

Cross-cultural application of a practice-oriented acquiescence measure

A practice-oriented
acquiescence
measure

Carolyn Krautz and Stefan Hoffmann

Department of Marketing, University of Kiel, Kiel, Germany

391

Received 21 March 2018
Revised 4 July 2018
7 October 2018
26 October 2018
Accepted 7 November 2018

Abstract

Purpose – Extant research shows that acquiescence response style (ARS) is culture-bound and may bias the results of comparative cross-cultural studies. Conventional measures of ARS are difficult to apply in practice. To overcome this limitation, the purpose of this paper is to propose an alternative, practice-oriented measure, namely, pARS. The authors apply Hofstede's cultural dimensions (Hofstede *et al.*, 2010) to test whether pARS is culture-bound. The cross-cultural study provides a high level of cross-cultural generalisability due to the extensive number of surveyed countries ($n = 30$) and subjects ($n = 236.089$). The authors run multi-level analysis to identify within- and between-country-level predictors.

Design/methodology/approach – On the individual level, the authors use data of a large-scale cross-cultural study, including 236.089 consumers from 30 countries worldwide. The authors apply several methods to test for the culture-boundness of pARS. First, they apply correlation analysis to replicate existing cross-cultural results and to ensure nomological validity. Second, applying ordinary least square regression, the authors simultaneously test the six Hofstede cultural dimensions (Hofstede *et al.*, 2010) and investigate interactions between the dimensions. Finally, they use multi-level analysis to confirm the stability of culture-bound results, controlling for individual- and country-level variability.

Findings – The paper introduces an alternative measure for acquiescence (pARS), which is particularly suitable for shorter questionnaires. A large-scale consumer study with 236.089 respondents in 30 countries supports the culture-bound validity of pARS. The authors confirm construct validity and the nomological network of pARS. Contrasting existing studies, multi-level analysis demonstrates that a high level of power distance majorly leads to ARS. Therefore, cross-cultural researchers need to control for ARS in countries high in power distance, especially when paired with high uncertainty avoidance.

Originality/value – A large-scale consumer study with 236.089 respondents from 30 countries shows that respondents from various countries differ significantly in their level of acquiescence. The study confirms that power distance is the most relevant cultural dimension to explain these differences. Although ARS may bias the results of comparative cross-cultural studies, it is rarely controlled by market research studies outside the academic realm. The present work proposes and establishes the validity of a practice-oriented measure of acquiescence, namely, pARS. pARS is particularly suitable for shorter questionnaires. In contrast to prior approaches, applying pARS does not require adding non-substantive items to the questionnaire.

Keywords Cross-cultural management, Hofstede, Hierarchical linear modelling, Response style, Acquiescence, ARS

Paper type Research paper

1. Introduction

Studies in the field of comparative cross-cultural marketing research aim at understanding different national and/or cultural markets (Malhotra *et al.*, 1996; Sojka and Tansuhaj, 1995). If enterprises build their global strategies on these studies, they need to be sure that study results are unbiased because they need to avoid budget misallocations. Yet, extant literature confirms that response styles may bias questionnaire data and that these contaminations are contingent on the culture of the respondents (Dolnicar and Grun, 2007; Hoffmann *et al.*, 2013). This (stable) systematic measurement error makes up to one quarter of the total variance (Baumgartner and Steenkamp, 2006a).

The term “response styles” refers to the respondents' systematic tendencies to answer to questionnaire items in a manner other than what the item intends to measure (Baumgartner and Weijters, 2015; Paulhus, 1991). Hence, response styles mask the substantive answer of the respondent and content validity is no longer given. These contaminations have major



effects on the magnitude of scale scores. They bias the results of statistical analyses, such as the magnitude of the correlation of two or more variables (Clarke, 2001; Dolnicar and Grun, 2007). Scholars have repeatedly warned that comparing biased data from different cultures leads to delusive conclusions for marketing practices (Cheung and Rensvold, 2000; Greenleaf, 1992; Singh, 1995).

The acquiescence response style (ARS), which is also referred to as agreement tendency, yeasaying or positivity (Baumgartner and Steenkamp, 2001), is the most pervasive and widely studied response pattern in cross-cultural marketing (Hoffmann *et al.*, 2013). By definition, it refers to the tendency to respond affirmatively to propositions regardless of their content (Paulhus, 1991; Winkler *et al.*, 1982). Affirmative responses (AR) may occur due to two reasons: true variance, which means that the respondent actually answers to the content of the item, and stylistic variance, which refers to the manner in which subjects respond in a biased way (Bentler *et al.*, 1971; Greenleaf, 1992; Harzing *et al.*, 2006).

Previous research introduced three measures to quantify stylistic variance of ARS in questionnaires (Baumgartner and Steenkamp, 2006b): an explicit measure for acquiescence (e.g. Couch and Keniston, 1960); acquiescence scores derived from balanced-worded scales including negatively and positively worded items (Billiet and McClendon, 2000; Baumgartner and Steenkamp, 2001; Welkenhuysen-Gybels *et al.*, 2003); or acquiescence scores derived from a number of content-wise heterogeneous items (Baumgartner and Steenkamp, 2001). Figure 1 gives an overview of different types of ARS measures. Each of the three approaches has some shortcomings, which limit the application in research practice. The first two approaches predominantly add non-substantive items to a questionnaire. We define non-substantive items as items which do not relate to the content of the study and the constructs which are to be measured. The third approach requires a large number of substantive items which are content-independent and measured on the same (rating) scale. Including additional (substantive and non-substantive) items raises the threat of common method variance (Podsakoff *et al.*, 2003), and is costly, both in terms of monetary resources and respondent's fatigue (Greenleaf, 1992). Following the discussion of rigour and relevance (e.g. Wolf and Rosenberg, 2012), the present paper aims to propose a new practice-oriented measure for acquiescence which is based on a test of truth. As a consequence, the new approach requires only a small set of items.

This study makes at least two contributions to the literature. First, we propose the practice-oriented measure for ARS (pARS), which accounts for affirmative responding independent of item content (Figure 1). We validate the new approach based on a large-scale consumer study. Second, we build on Hofstede's cultural dimensions (Hofstede *et al.*, 2010) to test whether ARS

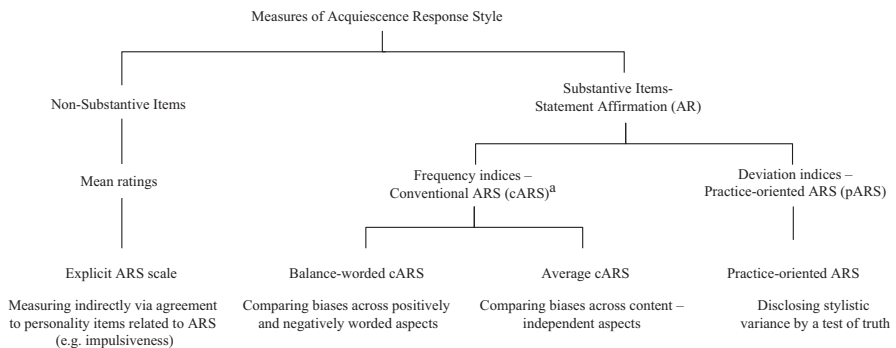


Figure 1.
Types of acquiescence measures

Note: ^aConventional ARS measures include model-based approaches

Sources: Billiet and McClendon (2000); de Jong *et al.* (2008); Welkenhuysen-Gybels *et al.* (2003)

is culture-bound. Our cross-cultural study makes several methodological improvements. In terms of the number of countries ($n = 30$) and subjects ($n = 236.089$) surveyed, our study is more comprehensive than previous approaches and allows for cross-cultural generalisability (e.g. Harzing *et al.*, 2006; Hoffmann *et al.*, 2013). Moreover, while previous studies consider only the main effects of Hofstede's cultural dimensions, we include interaction terms to account for holistic cultural profiles (Richter *et al.*, 2016). Additionally, we apply multi-level analysis to identify predictors for ARS on a within- and between-country level (see de Jong *et al.*, 2008 for a related analysis of extreme response styles).

Accordingly, we aim to answer the following four research questions:

- RQ1.* We consider whether an alternative to the conventional measure ARS exists which is more suitable for practical applications.
- RQ2.* We examine whether the pattern of cultural predictors of pARS supports nomological validity.
- RQ3.* We consider all cultural dimensions simultaneously and investigate whether this holistic approach enhances the prediction of ARS.
- RQ4.* We investigate whether these cultural dimensions interact in predicting ARS.

The remainder of the paper is organised as follows. First, with respect to our first research question, we describe and discuss conventional acquiescence response style measures (cARS) and develop the rationale of the new measure. Second, with respect to our second research question, we review extant studies on the impact of culture on ARS measured with cARS. We test whether these findings can be replicated with the new pARS measure. Additionally, we derive assumptions how the cultural dimensions long-term orientation and indulgence influence pARS. We also derive assumptions about interactions between cultural dimensions. We provide empirical support to answer *RQ1–RQ4*. Finally, we discuss the results.

2. Measuring acquiescence response style

2.1 Conventional measures of acquiescence response style (cARS)

Extant literature has identified several response styles, which may bias the results of survey data, including acquiescence, disacquiescence, mid-point and extreme response style (Hoffmann *et al.*, 2013). Previous studies demonstrate that ARS is a serious problem for cross-cultural research (see Table I; Hoffmann *et al.*, 2013; Van Herk *et al.*, 2004). Therefore, the present study focusses on ARS, which is the bias for the extreme positive end of (rating) scales. To date, researchers have two options to measure ARS: they can either include an explicit scale of non-substantive items in the research design or they can calculate ARS frequency scores *post hoc* over a number of items (Baumgartner and Steenkamp, 2006b). We differentiate between scales based on substantive items and those which add additional items which are not in the research focus (non-substantive). Furthermore, the rationale of the most conventional measures is to calculate indices of the frequency of specific responses to a number of items.

2.1.1 Explicit acquiescence scales. Researchers sometimes include a set of items which explicitly measures ARS. However, scholars have proposed only few explicit ARS scales. Couch and Keniston (1960) proposed a measurement that builds on the relationship between personality traits and ARS. They argue that ARS manifests in emotional responsiveness and impulsiveness. The authors composed a short scale with 20 items. However, this scale has rarely been applied (Paulhus, 1991).

