ARTICLE

<https://doi.org/10.1057/s41599-020-00665-x>

OPEN

Consistently biased: documented consistency in self-reported holiday healthfulness behaviors and associated social desirability bias

Courtney Bir ¹✉ & Nicole Olynk Widmar ²

Holiday healthfulness conversations are dominated by overindulgence of consumption and then, largely in reference to resolutions to do better, physical activity, and exercise aspirations. Consistency was found in self-reported agreement with a series of holiday healthfulness statements, across time, holidays (Thanksgiving versus Christmas), and samples of respondents. The largest proportion of respondents displaying social desirability bias (SDB) were found in response to two statements, namely “I will consume more alcohol during the holiday season than at other times of the year” at (63–66%) and “I make it a New Year’s Resolution to lose weight” (60–63%). Cheap talk was tested as a mechanism to reduce SDB in holiday healthfulness reporting, but showed only limited efficacy compared to the control group surveyed simultaneously. Nonetheless, the consistency across time in reporting and SDB are notable in both self-reporting of health-related data and in studying a unique consumption period around the holidays. Healthcare providers and researchers alike seek to improve the accuracy of self-reported data, making understanding of biases in reporting on sensitive topics, such as weight gain and eating over the holiday season, of particular interest.

¹Oklahoma State University, Stillwater, OK, USA. ²Purdue University, West Lafayette, IN, USA. ✉email: courtney.bir@okstate.edu

Introduction

Holiday eating is frequently associated with excess. An average holiday meal in the United States is between 3000 and 4500 calories (Jampolis, 2018), while Americans, on average, eat 2481 calories per day (Rehkamp, 2016). Ma et al. (2006) found that during the fall season, daily caloric intake was 86 kcal higher than the spring. Holiday season indulgence results in an average annual gain in bodyweight between mid-November to mid-January ranging from 0.4 to 0.7 kg (Schoeller, 2014). Recommendations for limiting or decreasing holiday weight gain include weighing yourself every day (Kaviani et al., 2019), decreasing consumption by reflecting on the exercise required to counteract the calorie count (Mason et al., 2018), and increasing exercise as part of New Year's resolutions (Hawkes, 2016). Conversely, Stevenson et al. (2013) found that exercise did not prevent holiday weight gain, and was not a significant predictor of body weight changes. Researchers are interested in the behaviors of people during the holiday season. However, due to social desirability bias (SDB), it can be difficult to assure that self-reporting on behaviors including eating, exercise, or even holiday spending, are reflective of reality.

SDB has long been recognized in psychology, being defined by Maccoby and Maccoby (1954). SDB occurs when in a sub-conscious effort to make themselves look better, a respondent answers a question in a way that deviates from their true behavior towards a real or perceived socially "correct" answer (Maccoby and Maccoby, 1954; Fisher, 1993). Many ways to combat this pervasive SDB phenomenon have been established. Post data collection calibration methods for SDB include the use of the Marlowe–Crowne Social Desirability Scale, which is a set of 13 questions chosen to establish how likely a person is to express SDB (Crowne and Marlowe, 1960). Based on the scale, researchers can set a threshold by which they can estimate models to adjust results; however, these methods are complicated (Crino et al., 1983). Furthermore, the additional questions involved in measuring SDB lengthen the survey instrument and may contribute to survey fatigue, which can result in decreased response rate and data quality (Galesic and Bosnjak, 2009). Another method to combat SDB is indirect questioning. By asking respondents what they believe the average person does, the respondent is likely to project their beliefs and evaluations when responding, without the social pressure associated with revealing one's own actions (Fisher, 1993). Fisher (1993) found that indirect questioning mitigates SDB without systematically affecting the means of questions that were not socially sensitive. This method of combating SDB has been used in a wide range of areas beyond psychology, such as the importance of environmental performance of automobiles (Johansson-Stenman et al., 2006), public goods (Lusk and Norwood, 2009), meat products (Olynk et al., 2010), and pet acquisition (Bir et al., 2018).

Various methods of combatting, or seeking to limit biases in data collection, have been developed, such as using a cheap talk statement to attempt to mitigate hypothetical bias. The term cheap talk was originally coined in game theory literature in reference to a costless transition of signals and information that does not affect the payoffs of the game (Farrell and Gibbons, 1989; Matthews et al., 1991; Farrell and Rabin, 1996). Cummings and Taylor (1999) built on that concept to determine a method to prevent hypothetical bias when asking respondents about hypothetical purchasing decisions in contingent valuation, as opposed to correcting for the bias post data collection. Their version of cheap talk is an explicit discussion of hypothetical bias through reference to budget constraints and budgetary substitutes prior to asking a respondent to make hypothetical choices (Cummings and Taylor, 1999). Inclusion of a cheap talk script in their experiments resulted in responses that were indistinguishable

between hypothetical valuation and valuation questions involving actual payments (Cummings and Taylor, 1999).

The objectives of this research are (1) to evaluate the prevalence of SDB as related to holiday eating habits, with data collected between the Thanksgiving and Christmas holidays in the U.S. and (2) evaluate the impact of a cheap talk statement designed to increase awareness of SDB, and mitigate its impact, prior to the presentation of SDB prone questions. There are likely differences in the prevalence of SDB and potential impact of a cheap talk statement for different subjects. Holiday eating habits were chosen for this analysis because there were two recent studies conducted using identical statements to measure SDB to which these results could be directly compared too. Additionally, holiday eating and New Years resolution setting in the U.S. remain topics with consistent social interest, providing a good first measure for the cheap talk experiment. The benefits of a cheap talk statement for SDB mitigation would include decreasing the need for post-data collection adjustments for SDB or the inclusion of SDB scale establishment questions which would lengthen survey instruments and contribute to fatigue of respondents.

