ASSESSING THE IMPACT OF ACQUIESCENCE RESPONSE BIAS ON MARKETING DATA

WILLIAM C. MARTIN, University of North Dakota
BRIAN T. ENGELLAND, The Catholic University of America
JOEL E. COLLIER, Mississippi State University

Although marketing researchers have long been aware of acquiescence response bias (ARB), there are conflicting views regarding its prevalence in and impact on marketing data. In this article, the effects of ARB on survey data are evaluated in four studies using a method of identifying ARB developed by the authors. Results of this investigation indicate that a significant portion of survey respondents may exhibit ARB, though there is wide variance in the incidence of ARB. However, even when ARB is present to a substantial extent, its effects on the psychometric properties of sound construct measures are minimal. This implies that ARB may not be as serious an issue for marketing researchers as has been suggested in prior research. Implications for marketing researchers are put forth, as are directions for future research.

INTRODUCTION

Almost from the inception of attitude measurement, it has been known that survey data are often plagued by respondent acquiescence (Lorge 1937). Acquiescence response bias (ARB) can be problematic for marketing researchers for a variety of reasons, foremost that it "causes respondents in interview or questionnaire situations to misrepresent their true attitudes" (Javeline 1999, p. 2). Like all forms of systematic bias, ARB has the potential to create problems with survey data. Various suggestions have been offered as to how to deal with ARB, though these are seldom heeded by marketing researchers, as noted by the dearth of research which makes any attempt to assess, control, or minimize ARB.

Acquiescence response bias is defined as the propensity of respondents to agree with items without regard to their content (Javeline 1999; Netemeyer, Bearden and Sharma 2003; Ray 1983). Disacquiescence response bias (DARB), the tendency of respondents to disagree with items without regard to their content (Couch and Keniston 1960), can be thought of as the inverse of ARB. Research regarding the extent and type of effects of ARB has produced widely disparate findings. For instance, Rorer (1965) and Gove and Geerken (1977) assert that ARB is neither widespread nor seriously problematic, though Ray (1983) and Heaven (1983) both state that ARB is present in most survey data and has adverse effects. In more recent research, Baumgartner and Steenkamp (2001) purport that ARB can significantly distort survey data, while Billiet and McClendon (2000) assert that ARB has minimal effects on construct measures.

Given this conflict, marketing researchers are left with very few, if any, insights regarding whether and how ARB should be handled. As such, there are two goals for this research: (1) to gauge the prevalence of ARB in marketing survey data and (2) to assess the impact of ARB on this data. We begin by giving a brief review of prior efforts aimed at controlling ARB. Next, we discuss some of the potential negative effects that ARB can have on survey data, along with a method we propose for identifying ARB. The results of four studies in which ARB is examined are then reviewed. Finally, we discuss implications of this research for marketing researchers and directions for future research efforts.

The Marketing Management Journal
Volume 21, Issue 1, Pages 31-46

Previous Efforts in Controlling ARB

Researchers concerned with controlling ARB have typically used combinations of both positively and negatively worded items concurrently in measures, a practice known as balancing a scale. This appears to be the predominant method of explicitly controlling for ARB (e.g., Baumgartner and Steenkamp 2001; Couch and Keniston 1960; Mirowsky and Ross 1991; Ray 1983). The argument for this procedure is that it 'cancels out' the bias of positively and negatively worded statements in that acquiescence to a positively worded item will be negated by acquiescence to a negatively worded item when the items are summed (Baumgartner and Steenkamp 2001; Mirowsky and Ross 1991; Ray 1983).

However, there are three problems with this approach. First, negatively worded items are not usually complete reversals with regard to content, an issue which is more serious when survey respondents are well educated or particularly discriminating (Rorer 1965; Schuman and Presser 1981). Second, responses to positively and negatively worded items tend to be reflective of different dimensions, an effect which often results in considerable degradation of the unidimensionality of scales which include both types of items (Herche and Engelland 1996). Third and most important of all, ARB is a systematic bias, not random error (Spector 1992). Unlike measurement errors, biases such as ARB are not random and do not come from a population whose mean is zero, meaning that the summation of positively and negatively worded items does not reduce the bias present in a scale. The seeming reduction of ARB from the use of balanced scales is highly deceptive in that it comes from simply changing the sign of one biased response against that of another so that the mere appearance of ARB is reduced. Consequently, balancing a scale may not even reduce the appearance of ARB unless there are an equal number of positively and negatively worded items in a scale (e.g., Baumgartner and Steenkamp 2001). In sum, ARB cannot be controlled for or eliminated by simply summating measures which contain both positively and negatively worded items (Javeline 1999; McClendon 1991; Spector 1992), particularly in situations where individual items are used as is often the case when structural equation modeling is used. Thus, another means of eliminating or controlling ARB must be employed.

Effects of ARB on Survey Data

Respondents' acquiescence has been acknowledged to potentially have several adverse effects on survey data. There are at least four ways that ARB can affect these data, one of which is that it falsely intensifies the correlation between similarly worded items (Winkler, Kanouse and Ware Jr. 1982). With the majority of marketing scales utilizing positively worded Likert-type items, ARB is very likely to bias a scale's reliability in an upward direction (herein referred to as Bias 1). Similarly, ARB is likely to increase the observed convergent validity of scales. To the extent that respondents answer a specific construct's items in a uniform pattern, this bias will increase the observed convergent validity for a set of items (Bias 2).