2.1.2 Post hoc calculated acquiescence scores. Building on Martin (1964), Baumgartner and Steenkamp (2001) suggest approaches which are based on substantive and non-substantive constructs and which rule out biased affirmation in the observed responses. They identify and apply two major methods to account for ARS, namely, the

	Balanced- worded	Baumgartner and Steenkamp (2001) ^b		Van Herk <i>et al.</i> (2004)	Average		Hoffmann <i>et al.</i> (2013)
	Johnson <i>et al.</i> (2005)	Johnson <i>et al.</i> (2005)	Johnson <i>et al.</i> (2005)	Johnson <i>et al.</i> (2005)	Smith (2004)	Harzing (2006)	Hoffmann <i>et al.</i> (2013)
Number of countries	10 (2)	11 (1)	6 (1)	/ ^c	26 (4)	8 (3)	
Measure	Balanced- worded scale	Weighted average, balance- worded scale	Arithmetic average corrected by number of disagreeing items	Arithmetic average	Arithmetic average corrected by number of disagreeing items	Weighted average	
Number of items	18 (9 random pairs)	60 (12 item sets)	114 (18 item sets)	/	69 (4 item sets)	11 (randomly selected)	
<i>Findings^a</i>							
Power distance	-	/	/	+	+	+	
Individualism/ collectivism	-	/	-	-	-	-	
Masculinity/ femininity	-	/	/	ns	ns	ns	
Uncertainty avoidance	-	/	/	ns (partly)	ns	ns	
Long-term orientation	/	/	/	/	/	/	
Indulgence	/	/	/	/	/	/	

Table I.

Overview of previous cross-cultural studies on cARS

Notes: ns, not significant; ns (partly), only partly significant. ^aCorrelation with Hofstede's cultural dimensions; ^bno cross-cultural prediction due to small cross-country differences; ^cnumber of countries not specified. +, significant positive correlation; -, significant negative; /, relationship not tested

balanced-worded approach and the average approach. We subsume these two approaches under the umbrella of cARS. Both methods are based on the calculation of a frequency index. Other scholars build indices on basis of measurement models (Billiet and McClendon, 2000; de Jong *et al.*, 2008; Welkenhuysen-Gybels *et al.*, 2003). However, in most cross-cultural studies, scholars have applied an adjusted form of the average approach. Table I provides an overview of the application of each approach in cross-cultural studies.

2.1.3 Balanced-worded approach. The balanced-worded approach suggests to calculate a score or method factor for the extent to which respondents agree with scales which comprise equally positively and negatively worded statements. Winkler *et al.* (1982) suggested a binary coding scheme that assigns "1" to each question pair the respondent agrees to (or tends to agree to) and "0" to pairs to which the respondent disagrees. Baumgartner and Steenkamp (2001), who applied the balanced-worded approach to validate the average approach, report that the two approaches are substantively correlated. Contrary, Swain *et al.* (2008) indicate that a misresponse to reversed items does not equate with ARS. Scholars have rarely applied the balanced-worded scales in cross-cultural research as the consistency of responses to positively and negatively worded items differ across cultures (Wong *et al.*, 2003).

2.1.4 Average approach. This approach suggests to average a number of content-wise heterogeneous items to which the respondent agrees on Likert-typed rating scales (Bachman and O'Malley, 1984; Marín *et al.*, 1992). The rationale is that unbiased subjects would not affirm to a range of content-wise heterogeneous items. Several scholars applied the average approach in comparative cross-cultural studies. Smith (2004) suggested to

average the scores for a range of unrelated items. Van Herk *et al.* (2004) and Harzing *et al.* (2006) also considered negative scores. They counted the number of clearly positive ticks, subtracted the number of negative ticks and calculated the average. Baumgartner and Steenkamp (2001) and Hoffmann *et al.* (2013) used a weighted mean with a stronger weight for extreme positive categories. The weighted mean is the most elaborate measure as it emphasises extreme positive answers without neglecting the less extreme ones.

2.2 Conventional measures of ARS in research practice

We consider all three approaches as feasible means to measure ARS in academic research. Nevertheless, all approaches suffer from shortcomings which limit the application in research practice. In practice, market researchers have to find a trade-off between including substantive content and the costs associated with questionnaire length. Practitioners often strive to design scales which are short and relevant rather than being precise with scientifically grounded multi-item scales (de Jong *et al.*, 2009). Adding non-substantive items for the sake of controlling ARS is costly, monetary-wise and with regard to respondent fatigue (Greenleaf, 1992). Remarkably, longer questionnaires are also a source of biases (Hinkin, 1995). Therefore, the practical relevance is at odds with the traditional approaches, such as including additional explicit scale approach or adding balanced positive and negative items. By contrast, the average approach requires a large number of content-independent substantive items (up to 100 items), which have to be measured on the same (rating) scale. Long item batteries raise the danger of common method bias (Podsakoff *et al.*, 2003). On the other hand, shorter surveys may not contain sufficient content-independent items to probe response styles (Greenleaf, 1992). Therefore, the present paper introduces pARS as a new measure of ARS, which satisfies both practical conditions: pARS only requires a small number of items which measure substantive content.

2.3 Introducing a practice-oriented measure of acquiescence response style (pARS)

The present paper proposes an alternative, practice-oriented approach to differentiate stylistic from true variance. In previous approaches, scholars used information derived from rating scales. Yet, by applying control measures which use the same method as the scale that has to be corrected they cannot completely disentangle true and stylistic effects. To overcome this limitation, our new measure comprises different question types, which detect the true level of ARS. Similarly, to conventional approaches, subjects firstly evaluate items on a scale which is vulnerable to ARS. The second question is a choice test, which subjects can answer correctly or incorrectly depending on their knowledge. We suggest a test of truth, which compares these objectively true or false answers to the ratings on the first scale. In this way, the new approach helps to quantify the level to which the respondent answered in a stylistically affirmative manner. The formula to calculate pARS is presented in the method section of this paper. In international marketing research, scholars may, for instance, apply this test of truth with regard to product knowledge (as done with regard to automotive knowledge in the present empirical study). The first question could cover the respondent's self-reported familiarity with technical and ecological ingredient brands to which we refer as sub-brands. The automotive sub-brands of the present empirical study indicate a specific technology of one car manufacturer (e.g. "Hybrid Synergy Drive" by Toyota). The second question asks for the identification of the corresponding parent brand. If data were unbiased (true), respondents of higher familiarity with a sub-brand are more likely to identify the corresponding parent brand. Thus, the deviation between the (weighted) average score of question 1 and the mean score of question 2 specifies the level of stylistic and true variance. If the deviation is high, the respondent's answers are stylistically biased. These respondents highly affirm familiarity with the presented sub-brands, but they

fail to identify the parent brand. Conversely, if respondents display affirmative answers and if the deviation is low, then we conclude on substantive knowledge. The above-mentioned field of cars is exemplary. In order to transfer this new method to other fields, researchers need to make a few adjustments. In general, the procedure is suitable for surveys in which the respondents can have any kind of objective knowledge. Above and beyond marketing research, pARS can be applied to inter-organisational (employee) or corporate partner (B-to-B business) studies also. Maintaining the format of a test of truth and the calculations described in the measurement section, adjustments are easy to make by simply changing the content of the indicators. cARS measures are equally valid if knowledge is not the primary study objective.

Conventional ARS measures, though defined as affirmations to propositions, were primarily applied to attitudinal Likert scales (e.g. Van Herk *et al.*, 2004). Our new approach uses the affirmation to familiarity scales. Familiarity scales are sensitive to stylistic answers as the response sets affect, among others, measures of belief and ability (Bachman and O'Malley, 1984; Couch and Keniston, 1960). Familiarity scales could be judged for objective truth which is in contrast to attitudinal Likert scales. Greenleaf's (1992) seminal work is based on a similar idea, contrasting subjective attitudinal scales to more objective self-reported behaviours. Greenleaf (1992) argues that acquiescence goes beyond attitudinal scales and is relevant in, for example, ratings of product attribute importance.

Conventional ARS measures require a wide range of content-independent items to split AR into true and stylistic parts. Our new measure of familiarity items neither requires to include content-independent items, nor does it require a large number of items. We suggest to include a separate knowledge question, which differentiates between the affirmative bias and substance.

3. Cultural predictors of acquiescence response style

3.1 *Extending knowledge on culture-bound ARS*

Having introduced the new measure, we intent to answer the research questions. To respond to *RQ1*, we apply the new approach to test stylistic bias in affirmative responding. For answering *RQ2*, we compare our findings with the cultural influences found in extant studies. In contrast to prior cross-cultural studies and with respect to *RQ3* and *RQ4*, we consider holistic cultural profiles in which we test all dimensions simultaneously and their interactions (Richter *et al.*, 2016). We advance research on culture-bound ARS by applying the updated cultural dimensions of Hofstede (Hofstede *et al.*, 2010).

3.2 *Previous findings*

This section reviews previous findings on how national-level culture influences individual ARS. According to Hofstede (1991, 1994), culture can be understood as "a collective programming of the mind" that one group shares and which sets the group apart. As most major cultural theories (e.g. House *et al.*, 2004; Schwartz, 1994; Triandis, 1988; Trompenaars, 1994), Hofstede suggests a multi-dimensional theory which distinguishes different national cultures. Cross-cultural studies, in particular those on culture-bound response styles, most widely apply Hofstede's cultural dimensions (Kirkman *et al.*, 2006). We use Hofstede's approach to ensure comparability. Hofstede's (1991) work describes the common unconscious socialisation which members of a society share. These normative values relate to how the members solve problems and organise their communal life. The original conceptualisation comprises four dimensions: individualism vs collectivism (IDV), power distance (PDI), masculinity vs femininity (MAS) and uncertainty avoidance (UAI). The work has been criticised for its methodology (Ailon, 2008; Venaik and Brewer, 2013), and for not being up-to-date with major economical and societal changes of the past decades

(Kirkman *et al.*, 2006; McSweeney, 2002). This critique led to an expansion and replication in other countries and with different samples (consumers, students and other professionals). Hofstede *et al.* (2010) updated the cultural dimensions. Our approach additionally includes the dimensions of long-term orientation vs short-term normative values (LTO) and indulgence vs restraint (ING). Both dimensions became relevant when Hofstede analysed a larger variety of countries and regions which represent the included nations of our work.

