Assessment of the General Response Style: A Cross-Cultural Validation

Evaluación de Estilos Generales de Respuesta: Una Validación Transcultural

Jia He[1] [a, b], and Fons J. R. van de Vijver [b, c, d, e, f]

[a] German Institute for International Educational Research, Germany
[b] Tilburg University, the Netherlands
[c] Higher School of Economics, Russia,
[d] Tilburg University, The Netherlands
[e] North-West University, South Africa
[f] University of Queensland, Australia

Abstract

Surveys that use self-reports are susceptible to response styles. The assessment and validation of response styles would benefit from a brief measure that captures the general tendency of responding. Going beyond the traditional view that response styles amount to deliberate impression management or lying, we propose to assess and conducted a validation of brief general response style (GRS) scales which are conceptualized as a trait-like communication style preference, in multiple cultural contexts. With university student data on the GRS measure, indirect measures of GRS, and personality and values, we found that (1) this direct measure of GRS showed acceptable internal consistency and comparable factor structure across countries (metric invariance); (2) although it did not correlate strongly with the indirect measure of GRS, it showed similar mean difference patterning across countries, with a lower level of GRS in China than Mexico and the Netherlands, and (3) in line with the trait-like conceptualization, the direct GRS was associated with being extravert, open, and valuing self-enhancement. We discuss the potential usefulness of the direct GRS measure in surveys involving different groups.

Resumen

Las encuestas que utilizan el auto reporte siempre son sensibles a los estilos de respuesta. La evaluación y validación de los estilos de respuesta se vería beneficiada de una medida corta que capture la tendencia general de respuesta. Este artículo va más allá de la visión tradicional sobre los estilos de respuesta que de manera deliberada afectan la conducta de mentir y el manejo de impresiones, ya que se propuso evaluar y llevar a cabo la validación de la versión corta de las escalas de Estilo de Respuesta General (*General Response Style*, GRS, por sus siglas en inglés), lo cual se conceptualizó como una preferencia tipo rasgo hacia un estilo de comunicación,

dentro de contextos culturales múltiples. Se aplicó una muestra de estudiantes, los datos en torno a la medida de GRS, una medida indirecta de GRS, personalidad y valores, se encontró que (1) la medida directa de GRS muestra consistencia interna aceptable y una composición factorial comparable a través de países (invarianza métrica); (2) a pesar de no correlacionar fuertemente con la medida indirecta de GRS, mostró patrones similares en diferencias de medias a través de los países, con un menor nivel de GRS en China que en México y los Países Bajos, y (3) en línea con la conceptuación tipo rasgo, el GRS directo se asoció con ser extrovertido, abierto, y valorar la autopromoción. Se discuten los usos potenciales de las medidas directas de GRS en encuestas que involucran diferentes grupos.

Palabras Clave: Estilos Generales de Respuesta, Auto-Reportes, Valores, Personalidad, Transcultural

Surveys that use self-reports are susceptible to response styles. Response styles refer to the systematic tendency to respond on some basis other than the targeted construct (Paulhus, 1991). The most common response styles include acquiescence, extremity, midpoint responding, and socially desirable responding. Early research integrated work on these specific response styles by showing that most of the variance in these styles is captured by a general underlying characteristic that was labeled the General Response Style (GRS) (e.g., He & van de Vijver, 2013, 2014). The previous research has garnered much attention, but the application of the approach is hampered by the lack of a simple, short measure of the GRS. The current paper extends this work by developing and validating short, simple measures for the GRS, which can be seen as novel, conceptually better-founded measures of response styles than those currently in use. We propose to assess and validate measures of general response style from different perspectives (i.e., behavioral response amplification, importance of response amplification, and suppression of expression) through measurement invariance testing of the scales, checking their convergence with indirect measures of response styles extracted from other Likert-scale responses, and linking them to self-reported personality traits and values in multiple cultural contexts.

