

Asking questions about vices really does increase vice behavior

Gavan J. Fitzsimons

Duke University, Durham, NC, USA

Lauren G. Block

Baruch College, New York, NY, USA

Patti Williams

University of Pennsylvania, Philadelphia, PA, USA

In a commentary in the previous issue of *Social Influence*, Schneider, Tahk, and Krosnick raise concerns about the analytical techniques and conclusions drawn in Williams, Block, and Fitzsimons (2006). In this response, we address a number of issues raised by their comment. We discuss their thorough reanalysis of our data, what we believe it implies, and what conclusions should be drawn. We also briefly present a number of replications of the Williams et al. experiment and discuss what the collective implications of this work are. Ultimately, we conclude that the original warning cast by Williams et al. still stands: when we ask questions about vice behaviors we should do so with great care as we may unintentionally be increasing the behavior we are asking about.

In our paper on the effects of asking questions about health-related behaviors (Williams, Block, & Fitzsimons, 2006) we reported the results of a relatively simple experiment. In their comment on that paper in the previous issue of *Social Influence*, Schneider, Tahk, and Krosnick have conducted a reanalysis of the data and concluded that the data do not support drawing a conclusion consistent with what we reported in Williams et al. (2006); namely that asking questions about seemingly negative behaviors such as using drugs can in fact lead to an increase in those vice behaviors. In this reply, we will review the concerns Schneider et al. (2007) raise in their commentary, and discuss areas in which we agree with their perspective as well as areas in which we disagree.

Address correspondence to: Gavan J. Fitzsimons, Professor of Marketing and Psychology, Duke University, 1 Towerview Drive, Durham, NC 27708, USA. E-mail: gavan@duke.edu

We will address a number of issues that Schneider et al. (2007) raise in turn, starting with whether Williams et al. (2006) made any inadvertent errors in reporting. We will then address the issue of whether a *t*-test is the appropriate test for examining differences between conditions like those we observed in the original experiment. Finally we will address the issue of whether one should continue to have faith in the conclusion we drew in Williams et al. (2006), that simply asking a question about a vice behavior can actually lead to increases in the respondents' engaging in the behavior. In the original experiment there was also a condition in which participants responded to a question about a positive behavior and showed increases in this behavior. Given this type of question-behavior effect (questions about positive behaviors increasing actual behavior) had been demonstrated many times prior to our experiment (see Sprott et al., 2006, for a brief review) we will focus this reply to the more novel, and controversial, finding that asking questions about a seemingly negative behavior (i.e., drug use) leads to increases in the actual performance of the behavior.

CORRECTING AN INADVERTENT ERROR

Due to the potential impact of the finding we reported in Williams et al. (2006), considerable attention was drawn to the published result and conclusion. One of the positives of this attention is that Jon Krosnick and his colleagues asked if they could examine the data that the paper was based upon. We happily shared the data with their research team and were somewhat dismayed when they discovered that we had inadvertently reported incorrect *t*-test statistics in the original paper. While the mean reported data for the treatment group asked a drug intent question 2 months prior ($X=2.8$) and the control group not asked a question ($X=1.0$) were correct, we had inadvertently reported the results of a one-sided *t*-test. The correct *t*-test statistic, which we reported in a corrigendum in the 1(3) issue of *Social Influence* (p.248), was 1.74 ($p=.085$) for the analysis including all participants. Further, the analysis using only the subset of respondents that reported non-zero drug usage was also incorrect. Again, while the mean reported data for the subset of the treatment group that were asked a drug intent question 2 months prior and reported non-zero drug use ($X=10.3$) and the control group not asked a question but reporting non-zero drug use ($X=4.0$) were correct, we had inadvertently reported the results of a one-sided *t*-test. The correct *t*-test statistic, again reported in a corrigendum in the 1(3) issue of *Social Influence* (p.248), was 2.02 ($p=.054$). We are certainly grateful that this error was caught by Schneider et al., and we corrected the record as soon as we possibly could.

WAS A *T*-TEST THE “RIGHT” TEST?

The use of a *t*-test to test for differences between groups assumes a number of basic elements, which Schneider et al. (2007) have very clearly laid out. The most basic assumption is of course that the data are normally distributed. This is a basic assumption that virtually all social scientists know to be required prior to using *t*-tests, analysis of variance, etc. In the case of the data reported in Williams et al. (2006), this is clearly not the case. The distribution of reported drug use is skewed, and the difference in mean reported drug use between the question and no-question conditions is heavily dependent on a small number of respondents. Given this, Schneider et al.’s (2007) basic argument that the data we collected do not meet the appropriate assumptions to use a *t*-test are compelling. Of course, the reason many researchers automatically utilize *t*-tests and analysis of variance is that they are relatively robust to small violations in the normality assumption, for example. However, in this case the normality assumption in the underlying data is not violated in a minor way, as the data are skewed fairly significantly.

