Understanding Acquiescence Bias: The Mediating Role of Careless Responding in Demographic Differences

Saram Han[1], Daria Mikhailova[1], Kyunghwa Chung[2] and Rohit Verma[2]

[1]College of Business and Technology, Seoul National University of Science and Technology, Nowon-gu, Seoul, South Korea

[2]Darla Moore School of Business, University of South Carolina, USA

saramhan@seoultech.ac.kr
physteshka@gmail.com
kyunghwa@email.sc.edu (Corresponding Author)
rohit.verma@moore.sc.edu

https://doi.org/10.34190/ejbrm.24.1.4315

Abstract: Acquiescence Response Style (ARS) results in unintended bias in survey research by systematically skewing respondents' ratings. ARS has been considered to be influenced by individual traits and thus commonly attributed to demographic characteristics. This study investigates the mediating role of Careless Responding (CR) in the manifestation of ARS. Using large-scale survey data, this study demonstrates that CR strengthens or weakens the influence of demographic characteristics on ARS, explaining a significant portion of the variance in ARS. This study also demonstrates that the effectiveness of the instructional manipulation check (IMC) in detecting CR is diminished when it is designed to match the effects of ARS. ARS triggered by CR decreases IMC failure rates, thereby reducing the IMC's ability to detect CR. The mediating role of CR remains underexplored in the extant literature. This study addresses this research gap, offering a deeper understanding of the underlying mechanisms driving ARS and CR.

Keywords: Acquiescence response style (ARS), Careless responding (CR), Measurement error, Instructional manipulation check (IMC), Survey

1. Introduction

Surveys are one of the most broadly adopted data collection methods in the social sciences. Surveys allow researchers to precisely design and manage the data collection process, thereby preventing external influences that could undermine data integrity (Lavrakas, 2008). Owing to these advantages, surveys have served as an indispensable data collection method in both academia and industry (Han, Anderson & Chung, 2023). In survey research, controlling for potential errors is essential. Researchers need to scrutinize the presence and sources of errors (Biemer, 2010). According to the total survey error framework, survey errors are classified into four categories: coverage, sampling, nonresponse, and measurement error (Groves & Lyberg, 2010; Weisberg, 2005). While coverage, sampling, and nonresponse stem from the inability to observe all selected cases, measurement error arises when inaccuracies exist in the data, despite the inclusion of all cases in the sample (Tourangeau, Rips & Rasinski, 2000). Among various sources of measurement error, response style (RS) refers to the systematic tendencies of respondents to answer survey questions in specific patterns, often regardless of the actual content of the questions (Harzing, 2006; Weijters, Geuens & Schillewaert, 2010). RS impedes measuring actual constructs, having a detrimental effect on survey outcomes (van Vaerenbergh & Thomas, 2013).

One prominent type of RS is acquiescence response style (ARS), commonly referred to as agreement bias, which means the tendency of respondents to agree with survey statements regardless of their true opinions (Billiet & McClendon, 2000; Knowles & Condon, 1999). ARS is often observed in surveys that use Likert-type scales, where agreement is frequently equated with endorsement (Greenleaf, 1992; Jackson & Messick, 1958). Therefore, ARS inflates agreement rates (Baumgartner & Steenkamp, 2001). In the true-score model of classical test theory, the basis of reliability measurement methodology, reliability is calculated as the ratio of true-score variance to observed-score variance, and systematic variance is regarded as true-score variance (Pedhazur & Schmelkin, 1991). As systematic variance includes reliable contamination, which refers to the portion of scores that measure something other than the defined construct (Schwab, 2005), measurement validation procedures such as reliability tests alone cannot fully detect and control systematic bias. ARS is a representative source of such systematic bias, distorting respondents' true attitudes or behaviors and undermining the validity of research findings (Krosnick, 1991; Podsakoff et al., 2003).

Cite this article: Han, S. et al. 2026. "Understanding Acquiescence Bias: The Mediating Role of Careless Responding in Demographic Differences", *Electronic Journal of Business Research Methods 24(1)*, pp.13-23, https://doi.org/10.34190/ejbrm.24.1.4315

Another major source of measurement error is careless responding (CR), which refers to responding with inattentive or insufficient effort during survey completion (Huang et al., 2012; Meade and Craig, 2012). As online and longitudinal surveys grow, control over the survey environment is limited, and respondent fatigue is heightened, increasing the likelihood of CR and making its control more important (Nichols & Edlund, 2020). Substantial studies have investigated the causes of ARS and CR, revealing that both are influenced by demographic factors. Based on the extant literature, this study examines a model in which the impact of demographic factors on ARS is mediated by CR. While correlation between ARS and CR has been evidenced (e.g., Alarcon & Lee, 2022; Henninger & Plieninger, 2021; Kam & Meyer, 2015), studies examining the mediating role of CR are limited. By examining both simultaneously, this study identifies and isolates the impact of CR from the variance in ARS that results from demographic characteristics. This study aims to provide a comprehensive and deeper understanding of the mechanisms by which demographic factors facilitate CR and ARS, and to offer actionable insights for managing data quality.