Results

General demographics and shopping behavior. The full sample, and the two subsamples (cheap talk and no cheap talk, henceforth control) all had higher proportions of respondents with residence in the South, and lower proportions of respondents from the Midwest or West, and did not graduate from high school when compared to the U.S. population via the U.S. Census (U.S. Census Bureau, 2016) (Table 1). The full sample had lower percentages of respondents who were male when compared to the U.S. population targets. The full sample and the cheap talk subsample had a lower proportion of people aged 25–34 when compared to the U.S. population. There was a higher percentage of respondents who attended college, Associates or Bachelor's Degree earned when compared to the U.S. population targets. Between the two subsamples, statistical differences were found between the percentage of respondents with an income of \$0–\$24,999 and those aged between 55 and 64 years of age.

Holiday shopping times reported by respondents were Cyber Monday (37%), Black Friday day (27%), and Black Friday morning (20%) (Table 2). On average, respondents estimated spending the most on holiday gifts (\$437.60 for the cheap talk sample, \$562.25 for the control sample). Similar spending, in terms of description selection was seen between the two categories of holiday meals and holiday travel. Most respondents (57–61%) indicated they anticipated spending the same amount in 2018 (the year the survey was administered) as they did in 2017 for holiday meals, gifts, and travel. A little over half of respondents (52%) indicated they gave to charities during the holiday season, with 74% indicating they were giving the same to charity as past planned holiday giving.

When comparing the mean response between self and the indirect question regarding the average American for both subsamples, differences were found for two statements between the subsamples (Table 3). For the cheap talk subsample, the average level of agreement between self and average American was statistically different with the exception of "I will maintain my workout schedule during the holiday season", "I will be vigilant about my weight during the holiday season", and "I watch what I eat during the holiday season". For the other statements, the average for self was statistically higher than the average for the average American, indicating the presence of SDB. Conversely, only the statement "I will be vigilant about my weight during the

Table 1 Respondent demographics (with proportion of respondents compared against the U.S. census targeted demographics).

Demographic variable	Full sample (%) <i>n</i> = 735	SDB cheap talk subsample (%) <i>n</i> = 368	SDB control subsample (%) <i>n</i> = 367	2016 U.S. Census (%)
<i>Gender</i>				
Male	44.354 [†]	44.022	44.687	49
<i>Age</i>				
18–24	13.197	13.133	13.079	13
25–34	14.830 [†]	13.859 [†]	15.803	18
35–44	16.190	17.935	14.444	16
45–54	18.911	20.109	17.711	17
55–64	18.503	14.946 ^ψ	22.071	17
65+	18.367	19.837	16.894	19
<i>Income</i>				
\$0–\$24,999	21.360	24.456 ^ψ	18.256	22
\$25,000–\$49,999	23.401	23.098	23.706	23
\$50,000–\$74,999	17.007	16.848	17.166	17
\$75,000–\$99,999	11.565	9.783	13.351	12
\$100,000 and higher	26.667	25.815	27.520	26
<i>Education</i>				
Did not graduate from high school	4.626 ^{††}	5.706 ^{††}	3.542 ^{††}	13
Graduated from high school, did not attend college	30.612	31.522	29.700	28
Attended College, no degree earned	22.721	22.554	22.888	21
Attended College, associates or bachelor's degree earned	29.796	27.174	32.425 [†]	27
Attended College, graduate or professional degree earned	12.224	13.043	11.444	12
<i>Region</i>				
Northeast	16.190	16.576	15.804	18
South	41.905 ^{††}	42.120 ^{††}	41.689 ^{††}	21
Midwest	22.312 ^{††}	23.913 ^{††}	20.708 ^{††}	38
West	19.592 ^{††}	17.391 ^{††}	21.798	24

[†]Percentage of respondents is statistically different than the percentage of the U.S. population per the US census (2016) at the 0.05 level.

^{††}Percentage of respondents is statistically different than the percentage of the U.S. population per the US census (2016) at the <0.001 level.

^ψSDB cheap talk subsample and SDB control subsamples are statistically different at the 0.05 level.

holiday season” was not statistically different between self and average American in the control sample. The average for the self-reported scores for the statements “I will maintain my workout schedule during the holiday season”, and “I watch what I eat during the holiday season” were both statistically lower than the average of the scores for the average American, indicating the presence of SDB. For the remaining questions, the average for self was statistically higher than the average American, indicating SDB for those questions in the control sample.

The distributions of difference between self and the average American were statistically compared using the Kolmogorov–Smirnov test (Kolmogorov, 1933) (Table 4). The distributions of the differences were statically different between the control group and the cheap-talk group, the cheap-talk group vs. Bir et al. (2020) and Widmar et al. (2016) for the statement “I will maintain my workout schedule during the holiday season”. Additionally, the distributions of the results for the cheap talk group and Bir et al. (2020) and the cheap talk group and Widmar et al. (2016) were statistically different for the statement “I will be vigilant about my weight during the holiday season”. Finally, the distribution of the cheap-talk and Bir et al. (2020) results were statistically different for the statement “I watch what I eat during the holiday season”. Although the analysis of the distributions can serve as a robustness check, it does not provide information regarding whether incidences of SDB have increased or decreased, but only that the distribution is different.

Therefore, incidences of SDB, defined as either having a difference of -1 or less or 1 or greater, depending on the

statement and associated directionality of SDB, are presented for the cheap talk sample in Fig. 1, and the control sample in Fig. 2, as well as Table 5. When comparing the proportion of respondents who exhibited SDB, only “I will maintain my workout schedule during the holiday season” had a statistically different percentage of SDB occurrences between the cheap talk (33%), and control samples (41%). For this particular statement, the cheap talk statement decreased the percentage of respondents who exhibited SDB in their agreement with the statement, while for all other statements, no statistical difference (either positive or negative) was found.