Current Measures of Response Styles and Their Limitations

Previous research has integrated specific response styles to a general response style factor, with positive loadings of extremity (and social desirability), and negative loadings of acquiescence and midpoint responding (He & van de Vijver, 2013, 2014). This GRS factor represents the continuum of response moderation to response amplification. Its stability and usefulness to provide a theoretical framework in studying different response styles and to create consistency in findings in response styles have been confirmed (1) both at individual level and at country level, (2) in various ethnic groups in the Netherlands and various countries in large-scale surveys, and (3) using both indirect measures of acquiescence, extremity, and midpoint responding, and direct self-reported measures of acquiescence, extremity, midpoint responding, and social desirability (He et al., 2014, He & van de Vijver, 2015, He et al., 2017). The GRS not only helps to understand individual and cultural differences in communication styles, but can also be used to partial out the effects of scale usage difference in responses in order to enhance data comparability across respondents and groups.

However, there is no brief and validated measure for this construct, as previous studies assessed different specific response styles and/or extracted statistical procedures to identify the commonality of these styles. Limitations of that approach a (1) too many items are needed to directly assess response styles (we previously administered 45 items to derive the GRS); (2) indirect assessment (e.g., using counting procedures with available data to approximate response style behaviors) is dependent on data availability, item content, and response formats. A self-reported measure that directly targets the GRS is needed. Moreover, the most popular self-reported measures of response styles, namely social desirability, measured either in the Marlowe-Crowne scale (Crowne

& Marlowe, 1960) or the Balanced Inventory of Desirable Responding (Paulhus, 2002) have been criticized as not targeting response bias but expression of honesty-humility (de Vries et al., 2013) and interpersonally oriented self-control (Uziel, 2010), thus a more direct assessment of communication styles may shed light on response style use. We propose such an assessment with adapted and refined items targeting communication styles with different operationalizations, and search for the most reliable and valid measure in different cultural contexts.

The Validation of the GRS

With a trait-like conceptualization, GRS can be measured with different item batteries (and response options). For instance, items tapping into the behavioral component (i.e., frequency of response style usage) without contextual cues can indicate stable response tendency, items targeting the attitudinal component (i.e., agreement on the importance of response amplification) can indicate the preference of GRS from the affective and cognitive perspective, whereas items about the likelihood of expressing opinions different to one's own in point to the suppression of expression. We term each of them behavioral GRS, attitudinal GRS, and suppression of GRS, respectively, and we explore their relevance to the reliably and validly measurement of GRS. Across cultures, we expect to find a similar structure and metric of each measure (i.e., metric invariance in measurement invariance testing). Scalar invariance may be difficult to find, given the different interpretations and scale usage preferences when responding to these measures.

Within cultures, respondents' self-reports of GRSs are expected to correlate positively with indirect measures of GRS (i.e., behavioral indicator of response amplification when responding to a heterogeneous set of Likert-scale items). Given the exploratory nature of the validation, we investigate the extent to which these three direct GRS measures show positive correlations with an indirect assessment of GRS.

For the nomological network of the GRS, previous research has shown that response amplification was related to various personality traits such as openness, intolerance of ambiguity, simplistic thinking, and decisiveness (e.g., Naemi, Beal, & Payne, 2009; Tsujimoto, 2003), and values such as self-enhancement (e.g., Uskul, Oyserman, & Schwarz, 2010). The GRS from integrated measures of acquiescence, extremity, midpoint response style and social desirability was found to be related to the "big one" factor of personality, which is the common variance of desirable traits (He & van de Vijver, 2013). We expect to find positive associations of the attitudinal, behavioral, and (reversed) suppression GRS with desirable personality traits (e.g., extraversion, openness in particular), and self-enhancement value and negative associations conservation value in different cultural contexts. The associations may differ for each direct GRS, and the empirical results are expected to shed light on the most reliable and valid GRS measure.

