Journal of Clinical Epidemiology 108 (2019) 102−109

Journal of Clinical Epidemiology

ORIGINAL ARTICLE

Randomized trial seeking to induce the Hawthorne effect found no evidence for any effect on self-reported alcohol consumption online

Jim McCambridge[a,b], Amanda Wilson[a], John Attia[a], Natasha Weaver[a], Kypros Kypri[a,*]

[a]*Centre for Clinical Epidemiology and Biostatistics, School of Medicine and Public Health, University of Newcastle, Newcastle, New South Wales, Australia*
[b]*Department of Health Sciences, University of York, UK*

Accepted 14 November 2018; Published online 17 November 2018

Abstract

Objective: We tested the hypothesis that participants who know the behavioral focus of a study and are thus aware that a particular behavior is being studied will modify that behavior, independently of any possible effect of assessment, thereby dismantling a Hawthorne effect into two putative components.

Study Design and Setting: We undertook a three-arm individually randomized trial online among students: group A (control) were told they were completing a lifestyle survey; group B were told the focus of the survey was alcohol consumption; and group C additionally answered 20 questions on their alcohol use and its consequences before answering the same lifestyle questions as Groups A and B. Non-drinkers were excluded, and all groups were aware they would be followed up after 1 month.

Results: Outcome data were obtained for 4,583 of 5,478 trial participants (84% follow-up rate). There were no differences between the three groups on primary (overall volume consumed) or secondary outcome measures (drinking frequency and amount per typical occasion) in the intervening 4 weeks.

Conclusions: There is no evidence that any form of Hawthorne effect exists in relation to self-reported alcohol consumption online among university students in usual research practice. Attention to study contexts is warranted for investigating research participation effects.

Keywords: Hawthorne effect; Research participation effects; Reactivity; Alcohol; Internet; Students

1. Introduction

Simply participating in a research study has the potential to modify behavior in ways that can bias the findings, a phenomenon widely known as the ''Hawthorne effect'' [1−3]. There is, however, little securely known about how any such effect works or the circumstances in which it may arise [4]. Better understanding how participants react to what they are asked to do in research and their experiences of the study conditions more broadly may help achieve less biased study designs and inform the interpretation of existing findings [5,6].

There is much uncertainty about what occurred in the original studies that inspired the naming of the Hawthorne effect, the various meanings attached to the term, and the resulting heterogeneity of objects of study in purposively designed evaluation studies of this phenomenon [4]. There are two principal types of studies of the Hawthorne effect, examining the effects of (1) answering questions in interviews or by completing questionnaires and (2) having behavior directly observed or otherwise being made aware of having a particular behavior studied [4]. These contrasting study types suggest two quite different mechanisms of effect: either that answering questions induces altered thinking about the behavior or that monitoring and surveillance in the context of research studies causes behavior change; and it could be the case that both can operate simultaneously [4]. Conformity to perceived norms or researcher expectations that may embody them potentially links these two putative mechanisms although a wide array of mechanisms have been proposed [7].

The effects of answering questions have been conceptualized in various ways in different topic areas and disciplines. Within social, health, and consumer psychology, there is a large volume of research [8−12] on what is now known as the *question-behavior effect*. Within the alcohol field, *assessment reactivity* has been the predominant conceptualization of essentially the same phenomenon

Conflicts of interest: None.

* Corresponding author. Tel.: +61 2 4042 0536; fax: +61 2 4042 0044.
E-mail address: kypros.kypri@newcastle.edu.au (K. Kypri).

https://doi.org/10.1016/j.jclinepi.2018.11.016
0895-4356/

What is new?

Key findings

- There is no evidence that any form of Hawthorne effect exists in relation to self-reported alcohol consumption online among university students in this large randomized trial.

What this adds to what was known?

- Study findings are interpreted as suggesting that Hawthorne effect phenomena are likely to be unimportant in online alcohol studies.

What is the implication and what should change now?

- There is a need to better understand the contexts in which research participation effects, which may threaten valid inferences, are more likely to occur.

[13,14]. An older and more recognizable construct known throughout the discipline of psychology is that of *demand characteristics* [15,16], which incorporates both the ways in which questions are asked and other implicit features of research situations that participants may form expectations about what is being required of them. Demand characteristics are thus somewhat analogous to the two main operationalizations of the Hawthorne effect [17]. Investigations of demand characteristics belong to a rich and largely discontinued tradition within social psychology (because of changing norms about the acceptability of deception in nonlaboratory research) of scrutinizing unintended artifacts of research studies [18−20].