In addition to reporting the percentage of respondents who exhibited SDB, Widmar et al. (2016) and Bir et al. (2020) reported the percentage of respondents with spreads of -4 to -3 , -2 to -1 , 0 , 1 to 2 , and 3 to 4 . The percentage of respondents in each sample who had spreads between self versus the average American were statistically compared to the cheap talk sample and the control sample from this study (Table 5). Focusing specifically on the direction of SDB, either negative or positive scores depending on the statement, there were 7 incidences of statistical differences between the proportion of respondent within a particular spread between the control sample and either the Widmar et al. (2016) sample or the Bir et al. (2020) sample. No discernable pattern emerged regarding whether there were higher or lower percentages of respondents in each SDB spread categories. When comparing the cheap talk sample to Widmar et al. (2016) and Bir et al. (2020), there were four SDB indicating spreads that were statistically different in the proportion of

Table 2 Respondent holiday spending behavior.**Days/times shopped in 2018 (multiple selections permitted) (%)**

	Cheap talk n = 368	Control n = 367		
Thanksgiving day	12.500	11.565		
Thanksgiving evening	12.500	13.605		
"Black Friday" morning	20.109	19.456		
"Black Friday" day	27.446	27.075		
"Black Friday" evening	15.761	16.463		
"Cyber Monday"	36.957	36.871		
None of the above	44.293	43.537		
Estimated anticipated dollar amount of spending for the December 2018 through January 2019 holiday period				
Event	Mean amount spent		Standard deviation	
	Cheap talk	Control talk	Cheap talk	Control talk
Spending on holiday meals	\$112.61 n = 363	\$120.58 n = 362	\$154.73 n = 363	\$161.58 n = 362 ^a
Spending on holiday gifts	\$437.60 n = 364	\$562.25 n = 362	\$548.96 n = 364	\$834.37 n = 362 ^a
Spending on holiday travel	\$122.19 n = 361	\$191.57 n = 361	\$428.20 n = 361	\$636.45 n = 361 ^a
Intended spending for the December 2018 through January 2019 holiday period n = 735 (%)				
Type of spending	I anticipate spending more this year than last year	I anticipate spending the same amount this year as last year	I anticipate spending less this year than last year	
Spending on holiday meals	20.544	60.952	18.503	
Spending on holiday gifts	23.810	54.150	22.041	
Spending on holiday travel	16.327	56.599	27.075	
Household gives to charities during the holiday season n = 735			52%	
	Less than past amounts (%)	The same as past amounts (%)	More than past amounts (%)	
Charitable giving in the 2018–2019 compared to past planned holiday giving	13.636	74.105	12.259	

^aOne respondent was not included in the analysis of this question. The respondent spent \$3000 on holiday meals (3 times the next highest amount), \$199,000 on holiday gifts (25 times the next highest amount), and \$13,000 on holiday travel (2 times the next highest amount) and thus were considered an outlier.

respondents that were in that spread. For the statement "I will maintain my workout schedule during the holiday season", lower percentages of respondents scored –4 to –3 when compared to the Widmar et al. (2016) and Bir et al. (2020) samples. Additionally, lower percentages of respondents scored –2 to –1 when compared to the Bir et al. (2020) sample.

Evaluating SDB occurrences more broadly, when considering the percentage of respondents who exhibit SDB, few differences are found across studies. A smaller proportion of respondents exhibited SDB, as defined as having an SDB score of less than –1 for the statement "I will maintain my workout schedule during the holiday season" in the current cheap talk sample when compared to Widmar et al. (2016) and Bir et al. (2020). The proportion of respondents who exhibited SDB was also lower in the current cheap talk sample for the statement "I watch what I eat during the holiday season" when compared to the Bir et al. (2020) sample. For the control sample, proportions of respondents were statistically different than Bir et al. (2020) and Widmar et al. (2016) for the statement "I anticipate gaining weight during the holiday season". A lower percentage of respondents in the control sample exhibited SDB for the statement "I will be vigilant about my weight during the holiday season" when compared to Bir et al. (2020). Overall, given that these studies span 4 years and 4 samples, and two holidays, the level of consistency in SDB occurrences is notable.

Discussion

Due to the holiday-associated nature of this study, the precision timing afforded by online surveys was instrumental. The comparison between demographics, shopping behavior, and holiday

spending indicates that it is unlikely that there are systematic differences between the two subsamples. Subsample characteristics were also statistically compared to determine if there were systematic differences by Rotko et al. (2000) in a study of European air pollution. The amount of money spent during the holidays is notoriously difficult to determine. The holiday shopping season, as defined in the U.S. as the time between Thanksgiving and Christmas, can vary from 26 to 32 days with large impacts on holiday spending (Basker, 2005). Each additional day results in ~\$6.50 in spending, mostly attributed to impulse purchases (Basker, 2005). Byrnes (2019) stated that on average people in the U.S. would spend nearly \$1050 on gifts, goodies, and travel. This is much higher than the combined averages found in this study. However, Byrnes (2019) was using data from the National Retail Federation, in self-stated data, such as the results of this study, people may be under-estimating the amount spent, or may have trouble remembering all purchases made. Although there is social pressure to spend during the holiday season spurred by the idea of gifts as expressions of love (Spector, 2018), there is also social pressure to mitigate spending, as exhibited by the many advice articles and books regarding not overspending (Spector, 2018; Epperson and Dickler, 2019; Karp, 2010).

Evidence of SDB in self-reporting holiday healthfulness-related behaviors was found, which is unsurprising in itself, although perhaps notable for the consistency with which it was documented over time in this analysis along with those of Widmar et al. (2016) and Bir et al. (2020). The literature provides ample evidence of SDB in self-reported health behaviors and outcomes, including both underreporting negatives and over-reporting positive behaviors. Hébert et al. (2001) found that women with

Table 3 Mean and standard deviations of self-reported agreement and beliefs regarding the average American for holiday healthfulness statements for both cheap talk and control samples.