2. Literature Review

2.1 Demographic Drivers, ARS, and CR

ARS has commonly been viewed as largely influenced by cultural and individual traits (e.g., Meisenberg & Williams, 2008; van Herk, Poortinga & Verhallen, 2004), and the relationship between demographic factors and ARS has been extensively investigated (e.g., Bachman & O'Malley, 1984; Marin, Gamba & Marin, 1992). The significant impacts of demographic characteristics, such as age, education level, gender, socioeconomic status, and cultural background on ARS, were demonstrated (van Vaerenbergh & Thomas, 2013). Education was one of the most frequently examined variables across diverse populations, showing an inverse relationship (e.g., Johnson et al., 2005; Narayan & Krosnick, 1996). Individuals with lower education levels exhibited higher ARS, often attributed to lower intellectual ability (e.g., Meisenberg & Williams, 2008). Socioeconomic status showed similar effects, indicating lower ARS when respondents' socioeconomic status was higher (Landsberger & Saavedra, 1967; Ross & Mirowsky, 1984). Studies on age generally reported a positive correlation with ARS. Older respondents were found to engage more in ARS due to cognitive conservatism and social desirability bias (Billiet & McClendon, 2000; Greenleaf, 1992; Weijters, Geuens & Schillewaert, 2010). In terms of gender, mixed results were found. Some studies found higher ARS among women (Weijters, Geuens & Schillewaert, 2010), others reported lower ARS among women (Greenleaf, 1992; Ross & Mirowsky, 1984), and some found no significant gender differences (Marin, Gamba & Marin, 1992). Cultural background was also found to be an important predictor of ARS. Compared to individualistic cultures that encouraged critical responses, collectivist cultures showed higher ARS, which was explained as resulting from social norms emphasizing harmony and respect for authority (Hamamura, Heine & Paulhus, 2008; Harzing, 2006; van Herk, Poortinga & Verhallen, 2004). Racial and ethnic differences also showed significant impacts, with higher ARS in racial minorities (Johnson & van de Vijver, 2003). Compared to White/Caucasian respondents, higher ARS was found among Black/African American, Latino/Hispanic, and Asian/Asian American respondents (Bachman & O'Malley, 1984; Marin, Gamba & Marin, 1992; Meisenberg & Williams, 2008).

Demographic factors drive CR as well. CR has been investigated under various terms, including random responding (Beach, 1989), inattentive responding (Meade & Craig, 2012), and insufficient effort responding (Huang et al., 2012). Satisficing, which refers to a behavior that respondents simply provide a satisfactory answer when optimally answering a survey question requires substantial cognitive effort (Krosnick, 1991), is also included in this research stream. As CR hinders measuring respondents' true attitudes or behaviors, great research efforts have been made to investigate its predictors, and it is demonstrated that respondents' attentiveness during surveys varies according to demographic characteristics. Similar to the pattern observed in ARS, respondents with higher education and higher socioeconomic status were found to be more attentive and less engaged in CR (Krosnick, 1991; Tourangeau, Rips & Rasinski, 2000). CR levels also appear to vary because some survey topics appeal more to certain demographic groups due to their personal relevance. For example, such topics related to recent consumer behaviors or service experiences were found to capture the attention of younger and more educated respondents (Groves, Singer & Corning, 2000). In online surveys, where less control is imposed on the survey environment, a significant number of respondents exhibited CR, which was more prevalent in Asian than North American samples (Nichols & Edlund, 2020). While CR is likewise influenced by demographic characteristics that drive ARS, what distinguishes them is that, rather than the respondent's inherent tendencies, situational factors such as motivation, task difficulty, and cognitive resources play more fundamental roles in shaping CR (Krosnick, 1991; Krosnick, Narayan & Smith, 1996; Roberts, 2019). When respondents lack sufficient motivation or face mental fatigue, cognitive effort decreases, and CR occurs (Bowling et al., 2016; DeSimone, Harms & DeSimone, 2015).

2.2 Mediating Role of CR in the Relationship Between Demographic Factors and ARS

Although ARS and CR have been considered as two major sources of measurement error, and substantial studies have examined each of them, research exploring their relationship in a systematic manner is limited (Alarcon & Lee, 2022). ARS and CR share demographic antecedents and, in some cases, can lead to the same outcomes (Chylíková, 2020). For instance, when survey motivation is low, a respondent who answers carelessly may simply choose the same number throughout the survey, rather than skipping items or selecting random, inconsistent responses. In such cases, CR leads to ARS. Indeed, the correlation between the two has often been reported. Henninger and Plieninger (2021) found that careless respondents who spent less time on the survey tended to overly select midpoint or high agreement options. In their research on the effects of ARS and CR on the construct dimensionality assessment, Kam and Meyer (2015) also found that respondents who engaged in CR showed higher ARS, which led to the conclusion that CR and ARS jointly distort factor analyses. Recently, Alarcon and Lee (2022) demonstrated that ARS and CR are significantly related and capable of influencing one another using an IRT model that assumes a sequence of decision processes in Likert-type responding.

There may be different perspectives on whether CR precedes ARS or vice versa, and which one between them serves as a moderator. This study proposes a model in which CR mediates the effect of demographic factors on ARS (Figure 1). In this study, we view CR as arising from situational factors, whereas ARS is understood as a dispositional response tendency associated with demographic factors. Respondents who consistently provide socially desirable responses because of their social beliefs or cultural backgrounds are less likely to become careless during surveys simply for those reasons. However, when responding to a survey that is particularly difficult or cognitively demanding, they may temporarily attempt to reduce cognitive effort by relying on heuristic shortcuts (Berinsky, Margolis & Sances, 2014; Greszki, Meyer & Schoen, 2015; Knowles & Condon, 1999; Oppenheimer, Meyvis & Davidenko, 2009). In such situations, CR can occur, and the temporary attentional lapse may lead to ARS, exacerbating the effects of demographics on ARS (Huang et al., 2012; Hui & Triandis, 1985; Meade & Craig, 2012). Inattentive respondents may acquiesce to minimize their cognitive effort when comprehending questions and formulating responses (Holbrook et al., 2007; Weijters, Baumgartner & Schillewaert, 2013). Survey fatigue can lead to inadvertent consent (Baumgartner & Steenkamp, 2001; Tourangeau, Rips & Rasinski, 2000; Zhang & Conrad, 2014). With this perspective, this study views CR as triggered by situational factors and posits that it mediates the effects of demographic characteristics on ARS, either amplifying their influence or attenuating it through negative mediation.