Research on culture-bound response styles is scarce. There are several other studies, which relate ARS to national-level indicators, including the rate of psychiatric disorders (Fischer *et al.*, 2009), types of communication style (Smith, 2011), socioeconomic development and the rate of religious people (He *et al.*, 2014). However, these studies do not model culture. In the following, we focus on studies, which explicitly consider cultural profiles (Table I). In Baumgartner and Steenkamp's (2001) fundamental work, response style variance is small across countries in comparison to the variance they exert on different scales. Other researchers attempt to prove the contrary (Smith, 2004). It is noteworthy that even though ARS measures differ, the basic findings are largely consistent (Table I). Hofstede's dimension of IDV is negatively related to ARS across all studies (e.g. Harzing *et al.*, 2006; Smith, 2004). The results for PDI are inconsistent. Most studies found a positive effect (e.g. Harzing *et al.*, 2006; Hoffmann *et al.*, 2013). In contrast, Johnson *et al.* (2005) measured acquiescence on an individual-level and revealed a negative relationship. The authors suggest that the relationship reverses when the ARS measure was aggregated per country and correlated with PDI country scores. The negative relationship might probably be due to the included control variable "gross national product" (Baumgartner and Weijters, 2015). These inconsistent findings call for a multi-level analysis. Johnson *et al.* (2005) were the only researchers to confirm a negative relation to UAI and MAS. Remarkably, researchers yet neither examined the cultural dimensions of LTO and ING nor did they consider cultural interactions. As detailed below, we expect that LTO might relate to impulsiveness and thus in turn to more affirmative answers. We also expect that being restraint may reduce the likelihood to answer in a highly affirmative way. In sum, we believe that including these cultural dimensions can help gain a better understanding of the cultural underpinning of ARS.

3.3 Assessing the nomological validity of pARS by Hofstede's original four dimensions

Prior works tested the above described cultural and national-level predictors of individual ARS with cARS. First of all, we will theoretically derive and test the relations of pARS to Hofstede's (1991, 1994) original four dimensions. This will establish the nomological validity for the new measure and thereby answer RQ2.

3.3.1 Power distance. Research on the relationship between PDI and ARS is inconsistent. While Johnson *et al.* (2005) find a negative correlation, most other studies report a positive relationship (Smith, 2004; Harzing *et al.*, 2006; Hoffmann *et al.*, 2013). The latter finding is plausible from a theoretical point of view, and we also expect a positive relationship. Individuals of countries high in PDI expect and accept hierarchy and an unbalanced distribution of power (Hofstede, 1994, 2001). The members of these societies behave deferential and fear the consequences of opposing the authorities. Submissiveness results in anxiety to reveal true opinions (true variance) in the surveys distributed by institutional or official authorities. This leads to stronger-biased affirmative responding (ARS). We expect for PDI to be positively related to ARS.

3.3.2 Individualism/collectivism. Previous research consistently supports a negative relationship for IDV (e.g. Harzing *et al.*, 2006; Smith, 2004; Van Herk *et al.*, 2004). This finding is consistent with theoretical considerations. Members of societies high in IDV strongly express their opinions, in terms of either strongly agreeing or disagreeing. They do not

consider consequences to others. Conversely, members of collectivistic cultures highly value the interests of the in-group and the harmony within the group (Triandis *et al.*, 2001). Collectivistic individuals subordinate individual beliefs which in turn steers in-group harmony (Hofstede, 1991). We presume that IDV is negatively related to ARS.

3.3.3 Masculinity/femininity. While Johnson *et al.* (2005) report a negative influence of MAS, most studies report insignificant results (e.g. Smith, 2004). Societies high in MAS are assertive on decisions and content-independent statement affirmation is low. Instead, societies high in feminine values are modest and cooperative, which leads to neither stylistic affirmative nor dishonest answering. Both society types tend to answer less stylistically affirmative, but with different motivations. The rectified effects leave us to expect no significant relationship between MAS and acquiescence.

3.3.4 Uncertainty avoidance. While Johnson *et al.* (2005) find a negative influence of UAI, most studies fail to detect a significant relationship (Harzing *et al.*, 2006; Hoffmann *et al.*, 2013). Yet, societies that are characterised by high UAI exhibit an anxiety of future consequences and ambiguity. Similar to the effect of PDI, we would expect that an aversion to prospective consequences would result in higher stylistic affirmation. Conversely, subjects who avoid ambiguity are expected to decide for the safe choice of less statement affirmation. The antagonistic effects make us expect no significant relationship.

3.4 Expectations about the two neglected cultural dimensions

Previous empirical works did not examine LTO and ING. In search for holistic cultural considerations and in order to answer RQ3, we will discuss the newer dimensions.

3.4.1 Short-term/long-term orientation. Empirical ARS works did not examine LTO. Nations high in LTO set their priorities with regard to the future (Hofstede, 1994; Hofstede *et al.*, 2010). They invest today to drive change for a different tomorrow and are aware of the prospective impact of current behaviour. Biased responses might have prospective consequences. Thus, long-term oriented societies consciously affirm statements of which they are certain and believe in instead of answering impulsively without reflection. Couch and Keniston (1960) found a relationship between impulsiveness and ARS on the individual level, and this may raise the question as to whether the same relationship can also be found on the country level. Given all these arguments, we expect that LTO is negatively related to acquiescence.

3.4.2 Indulgence. The neglected dimension of ING completes the cultural dimensions. We do not expect any influence on ARS. Societies high in ING expect and accept individuals to follow their human drives which relate to enjoying life and having fun (Hofstede *et al.*, 2010). We expect that the participation in social life and having control over how to create a lifestyle does not affect whether the responses to a questionnaire are in a biased/unbiased manner. There is supposedly no influence of expressing joy of life on acquiescence. We expect that in nations where consumers are generally more restrictive in their behaviour, there is usually a lower tendency to answer in a largely affirmative and biased way.

3.5 Full cultural profiles and interaction effects

A simultaneous consideration of all six dimensions at a time will answer RQ3. Finally, we expect that cultural dimensions interact in their effect on ARS (RQ4). Given the wide set of possible interactions, we suggest specific expectations for effects with theoretic reasoning.

3.5.1 Interaction effects of power distance and uncertainty avoidance. We expect that UAI is a relevant moderator. As outlined above, UAI may have ambivalent effects. It might depend on the combination with other cultural dimensions whether UAI amplifies or decreases ARS. In high PDI societies, individuals obey to authorities and affirm free of content. Conversely, individuals in high UAI societies avoid ambiguous answers, affirm less

and respond in a less biased manner. In societies of high PDI and UAI, individuals reduce their perceived uncertainties by searching orientation in authorities, which consequently amplifies affirmation to authority distributed surveys. In sum, we expect that PDI and UAI have a joint positive effect on acquiescence.

3.5.2 Interaction effect of individualism/collectivism and uncertainty avoidance. We presumed that IDV is negatively related to ARS. Individualistic cultures are freely and truthfully expressing their opinions without concerns about others. High UAI societies withdraw from ambivalent situation and answers. Individualistic societies express opinions with either unbiased statement agreement or disagreement. In combination with high UAI, those cultures prefer reserved responding which is safe and less ambiguous. Nevertheless, content-independent biases decrease to a minimum in societies high in IDV and UAI. We suggest that IDV and UAI have a joint negative effect on acquiescence.

4. Design

4.1 Individual-level data

4.1.1 Sample. We use data on the individual-level of a large-scale cross-cultural study, including nations from seven regions worldwide: 14 European, 3 North American, 2 South American, 6 Asian, 3 Near Eastern and 1 African country, as well as Australia. In total, 236,089 consumers from 30 countries participated between 2011 and 2012. The participants are members of large online panels. The selection criterion for the present automotive study is to own a new private car which is less than five years registered. Participants were recruited based on their purchased brand in order to represent market shares. The mean sample size per country is 7,616. On average, 54 per cent of the respondents are male with a mean age of 40.4 years. The respondents are adults of all ages, ranging from 18 (some countries 16) to 90 years. See Table II for sample descriptives.

4.1.2 Questionnaire design. The present study is embedded in a larger survey focussing on the preference of and loyalty to automotive brands, brand images and product satisfaction. It took approximately 45 min to respond to the 65 items. The questionnaire covered a range of different measurement scales. The respondents answered the questionnaire in their native language which probes more correct answers (Harzing *et al.*, 2005). We translated the English master into the official language(s) of each country using the process of translation/backtranslation (Brislin, 1970; Malhotra *et al.*, 1996).

4.2 Measures of AR and pARS

4.2.1 Rationale. The conventional approaches to measure ARS disentangle true and stylistic affirmative answers by either using negatively and positively worded items or averaging the affirmation to a large number of content-independent items. On the other hand, our new practice-oriented measure pARS includes a test of truth which compares subjective scale affirmation with an objectively judged measure. We build this approach on Paulhus *et al.*'s (2003) idea of over-claiming measures, but we include only questions about existing objects and compare the answers with explicit knowledge measures as controls. We propose a measure of content-independent affirmation. Answers to the first question (Q1) indicate on a Likert scale AR to content-wise dependent items (AR) which consists of true and stylistic variance. We calculate an index of AR on a set of five Q1-items similar to the weighted average approach. We add questions which measure true variance (Q2) and calculate an index. The deviation between both indices feeds into pARS. Figure 2 shows the items used for both indices. The respondents need about one minute to answer the two questions.