	Cheap talk sample <i>n</i> = 368			Control sample <i>n</i> = 367		
	Self		Average American		Self	
	Mean	Standard deviation	Mean	Standard deviation	Mean	Standard deviation
I anticipate gaining weight during the holiday season	3.261 [†]	1.417	2.177	1.160	3.275 [†]	1.438
I will gain more weight during the holiday season than during other times of the year	3.231 [†]	1.435	2.187	1.132	3.305 [†]	1.443
I make it a New Year's resolution to lose weight	3.625 [†]	1.543	2.307	1.242	3.632 [†]	1.546
I will maintain my workout schedule during the holiday season	3.255	1.495	3.348	1.119	3.087 [†]	1.530
I will be vigilant about my weight during the holiday season	3.408	1.388	3.408	1.191	3.338	1.464
I watch what I eat during the holiday season	3.361	1.428	3.410	1.228	3.223 [†]	1.450
I will consume more desserts during the holiday season than at other times of the year	2.883 [†]	1.495	2.100	1.152	2.880 [†]	1.442
I will consume more alcohol during the holiday season than at other times of the year	3.636 [†]	1.499	2.231	1.183	3.602 [†]	1.476

[†]Mean between self-reported scale and average American scale is statistically different within sample at the <0.001 level.

Table 4 Table of critical D values from the Kolmogorov-Smirnov test for two cumulative distributions.					
	Control vs. Cheap talk	Control vs. Bir et al. (2020)	Cheap talk vs. Bir et al. (2020)	Control vs. Widmar et al. (2016)	Cheap vs. Widmar et al. (2016)
I anticipate gaining weight during the holiday season	0.0328	0.041	0.037	0.064	0.033
I will gain more weight during the holiday season than during other times of the year	0.055	0.063	0.050	0.029	0.054
I make it a New Year's resolution to lose weight	0.038	0.053	0.028	0.048	0.026
I will maintain my workout schedule during the holiday season	0.083**	0.059	0.129**	0.029	0.098**
I will be vigilant about my weight during the holiday season	0.035	0.065	0.100**	0.045	0.080**
I watch what I eat during the holiday season	0.057	0.026	0.072**	0.021	0.036
I will consume more desserts during the holiday season than at other times of the year	0.013	0.031	0.037	0.028	0.041
I will consume more alcohol during the holiday season than at other times of the year	0.024	0.020	0.030	0.031	0.016
**Indicates statistical significance at the 0.05 level using a threshold of 0.071.					

college educations working in the health system tended to underreport caloric intake and Simons et al. (2015) found an underreporting of sedentary gaming hours among non-active videogame playing youths. Klesges et al. (2004) found that overestimates of self-reported activity, underestimates of sweetened beverage preferences, and lower ratings of weight concerns and dieting behaviors were related to SDB in 8–10-year-old girls. They suggested more research into the role of SDB in complicating relationships observed between self-reported diet and/or physical activity and health outcomes was needed. Adams et al. (2005) suggested that SDB led to an over reporting of physical activity among women in self-reported data.

The use of cheap talk to mitigate the percentage of respondents exhibiting SDB was only effective for one of the eight holiday statements studied. Additionally, the inclusion of the cheap talk statement resulted in fewer statistical differences between the average self-score of all respondents compared to the score assigned to the average American. This decrease in statistical difference between self and average American indicates that steps towards convergence of the average American and the self-reported score was occurring for more than just one statement. Despite this only mild success rate, the incidences of SDB did not increase due to the inclusion of the cheap talk statement. The use of cheap talk to prevent hypothetical bias as introduced by Cummings and Talyor (1999) experienced mixed results when being used with other products and scenarios. List (2001) found that the cheap talk script for hypothetical bias did not work on experienced bidders. Champ et al. (2009) found that the hypothetical script only worked for some offer amounts, and did not work on experienced bidders. Previous work has evaluated SDB in other health or food-related contexts for other regions of the world. Bergen and Labonte (2020) evaluated SDB in neonatal and child health care use in Ethiopia. Their identification strategy included using common cues, the nature of responses and choice patterns. They warned that SDB is influenced by accepted attitudes and behaviors, social position and affluence. Studying food-purchasing behaviors in Australia, Wheeler et al. (2019) found that SDB influenced responses regarding the purchase of organic food, increasing self-reported purchasing frequency. They found while accounting for SDB that respondents were motivated to purchase organic for non-selfish reasons including environmental and public good. The effectiveness of such tangentially related health and food-related questions in other countries would be an interesting extension of this work.

To further provide evidence of the prevalence and consistency of incidences of SDB in holiday eating and exercise-related statements, few differences were found between the SDB results of this study, Widmar et al. (2016) and Bir et al. (2020). The consistency of responses across samples and time is noteworthy for those studying holiday health. The few statistical differences found between the cheap talk subsample and the previous samples were mostly towards decreasing the percentage of respondents who exhibited SDB. The decrease in SDB also supports the idea that incorporating a cheap talk statement prior to SDB-sensitive questions may result in a mild decrease of incidences of SDB. Limited evidence of cheap talk reducing SDB in sensitive questions related to holiday eating and healthfulness was found. However, notable consistency across time, samples of respondents, and holidays (Christmas versus Thanksgiving) in terms of both responses and SDB exhibited was documented.

Limitations of research. Although the demographics of the full sample and subsamples closely mirrored the U.S. population, there were some statistical discrepancies. The samples of online survey respondents are often overeducated (Szolnoki and

Fig. 1 Holiday SDB cheap talk sample $n = 368$. Exhibits SDB is defined as having a score of less than -1 or >1 , depending on the statement. Black boxes indicate the direction of SDB, therefore the percentage of respondents who indicated SDB can be determined by adding the percentages within the box for that question. For example for the statement “I anticipate gaining weight during the holiday season”, the percentage who exhibit SDB can be calculated by adding $7 + 12 + 19 + 21 = 59$, and is given to the right of the box.