These considerations led us to develop the term *research participation effects* to refer broadly to the consequences arising from overlooked aspects of participating in research studies that may produce bias [6]. Although a rudimentary concept, it nonetheless connects a range of conventionally understood sources of bias (such as information and attrition bias) and provides a novel perspective that may guide dedicated studies, which extend our understanding of the nature of bias [6].

There is evidence from randomized trials showing that asking people about their drinking without subsequent intervention can produce small reductions in self-reported alcohol consumption [14]. The present study is the third in a series of large online methodological experiments undertaken with university students in which alcohol consumption was measured to assess the outcome of a dedicated experimental manipulation. In the first, we tested hypotheses about the effects of study design and allocation (ESDA), that is, if knowing whether the study was a cohort (i.e., observational) study vs. an intervention trial would affect university students' subsequent drinking [21]. We expected that being told they were in an intervention trial might have a greater self-focusing—and thereby moderating—effect than being told they were in an observational study. We also investigated how the participants who thought they were in an intervention trial reacted to being told they had been randomized to the control vs. intervention condition (allocation). We did not find support for either hypothesis, that is, there were no significant differences in alcohol consumption between the three experimental groups when their drinking was assessed a month later [21].

Because we had to rely on self-report of outcome, we were concerned that any differences observed might merely reflect effects on reporting. Accordingly, in the second study, we sought to induce socially desirable reporting through asking about drinking in ways that might stigmatize participants, comparing this to a neutral presentation of questions as part of an assessment of health behavior [22]. Again, we found no differences between the groups, and we concluded that in the Web-based context, measurement of drinking is robust to the effects of question framing, at least within university student populations [22]. Note that self-report of drinking and other behaviors is nonetheless subject to a range of validity concerns other than framing effects [23,24], and our focus is on usual research practice, not the possible effects of online surveillance. In the present study, using the same experimental paradigm, we investigated the effects of (1) participants knowing that the study's focus was alcohol consumption in the absence of significant alcohol assessment and (2) additionally, completing alcohol assessments on their self-reported drinking behavior 1 month later.

2. Methods

2.1. Design

The overarching aims were to explore the content of those research participation effects commonly labeled as the Hawthorne effect. We undertook a three-arm individually randomized parallel groups trial (allocation ratio 1:1:1) to dismantle a hypothesized Hawthorne effect into putative assessment and awareness of observation components among university students in respect of their alcohol consumption. The existing literature is methodologically limited and thus complex to interpret [4], and the Hawthorne, Assessment Reactivity, Bias Investigational trial was prospectively registered with the Australian New Zealand Clinical Trials Registry (ACTRN12612000255886).

2.2. Pilot work

The trial was preceded by pilot research to refine the content of the experimental manipulations and study procedures. In the ESDA study investigating related research participation effects, we found many participants spent little time with key Web pages open [21]. We were concerned that some participants may not have read the text carefully and thus were not fully exposed to the intended experimental manipulation.

For these reasons, the baseline instrument underwent two rounds of pilot testing at an Australian university. Students on campus were invited by a trained interviewer (A.W.) to assist with the design of a health behavior survey using methods we developed earlier [25]. Those providing informed consent completed the questionnaire on a computer, taking approximately 5 minutes to complete it, and then took part in a face-to-face interview for a further 5 minutes. In the first round of piloting, two-thirds of participants (20/29) indicated they were unlikely to carefully read the introductory information in such a survey, which would have undermined our intended experimental manipulation. We therefore added a mandatory manipulation check to the study procedures, confirming that the relevant experimental groups knew this study was about alcohol, involved repeated measurement, and was applicable to nondrinkers. Correct information was repeated in the event of incorrect answers to the questions presented to participants, which they had to complete before progressing to the next webpage. This was successful, with the second round of piloting yielding 100% (24/24) awareness that participation involved repeated surveys of alcohol consumption.