The respondents answered Q1 in a block by rating their familiarity with five automotive sub-brands on unidimensional five-point scales ranging from "1 – I know it very well" to "5 – I have never heard of it". Of course, it is also feasible to adjust this procedure to other

Table II.
Descriptive statistics

Country	<i>n</i>	Age		Sex % of male	PDI	IDV	Culture ^a			ING	AR	Acquiescence		
		<i>M</i>	<i>SD</i>				MAS	UAI	LTO			pARS	pARS-1	pARS-2
Total	236,089	40.4	13.5	54	58	53	52	65	52	52	0.47	0.19	0.11	0.06
<i>Europe</i>														
Austria	4,392	41.4	13.6	52.1	11	55	79	70	60	63	0.45	0.06	-0.05	-0.07
Belgium	8,329	43.5	13.7	51.9	65	75	54	94	82	57	0.45	0.13	0.01	0.00
France	8,286	45.0	12.7	43.2	68	71	43	86	63	48	0.44	0.18	0.09	0.04
Germany	8,483	42.1	12.9	50.8	35	67	66	65	83	40	0.46	0.08	-0.01	-0.04
Italy	8,388	38.7	11.7	51.5	50	76	70	75	61	30	0.49	0.22	0.11	0.11
The Netherlands	12,266	47.4	14.4	51.2	38	80	14	53	67	68	0.41	0.12	0.06	-0.03
Norway	4,191	52.8	14.0	68.7	31	69	8	50	35	55	0.45	0.13	-0.01	-0.03
Poland	4,271	37.6	11.9	54.3	68	60	64	93	38	29	0.46	0.19	0.16	0.07
Portugal	6,188	38.7	11.6	60.8	63	27	31	99	28	33	0.47	0.19	0.10	0.07
Russia	20,393	33.3	9.3	52.0	93	39	36	95	81	20	0.46	0.29	0.23	0.16
Spain	8,339	38.3	10.7	50.0	57	51	42	86	48	44	0.45	0.17	0.09	0.05
Sweden	8,376	52.3	14.2	56.9	31	71	5	29	53	78	0.44	0.12	-0.01	-0.01
Switzerland	4,543	41.1	13.3	46.6	34	68	70	58	74	66	0.43	0.09	0.02	-0.05
UK	9,497	46.3	14.7	45.6	35	89	66	35	51	69	0.43	0.15	0.09	0.00
<i>North America</i>														
Canada	4,796	48.1	15.1	42.2	39	80	52	48	36	68	0.40	0.21	0.19	0.06
Mexico	8,698	37.7	11.6	58.8	81	30	69	82	24	97	0.45	0.22	0.22	0.13
USA	16,327	47.7	15.1	42.4	40	91	62	46	26	68	0.43	0.23	0.19	0.09
<i>South America</i>														
Argentina	8,554	38.5	11.4	56.0	49	46	56	86	20	62	0.43	0.17	0.18	0.09
Brazil	8,218	38.5	12.0	54.2	69	38	49	76	44	59	0.45	0.18	0.15	0.08
<i>Asia</i>														
China	22,609	32.5	7.9	61.9	80	20	66	30	87	24	0.57	0.15	0.01	0.06
India	8,548	32.9	9.9	69.5	77	48	56	40	51	26	0.60	0.42	0.33	0.30
Japan	8,180	44.8	10.7	62.2	54	46	95	92	88	42	0.40	0.05	0.01	-0.09
Malaysia	4,481	32.6	9.1	52.9	100	26	50	36	41	57	0.51	0.18	0.07	0.05
Singapore	3,922	35.5	9.9	51.9	74	20	48	8	72	46	0.51	0.24	0.14	0.12
South Korea	3,069	36.2	9.7	55.3	60	18	39	85	100	29	0.46	0.26	0.19	0.15

(continued)

Country	<i>n</i>	Age		Sex % of male	PDI	Culture ^a			ING	AR	Accquiescence			
		<i>M</i>	<i>SD</i>			MAS	UAI	LTO			pARS	pARS-1	pARS-2	
<i>Near East</i>														
Kingdom of Saudi Arabia	1,276	30.7	9.1	84.4	95	25	60	80	36	52	0.52	0.43	0.37	0.24
Turkey	5,845	33.4	8.5	70.9	66	37	45	85	46	49	0.52	0.19	0.06	0.07
United Arab Emirates	1,905	33.4	8.8	66.8	90	25	50	80	36	52	0.53	0.28	0.18	0.14
<i>Africa</i>														
South Africa	5,336	38.2	12.0	45.7	49	65	63	49	34	63	0.48	0.14	0.02	0.01
<i>Australia</i>														
Australia	8,383	47.7	14.7	42.8	36	90	61	51	21	71	0.40	0.10	0.07	-0.06

Notes: PDI, power distance; IDV, individualism/collectivism; MAS, masculinity/femininity; UAI, uncertainty avoidance; LTO, long-term/short-term orientation; ING, indulgence/restraint. ^aCultural indices taken from Hofstede *et al.* (2010)

A practice-oriented
accquiescence
measure

Table II.

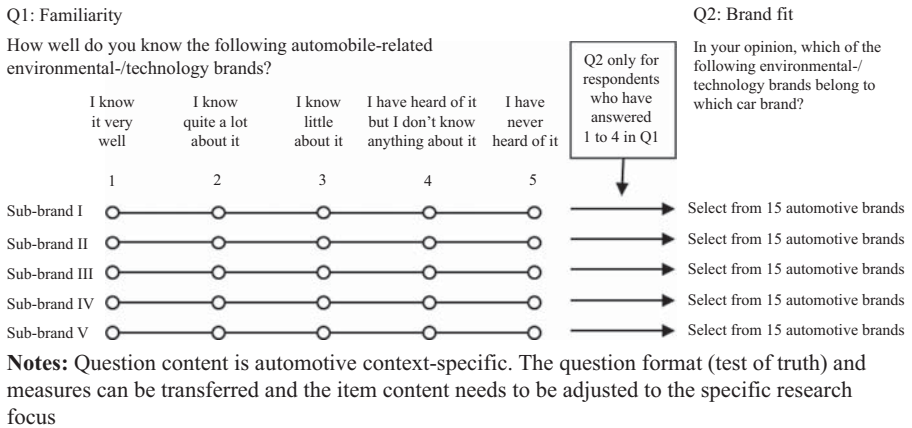


Figure 2. Testing truth via a combination of two questions

scaling formats (e.g. seven-point rating scales). Then, the subjects answered Q2 for each sub-brand, which they scored 1–4 (but not 5) in Q1. They had to select in Q2 the corresponding parent brand from a list of 15 country-specific automotive makes. Sub-brands are the same for all 30 countries and do not name-wise relate to the automotive makes, for instance the sub-brands “Hybrid Synergy Drive” or “EfficientDynamics” belong to Toyota and BMW, respectively. The test’s probability of failure, which means guessing the right brand from the list, is 6.7 per cent (1/15). The observed affirmative answers (AR) in Q1 mix up true and stylistic variances. We can measure this mix via the Q1-index. In order to estimate pure stylistic variance, we have to untangle from the mix true answers which are measured by the Q2-index. The design discriminates if the respondent answered sincerely in Q1 or if he/she agrees without knowledge and it is the rational for pARS. Different variations for calculation will ensure stability. All measures will be calculated for each respondent.

4.2.2 Transfer of pARS measure to other business contexts. The test of truth is a measure of ARS in all surveys in which the respondents have knowledge about the survey object, e.g. in consumer, employee or corporate partner studies. Respondents have to indicate in Q1 familiarity with the specific knowledge object. The specification option of non-knowledge is necessary to avoid uninformed responses (Hawkins and Coney, 1981). Research might provide answering options for the objective knowledge question, which is Q2 in the present case, or leave it as an open-ended question (Malhotra *et al.*, 1996). The decision depends on the difficulty of task and resources for coding answers. In difficult tasks, aided answers are advised, but they should contain a wide range of options to minimise the *a priori* probability of right guessing. Researchers have to manually code open-ended questions due to writing mistakes which the software will neglect.

4.2.3 Affirmative responses (AR). AR measures statement affirmation. It is a mix up of stylistic and true variability. We will use AR to discriminate it to pARS. The latter entangles the variance components. We use the weighted average approach in order to measure AR. This approach has been previously applied to content-independent items to measure cARS (Baumgartner and Steenkamp, 2001; Hoffmann *et al.*, 2013). We take five content-dependent items (Q1) to measure AR and assign higher weights for more extreme affirmation. It can be formally described as:

$$AR = \frac{\sum_{x=1}^5 f(x) \times w_x}{n}, \quad (1)$$

where $f(x)$ is the number of ticks of scale point x of the rating scale, w_x is the weight for scale point x and n is the total number of answers scored 1–4 in Q1. The weights are defined by $w_x = (k-x)/(k-1)$, where k is the total number of scale points and $k-1$ is the number of scale points (1–4) which lead to Q2.

4.2.4 *Practice-oriented measure of acquiescence response style (pARS)*. A clearance of AR by the true know-how (Q2) leads to the stylistic variance. Formally, pARS is a correction of AR by the proportion of rightly assigned brands:

$$pARS = \left[\frac{\sum_{x=1}^5 f(x) \times w_x}{n} \right] \frac{f(y)}{n}, \quad (2)$$

with $f(y)$ as the number of rightly assigned car brands to sub-brands in Q2. A value of $pARS = "1"$ indicates the highest stylistic responding.

4.2.5 *Alternative approaches to pARS*. We calculate two more alternatives for pARS to support robustness. Alternative 1 compares the (weighted) Q1 scores for falsely assigned brands (Q2) with the ones for truly assigned brands. Respondents, who score higher on the familiarity scale for wrongly assigned brands than for rightly assigned brands, respond content-independent and biased. Formally, this index is calculated as follows:

$$pARS-1 = \left[\frac{\sum_{x=1}^5 f_{false}(x) \times w_x}{n_{false}} \right] - \left[\frac{\sum_{x=1}^5 f_{true}(x) \times w_x}{n_{true}} \right], \quad (3)$$

where $f_{false}(x)/f_{true}(x)$ is the number of ticks of scale point x , for all falsely/truly assigned brands in Q2, and $n_{false}(x)/n_{true}(x)$ is the number of false/true assignments in Q2.

The second alternative is based on the probability theory. Each scale point has a probability for brand assignment failure and the complementary probability for success. For instance, we expect that a respondent, who ticks scale point "2" in Q1, has a probability of 75 per cent to assign the right brand in Q2 and a probability of 25 per cent to fail. Note that the *a priori* probability of guessing the right answer is 6.7 per cent as we presented 15 options. In case of biased and content-independent answers, the complementary probability, which is the probability of non-occurrence of the expected event, is true. For scale point "2", the complementary probability for wrong brand assignment would be 23.33 per cent [= $(1-0.75 \times (1.00-0.067))$]. Respondents, who respond to Q1 in a biased way, have a 69.98 per cent probability to tick scale point "2" even if they do not know the brand/sub-brand. Alternatively, two compares the survival function for false with true brand assignment as follows:

$$pARS-2 = \left[\frac{\sum_{x=1}^5 \overline{Pr}_{false}(x) \times f_{false}(x)}{n_{false}} \right] - \left[\frac{\sum_{x=1}^5 \overline{Pr}_{true}(x) \times f_{true}(x)}{n_{true}} \right], \quad (4)$$

where $\overline{Pr}_{false}(x)/\overline{Pr}_{true}(x)$ is the complementary probability (survival function) for scale point x for false/true brand assignment. The alternative indices range from -1 to 1. "-1" indicates low affirmation while all brands are rightly assigned. "0" points to either high agreement with all brands rightly assigned or low agreement with all brands wrongly assigned. "1" indicates the highest familiarity scores paired with all brands wrongly assigned.

