

The question-behaviour effect: A theoretical and methodological review and meta-analysis

Sarah Wilding, Mark Conner, Tracy Sandberg, Andrew Prestwich, Rebecca Lawton, Chantelle Wood, Eleanor Miles, Gaston Godin & Paschal Sheeran

To cite this article: Sarah Wilding, Mark Conner, Tracy Sandberg, Andrew Prestwich, Rebecca Lawton, Chantelle Wood, Eleanor Miles, Gaston Godin & Paschal Sheeran (2016) The question-behaviour effect: A theoretical and methodological review and meta-analysis, *European Review of Social Psychology*, 27:1, 196-230, DOI: [10.1080/10463283.2016.1245940](https://doi.org/10.1080/10463283.2016.1245940)

To link to this article: <http://dx.doi.org/10.1080/10463283.2016.1245940>



© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.



[View supplementary material](#)



Published online: 08 Nov 2016.



[Submit your article to this journal](#)




[View related articles](#)



[View Crossmark data](#)

The question-behaviour effect: A theoretical and methodological review and meta-analysis

Sarah Wilding ^a, Mark Conner^a, Tracy Sandberg^a, Andrew Prestwich^a,
Rebecca Lawton^a, Chantelle Wood^b, Eleanor Miles^c, Gaston Godin^d
and Paschal Sheeran^e

^aSchool of Psychology, University of Leeds, Leeds, UK; ^bDepartment of Psychology, University of Sheffield, Sheffield, UK; ^cSchool of Psychology, University of Sussex, Brighton, UK; ^dCanada Research Chair on Behavior and Health, Laval University, Ville de Quebec City, Canada; ^eDepartment of Psychology and Neuroscience, University of North Carolina at Chapel Hill, Chapel Hill, NC, USA

ABSTRACT


Research has demonstrated that asking people questions about a behaviour can lead to behaviour change. Despite many, varied studies in different domains, it is only recently that this phenomenon has been studied under the umbrella term of the question-behaviour effect (QBE) and moderators of the effect have been investigated. With a particular focus on our own contributions, this article: (1) provides an overview of QBE research; (2) reviews and offers new evidence concerning three theoretical accounts of the QBE (behavioural simulation and processing fluency; attitude accessibility; cognitive dissonance); (3) reports a new meta-analysis of QBE studies ($k = 66$, reporting 94 tests) focusing on methodological moderators. The findings of this meta-analysis support a small significant effect of the QBE ($g = 0.14$, 95% $CI = 0.11, 0.18$, $p < .001$) with smaller effect sizes observed in more carefully controlled studies that exhibit less risk of bias and (4) also considers directions for future research on the QBE, especially studies that use designs with low risk of bias and consider desirable and undesirable behaviour separately.

ARTICLE HISTORY Received 8 April 2016; Accepted 29 September 2016

KEYWORDS Question-behaviour effect; attitude accessibility; cognitive dissonance

The question-behaviour effect (QBE) refers to the influence of any sort of questioning—including questions about intentions, predictions of future behaviour and measures of satisfaction—on subsequent performance of that behaviour (Sprott, Spangenberg, Knuff, & Devezer, 2006b). This effect is also known as the mere-measurement effect, the self-prophecy effect, self-erasing errors of prediction and self-generated validity (Sprott et al., 2006a).

CONTACT Mark Conner  m.t.conner@leeds.ac.uk  School of Psychology, University of Leeds, Leeds LS2 9JT, UK

 Supplemental data for this article can be accessed [here](#).

© 2016 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The QBE was originally demonstrated by Sherman (1980). One group of participants was asked to self-predict how *likely* they would be to perform a socially desirable or socially undesirable behaviour; a second group made no such prediction about their behaviour. The results indicated that participants asked to predict their behaviour were subsequently more likely to report performing socially desirable (31% vs. 4%) and less likely to perform undesirable (40% vs. 68%) behaviours compared to participants making no prediction. Hence, mere *questioning* substantially affected subsequent *behaviour* (27–28% change in performance rates).

Many studies have now replicated Sherman's (1980) original demonstration in both laboratory and field settings. For example, intention questions have been shown to influence students' later brand choice in laboratory experiments of consumer behaviour (e.g., Morwitz & Fitzsimons, 2004), while measuring purchase intentions has been found to increase future purchases (Chandon, Morwitz, & Reinartz, 2004). Greenwald, Carnot, Beach and Young (1987) reported that students asked about their intentions to vote in the following day's elections were more likely to do so than were students who were not queried about their voting intentions. Methodological research has used the term "assessment reactivity" to refer to the impact of measuring behaviour at baseline on later behaviour. For example, measurement of baseline physical activity has been found to produce greater activity at follow-up compared to non-assessed controls (Spence, Burgess, Rodgers, & Murray, 2009).