Method

Sample and Procedure

This validation study is part of a larger project on enhancing data comparability of Likert-scale value and personality data (He et al., 2017) with university student samples from 16 countries. We made use of direct responese on response style items to form our GRS measures, indirect measures of response styles, and personality and values in 12 countries (we excluded countries with a sample size smaller than 100). These 12 countries show vastly different preferences of communication styles (e.g., Smith, 2011) and they differ in affluence level and value dimensions such as collectivism and uncertainty avoidance, which are relevant for scale usage differences. Particularly, they exemplify honor, dignity, and face cultures that may moderate the survey response processes with culturally transmitted response style perferences (Uskul, Oyserman, & Schwarz, 2010), a validation of the GRS measure with such diverse contexts can lend robustness to our conclusions.

University students were invited to take part in the survey. Administration procedures were standardized with slight variations across countries, given local contextual differences. In countries where English is not the mother tongue or language of instruction in the university, the questionnaire was translated by

two independent translators and convergence was sought to produce a final version. The demographics are presented in Table 1. Computerized assessment was employed in all countries but China, Indonesia, and Zambia where a paper and pencil survey was administered. There is evidence that mode effects are very small in self-reports of response styles (He et al., 2015); therefore, we treated the different modes as interchangeable. The participation of all students was voluntary.

Table 1
Demographics of Participants

Country	Sample Size	Mean Age (SD)	% of Males	Language
Canada	431	21.77 (2.54)	24.88	English
China	309	20.76 (1.01)	12.30	Chinese
Indonesia	403	22.32 (1.54)	30.02	English
Lithuania	259	23.07 (2.76)	13.13	Lithuanian
Mexico	163	21.68 (2.23)	28.83	Spanish
Netherlands	206	21.63 (1.84)	20.87	Dutch
Romania	215	22.46 (2.39)	27.10	Romanian
Singapore	275	23.03 (1.30)	33.58	English
South Africa	306	21.62 (2.03)	32.89	English
Spain	127	21.83 (1.44)	17.46	Spanish
Turkey	223	22.42 (2.46)	39.64	Turkish
Zambia	300	22.20 (2.38)	40.20	English

Measures

Self-report measures of specific response styles. Self-report measures of acquiescence, extremity, and midpoint response style developed and validated in He and van de Vijver (2013), were further adapted based on the pilot study. Each style, with 10 items, used balanced scales (i.e., half positively worded items and half negatively worded items) in an interrogative format (i.e., asking questions instead of rating on a statement) using five categories of semantic differentials. Each item had a different set of response options such as from never to always, not important at all to extremely important; this format has been shown to enhance cross-cultural comparability and to induce fewer response styles (e.g., Friborg, Martinussen, & Rosenvinge, 2006). For each style, item content included affective, cognitive and behavioral aspects involving the use of the style.

In the current study, we selected items from the specific response style scales that feature the three conceptualizations of GRS, respectively. Specifically, we sampled five items of context-free frequency of response style uses for the behavioral GRS, five items of agreement on importance of response amplification for the attitudinal GRS, and four items of avoiding expressions of own opinions as the suppression of GRS. The content and response options for the items are presented in the appendix.

Indirect Measures for Response Style Indexes. A total of 45 heterogeneous items randomly chosen from Measures of Personality and Social Psychological Attitude (Robinson, Shaver, & Wrightsman, 1991), from which behavioral indexes of response style could be extracted. These items covered different life domains and were answered on frequency- and agreement-based scales with three to seven options. For each style, 15 non-overlapping items were selected. Item responses were recoded to indicate the presence and absence of acquiescence (i.e., endorsement of agreeing options as 1 and other options as 0), extremity (endorsement the two end categories as 1 and other options as 0), and midpoint response style (endorsement of the middle category as 1 and other options as 0), respectively. The internal consistency of the recoded items for each style was checked, and it turned out that extremity and midpoint responding had moderate levels of internal consistency, whereas acquiescence had very low values of Cronbach's Alpha, Thus, acquiescence was excluded. A score of indirect GRS was computed as the sum of the two remaining response styles (with midpoint response style reverse-scored). The values of Cronbach's Alpha for the direct and indirect measures in each culture are presented in Table 2.

Personality. The Big Five personality scales (Agreeableness, Conscientiousness, Extroversion, Openness, and Emotional Stability) were measured with 50 items of the International Personality Item Pool (Goldberg et al., 2006) with response options ranging from 1 (very inaccurate) to 5 (very accurate).