4.2.6 *Structure Q1*. Finally, we calculated an index for heterogeneity within the individual Q1 answers. The index represents the average score deviation between each of the five sub-brand familiarity scores and ranges from 0 to 1. "0" indicates that the respondent ticks the same scale point for all five items, and "1" reflects extreme diverse ratings.

4.2.7 *Control for other types of measurement errors.* We controlled for uninformed responses and average scale usage with the following measures. Uninformed responses can be reduced by including the answering option “don’t know” (Hawkins and Coney, 1981). The familiarity scale in Q1 comprises such a scale point. The variable of “Structure Q1” controls for heterogeneity in average scale usage, but it does not indicate any specific scale usage pattern, e.g. mid-point or extreme ends.

4.3 *Validity checks*

We test the validity of the new pARS measure by comparing it to its alternatives and contrasting it to the measure of AR. Table III depicts the results. As expected, pARS strongly correlates with its two alternatives, pARS-1 ($r = 0.924$) and pARS-2 ($r = 0.901$), indicating the robustness of pARS. AR, the mix up measure of true and stylistic variance, and pARS, the measure of purely stylistic affirmation, are positively correlated but share less variability ($r = 0.599$).

4.4 *Data on the national-level: cultural indices*

We took cultural indices from Hofstede *et al.* (2010) (see Table II). These indices are from data collected in a comparable time period as our data, which prevents distortion due to time lags. We matched the two data sets country-wise. The integration of a second data set prevents common method bias and makes conclusions more valid (Kirkman *et al.*, 2006; Podsakoff *et al.*, 2003).

5. Results

We use a set of statistical methods to answer our research questions. Correlations analysis discriminates the substantive measure of pARS from the mix-up measure of AR. The correlations between national-level pARS and Hofstede’s (1991) original four cultural dimensions further replicate existing intercultural results and ensure nomological validity. This establishes pARS as an alternative measure for acquiescence (RQ1) and hints to construct validity based on the relations to cultural dimensions (RQ2). Two more analyses validate the holistic cultural effects. Ordinal least square regression simultaneously tests on a national-level all six dimensions and investigates in successive models interaction terms also (RQ3 and RQ4). Multi-level analysis confirms the stability of national culture-bound influences on individual pARS under the condition of individual- and country-level variability.

Table III.
Correlations of
different measures of
affirmative
responding on the
country level

	Acquiescence response style			Affirmative responses
	pARS	pARS-1	pARS-2	AR
<i>Acquiescence measures</i>				
pARS-1	0.92***			
pARS-2	0.97***	0.90***		
AR	0.60***	0.32***	0.65***	
<i>Cultural indicators</i>				
Individualism	-0.42*	-0.32****	-0.51**	-0.59***
Power distance	0.65***	0.58***	0.69***	0.59***
Masculinity	-0.09	0.03	-0.06	0.01
Uncertainty avoidance	0.05	0.15	0.07	-0.19
Long-term orientation	-0.18	-0.29	-0.15	0.08
Indulgence	-0.31	-0.16	-0.36*	-0.50**

Notes: $n = 30$. Pearson product moment correlation. * $p \leq 0.05$; ** $p \leq 0.01$; *** $p \leq 0.001$; **** $p \leq 0.1$

5.1 National differences in affirmative responding

We start our analyses on the national level. ANOVA reveals that the 30 aggregated national means of AR and pARS vary significantly across countries [AR: $F(29, 18.99) = 473.400, p \leq 0.001$, pARS: $F(29, 42.356) = 327.192, p \leq 0.001$]. See Table II for the descriptive statistics of the country means.

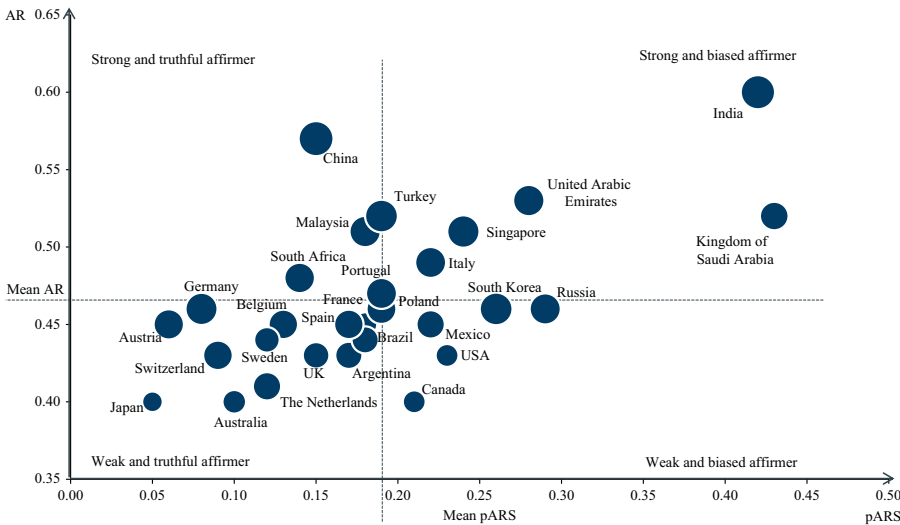
The aggregated means of AR and pARS are positively related ($r_{AR,pARS} = 0.599$), indicating that deviations from the “true” values (pARS) tend to raise with the level of affirmation (AR). Figure 3 visualises the relationship. The upper right corner consists of countries that are characterised by strong affirmation, but biased answers (strong and biased affirmer, e.g. India, Kingdom of Saudi Arabia). Those countries in which affirmative answers in Q1 are not validated in Q2 display the original concept of content-independent affirmation. In contrast, countries located in the lower left corner affirm less and their answers in Q2 truthfully confirm low knowledge (weak and truthful affirmer, e.g. Japan, Australia, Switzerland).

While the countries in the lower left and upper right corner reflect a positive relationship between the level of affirmation and the deviation from true answers, there are also nations that illustrate the imperfection of the relationship and the need for pARS measures. Some countries with higher agreement in Q1 display only a low discrepancy to the true Q2-score (strong and truthful affirmer, e.g. China). Hence, high agreement does not necessarily reflect content-independent, stylistic answering patterns. In particular, North American countries tend to affirm less although their score for stylistic variance is above average (weak and biased affirmer, e.g. Mexico, Canada and the USA).

These findings demonstrate that AR mixes up true (content-dependent) and stylistic (content-independent) variance in affirmations. The new measure of pARS overcomes this limitation and discriminates biased from unbiased responding.

5.2 Cultural influences

5.2.1 Main effects. We now examine how culture affects pARS. First, we replicate previous works by correlating different measures with each cultural dimension (Table III, lower panel).



Note: The bubble size indicates the percentage of answers affirming familiarity (number of answers 1–4 in Q1) with bigger bubbles reflecting more affirmation

Figure 3. Scatterplot of AR and pARS on the country-level

The subsequent Pearson correlation and regression analyses are based on the aggregated mean scores per country reducing the sample to 30 countries. Consistently, we found pARS to significantly vary as a function of IDV and PDI in correlation analyses. Acquiescence is negatively correlated to IDV (except for pARS-1) and positively correlated to PDI. Concerning the two newer cultural dimensions, only pARS-2 and ING correlate. The correlations neglect a control for individual variables as well as interactions between the cultural dimensions. We perform OLS regressions to test the effect stability for the simultaneously entered cultural dimensions. Table IV indicates several OLS models. Model 1 shows that for the 30 country scores pARS primarily relates to PDI, but no longer to IDV. Bootstrapping analysis confirms that these findings are stable (Table IV). In sum, the regression analysis shows that direct effects decrease or disappear when considering the full cultural profile.

5.2.2 Interaction effects. In subsequent regression models, we tested two-way interaction effects (Table IV, right side). We standardized all predictors before calculating the interaction term. Model 1 indicates for the 30 country scores a positive main effect of PDI on pARS. In addition, PDI tends to interact with UAI (see Model 2). Yet, the effect is only marginally significant ($p = 0.056$), which may be caused by a lack of power due to the small sample size ($n = 30$). Therefore, we use spotlight analysis to elaborate the interaction effect (Spiller *et al.*, 2013). The analysis confirms conditional effects for medium and high levels of UAI (UAI: -1 SD from mean: $\beta = 0.014$, $p = 0.62$; 0 SD: $\beta = 0.052$, $p = 0.01$; $+1$ SD: $\beta = 0.091$, $p = 0.00$). The floodlight analysis (Spiller *et al.*, 2013) reveals a Johnson–Neyman point for $p \leq 0.05$ at a value of -0.302 , indicating that when the standardised value of UAI is -0.302 or higher, the influence of PDI on pARS is significant. Individualism does not exert a main effect (see Model 1), but significantly interacts with UAI (see Model 3). Both interactions are visualised in Figure 4 by a simple slope one standard deviation above and below. The combination of high UAI and either low PDI or high IDV decreases pARS. The correlation of UAI with both dimensions is not surprising as PDI and IDV are known to be highly correlated (Hofstede, 1991).

5.3 Multi-level analysis

We apply hierarchical linear modelling to test the stability of cultural effects in a full model of individual-level and all six country-level indicators. Multi-level analysis is most appropriate for cross-cultural research (Chabowski *et al.*, 2016). Multi-level analysis simultaneously estimates parameters at all levels and avoids biases in parameters (Raudenbush and Bryk, 2002). We excluded the interaction terms to avoid overfitting and to support parsimonious models. Raudenbush and Bryk (2002) suggest limiting the number of country-level variables in case of few research entities ($n = 30$). Unlike in OLS regression, ARS scores are on an individual-level and not aggregated per country.

5.3.1 Fixed effects. The upper panel of Table V demonstrates that pARS increases with PDI and decreases for LTO. The two alternative measures confirm the effects which are in line with the expectations about LTO and prior investigations (PDI). As expected, ING does not relate to ARS. Note that the multi-level analysis controls for individual characteristics too and the findings are consistent with the country-level results reported by the OLS regression (Model 1).

Previous research investigated the relationship between person-related features and response tendencies with a focus on sociodemographic characteristics (Greenleaf, 1992). To control for the effects, we include age, gender, education and income in our model. Content-independent affirmative responding (pARS, -1 and -2) is higher for younger respondents. There is no relationship to gender, and pARS decreases with the education and income level. The two alternatives for pARS support the negative effect of education which is consistent with previous research (Greenleaf, 1992; Winkler *et al.*, 1982).