The QBE has been tested in a number of distinct areas, and most commonly in relation to health, consumer and prosocial behaviours. Within the health area, studies have demonstrated that the QBE can be harnessed as an effective intervention to increase uptake of health screening (Sandberg & Conner, 2009), health checks (Conner, Godin, Norman, & Sheeran, 2011; Spangenberg & Sprott, 2006; Sprott, Smith, Spangenberg, & Freson, 2004; Sprott, Spangenberg, & Fisher, 2003) and vaccinations (Conner et al., 2011). In relation to consumer purchase behaviours, numerous studies have supported the QBE (e.g., Chandon et al., 2004; Fitzsimons & Williams, 2000; Janiszewski & Chandon, 2007; Morwitz & Fitzsimons, 2004; Morwitz, Johnson, & Schmittlein, 1993; van Kerckhove, Geuens, & Vermeir, 2012). Similarly, the QBE has been tested in prosocial behaviours such as blood donation (e.g., Cioffi & Garner, 1998; Godin, Sheeran, Conner, & Germain, 2008). Table 1 provides examples of key studies testing the QBE for different behaviours.

There have been several narrative (Dholakia, 2010) and quantitative (Rodrigues, O'Brien, French, Glidewell, & Sniehotta, 2015; Spangenberg & Greenwald, 1999; Spangenberg, Kareklas, Devezer, & Sprott, 2016; Sprott et al., 2006a; Wood et al., 2016) reviews of the QBE that indicate a small but reliable effect (Rodrigues et al., 2015: Cohen's $d = 0.09$, 95% $CI = 0.04, 0.13$;

Table 1. Illustrative studies of the QBE in different behaviours.

Behaviours	Example studies
<u>Health behaviours</u>	
Alcohol consumption	Bendtsen, McCambridge, Bendtsen, Karlsson and Nilsen (2012); Bernstein et al. (2009); Kypri and McAnally (2005)
Drug use	Bernstein et al. (2009); Williams et al. (2006)
Flossing	Williams, Fitzsimons and Block (2004) study 1; Levav and Fitzsimons (2006) study 3
Health assessment	Sprott et al. (2004) study 1
Physical activity	Spence et al. (2009); Kypri and McAnally (2005)
Risky driving	Falk (2010) studies 1 and 2
Safe sex	Træen (2003); Kvaalem, Sundet, Rivø, Eilertsen and Bakketeig (1996)
Screening	Sandberg and Conner (2009); van Valkengoed, Morre, Meijer, van den Brule and Boeke (2002)
Vaccination	Conner et al. (2011) study 1
<u>Consumer behaviours</u>	
Purchasing	Dholakia and Morwitz (2002); Morwitz et al. (1993)
Brand choice	van Kerckhove et al. (2012) studies 1–3
<u>Prosocial behaviours</u>	
Blood donation	Godin et al. (2008, 2012)
Charity donation	Sherman (1980) study 1; Spangenberg and Sprott (2006) study 2
<u>Other behaviours</u>	
Voting	Greenwald et al. (1987) study 1; Nickerson and Rogers (2010)
Stereotyping	Spangenberg and Greenwald (1999) studies 1 and 2
Mailing letters	Chapman (2001)

Spangenberg et al., 2016: Cohen's $d = 0.28$, 95% $CI = 0.24, 0.32$ and Wood et al., 2016: Cohen's $d = 0.24$, 95% $CI = 0.18, 0.30$). The relative simplicity of creating a QBE (ensuring that respondents complete questions about the behaviour) has meant there has been interest in the phenomenon as a potentially cost-effective means to change socially important behaviours. The focus of this research has usually been on the moderating factors that maximise the size of the QBE. On the other hand, the widespread use of questions about behaviours within intervention studies means that many researchers are keen to ensure that the QBE does not bias their findings. Here, the focus of research has often been on minimising the QBE or gaining insights into where the QBE might interfere with other interventions. As a means to these ends, both bodies of research have also been concerned with exploring theoretical accounts of the QBE.