To measure CR, various techniques are used, which fall into three categories: response pattern analysis, self-reported measures, and embedded attention checks (Curran, 2016; DeSimone et al., 2018; Huang, Liu & Bowling, 2015; Meade & Craig, 2012). Response pattern analysis is performed in such ways as identifying long-string responses, detecting invariant response patterns, and assessing response time distributions (Meade & Craig, 2012; Niessen, Meijer & Tendeiro, 2016). Self-reported measures involve directly asking respondents about their attentiveness or effort during the survey (Maniaci & Rogge, 2014; Ward & Pond, 2015). Embedded attention checks ask respondents to follow specific instructions within the survey to confirm their engagement (Kung, Kwok & Brown, 2018; Oppenheimer, Meyvis & Davidenko, 2009). The instructional manipulation check (IMC) is a specific technique that belongs to embedded attention checks. The IMC assesses inattentiveness by checking respondents' success or failure in following given instructions (Abbey & Meloy, 2017; Meade & Craig, 2012). This study employs the IMC. IMCs are an effective and objective tool for distinguishing attentive respondents from those engaging in CR, as they require careful reading and compliance.

This study assumes that CR mediates the effects of demographics on ARS. Furthermore, this study posits that, because of the impact of CR on ARS, IMCs designed with correct answers consistent with the effect of ARS will be less effective in detecting CR. CR occurs when respondents experience the cognitive burden and limited cognitive resources associated with surveys and thus exert less cognitive effort (Krosnick, 1991). When experiencing CR, respondents use heuristic shortcuts to select easy and positive answers (Sears, 1983), which are typically convenient default responses. For respondents with ARS tendencies such as social desirability, the heuristics can lead them to select higher numbers because positive meanings (e.g., “strongly agree” and “very satisfied”) are commonly placed farthest to the right and associated with higher numerical values when survey items are measured on Likert-type scales. Once CR is triggered, it will facilitate the manifestation of ARS tendency, making people less concerned about the correct answers and more easily select higher numbers. If the task of the IMC is to select a number with relatively high values among the given numbers, the IMC may not screen respondents who are actually engaging in CR because they are likely to choose higher numbers due to ARS. This suggests that the impact of CR on ARS may reduce the power of the IMC to detect CR, which is worth investigating. This study examines how CR failure rates vary depending on the alignment of IMC options with the effect of ARS.

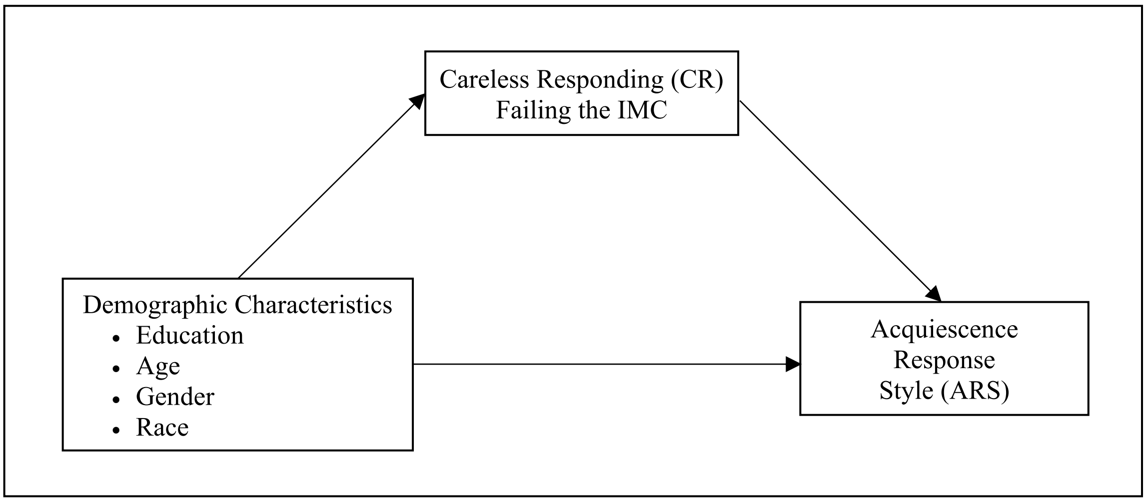

Figure 1: Research model

3. Method

3.1 Data

To test our model, we used survey data from a global market research firm that annually collects feedback on past brand experiences from a large sample of U.S. residents. This annual survey has a long history and is widely referenced, lending it credibility. We used data collected over a one-year period in 2017. By avoiding any interference from the pandemic, the data serves as a reliable resource to test respondents' behavioral survey patterns. The dataset comprises responses from 35,589 individuals who have experienced specific brands within one month prior to the survey. Among our data, answers for a set of six survey items were used. The first five items were to measure respondents' overall satisfaction with a primary service provider in five different industries, and the last item was to measure CR. These six items were selected because the five satisfaction items employed the same question, making them highly comparable, and also because there were no intervening items placed between the five satisfaction items and the IMC item, preventing context effects or unintended interference from other survey content.

To measure ARS, a widely used "Yea-Saying" approach was adopted, which counts the proportion of positive responses (e.g., "agree") to a set of heterogeneous items and considers a disproportionately high proportion of positive responses, relative to the sample average, as indicative of ARS (Baumgartner & Steenkamp, 2001; van Vaerenbergh & Thomas, 2013). The proportion of positive responses (i.e., high-value numeric responses) among responses to the first five survey items was tested (Greenleaf, 1992; Hui & Triandis, 1985). For these five items, the question "How would you rate your overall experience with your primary provider in the following industries?" was given. Each respondent rated their experience with five randomly assigned industries on a 10-point Likert scale (1= unacceptable, 10 = outstanding). A "Don't know" option was also available. A continuous measure was employed to capture the generalized tendency to agree for ARS. Each respondent's average score across the five items was calculated, and those with scores exceeding the mean score of the entire sample were identified as exhibiting ARS (Greenleaf, 1992). The mean score of all respondents was 6.96. A total of 19,320 respondents whose average scores exceeded this threshold were identified as exhibiting ARS, accounting for approximately 54.3% of the entire sample.