ARB may also result in the underestimation of a measure's discriminant validity. Research has suggested that this may be the case (Billiet and McClendon 2000), but has not empirically examined this possibility. In order for a measure to display discriminant validity, respondents must treat that measure differently than other measures. However, if respondents are significantly influenced by ARB, they may be apt to respond so similarly to disparate measures that the ability of each measure to yield substantially divergent results is reduced (Bias 3). Fourth, ARB may erroneously strengthen relationships between constructs whose indicator items are worded positively and weaken relationships between constructs whose indicator items are worded negatively (Winkler et al. 1982). Since the majority of marketing scales do not include negatively worded items, marketing researchers are most likely to observe construct relationships which

are positively biased when ARB is present (Bias 4).

One would expect, then, that a method of identifying ARB would, in general, result in the observation of these biases. This is particularly relevant for evaluating the method we propose for assessing ARB in the following section.

Proposed Method for Assessing ARB

To assess ARB, the authors recommend employing a method which uses difference scores between positively and negatively worded items of a construct. As stated above, the inclusion of negatively worded items in a psychometric scale has been shown to adversely affect the unidimensionality of the construct that the scale is intended to measure. Hence, such items should not be used for the measurement of latent constructs, but can be used for the control of ARB.

With this method of identifying ARB, the mean level of the positively worded items for a construct is first obtained. Next, negatively worded items related to the same construct are reverse-coded (i.e., when using seven-point response scales, '7' becomes '1', '6' becomes '2', etc.). Lastly, the mean score of the negatively worded items is subtracted from the mean score of the positively worded items. The absolute value of the resulting difference now represents an index of ARB. Small index values are consistent with random error effects, whereas large index values are suggestive of logically inconsistent responses. If a respondent is indeed agreeing with items without regard to their content, then only comparatively large values of this index are suggestive of ARB.

The key question is how large values of this index must be to be truly indicative of ARB. Though varying approaches have been taken with similar indices (e.g., Baumgartner and Steenkamp 2001; Winkler et al. 1982), we argue that two primary considerations must be made. First, the inconsistency of respondents who truly exhibit ARB should be fairly reliable. Accordingly, in order to ensure that inconsistent responses are not merely the result of a single potentially questionable item, we recommend that multiple negatively worded items be included in instruments when utilizing this approach to assess ARB. Second, some leeway is appropriate since a respondent may be able to logically agree with both types of items to a limited extent (Billiet and McClendon 2000; Javeline 1999; Schuman and Presser 1981), some asymmetry with regard to the strength of responses to positively and negatively worded items is to be expected (Herche and Engelland 1996), and no multi-item scale is perfectly reliable. Thus, low index values cannot be asserted to be evidence of ARB, but rather are likely to be due to one or more of the above referenced factors.

Given these needed considerations, potential responses to both positively and negatively worded items should be examined in order to determine what values of this index are likely due to ARB. For instance, when a seven point Likert-type response scale is used, if a respondent indicates moderately strong agreement (denotes '6') with multiple positively and negatively worded items (scoring an ARB index value of 4), such a response is highly illogical and very likely the result of ARB. Conversely, if a respondent indicates only slight agreement with both item types (denotes '5' for each; scoring an ARB index value of 2), such a response cannot conclusively be stated to be the result of ARB. In such an instance, the respondent's agreement with both item types is not strong enough to rule out the possibility that he or she may be logically agreeing with both types of items concurrently. Since the responses of individuals who are truly exhibiting ARB are the result of their acquiescing to the items and not their true attitudes, we recommend that such responses be, in general, eliminated from further analysis. No type or quantity of post hoc statistical analysis can help researchers determine what respondents' 'true' attitudes are in such a situation.

However, identifying a specific cut-off value for this index is inherently somewhat arbitrary, not unlike many of the psychometric standards frequently used in the social sciences (i.e., coefficient alpha, fit indices in structural equation modeling, etc.). Using relatively high cut-off values will result in fewer respondents being removed from a dataset, which helps to preserve the statistical power that marketing researchers obviously desire, while using lower cut-off values may result in fewer responses contaminated by ARB being included in the dataset and more valid data. As such, in Study 1 we empirically examine ARB with a seven point scale using cut-off values of both three and four to examine how varying cut-off values affect the data.

METHOD AND RESULTS

In order to examine the prevalence and effects of ARB, we conducted four studies in which we examine whether the method we developed identifies the biases identified above across various research contexts, samples, and constructs. In Study 1, we investigate the differences between using cut-off values of 3 and 4 for our ARB index. Studies 1, 2, and 3 use survey data primarily gathered from college students; Study 4 uses survey data collected from executives. Varying constructs are used in each study. In Study 3, a semantic differential response scale, rather than a Likert-type response scale, is used to assess ARB.