Hoffmann, 2013). However, the benefits of online data collection, including short completion time and affordable implementation, are often thought to outweigh this shortcoming (Louviere et al., 2000; Gao et al., 2009).

The use of cheap talk to mitigate the percentage of respondents exhibiting SDB was only effective for one of the eight holiday statements studied. The use of cheap talk to mitigate SDB may be more successful for other SDB prone questions, aside from the holiday-focused statements investigated here. It may simply be that the prevalence of SDB in the holiday eating and exercise statements is so engrained that the cheap talk statement had minimal effect. Or, perhaps there is an inherent difference in holiday-related reporting for the various cultural, economic, and social reasons which holiday spending and celebrations are so wrought with debate. Further research implementing cheap talk for other SDB prone questions could shed light on which situations this type of intervention works to mitigate SDB.

Methods

Survey instruments and data collection. The research project #60460205 was approved by the Purdue University institutional review board. Informed consent was obtained by all participants. Data collection took place during the peak of the 2018 holiday season, with data collection occurring December 18–26, 2018 to correspond with the winter holiday season surrounding Christmas Day (December 25) in the U.S. due to the holiday dining and health-related questions specific to this data collection effort. Kantar, a company which hosts a large opt-in panel database (Kantar, 2020), was used to obtain the survey respondents, who were required to be 18 years of age or older to participate. Quotas set within Qualtrics, an online survey tool (Qualtrics, 2020), were used to target the proportion of respondents to match the U.S.

census proportions for gender, age, education, income, and region of residence (U.S. Census Bureau, 2016). The test of proportions was used to evaluate if there were statistical differences between the subsamples employed in this study, as well as between each of the subsamples and the U.S. census. The one and two tailed tests of population proportion, assuming a normal distribution is calculated as

$$z = \frac{\hat{P} - p_0}{\sqrt{\frac{p_0(1-p_0)}{n}}}, \quad (1)$$

where p_0 is the hypothesized proportion (for example the census percentage), $\hat{P}$ is the sample proportion, and n is the sample size (Acocck, 2018). Equation (1) was used to compare each subsample to the U.S. population. A test of the difference of two proportions $\hat{P}_1$ and $\hat{P}_2$, for example comparing the demographics of the two subsamples, can be calculated as

$$z = \frac{\hat{P}_1 - \hat{P}_2}{\sqrt{\hat{P}_p \left(1 - \hat{P}_p\right) \left(\frac{1}{n_1} + \frac{1}{n_2}\right)}} \quad (2)$$

given

$$\hat{P}_p = \frac{x_1 + x_2}{n_1 + n_2}, \quad (3)$$

where x_1 and x_2 are the total number of successes in the two populations (Acocck, 2018). All calculations were conducted using STATA/SE16 (StataCorp, 2019).

Respondent demographics, holiday eating and shopping behavior, and responses regarding their level of agreement with holiday healthfulness statements designed to allow measurement of SDB were collected using a survey instrument deployed in

Fig. 2 Holiday SDB control sample $n = 367$. Exhibits SDB is defined as having a score of less than -1 or >1 , depending on the statement. Black boxes indicate the direction of SDB, therefore the percentage of respondents who indicated SDB can be determined by adding the percentages within the box for that question. For example for the statement “I anticipate gaining weight during the holiday season”, the percentage who exhibit SDB can be calculated by adding $10 + 10 + 15 + 20 = 55$, and is given to the right of the box.

Qualtrics. Data were checked for nonsensical answers and clear outliers, with special attention paid to write in answers. One respondent was not included in the analysis of the question regarding holiday spending. For this write in question the respondent spent \$3000 on holiday meals (3 times the next highest amount), \$199,000 on holiday gifts (25 times the next highest amount), and \$13,000 on holiday travel (2 times the next highest amount) and thus was considered an outlier.

The statements presented to respondents were previously employed in evaluations of SDB by Widmar et al. (2016), and Bir et al. (2020). Both Widmar et al. (2016) and Bir et al. (2020) conducted their national-scale U.S. data collection using online surveys surrounding the American November holiday Thanksgiving. Specifically the statements investigated were: “I anticipate gaining weight during the holiday season”, “I will gain more weight during the holiday season than during other times of the year”, “I make it a New Year’s resolution to lose weight”, “I will maintain my workout schedule during the holiday season”, “I will be vigilant about my weight during the holiday season”, “I watch what I eat during the holiday season”, “I will consume more desserts during the holiday season than at other times of the year”, “I will consume more alcohol during the holiday season than at other times of the year”. Respondents were asked to indicate on a scale from 1 (it describes you very well) to 5 (this statement does not describe you at all) how well the statements described them. Then, following indirect questioning (Fisher, 1993) respondents were asked to indicate on a scale from 1 (it describes the average American very well) to 5 (this statement does not describe the average American at all) how well the statements describe the average American. The order of the statements as seen by respondents was randomized for both rating self and the average American.

Two additional datasets are used in this analysis to facilitate robustness checks across samples, time, and holidays (Thanksgiving versus Christmas holiday time periods). Widmar et al. (2016) first used the set of holiday health statements employed in this work. Their targeted to be nationally representative sample of 620 U.S. respondents was collected November 17–19, 2014 (Widmar et al., 2016). Bir et al. (2020) used the same set of holiday health statements as part of a larger exploration of holiday eating with a focus on turkey consumption. Bir et al.’s (2020) nationally representative sample of 565 U.S. respondents was collected November 12–19, 2018.

Developing cheap talk for SDB. Given the propensity for respondents to under-report *bad* behaviors and over-report *good* behaviors, this study sought to measure and assess the impact of a cheap talk statement on incidences of SDB. Respondents were randomly assigned to two subsamples which each saw and responded with their level of agreement to identical statements about holiday healthfulness. One subsample was shown a cheap talk statement about SDB prior to providing their level of agreement to the statements for themselves and the average American, and one group was not shown any information prior to reporting their level of agreement to the same eight statements for themselves and the average American.