Following the five items, one IMC instruction was administered at the end. Each respondent was randomly assigned one option from 11 choices (i.e., numbers 1 to 10 and "Don't know") and instructed to select the designated option. If a respondent failed the IMC, for example, by choosing a number other than 1 or "Don't know" when the instruction was "Please select 1," the respondent was identified as having engaged in CR.

3.2 Sample

The descriptive statistics of the sample, including education level, age, gender, and race, which were independent variables, are presented in Table 1. Education level was measured using 8 categories, ranging from "8th grade or less" to "advanced degree." The largest group of respondents held a "4-year college degree," followed by "some college", "advanced degree", "high school graduate", "trade/technical school", and "some graduate courses." The proportion of respondents with an education level below high school graduation was relatively small. Age was calculated from the reported birth year by respondents, with an average of 47.29 as of 2017 (SD = 15.98). Female respondents outnumbered male respondents. The largest racial group was White/Caucasian, followed by Asian/Asian American, Black/African American, and Latino/Hispanic.

Table 1: Sample description

Variable	Count/Mean (SD)	%
Education level		
1: 8th grade or less	22	0.1
2: Some high school	291	0.8
3: High school graduate	3,928	11.1
4: Trade/technical school	2,027	5.7
5: Some college	8,485	24.0
6: 4-year college degree	11,203	31.7
7: Some graduate courses	1,972	2.6
8: Advanced degree	7,426	21.0
Age (as of 2017)	47.29 (15.98)	
Gender		
1: Male	14,352	40.4
2: Female	21,117	59.6
Race		
1: White/Caucasian	29,171	82.1
2: Black/African American	2,198	6.2
3: Latino/Hispanic	1,572	4.4
4: Asian/Asian American	2,600	7.3
Total sample N	35,589	

Note: Sample sizes vary across demographic categories due to non-responses.

3.3 Model

Two logistic regression models were developed. The binary variables CR and ARS were the outcome variables in the first and second models, respectively. CR functions as a mediator bridging demographic characteristics and ARS. The effects of demographic variables are sequentially propagated through CR's mediated path.

$$CR = \beta_0 + \beta_1 edu + \beta_2 age + \beta_3 male + \beta_4 aamerican + \beta_5 latino + \beta_6 asian + \sum_{k=1}^{10} \beta_{7k} CORECT_k + \epsilon_\beta \quad (1)$$

$$ARS = \alpha_0 + \alpha_1 edu + \alpha_2 age + \alpha_3 male + \alpha_4 aamerican + \alpha_5 latino + \alpha_6 asian + \alpha_7 CR + \epsilon_\alpha \quad (2)$$

edu represents education level. Because our primary interest was in comparing more educated vs. less educated respondents, rather than detailed comparisons among specific education levels, we coded it into a binary format. Respondents with "some college" or higher levels were coded as 1, and all others were coded as 0. *age* denotes the respondent's age. Both gender and race were dummy coded. *male* is a binary indicator for gender. *aamerican*, *latino*, and *asian* are binary indicators for race. Female and White/Caucasian were coded as 0, serving as the reference categories, and otherwise coded as 1. β_1, β_3, β_4, β_5, and β_6 represent differences in the

ISSN 1477-7029

likelihood of engaging in CR between the group coded as 1 and the reference group coded as 0 for each corresponding variable, respectively. β_2 captures the effect of age on the likelihood of CR. $CORECT_k$ is a vector that includes 10 dummy variables indicating which number was designated in the IMC as a correct option. For example, if 1 was designated, $CORECT_1$ was coded as 1, and if 2 was designated, $CORECT_2$ was coded as 1. The category in which the correct option was "Don't know" was coded as 0, serving as the reference group. Each coefficient β_{7k} represents the effect of designating number k, capturing how the failure rate changes relative to designating "Don't know." The baseline CR, the reference group's IMC failure level, is captured by β_0.

In the second model, where the outcome was ARS, α_1 through α_6 indicate demographics' effects analogous to β_1 through β_6 in the first model, capturing how each demographic variable influences the likelihood of ARS. α_7 represents the effect of CR on ARS, capturing how the likelihood of ARS changes as a function of CR. The intercept α_0 is the baseline ARS when all binary predictors are set to zero, and the continuous predictor, age, is set to the average, 47.29. The error terms, ϵ_β and ϵ_α, account for the remaining variance of CR and ARS, respectively. For the mediation analysis, we adopted a bootstrapping approach (Preacher & Hayes, 2008), which provides robust estimates of direct and indirect effects without requiring distributional assumptions. We performed 10,000 bootstrap resamples with 95 % confidence intervals (CIs).

4. Results

Tables 2 and 3 present the bootstrap parameter estimates and 95% CIs for the direct and indirect effects, respectively. All demographic variables had significant direct effects on CR with CIs excluding 0. Education and age had negative effects, indicating that respondents with higher education and older respondents were less likely to engage in CR. Male, African American, Latino, and Asian respondents showed significantly higher propensity for CR relative to their respective reference groups (i.e., female and White American). In the ARS model, all demographic variables had significant positive effects except for age. Age did not significantly affect ARS. Notably, CR demonstrated the largest direct effect on ARS (β = 1.06; 95% CI: 0.97, 1.15), indicating that engaging in CR was strongly associated with the likelihood of ARS when demographic conditions were held constant at baseline. All parameter estimates for CR's mediation effects were significant with CIs excluding 0, confirming the mediating role of CR. Given the negative direct effects of education and age on CR, their indirect effects on ARS via CR were significant and negative. Education and age decreased the likelihood of CR, which in turn lowered the likelihood of ARS. The indirect effects of male, African American, Latino, and Asian were positive and significant. These groups showed significantly higher propensity for CR, which in turn increased the likelihood of ARS through the mediating pathway.