Study 1

In Study 1, survey data regarding consumers' brand attitudes toward the Apple iPod were utilized. The rationale for choosing this brand is that the Apple iPod would be a brand that young adult respondents would have an active interest towards and would be less likely to acquiesce with questions regardless of content. Four brand related constructs were included: self-brand connection (Escalas and Bettman 2003), brand user identification, perceived connectedness to brand users, and brand commitment (Beatty, Kahle and Homer 1988; Yoo, Donthu and Lee 2000). Brand user identification refers to the degree to which an individual's identity is perceived to overlap with that of the users of a brand; perceived connectedness to brand users (PCBU) is defined as an individual's feeling of being linked to the users of a brand. Scales for these two constructs were developed by one of the authors, judged for face validity by six expert judges, and purified in a pilot study in which they and the two existing scales exhibited unidimensionality, convergent and discriminant validity, and high reliability by established standards. All four of these measures use a seven point Likert-type response scale anchored by "Strongly Disagree" (1) and "Strongly Agree" (7). Two negatively worded items related to PCBU were included for the purpose of assessing ARB.

Data were gathered via invitations to participate in the survey which were posted in online forums oriented around the Apple iPod and emailed to the students of a major university in the U.S. Entry into a drawing for $25 iTunes gift certificates was used as an incentive for participation. This resulted in 862 responses.

We used the above referenced method for assessing ARB to identify respondents exhibiting ARB. In doing so, we used two different cut-off values, three and four. Using a cut-off value of three, 223 respondents were identified as exhibiting ARB; using a cut-off value of four, 147 respondents were identified as such. We then split the sample into three groups: a 'total' group, which included all of the responses, a 'less restricted' group which included only the responses for which the absolute value of the ARB index was less than four, and a 'more restricted' group which included only the responses for which the absolute value of the ARB index was less than three. Consequently, the size of the total group was 862, the size of the less restricted group was 715, and the size of the more restricted group was 639. To investigate whether our method identified responses which biased the data as expected, the reliability, convergent validity, discriminant validity, and construct

relationships among all four scales were examined by comparing these groups.

First, the less restricted group and the more restricted groups were compared to one another to examine whether there were substantial differences between using different cut-off values of the ARB index. The reliability of the measures was assessed using coefficient alpha. For all four constructs, coefficient alpha was greater in the less restricted dataset, though the differences in each instance were relatively small. Thus, the data are exhibiting Bias 1 and support the validity of our method of assessing ARB. Next, the convergent validity of the scales was assessed by examining the differences in the standardized loading estimates for each of the items, as well as the average variance extracted (AVE) for each scale in three confirmatory factor analyses (CFAs), one among the total group (χ^2 = 667.23, df = 98, p <.001; RMSEA = .082; CFI = .97; NNFI = .96), one among the less restricted group (χ^2 = 575.73, df = 98, p <.001; RMSEA = .083; CFI = .96; NNFI = .95), and one among the more restricted group (χ^2 = 558.12, df = 98, p <.001; RMSEA = .086; CFI = .96; NNFI = .95). The results, including the coefficient alpha of each measure, are shown in Table 1.

Across the four construct measures investigated, the difference in coefficient alpha between the less restricted and the more restricted groups ranged from zero to .005. Differences in the standardized loading estimates for each item ranged from zero to .03; for eight of the sixteen items, there was no difference in these estimates. AVE was nearly identical in both groups; differences ranged from zero to .01. The discriminant validity of the measures was assessed by comparing the AVE of each measure to its squared correlation coefficient with each other measure (Fornell and Larcker 1981). To compare the differences in the discriminant validity of each construct between the two groups, the ratio of each construct's AVE to its squared correlation coefficient with the other constructs was examined for all four constructs. For instance, in the less restricted group, self-brand connection and brand user identification had ratios of 2.08 and 2.06. Eight of these ratios were larger in the more restricted group, two were equal, and two were smaller; the largest difference was .62.

Taken together, there are slight differences between these two groups, though we conclude that they are too small to justify eliminating such a large number of responses, as is the case with the more restricted group. The less restricted group consists of 82.9 percent of the original sample, while the more restricted group only consists of 74.1 percent of the total responses. Since some responses which are not truly exhibiting ARB may be eliminated from the more restricted group, further reducing the sample size by nearly nine percent is not justified by the miniscule differences such an action yields. As such, we only investigate differences between the total group and the less restricted group hereafter.

Coefficient alpha was smaller in the less restricted group than in the total group; differences ranged from .005 to .021. Standardized loading estimates were higher for fourteen of the sixteen items in the total group; differences ranged from zero to.04. AVE for each measure ranged from .02 to .07 higher in the total group. Finally, the measures' discriminant validity was investigated by examining the standardized construct correlation matrices for both of the CFAs. These matrices are shown in Table 2.

The AVE/squared correlation ratios were higher for every pair of constructs in the less restricted group. Though the AVE of each construct was smaller in the restricted group than in the total group, the relationships between the constructs were reduced to a greater extent. The size of the correlation estimates between constructs was diminished by .02 to .08 in the less restricted group.