Since its introduction to contingent valuation literature by Cummings and Taylor (1999), the use of cheap talk statements has expanded. Instructions to respondents, and specifically cheap talk statements have been employed by Lusk (2003) in WTP experiments to minimize hypothetical bias and by Blumenshien et al. (2007) as a comparison in field experiments, among others. The cheap talk statement provided to half of the randomly assigned respondents was developed for this study and read,

Table 5 Comparison of percentage of respondents with SDB scores between current sample and previous publications, box indicates direction of SDB.

	Spread of difference between self and the average American					Exhibits ¹ SDB
	-4 to -3	-2 to -1	0	1 to 2	3 to 4	
I anticipate gaining weight during the holiday season						
Current sample cheap talk ²	0.543	9.511	31.250	39.946	18.750	58.695
Current sample control	0.545	7.902	35.967	35.695	19.891	55.586
Widmar et al. (2016)	0	7	31 ^Ω	41 ^Ω	21	62 ^Ω
Bir et al. (2020)	1	11	33	39	16 ^Ω	55 ^Ω
I will gain more weight during the holiday season than during other times of the year						
Current sample cheap talk	0.543	10.870	30.435	40.761	17.391	58.152
Current sample control	0.545	6.812	33.515	36.785	22.343	59.128
Widmar et al. (2016)	1	5 ⁺	32	40	22 ⁺	62
Bir et al. (2020)	1	9	35	39	15 ^Ω	54
I make it a New Year's resolution to lose weight						
Current sample cheap talk	0.815	9.783	26.630	36.413	26.359	62.772
Current sample control	0.817	11.989	24.523	35.422	27.248	62.670
Widmar et al. (2016)	0	8 ^Ω	29	37	26	63
Bir et al. (2020)	1	10	28	33	27	60
I will maintain my workout schedule during the holiday season						
Current sample cheap talk	6.793 ^Δ	25.815	34.511	29.620 ^Δ	3.261	32.609 ^Δ
Current sample control	11.883 ^Δ	28.883	33.243	22.343 ^Δ	3.542	40.872 ^Δ
Widmar et al. (2016)	12 ⁺	29	36	21 ⁺	2 ^Ω	41 ⁺
Bir et al. (2020)	10 ⁺	32 ⁺	37	19 ⁺	2 ^Ω	43 ⁺
I will be vigilant about my weight during the holiday season						
Current sample cheap talk	5.707	26.630	33.696	28.261	5.707 ^Δ	32.337
Current sample control	8.719	20.981	39.782	28.065	2.452 ^Δ	29.700
Widmar et al. (2016)	6 ^Ω	27 ^Ω	41 ⁺	24	1 ⁺ Ω	33
Bir et al. (2020)	5 ^Ω	31 ^Ω	40 ⁺	22 ⁺ Ω	2 ⁺	36 ^Ω
I watch what I eat during the holiday season						
Current sample cheap talk	6.250	26.630	37.500	25.000	4.620	32.880
Current sample control	7.084	28.338	39.782	22.616	2.180	35.422
Widmar et al. (2016)	5	30	39	25	2 ⁺	35
Bir et al. (2020)	7	31	39	22	1 ⁺ Ω	38 ⁺
I will consume more desserts during the holiday season than at other times of the year						
Current sample cheap talk	0.272	13.859	38.043	33.152	14.674	47.826
Current sample control	1.090	11.717	38.965	34.332	13.896	48.229
Widmar et al. (2016)	0	10 ⁺	43	35	12	47
Bir et al. (2020)	1	11	40	36	11 ⁺	48
I will consume more alcohol during the holiday season than at other times of the year						
Current sample cheap talk	0.815	6.793	29.348	33.967	29.076	63.043
Current sample control	0.272	8.447	26.975	33.243	31.063	64.305
Widmar et al. (2016)	0	6 ^Ω	30	36	27	63
Bir et al. (2020)	0	11 ⁺	25	33	31	64

¹Exhibits SDB is defined as having a score of less than -1 or >1, depending on the statement. Black boxes indicate the direction of SDB, therefore the percentage of respondents who indicated SDB can be determined by adding the percentages within the box for that row. For example for the statement "I anticipate gaining weight during the holiday season" the percentage who exhibit SDB can be calculated by adding 40 + 19 = 59.

²Current sample cheap talk $n = 368$, current sample control talk $n = 367$, Widmar et al. (2016) $n = 620$, Bir et al. (2020) $n = 565$.

⁺Proportion of respondents is statistically different between non-cheap talk sample and previous work.

^ΩProportion of respondents is statistically different between cheap talk sample and previous work.

^ΔProportion of respondents is statistically different between non-cheap talk and cheap talk subsamples.

"Human inclination may be to answer questions in a way that deviates from your true behavior in an effort to improve the impression you make on others. This desire to give what is believed to be the socially "correct" or acceptable answer is often referred to as social desirability bias. Please keep this inclination in mind, and try to reflect on your true behavior when answering questions". Statistical comparisons between those who saw the cheap talk statement and those who did not were conducted in two ways.

Evaluating incidences of SDB, and comparing the results of indirect questioning across studies. First, the mean results for

the responses regarding self and the holiday healthfulness statements as well as average American for both the cheap talk and control subsamples were calculated. The mean scores for self and average American were statistically compared for each of the two subsamples using a paired t -test. The equation for paired observations is given by

$$t = \frac{\bar{x}_i - \bar{y}_i \sqrt{n}}{s_d}, \quad (4)$$

where x_i is observation i from the self score, y_i is observation i from the average American score, n is the number of observations, and S_d is the standard deviation (Dixon and Massey, 1983).