2.3. Participants

We conducted the trial in four New Zealand universities. All students (approximately $n = 60{,}000$) without any restriction by age or enrollment type, were invited to complete a Web survey in a message sent to their student

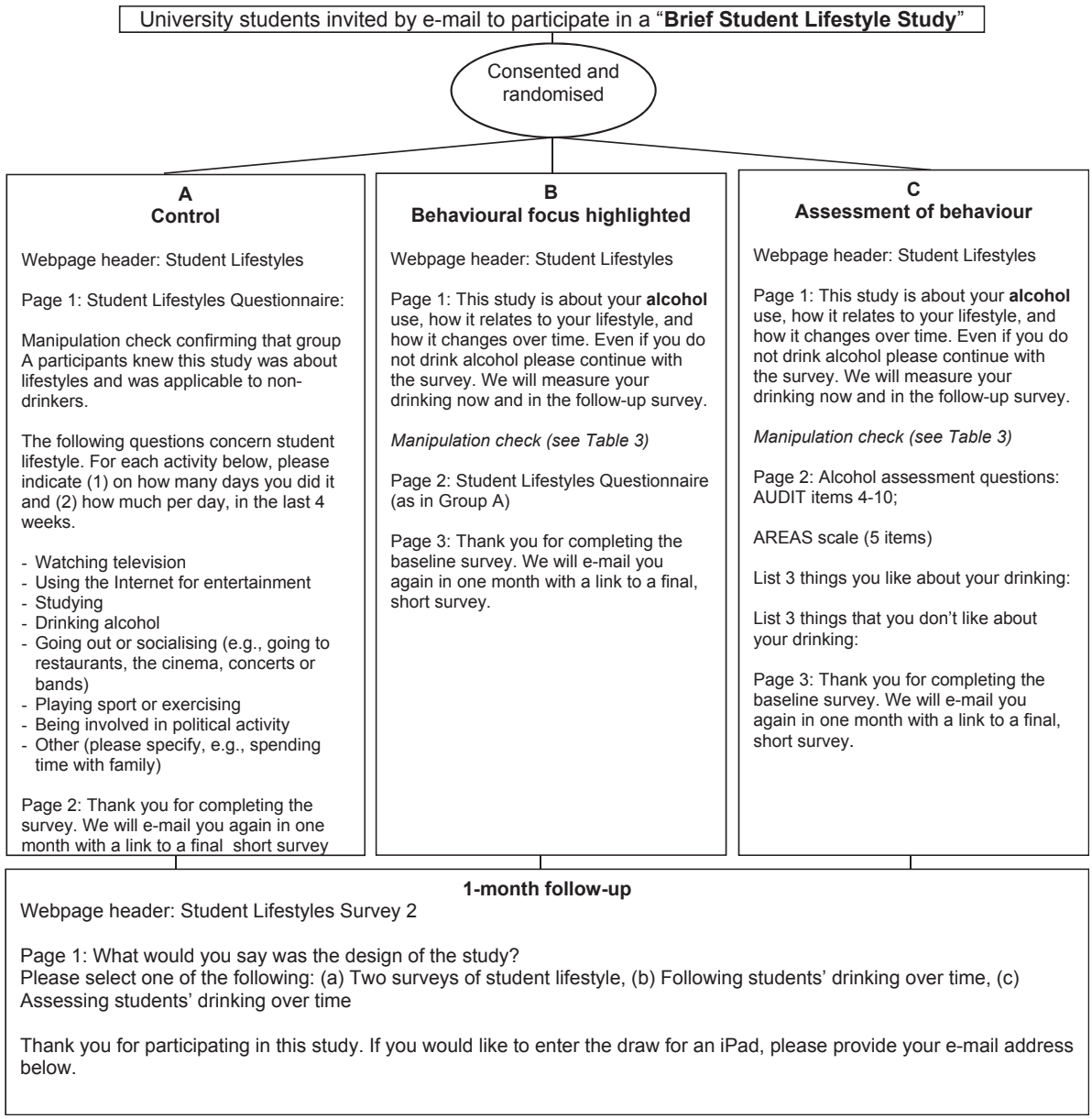

Fig. 1. Study design and experimental conditions.

e-mail address by the university administration. The message contained a hyperlink to a Participant Information Sheet, which invited participation in a brief student lifestyle study and presented a button to indicate consent and start the survey. Two alcohol questions were included within the student lifestyle questionnaire and used to identify nondrinkers who completed the research procedures but were not eligible for participation in the trial. All trial participants thus reported consuming alcohol in the previous 4 weeks.

2.4. Interventions/experimental manipulations

The detailed content of the interventions, that is, the text that comprised the experimental manipulation, is presented in Fig. 1. Briefly, group A (the control group) was asked to complete a questionnaire on "student lifestyles," group B was advised that the study focus was their alcohol use, and group C was also asked 20 questions about their drinking. All three groups were advised they would be followed up again in 1 month.

All surveys took less than 10 minutes to complete.