In sum, we find that the data exhibit all of the expected biases, which suggests that our method of identifying ARB is performing as

TABLE 1
Standardized Loading Estimates for Study 1

Scale	Total Group (N = 862)		Less Restricted Group (N = 715)		More Restricted Group (N = 639)	
	Coefficient Alpha	Standardized Loading Estimates	Coefficient Alpha	Standardized Loading Estimates	Coefficient Alpha	Standardized Loading Estimates
Self-Brand Connection	.937		.923		.921	
-This brand reflects who I am.		.86		.83		.81
-I can use this brand to communicate who I am to other people.		.87		.83		.83
-I think that this brand could help me become the type of person I want to be.		.90		.88		.88
-I consider this brand to be "me."		.93		.92		.92
Brand User Identification	.937		.918		.913	
-Most of the people who use this brand have a nature that is very much like mine.		.86		.83		.82
-The identity of the people who use this brand is almost identical to my own.		.92		.90		.89
-When I think of the people who use this brand, I think of myself.		.85		.80		.79
-My identity is very similar to that of the people who use this brand.		.93		.91		.91
PCBU	.967		.946		.946	
-I feel linked to those who use this brand.		.92		.88		.88
-I sense a bond with others who use this brand.		.92		.88		.85
-I sense a connection with those who use this brand.		.95		.92		.91
-When I think about this brand, I feel attached to those who use it.		.95		.92		.93
Brand Commitment	.929		.924		.926	
-I consider myself to be highly committed to this brand.		.96		.96		.96
-I feel strongly devoted to this brand.		.95		.95		.95
-Even if another brand were less expensive, I would always purchase this brand.		.80		.79		.79
-This brand would be my first choice of brands in this product category.		.75		.74		.75

theoretically expected. However, based on this study, the differences in the data caused by eliminating responses contaminated by ARB are very slight. Also, we conclude from this study that the differences produced in the data by utilizing an ARB index cutoff value of three, rather than four, are so minute that they do not justify eliminating additional responses from the dataset, nearly 9 percent of the original sample in this study, compared to the less restricted group.

Study 2

In Study 2 a different context was used to assess if our proposed method for capturing ARB would extend past the initial setting of Study 1 and to evaluate what effect ARB had on other data. In this study, survey data regarding Internet shopping behavior and price-related attitudes were used. Four constructs and measures were included in this study: consumer novelty seeking (Manning et al. 1995); market maveness (Feick and Price 1987): compulsive

TABLE 2
Construct Correlation Matrix across Groups for Study 1

	Self-Brand Connection	Brand User Identification	PCBU	Brand Commitment
Self-Brand Connection	**.79**	.46	.50	.36
Brand User Identification	.68 *14.56*	**.80**	.58	.21
PCBU	.71 *15.42*	.76 *16.03*	**.88**	.26
Brand Commitment	.60 *14.03*	.46 *11.40*	.51 *12.81*	**.77**

Standardized Construct Correlation Matrix – Total Group

	Self-Brand Con-nection	Brand User Identification	PCBU	Brand Commitment
Self-Brand Connection	**.75**	.38	.40	.34
Brand User Identification	.62 *12.12*	**.74**	.52	.14
PCBU	.63 *12.62*	.72 *13.58*	**.81**	.21
Brand Commitment	.58 *12.21*	.38 *8.80*	.46 *10.54*	**.75**

Standardized Construct Correlation Matrix – Less Restricted Group
Standardized Construct Correlation Matrix – More Restricted Group

Note: diagonal values represent the AVE for each construct; values below the diagonal are correlation coefficient estimates with t-values italicized on the line below; values above the diagonal are squared correlation coefficient estimates.

buying (Valence et al. 1988); and value consciousness (Lichtenstein et al. 1990). Two negatively worded items related to consumer novelty seeking were included for the purpose of assessing ARB.

Respondents consisted of a convenience sample of 518 subjects from two universities and included students enrolled in marketing and marketing communications courses. Students were offered extra credit in return for their participation in the study. Participants were required to view a specific website, seek out certain information, and then answer questions in an online format.

Using the ARB index value of four or greater as the criterion for identifying ARB, seven of the 518 responses were classified as such. The sample was separated into two groups: a total group in which all of the responses were included and a restricted group in which only the responses which were not classified as exhibiting ARB. Consequently, the size of the total group was 518 and the size of the restricted group was 511. The reliability, convergent validity, discriminant validity, and construct relationships among all four scales were examined by comparing the total and restricted groups. In addition to computing coefficient alpha, two CFAs were conducted, one in the total group (χ^2 = 1956.15, df = 399, p <.001; RMSEA = .087; CFI = .87; NNFI = .86) and one in the restricted group (χ^2 = 1925.90, df = 399, p <.001; RMSEA = .087; CFI = .87; NNFI = .86). The results are shown in Table 3.

Coefficient alpha was from .002 to .013 greater in the total group. Standardized loading estimates were between .01 and .03 higher for sixteen of the items in the total group, thirteen were identical in both groups, and one was higher in the restricted group. The AVE of three of the four scales was higher in the total group; differences ranged from .01 to .02. Lastly, the standardized construct correlation matrices for both of the CFAs were investigated. These matrices are shown in Table 4.

A comparison of the ratio of each construct's AVE to its squared correlation coefficient with the other constructs was then conducted. This revealed that the AVE/squared correlation ratios were higher for every pair of constructs in the restricted group. Further, the construct correlation estimates were stronger in every instance in the total group.