Table 2
Internal Consistency of Scales

Country	Direct General Response Style			Indirect Response Style	
	Behavioral	**Attitudinal**	**Suppression**	**Extremity**	**Midpoint Responding**
Canada	.643	.739	.637	.708	.565
China	.432	.481	.269	.791	.752
Indonesia	.509	.633	.023	.723	.412
Lithuania	.699	.729	.498	.616	.566
Mexico	.586	.717	.640	.706	.770
Netherlands	.626	.758	.557	.542	.285
Romania	.574	.779	.510	.717	.638
Singapore	.616	.662	.588	.637	.598
South Africa	.517	.736	.612	.735	.454
Spain	.664	.740	.615	.619	.377
Turkey	.429	.721	.556	.692	.628
Zambia	.413	.749	.344	.632	.398

Values. The four value dimensions (Self-enhancement, Self-transcendence, Openness to Change, and Conservation) were measured with the 21-item Portrait Values Questionnaire (Schwartz et al., 2001), with responses ranging from 1 (does not resemble me at all) to 5 (very much resembles me). The internal consistency of the personality and value scales was checked in the previous study (He et al, 2017) and all scales demonstrated acceptable values.

Results

We report the results in three parts: the factor structure and invariance of the GRS measures, convergence check with the indirect measure of GRS, and the nomological network of this measure (i.e., correlation with personality and values).

Factor Structure and Measurement Invariance

A principal component analysis of for each of the three direct GRS measures was conducted with the pooled sample. There was support for a one-factor solution (based on eigenvalues and the scree plot), with explained variance of 38%, 49% and 43%, respectively. In all three factors, items keying for higher extremity loaded positively and items on higher acquiescence and midpoint responding loaded negatively on the factor. The internal consistency of the scales differed (Table 2 first three columns) across countries. The attitudinal GRS showed the highest consistency across countries (except China), followed by the behavioral GRS (with problematic reliability in China, Indonesia, and Zambia), while the suppression factor showed the lowest internal consistency (with rather low values in China, Indonesian, Lithuania, and Zambia).

A measurement invariance testing of each scale in the multigroup confirmatory factor analysis across countries was performed in Mplus (Muthen & Muthen, 1998-2012). Three common levels of invariance were checked: (1) Configural invariance indicates that items measuring a construct cover facets of this construct adequately; (2) Metric invariance means that the items measuring a construct have the same factor loadings across groups. With metric invariance satisfied, associations between variables can be compared across groups; and (3) Scalar invariance implies that items have the same loadings and intercepts. Only with scalar invariance can mean scores be compared across cultures (van de Vijver & Leung, 1997). Items were treated as ordered categories and the WLSMV estimator was used. Due to some missing categories in the data, responses were collapsed to three categories from the original five to ensure non-zero observation in each category (a requirement for modelling data as categorical).

Table 3 presents the model fit of all three GRS measures across 12 countries. According to the model fit criteria including Comparative Fit Index (CFI: above .90), Root Mean Square Error of Approximation (RMSEA: below .055), and the change of CFI and RMSEA within .004 and .05 from the configural to metric model, and .004 and .01 respectively from the metric to scalar model as an indication of acceptance of a more restricted model (Rutkowski & Svetina, 2016), all three GRS measures across countries reached configural invariance, but not metric or scalar invariance across the 12 countries. The poor model fit could be due to the low internal consistency in a few countries (e.g., China). Therefore, the configural structure of these measures was supported across countries, but not the invariance of metrics or item intercepts. Therefore, caution is needed in interpreting the mean differences across countries.

Convergence between Direct and Indirect GRS
The country-specific correlations (Table 4) showed mixed results in different countries. Most correlations were weak, indicating low convergence. However, there was consistency in the positive correlation between the attitudinal GRS and the indirect GRS across countries. The behavioral GRS showed a weaker correlation, in comparison with the attitudinal

GRS, and China seemed to be an outlier in the self-reported response amplification behavior and the actual response style in the survey correlations between the GRS measure with the indirect indicators of GRS. The suppression of expression factor did not correlate with the indirect GRS in most cases, except for China (outlier as before), and in Turkey (a positive correlation).