Table 2: Direct effect results

Variable	Estimate	
	DV: CR (i.e., failing IMC)	DV: ARS
Education	-0.40 (-0.49, -0.30)	0.07 (0.01, 0.12)
Age	-0.04 (-0.04, -0.04)	0.00 (0.00, 0.00)
Male	0.35 (0.26, 0.43)	0.18 (0.14, 0.23)
African American	0.66 (0.52, 0.79)	0.26 (0.17, 0.35)
Latino	0.21 (0.05, 0.37)	0.12 (0.02, 0.23)
Asian	0.60 (0.47, 0.73)	0.16 (0.07, 0.24)
CR (i.e., failing IMC)		1.06 (0.97, 1.15)
Correct IMC Options		
Correct answer: 1	0.00 (-0.19, 0.17)	
Correct answer: 2	-0.13 (-0.31, 0.05)	
Correct answer: 3	-0.04 (-0.22, 0.15)	
Correct answer: 4	-0.06 (-0.24, 0.11)	
Correct answer: 5	-0.06 (-0.24, 0.12)	
Correct answer: 6	-0.27 (-0.46, -0.08)	
Correct answer: 7	-0.21 (-0.40, -0.02)	
Correct answer: 8	-0.29 (-0.49, -0.10)	

Variable	Estimate	
	DV: CR (i.e., failing IMC)	DV: ARS
Correct answer: 9	-0.26 (-0.45, -0.07)	
Correct answer: 10	-0.41 (-0.62, -0.22)	

Note: 95% CIs in parentheses.

Table 3: Indirect effect results

Demographic variables → CR → ARS	
Education	-0.42 (-0.53, -0.31)
Age	-0.04 (-0.05, -0.04)
Male	0.37 (0.27, 0.46)
African American	0.70 (0.55, 0.86)
Latino	0.22 (0.05, 0.40)
Asian	0.64 (0.49, 0.78)

Note: 95% CIs in parentheses.

The marginal effects of direct, indirect, and total effects were computed on a probability scale. The baseline ARS probability was calculated from the reference scenario in which binary variables were set to 0 and age was set to 47.29, the average. To calculate the total marginal effects of each variable, the value of the target variable was changed one at a time while holding all other variables at their baseline values. The binary variables were changed to 1, and age was changed to the 75th percentile age, which was 61. Then, the difference between the predicted ARS probability and the baseline ARS probability was calculated, which accounts for changes in CR. The direct marginal effects were calculated by holding the mediator CR at its baseline value. The indirect marginal effects were obtained by subtracting the direct marginal effects from the total marginal effects. Figure 2 visually illustrates whether the mediating path of CR strengthens or counteracts the direct effects and how these forces shape the total effects.

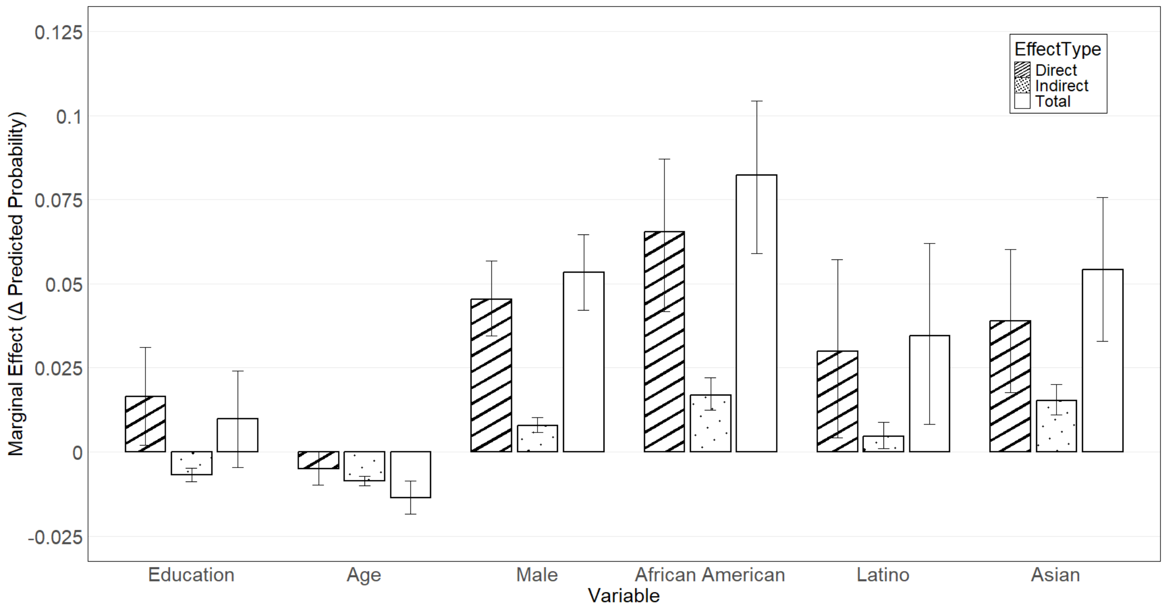

Figure 2: Marginal effects of demographic predictors on ARS probability: direct, indirect, and total effects

The decomposition, together with the bootstrap results, reveals several interesting patterns. For education, the bootstrap results showed a significant positive direct effect and a significant negative indirect effect on the likelihood of ARS. In the decomposition, the total effect of education was approximately 0.0087, which was decomposed into a direct effect of 0.0165 and an indirect effect of -0.0079, reflecting that education's positive direct effect was partially offset by a negative indirect effect through CR. This result suggests that the negative relationship between education and ARS, which has been commonly found in previous studies, may actually appear due to CR. Age also did not show consistent patterns across its direct and indirect effects. In the bootstrap results, age did not have a significant direct effect on the likelihood of ARS, while it had a small but significant indirect effect through CR. The decomposition showed a negative direction for all direct, indirect, and total effects. These results show that although age does not directly affect ARS, it can slightly decrease ARS via CR. These are meaningful findings that are revealed through a mediation model of this study. For the remaining variables, the direct and indirect effects were in the same direction. CR enhanced the effects of gender and race. Male, African American, Latino, and Asian respondents showed higher likelihoods of ARS, which were amplified through CR's mediating path, further increasing the predicted probability of ARS.

