 1939; Garner, 1960). The key argument between these opposing groups is whether additional answer options add precision to the measurement, or merely capture noise (such as response styles).

Garner (1960, p. 352) is representative of the opinion of multi-category proponents: “information transmission cannot be lost by increasing the number of rating categories. Therefore, it is better to err on the side of having too many categories than to err by having too few.” Peabody (1962, p. 73) characterizes the position of binary measure proponents: differences in responses using multi-category answer formats “primarily represent response sets, and only to a secondary degree actual differences in intensity.” This group believes that response sets represent contamination of data, rather than additional information. Avoiding response bias, according to Rossiter (2002, 2011), is a key requirement for any measure to be content valid, and content validity is the ultimate quality criterion for measures in the social sciences.

The body of literature on answer formats does not lead to any firm conclusion about what is ultimately the best answer format. This vagueness is attributable to how past researchers have conducted studies in a range of different contexts, using a range of different evaluation criteria for answer formats, and with many variations in how they word answer options or present them to respondents. Despite the significant body of research comparing answer formats, no work has yet been conducted comparing different formats of binary measures (e.g., pick any/n compared to forced full binary).

Only a very small number of studies are available that relate to translating responses from one answer format to another. Haley and Case (1979) provide the first study of this kind, evaluating 13 commonly used scales in brand image measurement with respect to answer patterns, measured content, concurrent validity, and discrimination between brands. They conclude that forced-choice answer formats, as well as answer formats with fully verbalized answer options, perform better. Hui and Triandis (1989) compare responses from five- and ten-point answer formats for Hispanic and non-Hispanic respondents. However, their research design, which is not longitudinal, does not permit mapping across answer formats. The chart they provide shows frequencies of use for each answer option for both formats, and indicates that more answer options reduce extreme response styles.

Dolnicar and Grün (2007a) and Dolnicar et al. (2011) examine transformations between a limited number of answer formats. Dolnicar and Grün (2007a) scrutinize measures of two different constructs (behavioral intentions and attitudes), employing a repeat measurement design on three different answer formats (full binary, metric and ordinal seven-point); while Dolnicar et al. (2011) investigate the mappings between a full binary and an ordinal six-point answer format.

2. Data and method

The experiment used a permission-based internet panel that asked respondents representative of the Australian adult population to complete two brand image questionnaires with an approximate two-week break between measurements. Both questionnaire versions were identical, except for the answer format. This design enabled the derivation of individual-level translations, because the collected data allowed mapping of how each respondent answered from one answer format to another. Any variation between the two measurements was not caused by inter-individual differences or changes in brand perception, because the time between measurements was short, and no changes in advertising campaigns or the marketplace occurred that could have changed respondents' brand evaluations.

Brand image measurements are not perfectly stable, even under unchanged market conditions or when the same answer format is used (Dolnicar & Grün, 2007b; Dolnicar & Rossiter, 2008; Rungie, Laurent, Dall'Olmo Riley, Morrison, & Roy, 2005). Therefore, also the present study will capture some of this instability. However, a reduction of this effect was achieved by following the measurement recommendations of Dolnicar and Rossiter (2008). Also, any variations due to instability in brand image measurement should affect all experimental conditions equally, with no bias toward any of the answer formats. In addition, base instability levels are reported for repeat measurements on the same answer format.

Respondents assessed two brands: McDonald's (very well known among Australians) and Red Rooster (less well known). The five attributes presented to respondents were *yummy*, *fast*, *cheap*, *healthy*, and *convenient*. These attributes were derived from a prior, extensive, qualitative study where interview respondents were asked about the relevant characteristics of fast food brands. Each item identified through the qualitative study was viewed by respondents as relevant to consumers, easy to understand, and formulated in consumer language.

The affirmative binary format is better known as the pick any/n format. Respondents were given a list of attributes and asked to select those that applied to a given brand. If they did not wish to assign an attribute to a brand then they were asked not to select the attribute. The full binary format version of the questions required respondents to state whether or not they believed that each of the listed attributes applied to any given brand. As with the affirmative binary format version, the information available in the data set was binary, but

Download English Version:

<https://daneshyari.com/en/article/10493036>

Download Persian Version:

<https://daneshyari.com/article/10493036>

[Daneshyari.com](https://daneshyari.com)