Despite the lack of scalar invariance for the direct GRSs and the weak convergence of these GRS measures, a MANCOVA was carried out with the all four GRS measures (three direct and one indirect) as the dependent variables, country as the grouping variable, and gender as a covariate. There was a significant main country effect, Wilks' Lambda = .631, $p < .01$, partial $\eta 2$ = .109. These GRS measures had differential sizes of cross-cultural differences, with partial $\eta 2$ of .171 for the attitudinal GRS, .128 for the suppression GRS, .124 for the indirect GRS, and .092 for the behavioral GRS.

Nomological Network
Table 5 presents correlations of the direct and indirect measures of GRS with all personality traits and

Table 3
Model Fit of the General Response Style Measure in Multigroup Confirmatory Factor Analysis

		$\chi 2$	df	RMSEA	CFI
Behavioral GRS	Configural	207.06**	60	.096	.954
	Metric	413.15**	104	.105	.904
	Scalar	838.94**	148	.132	.786
Attitudinal GRS	Configural	320.74**	60	.128	.936
	Metric	Non-convergence			
	Scalar	1138.20**	148	.159	.756
Suppression GRS	Configural	114.47**	24	.120	.940
	Metric	260.53**	57	.116	.865
	Scalar	480.86**	90	.128	.740

**$p < .01$.

Table 4
Correlations between the Indirect General Response Style (GRS) with the Three Direct GRS Measures

Country	Behavioral GRS	Attitudinal GRS	Suppression GRS (reversed)
Canada	.173**	.263**	.096
China	-.166**	.126*	-.174**
Indonesia	.014	.319**	-.084
Lithuania	.171**	.237**	.029
Mexico	.247**	.129	.100
Netherlands	.053	.070	.007
Romania	.195**	.309**	.104
Singapore	.155*	.298**	.077
South Africa	.089	.095	.072
Spain	.186*	.162	.160
Turkey	.089	.346**	.211**
Zambia	.258**	.201**	.074

*$p < .05$. **$p < .01$.

Table 5
Correlation of the General Response Style (GRS) Measures with Personality Traits and Values (Pooled Sample)

	Behavioral GRS	Attitudinal GRS	Suppression GRS (reversed)	Indirect GRS
Agreeableness	.009	.135**	.021	.273**
Consciousness	.093**	.176**	.100**	.205**
Extraversion	.303**	.278**	.201**	.098**
Openness	.264**	.324**	.181**	.294**
Emotional Stability	.037*	-.022	.130**	.012
Self-Transcendence	.055**	.184**	.064**	.366**
Self Enhancement	.245**	.345**	.021	.191**
Open to change	.246**	.338**	.161**	.324**
Conservation	-.075**	.101**	-.148**	.197**

$*p < .05.$ $**p < .01.$

values in the pooled sample (for the concise presentation). All measures of GRS were positively associated with extraversion, openness, conscientiousness, openness to change and self-transcendence. They differ in correlations with other traits and value dimensions. For instance, all except the suppression GRS had a positive association with self-enhancement; the behavioral and suppression GRS were negatively related to conservation, whereas the reverse was found for the attitudinal and indirect GRS. Attitudinal and indirect GRS were positively related to agreeableness, but not the other two. All in all, it seems that the attitudinal and indirect GRS were quite similar in their nomological network, whereas the behavioral and suppression GRS were more similar to each other.