Table 4 shows the failure rates for specific instructions of the IMC. The failure rate generally decreased as the value of the number for a correct answer increased. While the instruction "Please select 1" had the highest failure rate (9.33%), the instruction "Please select 10" had the lowest failure rate (6.83%). The parameter estimates representing the effect of each instruction on IMC failure are presented in Table 2, and their declining trend as the correct answer number increases is visualized in Figure 3. The estimated parameter value for 1 was 0.00, and all other estimated values were negative, with the lowest value for 10 (-0.41). The parameter estimates for the effects of correct answers 1-5 were not significant with CIs that included 0, and those for the effects of correct answers 6-10 were significant with CIs that excluded 0. Compared to designating "Don't know," designating higher value numbers significantly lowered the odds of IMC failure. The mean difference was tested using a Welch two-sample t-test between the parameter estimates for the lower numbers of correct answers and those for the higher numbers of correct answers, which yielded a significant difference (–0.0609 vs –0.2902, t = 103.51). These results suggest that respondents are predisposed to provide positive responses, and higher value numbers associated with positive meanings are less susceptible to IMC failure. Combining these with the significant mediation results, the conjecture of this study is supported: as CR triggers ARS and in turn, leads respondents to choose higher numbers, IMCs designed with correct answers consistent with the effect of ARS will be less effective in detecting CR and may fail to screen the actual occurrence of CR.

Table 4: IMC failure rates by correct response option

Instruction	Number of Respondents	Number of Respondents Failing	Percentage of Failure
Please select 1	3,193	298	9.33
Please select 2	3,333	283	8.49
Please select 3	3,225	281	8.71
Please select 4	3,288	293	8.91
Please select 5	3,266	270	8.27
Please select 6	3,229	237	7.34
Please select 7	3,276	254	7.75
Please select 8	3,197	230	7.19
Please select 9	3,246	236	7.27
Please select 10	3,135	214	6.83
Please select "Don't know"	3,201	291	9.09
Total	35,589	2,887	8.11

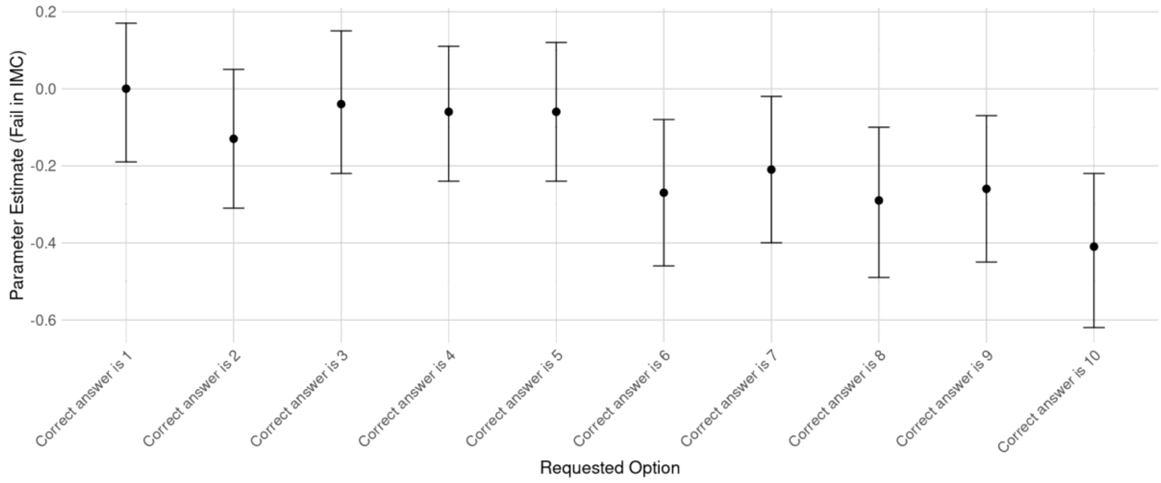

Figure 3: Bootstrapped parameter estimates for IMC failure by correct response option

5. Discussion

Our study demonstrated that CR partially mediates the relationship between demographic factors and ARS. A substantial portion of the variance in ARS was explained by CR. Although extensive research has demonstrated the impact of demographic characteristics on ARS, conflicting results have been found, suggesting a confounding effect of other factors. Our results show that while demographic characteristics predispose respondents to ARS, those characteristics also affect the extent to which respondents earnestly engage in surveys, and this attentiveness level plays a critical role in shaping ARS. This study fills a gap in the extant literature on ARS and CR through a nuanced approach.

The mediation analysis revealed an interesting pattern in which the directions of the direct effect and the indirect effect of education were opposite. Education's positive effect on ARS was offset by its indirect negative effect via CR. While the positive effect of education is not aligned with previous studies, we find clues to this confounding result in intellectual ability. In previous studies, the association between lower education and higher ARS has often been attributed to lower intellectual ability. However, lower intellectual ability is strongly related to a high propensity for CR that results from limited cognitive resources and mental fatigue in processing survey questions. Therefore, the negative impact of education on ARS may actually be driven by CR, and education may thus have a positive effect on ARS after controlling for CR.

We also found that respondents tend to default to positive responses and that the IMC designed with correct answers consistent with the effect of ARS tends to yield lower failure rates. Higher value numbers associated with positive meanings were less likely to result in IMC failure. This result suggests that the power of IMCs can be limited due to CR's unexpected effects. A poorly designed IMC may fail to capture the occurrence of CR, which necessitates the development of careful IMC strategies to measure and address CR.