All analysis was conducted in STATA (StataCorp, 2019). It was hypothesized that there would be fewer statistical differences for the subsample who viewed the cheap talk statement, as their self-assessment would be closer to that of their idea of the average American. This theory is supported by the construct and purpose of indirect questioning (Fisher, 1993).

Next, an index was created of the difference between the score respondents gave themselves and the score they gave the average American on the scale from 1 (describes well) to 5 (does not describe well) following Bir et al. (2020) and Widmar et al. (2016). For example, if a person indicated a score of 2 when asked about themselves, and a score of 4 when asked about the average American, the difference would be -2 . Depending on the specific statement, either a positive or negative difference between the score chosen for self and the average American indicated the presence of SDB. A positive difference would indicate the presence of SDB for the statement “I anticipate gaining weight during the holidays”, “I will gain more weight during the holiday season than during other times of the year”, “I make it a New Year’s resolution to lose weight”, “I will consume more desserts during the holiday season than at other times of the year”, and “I will consume more alcohol during the holiday season than at other times of the year”. A negative score would indicate the presence of SDB for the statement “I will maintain my workout schedule during the holiday season”, “I will be vigilant about my weight during the holiday season”, and “I watch what I eat during the holiday season”. Each respondent was evaluated as exhibiting or not exhibiting SDB for each of the 8 statements studied. Exhibiting SDB was defined as having a difference between the score for self and average American of 1 or greater, or a difference of -1 or less depending on the statement. Using the test of proportions Eqs. (2) and (3), the proportion of respondents who displayed SDB was statistically compared between the cheap talk and control samples. Additionally, using Eq. (1), the incidences of respondents who displayed SDB were statistically compared between the two subsamples from this study, Widmar et al. (2016) and Bir et al. (2020).

Finally, for a broader evaluation of potential incidences of SDB, Widmar et al. (2016) grouped the difference between self and average American scores as -4 to -3 , -2 to -1 , 0 , 1 to 2 , and 3 to 4 . In order to compare results between Widmar et al. (2016), Bir et al. (2020) and the results of this study, the same categories of differences between self and average American scores were determined for this study and Bir et al. (2020). These categories were then statistically compared between the different studies using Eqs. (1)–(3) to compare the individual differences between the categories. To better understand the distribution of the differences between self and average American scores for the holiday healthfulness questions the Kolmogorov–Smirnov test for cumulative distributions was used (Kolmogorov, 1933). The cumulative distribution function for the samples being compared were determined, and then the largest difference was identified. The largest difference between the CDFs were then compared to the critical value. The critical value (D) for a 0.05 statistical significance was calculated as

$$D = \frac{1.35810}{\sqrt{n}}. \quad (5)$$

Given the sample size of 368 the critical value was determined as 0.070796.

Data availability

The data that support the findings of this study are available from the corresponding author upon reasonable request.

Received: 25 March 2020; Accepted: 11 November 2020;

Published online: 10 December 2020

References

- Acock ACA (2018) Gentle introduction to stata, 6th edn. College Station, TX, STATA press
- Adams SA, Matthews CE, Ebbeling CB et al. (2005) The effect of social desirability and social approval on self-reports of physical activity. *Am J Epidemiol* 161 (4):389–398
- Basker E (2005) Twas four weeks before Christmas: retail sales and the length of the Christmas shopping season. *Econ Lett* 89(3):317–322
- Bergen N, Labonté R (2020) Everything is perfect, and we have no problems: detecting and limiting social desirability bias in qualitative research. *Qual Health Res* 30(5):783–792
- Bir C, Davis MK, Widmar NO et al. (2020) Holiday health consciousness and consumer demand for whole turkey attributes. *Poult Sci* 99(5):2798–2810
- Bir C, Widmar NO, Cronley C (2018) Exploring social desirability bias in perceptions of dog adoption: all’s well that ends well? Or does the method of adoption matter? *Animals* 8(9):154
- Blumenschein K, Blomquist GC, Johannesson M et al. (2007) Eliciting willingness to pay without bias: evidence from a field experiment. *Econ J* 118 (525):114–137
- Byrnes H (2019) Holiday spending bill: most of us spend \$1,000 or more on gifts travel and goodies. Available via USA Today. <https://www.usatoday.com/story/money/2019/11/29/this-is-how-much-the-average-person-spends-on-the-holidays-in-every-state/40688517/>. Accessed 23 Jan 2020
- Champ PA, Moore R, Bishop RC (2009) A comparison of approaches to mitigate hypothetical bias. *Agric Resour Econ Rev* 38(2):166–180
- Crino MD, Svoboda M, Rubinfeld S et al. (1983) Data on the Marlowe–Crowne and Edwards social desirability scales. *Psychol Rep* 53(3):963–968
- Crowne DP, Marlowe D (1960) A new scale of social desirability independent of psychopathology. *J Consult Clin Psychol* 24(4):349
- Cummings RG, Taylor LO (1999) Unbiased value estimates for environmental goods: a cheap talk design for the contingent valuation method. *Am Econ Rev* 89(3):649–665
- Dixon WJ, Massey FJ (1983) Introduction to statistical analysis, 4th edn. New York, STATA press
- Epperson S, Dickler J (2019) 4 ways to avoid the overspending trap this holiday. Available via CNBC. <https://www.cnbc.com/2019/11/16/how-to-avoid-overspending-this-holiday.html>. Accessed 11 Feb 2020
- Farrell J, Gibbons R (1989) Cheap talk can matter in bargaining. *J Econ Theory* 48 (1):221–37
- Farrell J, Rabin M (1996) Cheap talk. *J Econ Perspect* 10(3):103–18
- Fisher RJ (1993) Social desirability bias and the validity of indirect questioning. *J Consum Res* 20(2):303–315
- Gao Z, Schroeder T (2009) Effects of label information on consumer willingness-to-pay for food attributes. *Am J Agric Econ* 91(3):95–809
- Galesic M, Bosnjak M (2009) Effects of questionnaire length on participation and indicators of response quality in a web survey. *Public Opin Q* 73(2):349–360
- Hawkes N (2016) Sixty seconds on... New Year resolutions. *BMJ* 355:i6845
- Hébert JR, Peterson KE, Hurley TG et al. (2001) The effect of social desirability trait on self-reported dietary measures among multi-ethnic female health center employees. *Ann Epidemiol* 11(6):417–427
- Jampolis M (2018) Maintaining your weight through the holidays. Available via CNN. <https://www.cnn.com/2018/12/10/health/holiday-weight-diet-exercise-jampolis/index.html>. Accessed 16 Jan 2020
- Johansson-Stenman O, Martinsson PH (2006) What are you driving a BMW? *J Econ Behav Organ* 60(2):129–46
- Kantar (2020) About. Available via Kantar Website. <https://www.kantar.com/about>. Accessed 11 Feb 2020
- Karp G (2010) Seasonal gift strategies for spending smart. New York City, STATA press
- Kaviani S, vanDellen M, Cooper JA (2019) Daily self-weighing to prevent holiday-associated weight gain in adults. *Obesity* 27(6):908–916
- Klesges LM, Baranowski T, Beech B et al. (2004) Social desirability bias in self-reported dietary, physical activity and weight concerns measures in 8-to 10-year-old African-American girls: results from the Girls Health Enrichment Multisite Studies (GEMS). *Prev Med* 38:78–87
- Kolmogorov AN (1933) Sulla determinazione empirica di una legge di distribuzione. *Giornale dell’ Istituto Italiano degli Attuari* 4:83–91
- List JA (2001) Do explicit warnings eliminate the hypothetical bias in elicitation procedures? Evidence from field auctions for sports cards. *Am Econ Rev* 91 (5):1498–1507
- Louviere JJ, Hensher DA, Swait JD (2000) Stated choice methods: analysis and application. Cambridge, England, STATA press