2.5. Outcomes

The prespecified primary outcome was self-reported volume of alcohol consumption in the previous 4 weeks. This was computed as the product of the two secondary outcomes: frequency of days drinking and quantity (number of standard drinks) per drinking day, which were asked as follows: (1) on how many days in the last 4 weeks did you have a drink containing alcohol? (2) How many standard drinks containing alcohol did you have on a typical day when you were drinking in the last 4 weeks? Both questions were embedded within the lifestyle survey administered at baseline.

2.6. Sample size

The study was designed to detect between-group differences in effect size of $d = 0.1$. This small effect size was chosen following partitioning of previously observed effect sizes in a systematic review of alcohol assessment reactivity [14]. To achieve power of 0.8, with an alpha 0.05 and a two-sided test, required 1,946 participants per group (5,838 in total).

2.7. Randomization

Randomization without any restrictions was implemented using a computerized random number generator. All study procedures were computerized and automated ensuring that the researchers were blind to randomization, allocation, and assessment of outcome, and there was no direct contact with participants that could have compromised randomization.

2.8. Blinding

All three randomized groups were blinded to the true purpose of the study, as this would have interfered with hypothesis testing. We have offered ethical justifications for the use of deception in such studies elsewhere [26] and make clear that the University of Otago's Human Research Ethics Committee approved the conduct of this study. To conceal the alcohol focus of the study, no attempt was made to exclude nondrinkers from online participation in the surveys though they were not trial participants. All participants were given to understand that this was a brief student lifestyle study involving the completion of two surveys 4 weeks apart. No participant was aware of randomization.

2.9. Statistical methods

We used linear regression models with repeated measures fitted by generalized estimating equations for each of the three outcomes, testing for differences in outcomes between each pair of groups and across all three groups.

3. Results

The Consolidated Standards of Reporting Trials flowchart in Fig. 2 shows that 6,909 students (both drinkers and nondrinkers) initially consented, including $n = 5{,}478$ drinkers within the previous 4 weeks who were eligible for trial participation and were randomized to one of the three experimental groups. Trial enrollment took place in August to September 2012 and follow-up 1 month later. Table 1 shows that randomization worked well and attrition was unproblematic as the three experimental groups had similar distributions of demographic characteristics.

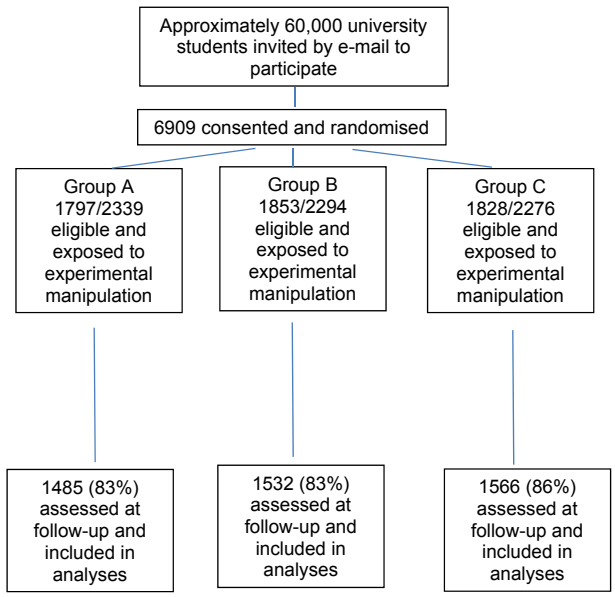

Fig. 2. Consolidated Standards of Reporting Trials diagram.

Table 1. Demographic characteristics of the experimental groups

Participant demographic characteristics	Experimental group			
	A (n = 1,797)	B (n = 1,853)	C (n = 1,828)	All (N = 5,478)
Age, mean (SD)	24.0 (7.9)	24.2 (8.1)	24.2 (8.1)	24.1 (8.1)
Gender, n (%)				
Men	624 (35)	622 (34)	653 (36)	1,899 (35)
Women	1,173 (65)	1,231 (66)	1,175 (64)	3,579 (65)
Living arrangements, n (%)				
Share a rented flat or house	966 (54)	998 (54)	992 (54)	2,956 (54)
Living in a residential college or hall	228 (13)	252 (14)	236 (13)	716 (13)
Living with parent(s) or guardian(s)	238 (13)	228 (12)	219 (12)	685 (13)
Living in my own home	333 (19)	327 (18)	342 (19)	1,002 (18)
Boarding	20 (1.1)	29 (1.6)	22 (1.2)	71 (1.3)
Living elsewhere	8 (0.4)	7 (0.4)	13 (0.7)	28 (0.5)
Prefer not to answer	4 (0.2)	12 (0.6)	4 (0.2)	20 (0.4)

Table 2 summarizes alcohol consumption at baseline and follow-up by experimental group. This shows all three groups were similar in drinking at baseline, and that there were no differences in outcomes, that is, the groups were not found to change differently over time. None of the pairwise comparisons showed a significant difference at follow-up. In contrast to the pilot study, Table 3 indicates that the experimental manipulation was similarly unsuccessful for groups B and C in approximately 24% of participants.