In retrospect, a more careful assessment of the underlying distribution would have been a wise course of action to have followed. Perhaps due to the simplicity of the design and the relatively straightforward continuous nature of the dependent variable, we neglected to do so. As Schneider et al. (2007) have laid out, a number of reasonable courses of action are available to the researcher when their data do not meet the assumptions required to perform a *t*-test. Without going through in detail all of the various approaches (which Schneider et al. have very nicely and clearly laid out) the results of the various tests were as follows: two of the tests (i.e., the *Z*-score and the negative binomial regression) yielded significant differences ($p < .05$) between the group asked a question about drug use and the control group not asked a question. Three tests (i.e., the OLS *t*-test, the permutation test, and the bootstrap test) yielded marginally significant differences between the question and the control group ($p < .10$). The remaining eight tests conducted by Schneider et al. (2007) indicated that the effect of the behavior prediction question was non-significant. In a situation such as this the reader could reasonably (i) conclude that there is mixed evidence for whether the difference between the groups was significant, or (ii) determine for themselves which test they felt was most appropriate and put increased weight on that test. We leave it to the reader to determine which of these two conclusions they believe to be most reasonable. Our own position is that there is substantial statistical support for differences between the question and no-question groups using statistical techniques that many would argue are reasonable approaches to analyzing the types of data that we have (e.g., negative binomial regression, permutation and bootstrap tests).

Upon concluding that the results of their reanalysis of the Williams et al. (2006) might be frustrating to a reader seeking a definitive answer as to whether the group asked a drug intent question really did report significantly different drug use than the controls, Schneider et al. (2007) engaged in further analysis in which they selectively eliminated either one or two respondents they classified as outliers. Their conclusion after excluding these respondents was that the various tests employed in their re-analysis would each conclude no significant differences occurred. We are reluctant to conclude, as they did, that these respondents should be excluded from the analysis. As they point out, many people are highly skeptical of the removal of participants from analyses so that previously non-significant differences become significant. We are similarly skeptical of conclusions drawn by removing legitimate responses from the current data set.

Further, as drug behavior is likely to have substantial variance in the undergraduate population studied in Williams et al. (2006), we do not feel that exclusion of reported drug use of less than once a day over the 2-month period seems appropriate. Drug users are reasonably likely to engage in drug behavior on a relatively frequent basis. Common definitions of “heavy” marijuana users, for example, classify the participants as heavy users only if they used marijuana at least once a day (Block & Ghoneim, 1993; Mirin, Shapiro, Meyer, Pillard, & Fisher, 1971). Thus reports of using drugs 50 times over a 2-month period (as did the two participants Schneider et al. excluded from their second stage of analyses) would simply indicate that the two “outliers” were intermediate to heavy drug users, not that their responses should be excluded from subsequent analyses.

Schneider et al. (2007) also suggest that because the two respondents they argue should be excluded from the analysis reported low stated intent to use drugs, it should not be the case that they would actually be heavy drug users. On the surface, this seems reasonable. However, previous work examining the question-behavior effect has found mixed results in terms of the relationship between stated intent and actual behavior. While in some cases a positive correlation has been observed, in many cases no significant correlation was obtained. That is, there was not a significant link between the stated intent (in this case stated intent to use drugs in the coming 2 months) and actual behavior (in this case reported drug use) in a number of previous studies. Given the complicated social norms surrounding drug use, it is not surprising to us that this lack of correlation would be observed in the Williams et al. (2006) experiment. Ultimately, the reader must determine whether they prefer to interpret analyses based on all the participants or a subset of participants.

DOES THIS EFFECT REPLICATE?

Ultimately, the reason Schneider et al. have written their response to Williams et al. (2006) is that they are concerned the results obtained and reported may not be robust. Quoting from their argument for excluding outliers from their reanalysis they state (p. 189):

Similarly, Cohen et al. (2003) argued that significant findings that are generated only by the presence of one or two outliers will most likely not hold up upon replication and should therefore not be reported.

Thus, replication of the basic results of the Williams et al. (2006) experiment should go a long way to reassuring the reader that the conclusions drawn were not the result of underlying distributional assumptions or of using all of the respondents instead of excluding some characterized as outliers. In a forthcoming paper to appear in the June 2007 issue of the *Journal of Consumer Research* we (Fitzsimons, Nunes, & Williams, 2007) replicated the basic effect reported in the Williams et al. (2006) paper in three separate experiments. In each experiment, we either asked participants a question about a vice behavior or did not ask such a question, and then later observed their behavior. In each case, asking a question about a vice behavior led to subsequent increases in that behavior. (In a fourth experiment in the Fitzsimons et al., 2007, paper we showed that asking a question about a vice behavior led to facilitated access to a positive implicit attitude.)