Predictor	Model 1			Model 2			Model 3			
	β	t		β	t		β	t		
Power distance (PDI)	0.62	2.89***	0.03	0.10	0.10	0.03	0.10	0.10	0.04	0.12
Individualism (IDV)	0.06	0.30	-0.02	0.04	0.03	-0.04	0.03	0.03	-0.03	0.04
Masculinity (MAS)	-0.06	-0.44	-0.03	0.01	0.03	-0.02	0.03	0.00	-0.02	0.02
Uncertainty avoidance (UAI)	-0.12	-0.80	-0.05	0.02	0.00	-0.06	0.00	0.00	-0.07	0.00
Long-term orientation (LTO)	-0.30	-1.77*	-0.06	0.00	0.00	-0.06	0.00	-0.20	-0.05	0.01
Indulgence (IDG)	-0.28	-1.51	-0.06	0.01	0.00	-0.07	0.00	-0.33	-0.07	0.00
PDI × UAI										
IDV × UAI										
F		4.19***								
R^2		0.52								
R^2_{adj}		0.40								
ΔR^2										

Notes: $n = 30$. DV, practice-oriented acquiescence style (pARS); Bootstrapping (1,000 samples; CI, confidence interval 90%), LL, lower level of CI; UL, upper level of CI. OLS regressions: standardised regression coefficients. Cultural indicators are standardised. * $p < 0.10$; ** $p < 0.05$; *** $p < 0.01$

Table IV.
Linear regression for
pARS (country-level)

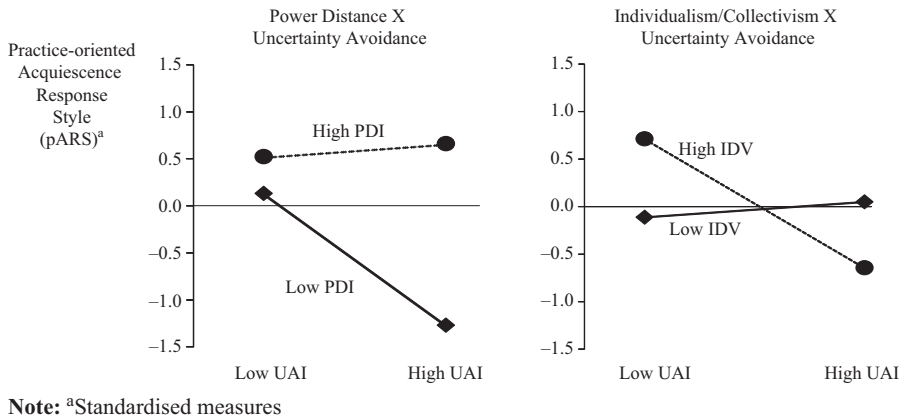


Figure 4. Simple slopes of interaction effects (country-level)

Note: ^aStandardised measures

	pARS			pARS-1		pARS-2	
Fixed effects	Coefficient (SE)	t-ratio	p	t-ratio	p	t-ratio	p
Intercept	0.2039 (0.0872)	2,338	*	0.520		1,266	
<i>Person-level predictors</i>							
Gender (male)	-0.0019 (0.0053)	-0.357		-7,175	***	4,010	***
Age	-0.0008 (0.0003)	-2,361	*	-1,106		-5,078	***
Education	-0.0101 (0.0021)	-4,799	***	-7,777	***	-1,990	*
Income	-0.0089 (0.0024)	-3,769	***	-3,891	***	-0.728	
Structure Q1	0.0118 (0.0068)	1,729		-4,124	***	-6,003	***
<i>Country-level predictors</i>							
Power distance	0.0024 (0.0007)	3,285	**	2,715	**	3,117	**
Individualism	0.0003 (0.0006)	0.493		0.397		-0.035	
Masculinity	-0.0004 (0.0004)	-0.900		0.156		-0.735	
Uncertainty avoidance	-0.0004 (0.0006)	-0.752		0.235		-0.794	
Long-term orientation	-0.0012 (0.0006)	-2,092	*	-2,164	*	-2,049	*
Indulgence	-0.0013 (0.0008)	-1,554		-0.678		-1,480	
<i>Random effects</i>							
Person-level variance	Variance components (df)	χ^2	p	χ^2	p	χ^2	p
	0.12744		***				
<i>Country-level variance (between country)</i>							
Average response style	0.00441 (23)	2,889.75	***	5,788.24	***	3,327.91	***
Effect of gender (male)	0.00088 (29)	255.35	***	291.17	***	252.58	***
Effect of age	0.00000 (29)	391.64	***	323.95	***	326.03	***
Effect of education	0.00015 (29)	184.31	***	195.69	***	167.25	***
Effect of income	0.00017 (29)	214.33	***	236.46	***	195.03	***
Effect of structure Q1	0.00137 (29)	405.84	***	5,361.54	***	963.74	***

Notes: n = 236,089. Gender, age and structure Q1 are grand mean centred; education and income are standardised per country. *p<0.05; **p<0.01; ***p<0.001

Table V. Effects of national and individual indicators on pARS

In summary, the multi-level analysis confirms our expectation about the, before neglected, cultural influence of LTO. Consistent with previous investigations, we could replicate the relationship of ARS to PDI and the insignificant effects of MAS and UAI. Against our expectations, we found no support for the influence of IDV. Even though isolated analyses confirm the predicted relationship, it disappears when PDI is included. In line with RQ4, this indicates that holistic cultural profiles more realistically predict ARS.

5.3.2 Random effects. Individual-level predictors vary across nations as do the individual indices of pARS (Table V). We conclude that highly significant differences exist among the country means of affirmative responding, gender, age education, income and structure in Q1.

5.3.3 Model fit and explained variance. The comparison of the χ^2 test statistics, for a simple and for a full model, points to model fit (Raudenbush and Bryk, 2002). Compared to the zero model, in which the individual-level variables are constrained to zero, the difference in χ^2 s is statistically significant for pARS ($\Delta\chi^2 = 13.563$, $df = 20$, $p < 0.001$) and for the two alternative measures (pARS-1: $\Delta\chi^2 = 23.566$, pARS-2: $\Delta\chi^2 = 41.788$, $df = 20$, $p < 0.001$). The model including country- and individual-level variables improves the explanation of underlying data. The intra-class coefficient (ICC) indicates how much shared variance is at each level. The respondent's culture accounts for 5.9 per cent of pARS, for 5.1 per cent of pARS-1 and for 6.4 per cent of pARS-2. The remaining variance is at the individual level (e.g. pARS: $100 - 5.1 = 94.1$ per cent). A comparison of the full intercept- and slopes-as-outcomes model to various unconditional models points to partial explanatory variances (Raudenbush and Bryk, 2002). All six cultural dimensions explain 38.2 per cent of the between-country variance of pARS (5.9 per cent) and 28.0/40.3 per cent of pARS-1/-2. The individual-level indicators explain only 0.8 per cent of the individual variance of pARS (94.1 per cent) and 3.6/0.8 per cent of pARS-1/-2.

6. Discussion

This paper investigated four research questions which contribute to the literature. *RQ1* investigates an alternative ARS measure which is suitable for application in practical research and contributes methodological-wise. We introduce an alternative measure, namely, pARS, which discriminates between stylistic and true variance. With respect to *RQ2*, the present study replicated the pattern of cultural influences of Hofstede's (1991) original four dimensions, shown in previous studies to predict ARS. This nomological network establishes construct validity for pARS. pARS is particularly suitable for shorter questionnaires. In contrast to prior approaches, it does not add non-substantive items. To measure pARS, only a small set of substantive items is needed (e.g. product knowledge). Therefore, pARS is appropriate for practice-oriented research. Unlike previous research, pARS used neither Likert-typed scaling nor attitudinal items to measure pure affirmative vs profound stylistic responding (Greenleaf, 1992; Baumgartner and Steenkamp, 2001). For the first time, we apply knowledge-based familiarity scales, which are the basis for a test of truth. Therefore, our research design enables us to contrast the stylistic and true parts in affirmative responding. Note that the familiarity measures are flexible in such a way that the measures could apply Likert scales with different numbers of points or with binary answers (yes or no).

RQ3 and *RQ4* investigate the relations of ARS to holistic cultural profiles and contribute by theoretical development. *RQ3* tests for the superiority of including all dimensions at a time to account for holistic cultural profiles (Richter *et al.*, 2016). *RQ4* seeks for interaction terms between cultural dimensions. The empirical results highlight the importance of including multiple cultural predictors and interaction terms rather than only main effects of single dimensions. Our study investigated the relations to long-term orientation and indulgence which are the more current cultural dimensions of Hofstede (Hofstede *et al.*, 2010) and which were neglected in extant research. In the subsequent discussion, we focus on the theoretical implications and interpret the findings on how culture affects stylistic, affirmative answers. The large number of countries involved in the present study establishes cross-country generalisability of the culture-bound results.

In line with previous studies (Harzing *et al.*, 2006; Hoffmann *et al.*, 2013; Smith, 2004), Hofstede's dimension of PDI is the most relevant predictor of the new acquiescence measure.

Multi-level analysis confirms that, if controlled for individual-level (age, gender, education, income and structure Q1) and country-level effects (six cultural dimensions), PDI (along with LTO) mainly predicts ARS. The two alternative measures of pARS confirm robustness. In countries high in PDI, the powerless accept the unbalanced power distribution as well as authority-presented statements. They agree independent of item content and own opinions to authorities. The effect of PDI on free-of-topic affirmation interacts with UAI. In cultures characterised by high UAI and high PDI, the content-independent affirmation is the highest.

In isolated correlations, we found a consistent negative relationship between IDV and ARS (e.g. Hoffmann *et al.*, 2013; Smith, 2011). However, in holistic cultural profiles, the effect disappears in favour of a positive main effect of PDI. It is known that both dimensions share some common variability and are in a negative relationship (Hofstede, 1991). Accordingly, the effect of IDV reported in previous studies is overstated, and cross-cultural researchers should prefer statistical methods which integrate all predictors simultaneously over correlations. Remarkably, IDV interacts negatively with high UAI. Individualists freely express their opinions without concerns about others. Individualists in high uncertainty avoiding societies are more concerned about the consequences of their behaviours. Their responding is well-considered and less lavish. They answer less stylistically biased. However, the before neglected dimension of LTO negatively affects stylistic answers. Long-term oriented societies are aware of their future behavioural consequences and securely answer less biased (He *et al.*, 2014). The other cultural dimension disregarded by prior studies, indulgence, does not influence ARS.