4. Discussion

This study found no evidence that any form of Hawthorne effect exists in relation to self-reported alcohol consumption online among university students. Previous findings on assessment of this behavior in this population online are mixed; some randomized trials have identified small effects for some outcomes [27–29], whereas others find no such effects [30,31]. An earlier systematic review of randomized studies of alcohol assessment found some evidence of small effects and concluded that answering questions "appears to exert a subtle influence on subsequent self-reported drinking behavior among students" [14], partly as there were no effects in randomized studies in nonstudent populations. It is noteworthy that in this earlier review [14], there were effects observed in trials investigating self-completion of pen-and-paper questionnaires [32] and interview administration [33] among students respectively. There are no prior randomized trials of simply being aware of the possible effects of a study focus on alcohol as investigated here.

Beyond alcohol, the larger literature on question-behavior effects provides evidence of small effects on both self-reported and objectively ascertained outcomes, with much unexplained heterogeneity among studies, in a literature vulnerable to bias [8,10,12]. French and Sutton [34] identify larger assessment effects on emotion and cognition than on

Table 2. Alcohol consumption at baseline and follow-up by experimental group

Outcome	Time	Statistic	Experimental group				*P* value		
			A (n = 1,485)	B (n = 1,532)	C (n = 1,566)	Overall	B vs. A	C vs. A	B vs. C
Total volume consumed	Baseline	Mean (SD)	24.6 (31.8)	23.8 (31.8)	23.0 (30.6)				
		Median (min, max)	15 (1, 336)	14 (1, 459)	13 (1, 476)				
	Follow-up	Mean (SD)	18.4 (23.3)	17.5 (23.8)	17.5 (23.2)	0.66	0.89	0.51	0.39
		Median (min, max)	12 (0, 240)	10 (0, 360)	10 (0, 308)				
Days drinking alcohol	Baseline	Mean (SD)	5.9 (5.2)	5.8 (4.9)	5.5 (5.0)				
		Median (min, max)	4 (1, 28)	4 (1, 28)	4 (1, 28)				
	Follow-up	Mean (SD)	5.2 (5.5)	5.0 (5.1)	4.9 (5.0)	0.26	0.22	0.80	0.12
		Median (min, max)	4 (0, 28)	4 (0, 28)	4 (0, 28)				
Drinks per drinking day	Baseline	Mean (SD)	4.3 (3.6)	4.2 (3.6)	4.2 (3.6)				
		Median (min, max)	3 (1, 30)	3 (1, 27)	3 (1, 22)				
	Follow-up	Mean (SD)	3.5 (3.3)	3.4 (3.4)	3.4 (3.3)	0.87	0.72	0.86	0.60
		Median (min, max)	2 (0, 24)	2 (0, 30)	2 (0, 30)				

Table 3. Experimental manipulation checks for Hawthorne effect groups

Question	Class	Experimental group		
		B (n = 1,853), n (%)	C (n = 1,828), n (%)	P value
Study purpose	Your drinking (correct)	1,418 (77)	1,372 (75)	0.2035
	Health behavior (incorrect)	433 (23)	456 (25)	
	Other (incorrect)	2 (0.1)		
Study requirements	One now and another in 4 wk (correct)	1,779 (96)	1,741 (95)	0.0424
	Just one now (incorrect)	59 (3.2)	80 (4.4)	
	Other (incorrect)	15 (0.8)	7 (0.4)	
Study eligibility	All students (correct)	1,804 (97)	1,773 (97)	0.7826
	Only students who drink alcohol (incorrect)	43 (2.3)	49 (2.7)	
	Other (incorrect)	6 (0.3)	6 (0.3)	

behavior in a narrative review. Some meta-analytic reviews are restricted to trials [8], whereas others also synthesize nonrandomized studies [9], with the latter identifying much larger effects than the former. These meta-analytic reviews identify larger effect sizes among students and do not generally suggest that alcohol is clearly different than other behaviors although effects are smaller with more challenging behaviors to perform. Effects are larger in promoting socially desirable behaviors than in reducing socially undesirable behaviors [9], as was the nature of the original Hawthorne factory research artifact. Wilding et al. [9] identify larger effects in laboratory settings and with face-to-face questioning, with online questions similar to mail and phone. These studies were largely undertaken by psychologists, and there has been a strong emphasis on question types and content [35], as well as a search for putative mechanisms underlying effects, which has proven difficult to progress [11]. The prior systematic review of the Hawthorne effect [4] also found the heterogeneity in previous studies of this construct challenging to interpret, as did another systematic review of Solomon four group designs [36].