Discussion

Response styles present a persistent challenge in surveys, as they can invalidate the measurement, the structural and mean comparisons of Likert-scale measures (e.g., van Vaerenbergh & Thomas, 2013), therefore its measurement and validation has important implications for improving the quality of survey methodology. In this study, we made use of data from 12 countries with distinctive cultural values to explore the validity of several brief self-reported measures of GRS conceptualized from behavioral, attitudinal, and suppression of expression perspectives. Our approach takes response styles as trait-like communication styles that can be perceived and

reported by individual respondents, and it goes beyond the traditional view that response styles amount to deliberate impression management or even lying, which has prevailed in the literature. The main findings include that (1) These three direct measures of GRS showed largely acceptable internal consistency (except for the suppression GRS); they demonstrated configural invariance across cultures but not metric or scalar invariance; (2) Although they did not correlate strongly with the indirect measure of GRS, there was consistent, stronger convergence with the attitudinal GRS and the indirect measure than the other two direct measures, and (3) different GRS measures (both direct and indirect) were consistently associated with being extravert, open, conscientious, and valuing self-transcendence, but different patterning was observed with specific GRS measures and self-enhancement, agreeableness, and conservation. We discuss the measurement and the potential use of GRS measures in surveys involving different groups.

The three direct GRS measures consist of items tapping into the use of different response styles (behavioral), importance and preference of response styles (attitudinal), and tendency to express opinions different to one's own (suppression), respectively. In line with previous research on the integration of different response styles, in all three direct measures, items on high extremity loaded positively on the factor, and items on high acquiescence and midpoint responding loaded negatively on the factor. These three direct

measures are moderately, positively related to each other, pointing to certain convergence in looking at response amplification versus moderation from different perspectives. Across countries, the configural model in MGCFA was supported, but factor loadings and item intercepts vary, possibly due to low internal consistency in certain countries, poor translation for some items, and the difficulty in responding to the items with differing response options from item to item.

We did not find much support for strong convergence of the direct and indirect measures of GRS, which was not entirely unexpected. Previous research has often reported weak or lack of correlation of attitudinal and behavioral measures of psychological constructs such as impulsivity, distress tolerance, risk taking, and self-control, and comparisons based on either type tend to lead to divergent conclusions (e.g., de Ridder et al, 2011; Malesza & Ostaszewski, 2016; McHugh et al., 2011). This may be due to measurement bias in each type of measure that reduces shared common variance, and individual differences in self-related attitude stability, accessibility, affective-cognitive consistency, and self-regulation (e.g., Fazio, 1990). Nevertheless, this low convergence also speaks to the need to use both types of measure complementarily.

Among the three direct measures, the attitudinal measure had the highest internal consistency, it showed the most consistent convergence when related to the indirect measure, and the nomological network for the attitudinal GRS and the indirect GRS was more similar than the other two direct measures, indicating that the attitudinal component of GRS captures the actual response tendency more accurately than the behavioral GRS and suppression GRS. Moreover, there seems to be more cross-cultural differences (indicated by partial eta-squared) in this attitudinal GRS than any other GRS measures. We see advantages in measuring the attitudinal GRS with the self-reported scale, because it is brief and reliable (indirect GRS requires many more items and is sensitive to the data source used for its construction), and it captures the core, trait-like variations across individuals and cultural groups. Thus, this direct GRS measure may hold promise in better understanding communication styles and in correcting for the differences in Likert-scale scores due to the communication styles.

Conclusions

We provided a search and validation of new measures to access the trait-like communiation styles of survey respondents, namely the direct GRS measures. With the validation of behavioral, attitudinal and suppresion GRS in 12 different cultural contexts, we showed that the attitudinal component can be reliably and validly measured, and it converged better with the indirect measure. Our study is not without limitation. The student sample may not be representative, the selected items could be refined, and more nomological network measures are in need to check its convergent and discriminative validity of each measure. Future studies with more varied samples (and more representative samples in more cultural groups) can further validate and refine these measures, and their correction effects in various cross-cultural survey data are to be examined.

Appendix: Items for the General Response Style Measures
Instruction: Below are questions on your communication preferences. Please answer each of them based on your own experience. Please note that the response options are different from item to item.