The pattern of significant and non-significant correlations between pARS and Hofstede's four original cultural dimensions is identical to those of previous works using conventional ARS measures (e.g. Harzing *et al.*, 2006; Hoffmann *et al.*, 2013). This supports nomological validity for pARS.

In sum, the multi-level analysis demonstrates that national culture explains a small proportion of total variance in ARS (ICC: 5.1–6.4 per cent). While individual-level influences are stronger, culture still determines a substantial part of the (stylistic) variance in AR and these systematic biases need to be controlled (Singh, 1995). Other large-scale cross-cultural studies found a rather small portion of explained variance also (Gelfand *et al.*, 2011; Kirkman *et al.*, 2006). However, systematic cultural biases are of strong practical relevance. Scholars can feasibly predict country-based distortions on the basis of available cultural indices. Thus, researchers who conduct studies in countries high in PDI and UAI could and should be aware of stylistic bias. Additionally, we found on the individual level, that content-independent affirmative responding (pARS, -1 and -2) is higher for younger respondents. This finding might be due to a different age structure in the samples. Respondents of countries high in pARS have the lowest average ages, such as India and Saudi Arabia (Table IV). Hence, we call for future research that elaborates on the drivers of ARS at the individual level (e.g. age).

Independent of cross-country effects, researchers should note the impact of individual response styles. Practitioners and scholars striving to avoid the contamination should either adopt the research design (pre-data collection) or correct scales post-data collection (e.g. Welkenhuysen-Gybels *et al.*, 2003). Pre-data collection techniques are the application of ranking rather than rating scales (Fischer, 2004; Harzing *et al.*, 2009), interrogative scaling formats (Wong *et al.*, 2003), balanced constructs with positive and negative statements (Baumgartner and Steenkamp, 2001, 2006a; Podsakoff *et al.*, 2003), or eliminations of ARS inducing factors, such as vague or ambiguous wording (Paulhus, 1991).

7. Limitations and future research directions

We hope that this work stimulates further research to apply an adjusted form of the new measure. Additionally, like every empirical investigation, the present study has methodological limitations which could be addressed in future research.

While the new measure empirically differentiates substance from style through an objective test of truth, the line between substance and style is fuzzy from a conceptual point of view. Steenkamp, de Jong and Baumgartner (2010) argue that the “positive illusions”, which the respondent believes to be true, might be unbiased variance even though objective measures would categorise them as stylistic. This reflects the discussions about objective and subjective knowledge (Carlson *et al.*, 2009) and over-claiming (Paulhus *et al.*, 2003). Subjective knowledge reflects an overestimation of one’s know-how and is associated with higher levels of self-confidence (Raju, Lonial and Mangold, 1995). Over- or under-confidence occurs in know-how tasks which are of high difficulty (Dunning *et al.*, 1990). This does not apply for the presently used familiarity with sub-brands. Moreover, objective and subjective types of knowledge are essentially correlated, especially for durable products, such as cars (Carlson *et al.*, 2009). That means for the present work: if the respondent is convinced that he/she knows the sub-brand and affirms familiarity, but does fail to indicate the parent brand, his/her affirmation is specified as stylistic.

Moreover, this study tested and validated pARS in the car industry. The approach is very flexible and can be applied to any other field of studies, where the objective tests of the respondents’ knowledge can be feasibly applied. This includes other product categories (e.g. electronics, food, cosmetics etc.), other fields of business research (e.g. taxes), and even other fields of social science (e.g. politics, climate change). We call for further research to adjust and validate the new measure in different fields.

Finally, the present study is based on a uniquely large sample with more than 200,000 participants from 30 countries. Although, we could not include all nations, the cultural variations are quite large (PDI 11-100, IDV 18-91, MAS 5-95, UAI 8-94, LTO 20-100, ING 24-97). As we applied car ownership as the inclusion criteria, the selection might be a bit more restrictive in the less wealthy nations. Further studies may replicate our approach with large-scale data sets in another product category, which is less dependent on individual income, to overcome this limitation and increase generalisation. In a similar vein, our regression analysis at the country level is based on rather small samples. Although the robustness of these findings is supported by the multi-level analysis and by the bootstrapping analysis (see Table IV), we hope that further research will replicate our findings with large data sets on the country level.

References

- Ailon, G. (2008), “Mirror, mirror on the wall: culture’s consequences in a value test of its own design”, *The Academy of Management Review*, Vol. 33 No. 4, pp. 885-904.
- Bachman, J.G. and O’Malley, P.M. (1984), “Yea-saying, nay-saying, and going to extremes: black-white differences in response styles”, *Public Opinion Quarterly*, Vol. 48 No. 2, pp. 491-509.
- Baumgartner, H. and Steenkamp, J.-B.E.M. (2001), “Response styles in marketing research: a cross-national investigation”, *Journal of Marketing Research*, Vol. 38 No. 2, pp. 143-156.
- Baumgartner, H. and Steenkamp, J.-B.E.M. (2006a), “An extended paradigm for measurement analysis of marketing constructs applicable to panel data”, *Journal of Marketing Research*, Vol. 43 No. 3, pp. 431-442.
- Baumgartner, H. and Steenkamp, J.-B.E.M. (2006b), “Response biases in marketing research”, in Grover, R. and Vriens, M. (Eds), *The Handbook of Marketing Research – Uses, Misuses, and Future Advances*, Sage, Thousand Oaks, CA, pp. 95-109.
- Baumgartner, H. and Weijters, B. (2015), “Response biases in cross-cultural measurement”, in Ng, S. and Lee, A.Y. (Eds), *Handbook of Culture and Consumer Behavior*, Oxford University Press, Oxford, pp. 150-180.
- Bentler, P.M., Jackson, D.N. and Messick, S. (1971), “Identification of content and style: a two-dimensional interpretation of acquiescence”, *Psychological Bulletin*, Vol. 76 No. 3, pp. 186-204.

- Billiet, J.B. and McClendon, M.J. (2000), "Modeling acquiescence in measurement models for two balanced sets of items", *Structural Equation Modeling*, Vol. 7 No. 4, pp. 608-628.
- Brislin, R.W. (1970), "Back-translation for cross-cultural research", *Journal of Cross-Cultural Psychology*, Vol. 1 No. 3, pp. 185-216.
- Carlson, J.P., Vincent, L.H., Hardesty, D.M. and Bearden, W.O. (2009), "Objective and subjective knowledge relationships: a quantitative analysis of consumer research findings", *Journal of Consumer Research*, Vol. 35 No. 5, pp. 864-876.
- Chabowski, B.R., Saeed Samiee, S. and Hult, G.T.M. (2016), "Cross-national research and international business: an interdisciplinary path", *International Business Review*, Vol. 26 No. 1, pp. 89-101.
- Cheung, G.W. and Rensvold, R.B. (2000), "Assessing extreme and acquiescence response sets in cross-cultural research using structural equations modeling", *Journal of Cross-Cultural Psychology*, Vol. 31 No. 2, pp. 187-212.
- Clarke, I. III (2001), "Extreme response style in cross-cultural research", *International Marketing Review*, Vol. 18 No. 3, pp. 301-324.
- Couch, A. and Keniston, K. (1960), "Yeasayers and naysayers: agreeing response set as a personality variable", *Journal of Abnormal and Social Psychology*, Vol. 60 No. 2, pp. 151-172.
- de Jong, M.G., Steenkamp, J.-B.E.M. and Veldkamp, B.P. (2009), "A model for the construction of country-specific yet internationally comparable short-form marketing scales", *Marketing Science*, Vol. 28 No. 4, pp. 674-689.
- de Jong, M.G., Steenkamp, J.-B.E.M., Fox, J.P. and Baumgartner, H. (2008), "Using item response theory to measure extreme response style in marketing research: a global investigation", *Journal of Marketing Research*, Vol. 45 No. 1, pp. 104-115.
- Dolnicar, S. and Grun, B. (2007), "Cross-cultural differences in survey response patterns", *International Marketing Review*, Vol. 24 No. 2, pp. 127-143.
- Dunning, D., Griffin, D.W., Milojkovic, J.D. and Ross, L. (1990), "The overconfidence effect in social prediction", *Journal of Personality and Social Psychology*, Vol. 58 No. 4, pp. 568-581.
- Fischer, R. (2004), "Standardization to account for cross-cultural response bias: a classification of score adjustment procedures and review of research in JCCP", *Journal of Cross-Cultural Psychology*, Vol. 35 No. 3, pp. 263-282.
- Fischer, R., Fontaine, J., van de Vijver, F.J.R. and van Hemert, D.A. (2009), "An examination of acquiescent response styles in cross-cultural research", in Gari, A. and Mylonas, K. (Eds), *Quod Erat Demonstrandum: From Herodotus' Ethnographic Journeys to Cross-Cultural Research*, Pedio Books Publishing, Athens, pp. 137-147.
- Gelfand, M.J. et al. (2011), "Differences between tight and loose cultures: a 33-nation study", *Science*, Vol. 332 No. 6033, pp. 1100-1128.
- Greenleaf, E.A. (1992), "Improving rating scale measures by detecting and correcting bias components in some response styles", *Journal of Marketing Research*, Vol. 29 No. 2, pp. 176-188.
- Harzing, A.-W. et al. (2005), "Does the use of English-language questionnaires in cross-national research obscure national differences?", *International Journal of Cross Cultural Management*, Vol. 5 No. 2, pp. 213-224.
- Harzing, A.-W. et al. (2006), "Response style in cross-national survey research: a 26-country study", *International Journal of Cross-Cultural Management*, Vol. 6 No. 2, pp. 243-266.
- Harzing, A.-W. et al. (2009), "Rating versus ranking: what is the best way to reduce response and language bias in cross-national research?", *International Business Review*, Vol. 18 No. 4, pp. 417-432.
- Hawkins, D.I. and Coney, K.A. (1981), "Uninformed response error in survey research", *Journal of Marketing Research*, Vol. 18 No. 3, pp. 370-374.
- He, J., van de Vijver, F.J.R., Domínguez, A.d.C. and Mui, P.H.C. (2014), "Toward a unification of acquiescent, extreme, and midpoint response styles: a multilevel study", *International Journal of Cross-Cultural Management*, Vol. 14 No. 3, pp. 306-322.