The existing literature arguably requires much greater attention to both conceptualization and empirical investigation of the effects of study contexts [37]. For example, it has been shown that completion of the same alcohol questionnaires in different contexts will produce different findings [38]. It has previously been suggested that question-behavior effect studies need to pay greater attention to study contexts [39], and we suggest this is a particularly fruitful direction for studies of Hawthorne effect phenomena more generally. Indeed, study context is what shapes the original emphasis on observation, monitoring, and surveillance [1–3,40]. We have previously suggested that how participants interact with what they are asked to do in the specific contexts of research studies, that is, participant engagement with study context may be a fruitful line of enquiry [6]. Qualitative studies already suggest that there is much value in asking participants about their experiences in research to better understand the data they provide and the potential implications for bias [41,42].

The interpretation of this study's findings needs to address study limitations. We fell a little below the target sample size and suggest that reaching it would not have altered the study findings. We originally were attracted to the online university context as offering a pragmatic opportunity to recruit large numbers of research participants. An inherent limitation of this context is reliance on self-report of drinking behavior [24]. Although in treatment contexts, the conclusion has been reached that self-report of drinking is valid [43]' this is not clear in other settings [44] although computerized assessments of sensitive behaviors may be more valid [45]. Note also that invalid self-report may itself be conceptualized as a research participation effect [6].

Another substantial study limitation lies in the lack of optimization of the experimental manipulations, where despite careful pilot work, approximately one-quarter of those randomized to the experimental groups, who were all told that the study focus was alcohol, did not confirm this approximately 5 minutes later. Interestingly, the asking of 20 alcohol questions did not make any difference in the randomized comparison to the simple study focus manipulation, and it may be that the asking of 18 nonalcohol-related questions was responsible.

The earlier literature on assessment reactivity in alcohol treatment contexts where assessments were typically lengthy did consider that the overall burden of assessment was likely to be important to the stimulation of unintended effects [13,46,47]. It is possible that a stronger contrast between groups B and C achieved by asking many more questions may show some effects although this would be untypical of how online research in this area is conducted, where participant burden is a consideration in relation to preventing attrition [48].

It also appears that the online context is not conducive to the elicitation of a Hawthorne effect sense of being observed perhaps because of the ways that participants engage with study and nonstudy requests for information and/or the nature of the medium. This suggests a much stronger surveillance manipulation would be needed to evoke a sense of being observed in brief contacts online.

This would be challenging to investigate, for example, without compromising study retention. It may be that, despite the capacity for access to large numbers of participants, the online study context imposes limitations to further investigations of surveillance or that a different kind of manipulation is needed than was evaluated here. The effects of online surveillance could become the focus of social psychology investigations although our primary focus is usual practices in research. Strengths of this study include security against any compromise of randomization because of the automated nature of the study.

We conclude on the basis of the evidence examined here that there is no particular need for concern about Hawthorne type phenomena in online alcohol studies among university students in the way they are usually conducted and perhaps also in other online populations. On the basis of our studies, we suggest that future research in this area may be most usefully advanced in research studies in which there is direct person-to-person contact, which evokes social desirability concerns. Future studies of research participation effects should also identify other features of contexts in which there is a basis for concern about such effects, be explicit about the underlying reasoning to further develop appropriate concepts, and use high-quality study designs to identify where these issues pose threats to valid inference.

Acknowledgments

Authors' contributions: J.M. and K.K. conceived of and designed the trial and together with J.A. obtained the funding to undertake it. A.W. conducted the pilot study and developed and implemented the trial procedures under the direction of K.K. N.W. analyzed the data. J.M. led the drafting of the article with inputs from K.K. All authors approved the final article for submission.

The authors are grateful to the participating universities.

This work was supported by the Australian Research Council, Australia (Discovery Grant DP1093809). The funder had no role in the study design or conduct or in the decision to publish the results.