• **Behavioral GRS**

1. In general, how often do you rather agree than disagree with others?

Never agreeing	Rarely agreeing	Sometimes agreeing	Often agreeing	Always agreeing
1	2	3	4	5

2. How often do you prefer to express a strong opinion to no opinion at all?

Never strong	Rarely strong	Sometimes strong	Often strong	Always strong
1	2	3	4	5

3. How often do you take an extreme stance?

Never	Rarely	Sometimes	Often	Always
1	2	3	4	5

4. How often do you ever prefer neutral opinions to strong opinions?

Never neutral	Rarely neutral	Sometimes neutral	Often neutral	Always neutral
1	2	3	4	5

5. How often do you give neutral opinions?

Never	Rarely	Sometimes	Often	Always
1	2	3	4	5

• **Attitudinal GRS**

6. How important is it for you to have strong opinions?

Not at all	Slightly	Somewhat	Moderately	Extremely
1	2	3	4	5

7. Do you like to be viewed as a person with strong opinions?

Extremely so	Moderately so	Somewhat so	Slightly so	Not at all so
1	2	3	4	5

8. When you have an opinion on something, how is this opinion usually best described?

Very mild	Mild	Neither mild nor strong	Strong	Very strong
1	2	3	4	5

9. In what way are your opinions best described?

Very strong	Strong	Somewhat strong	Slightly strong	Not strong at all
1	2	3	4	5

10. How strongly do you agree with the statement below?
"I think it is good to have strong opinions"

Strongly agree	Agree	Neither agree nor disagree	Disagree	Strongly disagree
1	2	3	4	5

• **Suppression of Expression**

11. How often do you express agreement to make everyone feel at ease?

Always	Often	Sometimes	Rarely	Never
1	2	3	4	5

12. How often do you agree with others to show your empathy?

Never	Rarely	Sometimes	Often	Always
1	2	3	4	5

13. Have you ever given neutral opinions when you do not understand a question very well?

Never	Rarely	Sometimes	Often	Always
1	2	3	4	5

14. How often do you hide your true opinions by remaining neutral?

Always	Often	Sometimes	Rarely	Never
1	2	3	4	5

References

1. Crowne, D. P., & Marlowe, D. (1960). A new scale of social desirability independent of psychopathology. *Journal of Consulting Psychology, 24*, 349-354.

2. de Ridder, D. T. D., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., & Baumeister, R. F. (2011). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors. *Personality and Social Psychology Review, 16*, 76-99. doi:10.1177/1088868311418749

3. de Vries, R. E., Zettler, I., & Hilbig, B. E. (2013). Rethinking trait conceptions of social desirability scales: Impression management as an expression of honesty-humility. *Assessment, 21*, 286-299.

4. Fazio, R. H. (1990). Multiple processes by which attitudes guide behavior: The mode model as an integrative framework. In M. P. Zanna (Ed.), *Advances in Experimental Social Psychology* (Vol. 23, pp. 75-109). New York, NY: Academic Press.

5. Friborg, O., Martinussen, M., & Rosenvinge, J. H. (2006). Likert-based vs. semantic differential-based scorings of positive psychological constructs: A psychometric comparison of two versions of a scale measuring resilience. *Personality and Individual Differences, 40*, 873-884. doi:10.1016/j.paid.2005.08.015

6. Goldberg, L. R., Johnson, J. A., Eber, H. W., Hogan, R., Ashton, M. C., Cloninger, C. R., & Gough, H. G. (2006). The international personality item pool and the future of public-domain personality measures. *Journal of Research in Personality, 40*, 84-96. doi:10.1016/j.jrp.2005.08.007

7. He, J., & van de Vijver, F. J. R. (2013). A general response style factor: Evidence from a multi-ethnic study in the Netherlands, *Personality and Individual Difference, 55,* 794-800.

8. He, J., & van de Vijver, F. J. R. (2014). Self-presentation styles in self-reports: Linking the general factors of response styles, personality traits, and values in a longitudinal study. *Personality and Individual Differences*, *31,* 129-134.

9. He, J., & van de Vijver, F. J. R. (2015). Effects of a general response style on cross-cultural comparisons: Evidence from the Teaching and Learning International Survey. *Public Opinion Quarterly, 79*, 267-290.