As CR can be controlled and mitigated through improved survey design, its effects need to be disentangled from the effects of demographics and need to be controlled. While survey designers have little control over respondents' inherent demographic characteristics, better strategies can be implemented to reduce CR. Such tactics, incorporating clear instructions, optimizing question wording, and employing attention checks, can be helpful in balancing respondents' intellectual abilities with survey demands and preventing high cognitive load.

Our findings call for further investigations into the interplay between ARS and CR. Future research needs to consider the interactions among a wider range of factors, such as personality traits, survey topics, survey administration context, and survey environment. It is necessary to explore how these factors intervene in respondents' cognitive processes and thus affect their survey engagement. The mechanisms by which these factors contribute to behavioral response tendencies need to be elucidated.

Acknowledgements

This study was supported by the Research Program funded by the SeoulTech (Seoul National University of Science and Technology).

Artificial intelligence statement: The authors confirm that no generative artificial intelligence tools were used in the conception, writing, analysis, or interpretation of this study.

Ethical statement: This article does not contain any studies with human participants performed by any of the authors. The data in this article was provided by a third-party company that collected the data anonymously and provided participants with a detailed privacy policy and rights. Only individuals who agreed to the policy participated in the survey.

Competing interests: The authors declare that there are no competing interests associated with this publication.

References

Abbey, J.D. and Meloy, M.G. (2017) 'Attention by design: Using attention checks to detect inattentive respondents and improve data quality', *Journal of Operations Management*, 53, pp. 63–70.

Alarcon, G.M. and Lee, M.A. (2022) 'The relationship of insufficient effort responding and response styles: An online experiment', *Frontiers in Psychology*, 12, p. 784375.

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*, 48(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*, 38(2), pp. 143–156.

Beach, D.A. (1989) 'Assessing response patterns: The impact of random responding', *Educational and Psychological Measurement*, 49(4), pp. 897–908.

Berinsky, A.J., Margolis, M.F. and Sances, M.W. (2014) 'Separating the shirkers from the workers? Making sure respondents pay attention on self-administered surveys', *American Journal of Political Science*, 58(3), pp. 739–753.

Biemer, P.P. (2010) 'Total survey error: Design, implementation, and evaluation', *Public Opinion Quarterly*, 74(5), pp. 817–848.

Billiet, J. and McClendon, M.J. (2000) 'Modeling acquiescence in measurement models for two balanced sets of items', *Structural Equation Modeling*, 7(4), pp. 608–628.

Bowling, N.A., Huang, J.L., Bragg, C.B., Khazon, S., Liu, M. and Blackmore, C.E. (2016) 'Who cares and who is careless? Insufficient effort responding as a reflection of respondent personality', *Journal of Personality and Social Psychology*, 111(2), pp. 218–229.

Chylíková, J. (2020) '(Un)expected exception: Validating acquiescent response style factor in the Czech Republic', *International Journal of Social Research Methodology*, 23(6), pp. 677–694.

Curran, P.G. (2016) 'Methods for the detection of carelessly invalid responses in survey data', *Journal of Experimental Social Psychology*, 66, pp. 4–19.

DeSimone, J.A., Davison, H.K., Schoen, J.L. and Bing, M.N. (2018) 'Insufficient effort responding as a meaningful construct and partial function of latent aggression', *Journal of Business and Psychology*, 33(4), pp. 547–561.

DeSimone, J.A., Harms, P.D. and DeSimone, A.J. (2015) 'Best practice recommendations for data screening', *Journal of Organizational Behavior*, 36(2), pp. 171–181.

Greenleaf, E.A. (1992) 'Measuring extreme response style', *Public Opinion Quarterly*, 56(3), pp. 328–351.

Greszki, R., Meyer, M. and Schoen, H. (2015) 'Exploring the effects of removing "too fast" responses and respondents from web surveys', *Public Opinion Quarterly*, 79(2), pp. 471–503.

Groves, R.M. and Lyberg, L. (2010) 'Total survey error: Past, present, and future', *Public Opinion Quarterly*, 74(5), pp. 849–879.

Groves, R.M., Singer, E. and Corning, A. (2000) 'Leverage-salience theory of survey participation: Description and an illustration', *Public Opinion Quarterly*, 64(3), pp. 299–308.

Hamamura, T., Heine, S.J. and Paulhus, D.L. (2008) 'Cultural differences in response styles: The role of dialectical thinking', *Personality and Individual Differences*, 44(4), pp. 932–942.

Han, S., Anderson, C.K. and Chung, K. (2023) 'Do managerial communications improve customer satisfaction and eWOM? The moderating effect of response authenticity', *Humanities and Social Sciences Communications*, 10(1), pp.1-12.

Harzing, A.W. (2006) 'Response styles in cross-national survey research: A 26-country study', *International Journal of Cross Cultural Management*, 6(2), pp. 243–266.

Henninger, M. and Plieninger, H. (2021) 'Different styles, different times: How response times can inform our knowledge about the response process in rating scale measurement', *Assessment*, 28(5), pp. 1301–1319. https://doi.org/10.1177/1073191119900003

Holbrook, A.L., Krosnick, J.A., Moore, D. and Tourangeau, R. (2007) 'Response order effects in dichotomous categorical questions presented orally: The impact of question and respondent attributes', *Public Opinion Quarterly*, 71(3), pp. 325–348.

Huang, J.L., Curran, P.G., Keeney, J., Poposki, E.M. and DeShon, R.P. (2012) 'Detecting and deterring insufficient effort responding to surveys', *Journal of Business and Psychology*, 27(1), pp. 99–114.

Huang, J.L., Liu, M. and Bowling, N.A. (2015) 'Insufficient effort responding: Examining an insidious confound in survey data', *Journal of Applied Psychology*, 100(3), pp. 828–845.

Hui, C.H. and Triandis, H.C. (1985) 'The instability of response sets', *Public Opinion Quarterly*, 49(2), pp. 253–260.