- Lusk JL (2003) Effects of cheap talk on consumer willingness-to-pay for golden rice. *Am J Agric Econ* 85(4):840–856
- Lusk JL, Norwood FB (2009) An inferred valuation method. *Land Econ* 85(3):500–514
- Ma Y, Olendzki BC, Li W et al. (2006) Seasonal variation in food intake, physical activity, and body weight in a predominantly overweight population. *Eur J Clin Nutr* 60:519–528
- Maccoby EE, Maccoby N (1954) The interview: a tool of social science. *Handb Soc Psychol* 1(1):449–487
- Mason F, Pallan M, Sitch A et al. (2018) Effectiveness of a brief behavioural intervention to prevent weight gain over the Christmas holiday period: randomised controlled trial. *BMJ* 363:k4867
- Matthews S, Okuno-Fujiwara M, Postlewaite (1991) A refining cheap-talk equilibrium. *J Econ Theory* 55(2):247–73
- Olynk NJ, Tonsor GT, Wolf CA (2010) Consumer willingness to pay for livestock credence attribute claim verification. *J Agr Resour Econ* 35(2):261–280
- Rehkamp S (2016) A look at Calorie sources in the American diet. Available via USDA. <https://www.ers.usda.gov/amber-waves/2016/december/a-look-at-calorie-sources-in-the-american-diet>. Accessed 10 Feb 2020
- Rotko T, Oglesby L, Künzli N et al. (2000) Population sampling in European air pollution exposure study. EXPOLIS: comparisons between the cities and representativeness of the samples. *J Expo Sci Environ Epidemiol* 10(4):355–364
- Schoeller DA (2014) The effect of holiday weight gain on body weight. *Physiol Behav* 134:66–69
- Simons M, Chinapaw MJ, Brug J et al. (2015) Associations between active video gaming and other energy-balance related behaviours in adolescents: a 24-hour recall diary study. *Int J Behav Nutr Phys Act* 12(1):192
- Spector N (2018) Amid pressure to overspend on holidays, consumers embrace the gift of minimalism. Available via NBC News. <https://www.nbcnews.com/business/consumer/amid-pressure-overspend-holidays-consumers-embrace-gift-minimalism-n938901>. Accessed 21 Jul 2019
- StataCorp (2019) Stata statistical software: release 16. StataCorp, College Station, TX.
- Stevenson JL, Krishnan S, Stoner MA, Goktas Z et al. (2013) Effects of exercise during the holiday season on changes in body weight, body composition and blood pressure. *Eur J Clin Nutr* 67(9):944–949
- Szolnoki G, Hoffmann D (2013) Online, face-to-face and telephone surveys—comparing different sampling methods in wine consumer research. *Win Econ Policy* 2:57–66
- U.S. Census Bureau (2016) Annual Estimates of the Resident Population for the United States, Regions, States, and Puerto Rico: April 1, 2010 to July 1, 2015 (NST-EST2015-01). Available via U.S. Census Bureau, Population Division. <https://www.census.gov/data/tables/time-series/demo/popest/2010s-state-total.html>. Accessed 16 Jan 2018
- Qualtrics (2020) About. Available via Qualtrics Website. <https://www.qualtrics.com/>. Accessed 11 Feb 2020
- Wheeler SA, Gregg D, Singh M (2019) Understanding the role of social desirability bias and environmental attitudes and behaviour on South Australians' stated purchase of organic foods. *Food Qual Prefer* 74:125–134
- Widmar NJ, Byrd ES, Dominick SR et al. (2016) Social desirability bias in reporting of holiday season healthfulness. *Prev Med Rep* 4:270–276

Acknowledgements

This work was supported partially by the USDA National Institute of Food and Agriculture, Hatch project IND0004133 entitled Changing Preferences for Meat Proteins by US Residents.

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to C.B.

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2020