10. He, J., van de Vijver, F. J. R., Domínguez Espinosa, A., Abubakar, A., Dimitrova, R., Adams, B., . . . Villieux, A. (2015). Socially desirable responding: Enhancement and denial in 20 countries. Cross-Cultural Research, 49, 227–249.

11. He, J., van de Vijver, F. J. R., Fetvadjiev, V. H., ...& Zhang, R. (2017). On enhancing the cross-cultural comparability of Likert-scale personality and value measures: A comparison of common procedures. *European Journal of Personality, 31,* 642-657.

12. Malesza, M., & Ostaszewski, P. (2016). Dark side of impulsivity — Associations between the Dark Triad, self-report and behavioral measures of impulsivity. *Personality and Individual Differences, 88*, 197-201. doi:10.1016/j.paid.2015.09.016

13. McHugh, R. K., Daughters, S. B., Lejuez, C. W., Murray, H. W., Hearon, B. A., Gorka, S. M., & Otto, M. W. (2011). Shared variance among self-report and behavioral measures of distress intolerance. *Cognitive Therapy and Research, 35*, 266-275. doi:10.1007/s10608-010-9295-1

14. Muthen, L. K., & Muthen, B. O. (1998-2012). *Mplus user's guide* (7th ed.). Los Angeles, CA: Muthen & Muthen.

15. Naemi, B. D., Beal, D. J., & Payne, S. C. (2009). Personality predictors of extreme response style. *Journal of Personality, 77*, 261-286. doi:10.1111/j.1467-6494.2008.00545.x

16. Robinson, J. P., Shaver, P. R., & Wrightsman, L. S. (Eds.). (1991). *Measures of personality and social psychological attitudes*. San Diego, CA: Academic Press.

17. Rutkowski, L., & Svetina, D. (2016). Measurement invariance in international surveys: Categorical indicators and fit measure performance. *Applied Measurement in Education*, 30, 39-51. doi:10.1080/08957347.2016.1243540

18. Paulhus, D. L. (1991). Measurement and control of response biases. In J. Robinson, P. Shaver & L. Wrightsman (Eds.), *Measures of personality and social psychological attitudes* (Vol. 1, pp. 17-59). San Diego, CA: Academic Press.

19. Paulhus, D. L. (2002). Socially desirable responding: The evolution of a construct. In H. I. Braun, D. N. Jackson & D. E. Wiley (Eds.), *The role of constructs in psychological and educational measurement* (pp. 49-69). Mahwah, NJ: Erlbaune.

20. Schwartz, S. H., Melech, G., Lehmann, A., Burgess, S., Harris, M., & Owens, V. (2001). Extending the cross-cultural validity of the theory of basic human values with a different method of measurement. *Journal of Cross-Cultural Psychology, 32*, 519-542. doi:10.1177/0022022101032005001

21. Tsujimoto, H. (2003). Extreme response style and openness, playfulness, and extraversion. *The Japanese Journal of Personality, 12*, 14-26. doi:10.2132/personality.2003.14

22. Uskul, A. K., Oyserman, D., & Schwarz, N. (2010). Cultural emphasis on honor, modesty or self-enhancement: Implications for the survey response process. In J. A. Harkness, M. Broun, B. Edwards, T.

P. Johnson, L. Lyberg, P. P. Mohler, B.-E. Pennell, & T. W. Smith (Eds.), *Survey methods in multinational, multiregional and multicultural contexts* (pp. 191-201). New York, NY: Wiley.

23. Uziel, L. (2010). Rethinking social desirability scales: From impression management to interpersonally oriented self-control. *Perspectives on Psychological Science, 5*, 243-262.

24. van de Vijver, F. J. R., & Leung, K. (1997). *Methods and data analysis of comparative research*. Thousand Oaks, CA: Sage.

25. van Vaerenbergh, Y., & Thomas, T. D. (2013). Response styles in survey research: A literature review of antecedents, consequences, and remedies. *International Journal of Public Opinion Research, 25*, 195-217. doi:10.1093/ijpor/eds021