As with Study 1, our method of identifying ARB is performing as expected, though again there are only minimal differences in the data when responses contaminated by ARB are removed from the dataset, likely due in large part to the samples being almost identical since only 1 percent of the original sample was categorized as displaying ARB, a much smaller portion than was observed in Study 1.

Study 3

To broaden the generalizability of our method of assessing ARB, different scale formats were utilized in Study 3. Survey data regarding consumers' word-of-mouth were utilized. Four constructs were included in this study: source trustworthiness (Pornpitakpan 2004), attitude toward the product (Iyer 1988; MacKenzie and Lutz 1989), purchase intentions (Putrevu and Lord 1994), and purchase involvement (Zaichowsky 1985). With the exception of purchase intentions, which was measured using a seven-point Likert-type response scale anchored by "Strongly Disagree" and "Strongly Agree," all of the measures utilized a seven-point semantic differential response scale. Two negatively worded items related to purchase involvement were included in order to make use of our method of assessing ARB.

Data were gathered via the student referral method (Babin et al. 2003) from students at a major university in the U.S., resulting in the collection of 589 responses. Randomly contacted respondents verified the validity of the responses. Six responses with missing data were eliminated from the dataset.

Using the method for assessing ARB, thirty-seven responses were identified as exhibiting ARB. The sample was again separated into two groups: a total group which included all of the responses, and a restricted group which included only the responses for which the absolute value of the ARB index was less than four. As a result, the size of the total group was 583 and the size of the restricted group was 546. Coefficient alpha was observed and two CFAs were conducted, one in the total group (χ^2 = 575.62, df = 129, p <.001; RMSEA = .077; CFI = .96; NNFI = .95) and one in the restricted group (χ^2 = 571.44, df = 129, p <.001; RMSEA = .079; CFI = .95; NNFI = .94). The results are shown in Table 5.

TABLE 3
Coefficient Alpha and Standardized Loading Estimates across Groups for Study 2

Scale	Total Group (N = 518)		Restricted Group (N = 511)	
	Coefficient Alpha	Standardized Loading Estimates	Coefficient Alpha	Standardized Loading Estimates
Consumer Novelty Seeking	.877		.864	
- I often seek out information about new products and brands.		.66		.63
- I like to go places where I will be exposed to information about new products and brands.		.86		.85
- I like magazines that introduce new brands. I frequently look for new products and services.		.83		.82
- I am continually seeking new product experiences.		.81		.80
- I take advantage of the first available opportunity to find out about new and different products.		.69		.67
Market Maveness	.913		.911	
-I like introducing new brands and products to my friends.		.73		.73
-I like helping people by providing them with information about many kinds of products.		.78		.78
-People ask me for information about products, places to shop, or sales.		.82		.82
-If someone asked where to get the best buy on several types of products, I could tell him or her where to shop.		.84		.83
-My friends think of me as a good source of information when it comes to new products and sales.		.87		.87
-Think about a person who has information about a variety of products and likes to share this information with others. This person knows about new products, stores, sales, and so on, but does not necessarily feel that he or she is an expert on one particular brand. Do you think this description fits you very well?		.74		.73
Compulsive Buying	.901		.899	
-When I have money, I can't help but spend part or all of it.		.63		.63
-I am often impulsive in my behavior.		.63		.63
-For me, shopping is a way of facing the stress of my daily life and of relaxing.		.83		.83
-I sometimes feel that something inside me pushed me to go shopping.		.86		.85
-There are many times when I have a strong urge to buy.		.77		.77

Continued

TABLE 3 (*continued)*
Coefficient Alpha and Standardized Loading Estimates across Groups for Study 2

Scale	Total Group		Restricted Group	
	Coefficient Alpha	Standardized Loading Estimates	Coefficient Alpha	Standardized Loading Estimates
Compulsive Buying (*continued)*	.901		.899	
-At times I have felt guilty after buying because it seemed unreasonable.		.52		.51
- There are some things I buy that I do not show to anybody for fear of being perceived as irrational in my buying behavior.		.45		.45
- I often have an unexplainable urge, a sudden and spontaneous desire, to go and buy something in a store.		.84		.84
- As soon as I enter a shopping center, I have an irresistible urge to buy something.		.84		.85
- I am one of those people who often respond to direct mail offers.		.40		.40
- I have often bought a product I don't need, while knowing that I have very little money left.		.65		.65
- I have sometimes thought, "if I had to do it all over again, I would.." and felt sorry for something I have done or said.		.32		.31
Value Consciousness	.910		.904	
- I am very concerned about low prices, but I am concerned about product quality.		.76		.76
- When grocery shopping, I compare the prices of different brands to be sure I get the best value for the money.		.81		.80
- When purchasing a product, I always try to maximize the quality I get for the money I spend.		.86		.85
- When I buy products, I like to be sure that I am getting my money's worth.		.87		.86
- I generally shop around for lower prices on products, but they still must meet certain quality requirements before I will buy them.		.85		.84
- When I shop, I usually compare the "price per ounce" information for brands I normally buy.		.51		.50
- I always check prices at the grocery store to be sure I get the best value for the money I spend.		.78		.77

Coefficient alpha was from .001 to .005 greater in the total group for three of the four measures; there was no difference in coefficient alpha for one of the measures. Standardized loading estimates were .01 higher for eight of the eighteen items in the total group, identical for nine of the items, and .01 higher for one item in the restricted group. AVE was equal for three of the four measures and was .01 higher for one of the scales in the total group. The standardized construct correlation matrices for both of the CFAs were then examined. These matrices are shown in Table 6.