To briefly summarize the studies, below we will highlight the means between the question and no-question group for each of the three experiments. All tests of statistical significance were significant ($p < .05$) whether one used parametric (e.g., *t*-test, ANOVA) or non-parametric tests (e.g., Wilcoxon-Mann-Whitney; chi-square), hopefully reassuring the reader that these results are not vulnerable to similar debate as might be the Williams et al. (2006) results. To further ensure that these results were not being driven by “outliers”, we eliminated respondents more than two standard deviations above and below the mean on the dependent variable and reran the analyses, although we would not encourage this as a common course of action. The results remained significant and consistent with the “outliers” removed using both parametric and non-parametric tests. (Note that outlier removal is not possible in Experiment 3 described below, as the dependent variable is a binary response.)

In Experiment 4 of Fitzsimons et al. (2007), participants were asked in a first session about drinking more than two drinks in a sitting in the coming week or about watching television instead of studying. In a second session 1 week later participants reported both their drinking and television-watching

behavior from the prior week. As expected, asking a question about drinking behavior increased drinking behavior from 1.2 times to 3.2 times over the week. Similarly, participants asked about watching television instead of studying increased reported television watching when they should have been studying from 2.7 to 3.9 times.

As both the drug data in Williams et al. (2006) and the drinking and television watching instead of studying were based on self-report data, we wanted to be sure to replicate the basic effect using actual behavior rather than self-reported behavior. In Experiment 3 of Fitzsimons et al. (2007) participants were asked a question about their being distracted from studying in the coming week or were asked a control question. After filling out a number of other measures, participants were provided with an opportunity to sign up for a (“vice”) behavior that would result in considerable distraction from studying (i.e., going to four movie screenings in a single week) by providing their email address, ostensibly to be passed along to the movie studio. Compared to the control group, those asked the question about the vice behavior were substantially more likely to sign up for the vice behavior (76.6% versus 53.1%).

Similarly, in Experiment 1 in Fitzsimons et al. (2007) students were asked a series of questions about their curriculum, plans for graduation, etc., on the first day of a semester-length class. Half the group was also asked a question about their future class-skipping behavior. Actual class attendance was monitored over the course of the semester. Those that answered a skipping question ended up missing 3.78 classes, while a control group not asked a question about skipping missed only 2.95 classes.

CONCLUSIONS

The response that Schneider et al. prepared certainly raises a number of excellent points that we as social scientists should all take to heart. They have clearly reminded at least the authors of the current article that it is extremely important to check the various underlying distributional, etc., assumptions of one’s data prior to running one type of analysis versus another. We are also grateful to them for helping us correct an inadvertent error in the reporting of our original paper.

However, they ultimately reach very different conclusions than we do about what one should take away from reading the Williams et al. (2006) article. After a careful reassessment of the Williams et al. (2006) data we have concluded that the pattern of results obtained in that experiment is quite consistent with the pattern obtained in three subsequent experiments of asking about vice behavior reported in Fitzsimons et al. (2007). As a result, we believe the basic conclusions drawn by Williams et al. to be appropriate. When researchers ask questions about vice behaviors such as

using drugs, drinking, skipping class, etc., it is quite possible that simply as a function of asking the question the research may well be increasing the likelihood that the respondent will engage in the behavior subsequently. This is not to suggest that this will occur in every situation to every respondent. As Fitzsimons et al. (2007) point out, it is likely to only be the case when the respondent to the question about the vice behavior has a positive attitude toward it, despite the fact that societal norms dictate that one should hold a negative attitude. Collectively, however, we hope that our research on the effects of questions about vice behaviors will lead researchers to exercise caution in their use of such questions.

Manuscript received 12 March 2007

Manuscript accepted 16 April 2007

REFERENCES

- Block, R. I., & Ghoneim, M. M. (1993). Effects of chronic marijuana use on human cognition. *Psycho-pharmacology*, *110*, (1–2), 219–228.
- Fitzsimons, G. J., Nunes, J. C., & Williams, P. (2007). License to sin: The liberating role of reporting expectations. *Journal of Consumer Research*, *34*(1), 22–31.
- Mirin S, M., Shapiro, L. M., Meyer, R. E., Pillard, R. C., & Fisher, S. (1971). Casual versus heavy use of marijuana: A redefinition of the marijuana problem. *American Journal of Psychiatry*, *127*, 1134–1140.
- Schneider, D., Tahk, A., & Krosnick, J. A. (2007). Reconsidering the impact of behavior prediction questions on illegal drug use: The importance of using proper analytic methods in social psychology. *Social Influence*, *2*, 178–196.
- Sprott, D. E., Spangenberg, E. R., Block, L. G., Fitzsimons, G. J., Morwitz, V. G., & Williams, P. (2006). The question–behavior effect: What we know and where we go from here. *Social Influence*, *1*(2), 129–137.
- Williams, P., Block, L. G., & Fitzsimons, G. J. (2006). Simply asking questions about health behaviors increases both healthy and unhealthy behaviors. *Social Influence*, *1*(2), 117–127.

Copyright of Social Influence is the property of Psychology Press (UK) and its content may not be copied or emailed to multiple sites or posted to a listserv without the copyright holder's express written permission. However, users may print, download, or email articles for individual use.