References

[1] French JRP. Experiments in field settings. In: Festinger L, Katz D, editors. Research methods in the behavioral sciences. New York: Holt, Rinehart & Winston; 1953.
[2] Parsons HM. What happened at Hawthorne? Science 1974;181:922−32.
[3] Gillespie R. Manufacturing knowledge: a history of the Hawthorne experiments. Cambridge: Cambridge University Press; 1991.
[4] McCambridge J, Witton J, Elbourne D. Systematic review of the Hawthorne effect: new concepts are needed to study research participation effects. J Clin Epidemiol 2014;67:267−77.
[5] McCambridge J, Kypri K, Elbourne D. In randomisation we trust? There are overlooked problems in experimenting with people in behavioural intervention trials. J Clin Epidemiol 2014;67:247−53.
[6] McCambridge J, Kypri K, Elbourne D. Research participation effects: a skeleton in the methodological cupboard. J Clin Epidemiol 2014;67:845−9.
[7] Chiesa M, Hobbs S. Making sense of social research: how useful is the Hawthorne effect? Eur J Soc Psychol 2008;38(1):67−74.
[8] Rodrigues AM, O'Brien N, French DP, Glidewell L, Sniehotta FF. The question-behavior effect: genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. Health Psychol 2015;34:61−78.
[9] Wilding S, Conner M, Sandberg T, Prestwich A, Lawton R, Wood C, et al. The question-behaviour effect: a theoretical and methodological review and meta-analysis. Eur Rev Soc Psychol 2016;27(1):196−230.
[10] Wood C, Conner M, Miles E, Sandberg T, Taylor N, Godin G, et al. The impact of asking intention or self-prediction questions on subsequent behavior: a meta-analysis. Pers Soc Psychol Rev 2016;20(3):245−68.
[11] Spangenberg ER, Kareklas I, Devezer B, Sprott DE. A meta-analytic synthesis of the question-behavior effect. J Consumer Psychol 2016;26(3):441−58.
[12] Rodrigues AM, French DP, Sniehotta FF. Commentary: the impact of asking intention or self-prediction questions on subsequent behavior: a meta-analysis. Front Psychol 2016;7:879.
[13] McCambridge J. Research assessments: instruments of bias and brief interventions of the future? Addiction 2009;104(8):1311−2.
[14] McCambridge J, Kypri K. Can simply answering research questions change behaviour? Systematic review and meta analyses of brief alcohol intervention trials. PLoS One 2011;6:e23748.
[15] Orne MT. The nature of hypnosis: artifact and essence. J Abnorm Soc Psychol 1959;58:277−99.
[16] Orne MT. On the social psychology of the psychological experiment: with particular reference to demand characteristics and their implications. Am Psychol 1962;17:776−83.
[17] McCambridge J, de Bruin M, Witton J. The effects of demand characteristics on research participant behaviours in non-laboratory settings: a systematic review. PLoS One 2012;7:e39116.
[18] Webb EJ, Campbell DT, Schwartz RD, Sechrest L. Unobtrusive measures: nonreactive research in the social sciences. Oxford: Rand McNally; 1966.
[19] Rosnow RL, Rosenthal R. People studying people: artifacts and ethics in behavioral research. New York: Freeman; 1997.
[20] Sharpe D, Whelton WJ. Frightened by an old scarecrow: the remarkable resilience of demand characteristics. Rev Gen Psychol 2016;20(4):349−68.
[21] Kypri K, Wilson A, Attia J, Sheeran PJ, McCambridge J. Effects of study design and allocation on self-reported alcohol consumption: randomized trial. Trials 2015;16:127.
[22] Kypri K, Wilson A, Attia J, Sheeran P, Miller P, McCambridge J. Social desirability bias in the reporting of alcohol consumption: a randomized trial. J Stud Alcohol Drugs 2016;77(3):526−31.
[23] Tourangeau R, Yan T. Sensitive questions in surveys. Psychol Bull 2007;133(5):859−83.
[24] Davis CG, Thake J, Vilhena N. Social desirability biases in self-reported alcohol consumption and harms. Addict Behav 2010;35(4):302−11.
[25] Hallett J, Maycock B, Kypri K, Howat P, McManus A. Development of a Web-based alcohol intervention for university students: processes and challenges. Drug Alcohol Rev 2009;28(1):31−9.
[26] McCambridge J KK, Bendtsen P, Porter J. The use of deception in public health behavioural intervention trials: a case study of three online alcohol trials. Am J Bioeth 2013;13(11):39−47.
[27] Kypri K, Langley JD, Saunders JB, Cashell-Smith ML. Assessment may conceal therapeutic benefit: findings from a randomized controlled trial for hazardous drinking. Addiction 2007;102(1):62−70.
[28] McCambridge J, Bendtsen M, Karlsson N, White IR, Nilsen P, Bendtsen P. Alcohol assessment and feedback by e-mail for university students: main findings from randomised controlled trial. Br J Psychiatry 2013;203:334−40.