A comparison of the ratio of each construct's AVE to its squared correlation coefficient with the other constructs revealed that of the six AVE/squared correlation ratios present, only

TABLE 4
Construct Correlation Matrix across Samples for Study 2

	Consumer Novelty Seeking	Market Maveness	Compulsive Buying	Value Consciousness
Consumer Novelty Seeking	**.60**	.36	.13	.09
Market Maveness	.60 *18.11*	**.64**	.06	.07
Compulsive Buying	.39 *9.39*	.25 *5.57*	**.45**	.00
Value Consciousness	.30 *6.73*	.26 *5.89*	.05 *1.06*	**.61**

Standardized Construct Correlation Matrix in Total Group
Standardized Construct Correlation Matrix in Restricted Group

	Consumer Novelty Seeking	Market Maveness	Compulsive Buying	Value Consciousness
Consumer Novelty Seeking	**.58**	.34	.14	.07
Market Maveness	.58 *16.80*	**.63**	.05	.05
Compulsive Buying	.37 *8.60*	.23 *5.02*	**.45**	.00
Value Consciousness	.27 *5.80*	.23 *5.02*	.06 *.56*	**.60**

Note: diagonal values represent the AVE for each construct; values below the diagonal are correlation coefficient estimates with t-values italicized on the line below; values above the diagonal are squared correlation coefficient estimates.

one was greater in the restricted group; the others were higher in the total group, contradictory to our expectations. In addition, only one of the construct correlation estimates was stronger in the total group; the rest were stronger in the restricted group.

In this study, we find further support for our method in that ARB is increasing the measures' reliability and convergent validity. Six percent of the respondents in this study were found to exhibit ARB. Interestingly though, we do not find in this study that ARB weakens the measures' discriminant validity or strengthens the relationships between constructs. In all cases, however, the effect of ARB on the data is very minimal. Many of the indicators we examined remained unchanged whether responses exhibiting ARB were present in the data or not.

Study 4

In Study 4, the proposed method for assessing and controlling ARB was examined in a sample that did not include students. This study included survey data on product development executives' views regarding innovativeness in new product development. Five construct measures were included in this study: organizational new product innovativeness, using a measure developed by one of the authors; organizational innovativeness (Miller and Friesen 1982); personal innovativeness (Oliver and Bearden 1985); organizational analysis (Venkatraman 1989); and performance satisfaction (Venkatraman 1989). Six negatively worded items related to organizational new product innovativeness were included for the purpose of testing for ARB.

TABLE 5
Coefficient Alpha and Standardized Loading Estimates across Groups for Study 3

Scale	Total Group (N = 583)		Restricted Group (N = 546)	
	Coefficient Alpha	Standardized Loading Estimates	Coefficient Alpha	Standardized Loading Estimates
Source Trustworthiness	.855		.853	
-Very Untrustworthy/Very Trustworthy		.80		.80
-Very Insincere/Very Sincere		.86		.86
-Very Dishonest/Very Honest		.85		.84
-Very Unpredictable/Very Predictable		.61		.61
Attitude Toward the Product	.905		.905	
-Bad/Good		.80		.81
-Unfavorable/Favorable		.91		.91
-Unpleasant/Pleasant		.90		.89
-Harmful/Beneficial		.78		.77
Purchase Intentions	.959		.958	
-It is very likely that I will buy this product.		.92		.91
-I will definitely try this product.		.88		.88
-I plan on buying this product.		.97		.97
-I will certainly purchase this product.		.94		.94
Purchase Involvement	.960		.955	
-Unimportant/Important		.83		.82
-Of No Concern/Of Concern to Me		.89		.88
-Irrelevant/Relevant		.92		.91
-Meant Nothing to Me/Meant a Lot to Me		.87		.86
-Didn't Matter/Mattered to Me		.92		.92
-Insignificant/Significant to Me		.93		.93

Data were gathered by mailing a questionnaire booklet to 680 product development executives working for major corporations in the electronics, transportation, and food products industries. After reminders, 200 usable responses were received, resulting in a 29.4 percent response rate.

Surprisingly, no responses had an ARB index value greater than four. This means that apparently no responses were contaminated by ARB. While the precise reason for this is unknown, there are some interesting possibilities. For one, the tendency for ARB to occur may be less among paper-and-pencil questionnaires than online questionnaires. Another factor may be the previously cited relationship between ARB and education level (Heaven 1983; Schuman and Presser 1981; Ware 1978) or income (Ware 1978). Consequently, students with less education and lower income may be more likely to exhibit ARB than business executives. Business executives' interest and involvement with relevant research topics may also reduce ARB.