Jackson, D.N. and Messick, S. (1958) 'Content and style in personality assessment', *Psychological Bulletin*, 55(4), pp. 243–252.

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*, 36(2), pp. 264–277.

Johnson, T.P. and Van de Vijver, F.J. (2003) 'Social desirability in cross-cultural research', in Harkness, J.A., van de Vijver, F.J.R. and Mohler, P.P. (eds.) *Cross-cultural survey methods*. Hoboken, NJ: Wiley, pp. 195–204.

Kam, C.C.S. and Meyer, J.P. (2015) 'How careless responding and acquiescence response bias can influence construct dimensionality: The case of job satisfaction', *Organizational Research Methods*, 18(3), pp. 512–541.

Knowles, E.S. and Condon, C.A. (1999) 'Why people say "yes": A dual-process theory of acquiescence', *Journal of Personality and Social Psychology*, 77(2), pp. 379–386.

Krosnick, J.A. (1991) 'Response strategies for coping with the cognitive demands of attitude measures in surveys', *Applied Cognitive Psychology*, 5(3), pp. 213–236.

Krosnick, J.A., Narayan, S. and Smith, W.R. (1996) 'Satisficing in surveys: Initial evidence', *New Directions for Evaluation*, 1996(70), pp. 29–44.

Kung, F.Y., Kwok, N. and Brown, D.J. (2018) 'Are attention check questions a threat to scale validity?', *Applied Psychology*, 67(2), pp. 264–283.

Landsberger, H.A. and Saavedra, A. (1967) 'Response set in developing countries', *Public Opinion Quarterly*, 31(2), pp. 214–229.

Lavrakas, P.J. (2008) *Encyclopedia of survey research methods*. Thousand Oaks, CA: SAGE Publications.

Maniaci, M.R. and Rogge, R.D. (2014) 'Caring about carelessness: Participant inattention and its effects on research', *Journal of Research in Personality*, 48, pp. 61–83.

Marin, G., Gamba, R.J. and Marin, B.V. (1992) 'Extreme response style and acquiescence among Hispanics: The role of acculturation and education', *Journal of Cross-Cultural Psychology*, 23(4), pp. 498–509.

Meade, A.W. and Craig, S.B. (2012) 'Identifying careless responses in survey data', *Psychological Methods*, 17(3), pp. 437–455.

Meisenberg, G. and Williams, A. (2008) 'Are acquiescent and extreme response styles related to low intelligence and education?', *Personality and Individual Differences*, 44(7), pp. 1539–1550.

Narayan, S. and Krosnick, J.A. (1996) 'Education moderates some response effects in attitude measurement', *Public Opinion Quarterly*, 60(1), pp. 58–88.

Nichols, A.L. and Edlund, J.E. (2020) 'Why don't we care more about carelessness? Understanding the causes and consequences of careless participants', *International Journal of Social Research Methodology*, 23(6), pp. 625–638.

Niessen, A.S.M., Meijer, R.R. and Tendeiro, J.N. (2016) 'Detecting careless respondents in web-based questionnaires: Which method to use?', *Journal of Research in Personality*, 63, pp. 1–11.

Oppenheimer, D.M., Meyvis, T. and Davidenko, N. (2009) 'Instructional manipulation checks: Detecting satisficing to increase statistical power', *Journal of Experimental Social Psychology*, 45(4), pp. 867–872.

Pedhazur, E.J. and Schmelkin, L.P. (1991) *Measurement, design, and analysis: An integrated approach*. Hillsdale, NJ: Lawrence Erlbaum Associates.

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*, 88(5), pp. 879–903.

Preacher, K.J. and Hayes, A.F. (2008) 'Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models', *Behavior Research Methods*, 40(3), pp. 879–891.

Roberts, C., Gilbert, E., Allum, N. and Eisner, L. (2019) 'Satisficing in surveys: A systematic review of the literature', *Public Opinion Quarterly*, 83(3), pp. 598–626.

Ross, C.E. and Mirowsky, J. (1984) 'Socially desirable response and acquiescence in a cross-cultural survey of mental health', *Journal of Health and Social Behavior*, 25(2), pp. 189–197.

Schwab, D.P. (2005) *Research methods for organizational studies*. Mahwah, NJ: Lawrence Erlbaum Associates.

Sears, D.O. (1983) 'The person-positivity bias', *Journal of Personality and Social Psychology*, 44(2), pp. 233–250.

Tourangeau, R., Rips, L.J. and Rasinski, K. (2000) *The psychology of survey response*. Cambridge: Cambridge University Press.

van Herk, H., Poortinga, Y.H. and Verhallen, T.M.M. (2004) 'Response styles in rating scales: Evidence of method bias in data from six EU countries', *Journal of Cross-Cultural Psychology*, 35(3), pp. 346–360.

van Vaerenbergh, Y. and Thomas, T.D. (2013) 'Response styles in survey research: A literature review of antecedents, consequences, and remedies', *International Journal of Public Opinion Research*, 25(2), pp. 195–217.

Ward, M.K. and Pond, S.B. (2015) 'Using virtual presence and survey instructions to minimize careless responding on Internet-based surveys', *Computers in Human Behavior*, 48, pp. 554–568.

Weijters, B., Baumgartner, H. and Schillewaert, N. (2013) 'Reversed item bias: An integrative model', *Psychological Methods*, 18(3), pp. 320–334.

Weijters, B., Geuens, M. and Schillewaert, N. (2010) 'The individual consistency of acquiescence and extreme response style in self-report questionnaires', *Applied Psychological Measurement*, 34(2), pp. 105–121.

Weisberg, H.F. (2005) *The total survey error approach: A guide to the new science of survey research*. Chicago, IL: University of Chicago Press.

Zhang, C. and Conrad, F.G. (2014) 'Speeding in web surveys: The tendency to answer very fast and its association with straightlining', *Survey Research Methods*, 8(2), pp. 127–135.