-
- Hinkin, T.R. (1995), "A review of scale development practices in the study of organizations", *Journal of Management*, Vol. 21 No. 5, pp. 967-988.
- Hoffmann, S., Mai, R. and Cristescu, A. (2013), "Do culture-dependent response styles distort substantial relationships?", *International Business Review*, Vol. 22 No. 5, pp. 814-827.
- Hofstede, G. (1991), "Cultures and organizations", *Software of the Mind*, McGraw-Hill, London.
- Hofstede, G. (1994), "The business of international business is culture", *International Business Review*, Vol. 3 No. 1, pp. 1-14.
- Hofstede, G. (2001), *Culture's Consequences: Comparing Values, Behaviors, Institutions and Organizations Across Nations*, Sage, Thousand Oaks, CA.
- Hofstede, G., Hofstede, G.J. and Minkov, M. (2010), *Cultures and Organizations – Software of the Mind: Intercultural Cooperation and its Importance for Survival*, McGraw-Hill, London.
- House, R.J., Hanges, P.J., Javidan, M., Dorfman, P.W. and Gupta, V. (2004), *Culture, Leadership, and Organizations: The Globe Study of 62 Societies*, Sage, Thousand Oaks, CA.
- Johnson, T., Kulesa, P., Cho, Y.I. and Shavitt, S. (2005), "The relation between culture and response styles: evidence from 19 countries", *Journal of Cross-Cultural Psychology*, Vol. 36 No. 2, pp. 264-277.
- Kirkman, B.L., Lowe, K.B. and Gibson, C.B. (2006), "A quarter century of Culture's consequences: a review of empirical research incorporating Hofstede's cultural values framework", *Journal of International Business Studies*, Vol. 37 No. 3, pp. 285-320.
- McSweeney, B. (2002), "Hofstede's model of national differences and their consequences: a triumph of faith – a failure of analysis", *Human Relations*, Vol. 55 No. 1, pp. 89-118.
- Malhotra, N.K., Agarwal, J. and Peterson, M. (1996), "Methodological issues in cross-cultural marketing research: a state-of-the-art review", *International Marketing Review*, Vol. 13 No. 5, pp. 7-43.
- Marín, G., Gamba, R.J. and Marín, B.V. (1992), "Extreme response style and acquiescence among Hispanics", *Journal of Cross-Cultural Psychology*, Vol. 23 No. 4, pp. 498-509.
- Martin, J. (1964), "Acquiescence – measurement and theory", *British Journal of Social and Clinical Psychology*, Vol. 3 No. 3, pp. 216-225.
- Paulhus, D.L. (1991), "Measurement and control of response bias", in Robinson, J.P., Shaver, P.R. and Wright, L.S. (Eds), *Measures of Personality and Social Psychological Attitudes*, Academic Press, San Diego, CA, pp. 17-59.
- Paulhus, D.L., Harms, P.D., Bruce, M.N. and Lysy, D.C. (2003), "The over-claiming technique: measuring self-enhancement independent of ability", *Journal of Personality and Social Psychology*, Vol. 84 No. 4, pp. 890-904.
- Podsakoff, P.M., MacKenzie, S.B., Lee, J.-Y. and Podsakoff, N.P. (2003), "Common method biases in behavioral research: a critical review of the literature and recommended remedies", *Journal of Applied Psychology*, Vol. 88 No. 5, pp. 879-903.
- Raju, P.S., Lonial, S.C. and Mangold, W.G. (1995), "Differential effects of subjective knowledge, objective knowledge, and usage experience on decision making: an exploratory investigation", *Journal of Consumer Psychology*, Vol. 4 No. 2, pp. 153-180.
- Raudenbush, S.W. and Bryk, A.S. (2002), *Hierarchical Linear Models: Applications and Data Analysis Methods*, Sage, Thousand Oaks, CA.
- Richter, N.F., Hauff, S., Schlaegel, C., Gudergan, S., Ringle, C.M. and Gunzel, M. (2016), "Using cultural archetypes in cross-cultural management studies", *Journal of International Management*, Vol. 22 No. 1, pp. 63-83.
- Schwartz, S.H. (1994), "Beyond individualism/collectivism: new cultural dimensions of value", in Kim, U., Triandis, H.C., Kagitçibaşı, C., Choi, S.C. and Yoon, G. (Eds), *Individualism and Collectivism: Theory, Method, and Applications*, Sage, Thousand Oaks, CA, pp. 85-115.
- Singh, J. (1995), "Measurement issues in cross-national research", *Journal of International Business Studies*, Vol. 26 No. 3, pp. 597-619.

- Smith, P.B. (2004), "Acquiescent response bias as an aspect of cultural communication style", *Journal of Cross-Cultural Psychology*, Vol. 35 No. 1, pp. 50-61.
- Smith, P.B. (2011), "Communication styles as dimensions of national culture", *Journal of Cross-Cultural Psychology*, Vol. 42 No. 2, pp. 216-233.
- Sojka, J.Z. and Tansuhaj, P.S. (1995), "Cross-cultural consumer research: a twenty-year review", *Advances in Consumer Research*, Vol. 22 No. 1, pp. 461-474.
- Spiller, S.A., Fitzsimons, G.J., Lynch, J.G. Jr and McClelland, G.H. (2013), "Spotlights, floodlights, and the magic number zero: simple effects tests in moderated regression", *Journal of Marketing Research*, Vol. 7 No. 2, pp. 277-288.
- Steenkamp, J.-B.E.M., de Jong, M.J. and Baumgartner, H. (2010), "Socially desirable response tendencies in survey research", *Journal of Marketing Research*, Vol. 47 No. 2, pp. 199-214.
- Swain, S.D., Weathers, D. and Niedrich, R.W. (2008), "Assessing three sources of misresponse to reversed Likert items", *Journal of Marketing Research*, Vol. 45 No. 1, pp. 116-131.
- Triandis, H.C. (1998), "Collectivism vs individualism: a reconceptualization of the basic concept in cross-cultural psychology", in Verma, G. and Bagley, C. (Eds), *Crosscultural Studies of Personality, Attitudes and Cognition*, MacMillan, London, pp. 60-95.
- Triandis, H.C., Kim, U., Carnevale, P.J., de Dreu, C., Gelfand, M.J., Robert, C., Wasti, S.A., Probst, T., Kashima, E.S., Dragonas, T., Chan, D., Chen, X.P., Van De Vliert, E., Iwao, S., Ohbuchi, K.-I. and Schmitz, P. (2001), "Culture and deception in business negotiations: a multilevel analysis", *International Journal of Cross-Cultural Management*, Vol. 1 No. 1, pp. 73-90.
- Trompenaars, F. (1994), *Riding the Waves of Culture: Understanding Cultural Diversity in Business*, Economist Books, London.
- Van Herk, H., Poortinga, Y.P. and Verhallen, T.M.M. (2004), "Response style in rating scales: evidence of method bias in data from six EU countries", *Journal of Cross-Cultural Psychology*, Vol. 35 No. 3, pp. 346-360.
- Venaik, S. and Brewer, P. (2013), "Critical issues in the Hofstede and GLOBE national culture models", *International Marketing Review*, Vol. 30 No. 5, pp. 469-482.
- Welkenhuysen-Gybels, J., Billiet, J. and Cambré, B. (2003), "Adjustment for acquiescence in the assessment of the construct equivalence of Likert-type score items", *Journal of Cross-Cultural Psychology*, Vol. 34 No. 6, pp. 702-722.
- Winkler, J.D., Kanouse, D.E. and Ware, J.E. Jr (1982), "Controlling for acquiescence response set in scale development", *Journal of Applied Psychology*, Vol. 67 No. 5, pp. 555-561.
- Wolf, J. and Rosenberg, T. (2012), "How individual scholars can reduce the rigor-relevance gap in management research", *BuR – Business Research*, Vol. 5 No. 2, pp. 178-196.
- Wong, N., Rindfleisch, A. and Burroughs, J.E. (2003), "Do reverse-worded items confound measures in cross-cultural consumer research? The case of the material values scale", *Journal of Consumer Research*, Vol. 30 No. 1, pp. 72-79.

Corresponding author

Carolin Krautz can be contacted at: c.krautz@outlook.com

Q 1: And how well do you know the following automobile-related environmental-/ technology brands?

	1	2	3	4	5
	I know it very well	I know quite a lot about it	I know little about it	I have heard of it but I don't know anything about it	I have never heard of it
Label I				X	
Label II	X				
Label III		X			
Label IV		X			
Label V		X			

Q 2: In your opinion, which of the following environmental-/ technology brands belong to which car brand

Label I	False
Label II	False
Label III	Right
Label IV	False
Label V	False

$$AR = \frac{\sum_{x=1}^5 f(x) * w_x}{n}$$

$$AR = \frac{1 * \frac{4}{4} + 3 * \frac{3}{4} + 1 * \frac{1}{4}}{5} = 0.7$$

e.g. $w_2 = \frac{5-2}{5-1} = \frac{3}{4}$

$$pARS = \left[\frac{\sum_{x=1}^5 f(x) * w_x}{n} \right] - \frac{f(y)}{n}$$

$$pARS = \frac{1 * \frac{4}{4} + 3 * \frac{3}{4} + 1 * \frac{1}{4}}{5} - \frac{1}{5} = 0.5$$

$$pARSO1 = \left[\frac{\sum_{x=1}^5 f_{false}(x) * w_x}{n_{false}} \right] - \left[\frac{\sum_{x=1}^5 f_{true}(x) * w_x}{n_{true}} \right]$$

$$pARSO1 = \left[\frac{1 * \frac{4}{4} + 2 * \frac{2}{4} + 1 * \frac{1}{4}}{4} \right] - \left[\frac{1 * \frac{3}{4}}{1} \right] = -0.0625$$

$$pARS-2 = \left[\frac{\sum_{x=1}^5 PR_{false}(x) * f_{false}(x)}{n_{false}} \right] - \left[\frac{\sum_{x=1}^5 PR_{true}(x) * f_{true}(x)}{n_{true}} \right]$$

$$pARS-2 = \left[\frac{1 * 0 + 2 * \frac{1}{4} + 1 * \frac{3}{4}}{4} \right] - \left[\frac{1 * \frac{3}{4}}{1} \right] \approx -0.44$$

e.g. $\overline{PR}_{false}(2) = 1 - \frac{3}{4} = \frac{1}{4}$; $\overline{PR}_{true}(2) = 1 - \frac{1}{4} = \frac{3}{4}$

Figure A1. Example on the construction of indices