CONCLUSION

Prior research regarding acquiescence response bias has produced a gamut of findings. Some researchers assert that ARB is widespread, while some claim that it is virtually nonexistent. Its effects on survey data have been purported by some to be extensive, whereas others state that its effects are only negligible. This has left marketing researchers wondering what the effects of ARB are on their data and whether and how they should attempt to assess and control it. In this article, we investigate the prevalence of acquiescence response bias in

TABLE 6
Construct Correlation Matrix across Samples for Study 3

	Source Trustworthiness	Attitude Toward the Product	Purchase Intentions	Purchase Involvement
Source Trustworthiness	**.62**	.34	.13	.12
Attitude Toward the Product	.58 *17.91*	**.72**	.30	.24
Purchase Intentions	.36 *9.26*	.55 *17.73*	**.86**	.18
Purchase Involvement	.35 *8.91*	.49 *14.42*	.43 *12.04*	**.80**

Standardized Construct Correlation Matrix in the Total Group
Standardized Construct Correlation Matrix in the Restricted Group

	Source Trustworthiness	Attitude Toward the Product	Purchase Intentions	Purchase Involvement
Source Trustworthiness	**.62**	.38	.14	.15
Attitude Toward the Product	.62 *19.80*	**.72**	.31	.26
Purchase Intentions	.38 *9.62*	.56 *17.72*	**.86**	.18
Purchase Involvement	.39 *9.66*	.51 *14.73*	.42 *11.56*	**.78**

Note: diagonal values represent the AVE for each construct; values below the diagonal are correlation coefficient estimates with t-values italicized on the line below; values above the diagonal are squared correlation coefficient estimates.

marketing data and examine its impact on those data. A method of identifying ARB among survey responses is developed and shown to perform as theoretically expected. Further, a variety of marketing measures, response scale formats, and sample populations are investigated across four studies.

Based on these four studies, we make a number of conclusions. First, the method that we have developed to identify ARB, which makes use of differences between responses to positively and negatively worded items, appears to be performing as theoretically expected. Using this method, we find, in general, that the presence of ARB appears to increase the reliability and convergent validity of measures, as well as the relationships between constructs, while it diminishes measures' discriminant validity.

Second, the frequency of ARB among survey respondents varies considerably. In our four studies, the rate of ARB's occurrence varied from zero to 17 percent of participants. Interestingly, we observed ARB in each of our studies which included data obtained from college students, though we found no ARB among survey data gathered from business executives. While some efforts can undoubtedly be made to reduce this bias, they are much more apt to be related to the design of the survey than the content of the items since ARB is, by definition, unaffected by item content. However, even when the survey's topic is of presumed keen interest to respondents, as was the case in Study 1, ARB may still arise.

Third, even when ARB is present to a sizeable extent, its effects on the data are quite small. When responses which exhibited ARB were

present, coefficient alpha increased in size, but by a mean level of only .008. Similarly, standardized loading estimates averaged .01 higher when ARB was present. Discriminant validity was only slightly lessened in the presence of ARB; correlations between constructs were stronger by a mean level of .018. Such trivial differences call into serious question whether removing as many as 17 percent of respondents from a sample is justified by the resultant loss of statistical power. Given these findings, we conclude that while ARB appears to be present to a varying extent in marketing survey data, it does not seem likely to result in the distortion of psychometrically sound construct measures such that eliminating responses which appear to exhibit ARB can be rationalized.

FUTURE RESEARCH AND LIMITATIONS

Future research should seek to determine the generalizability of our findings regarding ARB, particularly regarding whether ARB has a significant impact on marketing data in instances other than those we examined. In this research endeavor, we did not assess the factors which may lead respondents to engage in ARB, such as their motivation to participate in the surveys. It may well be that ARB stemming from different sources produces varying effects. An intriguing possibility is that two different types of ARB exist. For instance, it is possible that one form of ARB may indeed stem from respondent-based characteristics, such as education, age, income, or culture, while another form may be attributed to item-based characteristics, such as an item's form or content. This may be of keen interest to marketing researchers since the features of items can be controlled with relative ease, though controlling ARB which emanates from respondents may be exceedingly difficult.

In our examination of ARB, we only included negatively worded items that were relevant to a single construct measure. Future research should examine how the inclusion of negatively worded items related to multiple construct measures might affect the appraisal of the ARB present in data. In addition, we did not investigate the potential impact that placement of the negatively worded items in the questionnaire might have on ARB. It is possible that a respondent might not exhibit ARB in the early stages of completing a questionnaire, but may do so during the course of completing the questionnaire. This may be more likely for lengthy questionnaires since respondents may grow weary and become disinterested in completing such questionnaires (DeVellis 2003). Further research is needed to examine how the placement of negatively worded items in a questionnaire as well as questionnaire length may affect the assessment of ARB.

REFERENCES

Baumgartner, Hans and Jan-Benedict E. M. Steenkamp (2001), "Response Styles in Marketing Research: A Cross-National Investigation," *Journal of Marketing Research*, Vol. 38, May, 143-56.

Beatty, Sharon E., Lynn R. Kahle, and Pamela Homer (1988), "The Involvement-Commitment Model: Theory and Implications," *Journal of Business Research*, Vol. 16, March, 149-67.

Billiet, Jaak B. and McKee J. McClendon (2000), "Modeling Acquiescence in Measurement Models for Two Balanced Sets of Items," *Structural Equation Modeling*, Vol. 7, 4, 608-28.