[29] Walters ST, Vader AM, Harris TR, Jouriles EN. Reactivity to alcohol assessment measures: an experimental test. Addiction 2009;104(8):1305−10.
[30] Bendtsen P, McCambridge J, Bendtsen M, Karlsson N, Nilsen P. RCT of the effectiveness of proactive mail based alcohol Internet intervention with university students: dismantling the assessment and feedback components. J Med Internet Res 2012;14(5):e142.
[31] Kypri K, McAnally HM. Randomized controlled trial of a web-based primary care intervention for multiple health risk behaviors. Prev Med 2005;41:761−6.
[32] McCambridge J, Day M. Randomized controlled trial of the effects of completing the alcohol use disorders identification test questionnaire on self-reported hazardous drinking. Addiction 2008;103:241−8.
[33] Carey KB, Carey MP, Maisto SA, Henson JM. Brief motivational interventions for heavy college drinkers: a randomized controlled trial. J Consult Clin Psychol 2006;74(5):943−54.
[34] French DP, Sutton S. Reactivity of measurement in health psychology: how much of a problem is it? What can be done about it? Br J Health Psychol 2010;15(Pt 3):453−68.
[35] Mankarious E, Kothe E. A meta-analysis of the effects of measuring theory of planned behaviour constructs on behaviour within prospective studies. Health Psychol Rev 2015;9(2):190−204.
[36] McCambridge J, Butor-Bhavsar K, Witton J, Elbourne D. Can research assessments themselves cause bias in behaviour change trials? A systematic review of evidence from Solomon 4-group studies. PLoS One 2011;6:e25223.
[37] McCambridge J. From question-behaviour effects in trials to the social psychology of research participation. Psychol Health 2015;30(1):72−84.
[38] Cooke R, French DP. The role of context and timeframe in moderating relationships within the theory of planned behaviour. Psychol Health 2011;26(9):1225−40.
[39] Sprott DE, Spangenberg ER, Block LG, Fitzsimons GJ, Moritz VG, Williams P. The question-behavior effect: what we know and where we go from here. Soc Infl 2006;1(2):128−37.
[40] Gale EAM. The Hawthorne studies − a fable for our times? Q J Med 2004;97:439−49.
[41] Locock L, Smith L. Personal experiences of taking part in clinical trials - a qualitative study. Patient Educ Couns 2011;84(3):303−9.
[42] McCambridge J, Sorhaindo A, Quirk A, Nanchahal K. Patient preferences and performance bias in a weight loss trial with a usual care arm. Patient Educ Couns 2014;95(2):243−7.
[43] Babor TF, Steinberg K, Anton R, Del Boca F. Talk is cheap: measuring drinking outcomes in clinical trials. J Stud Alcohol 2000;61(1):55−63.
[44] Noknoy S, Rangsin R, Saengcharnchai P, Tantibhaedhyangkul U, McCambridge J, et al. RCT of effectiveness of motivational enhancement therapy delivered by nurses for hazardous drinkers in primary care units in Thailand. Alcohol Alcohol 2010;45(3):263−70.
[45] Turner CF, Ku L, Rogers SM, Lindberg LD, Pleck JH, Sonenstein FL. Adolescent sexual behavior, drug use, and violence: increased reporting computer survey technology. Science 1998;280:867−73.
[46] Gastfriend DR, Donovan D, Lefebvre R, Murray KT. Developing a baseline assessment battery: balancing patient time burden with essential clinical and research monitoring. J Stud Alcohol Suppl 2005;(15):94−103. discussion 92−93.
[47] Orford J, Hodgson R, Copello A, John B, Smith M, Black R, et al. The clients' perspective on change during treatment for an alcohol problem: qualitative analysis of follow-up interviews in the UK Alcohol Treatment Trial. Addiction 2006;101(1):60−8.
[48] McCambridge J, Kalaitzaki E, White IR, Khadjesari Z, Murray E, Linke S, et al. Impact of length or relevance of questionnaires on attrition in online trials: randomized controlled trial. J Med Internet Res 2011;13(4):e96.