Couch, Arthur and Kenneth Keniston (1960), "Yeasayers and Naysayers: Agreeing Response Set as a Personality Variable," *Journal of Abnormal and Social Psychology*, Vol. 60, 2, 151-72.

DeVellis, Robert F. (2003), *Scale Development: Theory and Applications*. Thousand Oaks: Sage Publications, Inc.

Feick, Lawrence F. and Linda L. Price (1987), "The Market Maven: A Diffuser of Marketplace Information," *Journal of Marketing*, Vol. 51, January, 83-97.

Fornell, Claes and David F. Larcker (1981), "Evaluating Structural Equation Models with Unobservable Variables and Measurement Error," *Journal of Marketing Research*, 18 (February), 39-50.

Gove, Walter R. and Michael R. Geerken (1977), "Response Bias in Surveys of Mental Health: An Empirical Investigation," *American Journal of Sociology*, Vol. 82, 6, 1289-317.

Heaven, Patrick C. L. (1983), "Authoritarianism or Acquiescence? South African Findings," *Journal of Social Psychology*, Vol. 119, 11-15.

Herche, Joel and Brian T. Engelland (1996), "Reversed-Polarity Items and Scale Unidimensionality," *Journal of the Academy of Marketing Science*, Vol. 24, 4, 366-74.

Javeline, Debra (1999), "Response Effects in Polite Cultures: A Test of Acquiescence in Kazakhstan," *Public Opinion Quarterly*, Vol. 63, 1-28.

Krosnick, J.A. (1991), "Response Strategies for Coping with the Cognitive Demands of Attitude Measures in Surveys," *Applied Cognitive Psychology*, Vol. 5, 213-36.

Lichtenstein, Donald R., Richard G. Netemeyer, and Scot Burton (1990), "Distinguishing Coupon Proneness from Value Consciousness: An Acquisition-Transaction Utility Theory Perspective," *Journal of Marketing*, Vol. 54, 54-67.

Lorge, Irving (1937), "Gen-Like: Halo or Reality," *Psychological Bulletin*, Vol. 34, 545-46.

Manning, Kenneth C., William O. Bearden and Thomas J. Madden (1995), "Consumer Innovativeness and the Adoption Process," *Journal of Consumer Psychology*, Vol. 4, 4, 329-345.

McClendon, McKee (1991), "Acquiescence and Recency Response-Order Effects in Interview Surveys," *Sociological Methods and Research*, Vol. 20, August, 60-103.

Miller, Danny and Peter H. Friesen (1982), "Innovation in Conservative and Entrepreneurial Firms: Two Models of Strategic Momentum," *Strategic Management Journal*, Vol. 3, 1-25.

Mirowsky, John and Catherine E. Ross (1991), "Eliminating Defense and Agreement Bias from Measures of Sense of Control: A 2 X 2 Index," *Social Psychology Quarterly*, Vol. 54, 2, 127-45.

Netemeyer, Richard G., William O. Bearden, and Subhash Sharma (2003), *Scaling Procedures: Issues and Applications*. Thousand Oaks, CA: Sage Publications, Inc.

Oliver, Richard L. and William O. Bearden (1985), "Crossover Effects in the Theory of Reasoned Action: A Moderating Influence Attempt," *Journal of Consumer Research*, Vol. 12, December, 324-340.

Ray, John J. (1979), "Is the Acquiescent Response Style Not So Mythical After All? Some Results From a Successful Balanced F Scale," *Journal of Personality Assessment*, Vol. 43, 638-43.

_____ (1983), "Reviving the Problem of Acquiescent Response Bias," *Journal of Social Psychology*, Vol. 121, 81-96.

Rorer, Leonard G. (1965), "The Great Response-Style Myth," *Psychological Bulletin*, Vol. 63, March, 129-59.

Schriesheim, Chester A. and Kenneth D. Hill (1981), "Controlling Acquiescence Response Bias By Item Reversals: The Effect on Questionnaire Validity," *Educational and Psychological Measurement*, Vol. 41, 4, 1101-14.

Schuman, Howard and Stanley Presser (1981), *Questions and Answers in Attitude Surveys: Experiments on Question Form, Wording and Content*. New York: Academic Press.

Spector, Paul E. (1992), *Summated Rating Scale Construction*. Newbury Park, CA: Sage Publications, Inc.

Valence, Gilles, Alain d'Astous and Louis Fortier (1988), "Compulsive Buying: Concept and Measurement," *Journal of Consumer Policy*, Vol. 11, 419-433.

Venkatraman, N. (1989), "Strategic Orientation of Business Enterprises: The Construct, Dimensionality and Measurement," *Management Science*, Vol. 35, August, 942-962.

Ware, John E. Jr. (1978), "Effects of Acquiescent Response Set on Patient Satisfaction Ratings," *Medical Care*, Vol. 16, 327-36.

Winkler, John D., David E. Kanouse, and John E. Ware Jr. (1982), "Controlling for Acquiescence Response Set in Scale Development," *Journal of Applied Psychology*, Vol. 67, 5, 555-61.

Yoo, Boonghee, Naveen Donthu, and Sungho Lee (2000), "An Examination of Selected Marketing Mix Elements and Brand Equity," *Journal of the Academy of Marketing Science*, Vol. 28, 2, 195-211.