The present review aims to provide the most comprehensive review of work conducted in this area to date. The key contribution of the present paper is that it provides a general overview of the QBE literature *and* meta-analytic tests of a range of different moderators and mediators. In relation to mechanisms underlying the QBE, the present review considers existing findings and presents new evidence from a range of unpublished work conducted by the authors. The meta-analytic review focuses on the influence of methodological moderators that have received little attention from previous reviews. Whereas previous meta-analyses focused on health behaviours (Rodrigues et al., 2015) or tested underlying mechanisms (Wood

et al., 2016), the present review considers how methodological factors influence the magnitude of the QBE literature as a whole. In addition, the review focuses attention on differences between studies of desirable and undesirable behaviours in relation to the QBE. While there is an overlap between the studies included in the present and other meta-analytic reviews, the present meta-analysis considers a larger number of studies and a different set of moderators (Table 2). Table 2 provides a summary of the key contributions of the present meta-analytic review compared to previous meta-analyses. The present meta-analysis is distinct from our previous meta-analysis (Wood et al., 2016) in terms of the 14 moderators considered: 10 of these were not previously considered, while a further 4 use a more refined coding of categories. Another seven moderators (experience with behaviour, whether questions were based on the theory of planned behaviour, degree of correspondence between cognition and behaviour measures, type of behaviour, frequency of the behaviour, difficulty of the behaviour, objective vs. subjective measurement of behaviour) that

Table 2. Summary the contributions of the four recent QBE meta-analyses.

	Wood et al. (2016)	Rodrigues et al. (2015)	Spangenberg et al. (2016)	The present review ^a
Includes Randomised Controlled Trials	✓	✓	✓	✓
Includes Non-Randomised Controlled Trials	✓		✓	✓
Include studies on health behaviour	✓	✓	✓	✓
Include studies on non-health behaviours	✓		✓	✓
Include questions assessing behaviour as well as prediction/intention		✓		✓
Number of papers (studies/ tests) included	35 (55 studies)	33 (41 studies)	51 (104 studies)	65 (94 studies)
Number of moderators analysed	18	6	13	13 ^b
Risk of bias included as moderator		✓ ¹		✓ ¹
Includes new, unpublished studies as part of review				✓
Includes discussion of mechanisms	✓		✓	✓
Includes mechanism meta- analysis	✓			✓

^a A further seven moderators were also considered in the present analysis although these had previously been considered by Wood et al. (2016) and are therefore not included in the main text. Results of these additional moderators are available in on-line materials (Table D1).¹ The coding and analysing of risk of bias in the present review categorised bias into low and high risk studies, it also considered overall bias score (i.e., as a continuous variable predicting effect size). This was more rigorous than the simple approach conducted in Rodrigues et al. (2015) which grouped the high risk studies together and the low risk studies together.

overlap with Wood et al. (2016) and produced similar effects (only difficulty of the behaviour was a significant moderator) are reported in Appendix D.

In the following sections, we: (1) consider the evidence concerning three different theoretical accounts of the QBE (behavioural simulation and processing fluency; attitude accessibility and cognitive dissonance); (2) report a new meta-analysis of QBE studies focusing on methodological moderators including risk of bias and the difference between desirable and undesirable behaviours and (3) discuss future directions for research on the QBE in relation to maximising or minimising the effect.

Theoretical accounts of the question-behaviour effect

The most prominent theoretical accounts of the QBE consider processes related to behavioural simulation and processing fluency (Janiszewski & Chandon, 2007; Levav & Fitzsimons, 2006; Sherman, 1980), impacts on the accessibility of attitudes (Morwitz & Fitzsimons, 2004; Morwitz et al., 1993) and cognitive dissonance processes (Spangenberg & Greenwald, 1999; Spangenberg & Sprott, 2006; Spangenberg, Sprott, Grohmann, & Smith, 2003). The following subsections discuss these key explanations of the QBE in turn, assess the evidence for each, and identify gaps in the literature.

Behavioural simulation and processing fluency

A first theoretical account of the QBE focuses on the processes involved in the simulation of behaviour, and related effects on processing fluency. Sherman (1980), in his demonstration of the QBE, suggested that the effect was driven by the formation of cognitive representations or behavioural scripts during questioning (i.e., mental simulation) which become reactivated when the individual has the opportunity to perform the behaviour. Sherman (1980) hypothesised that this mental simulation increases the accessibility of the behavioural script or the perceived likelihood of behaviour and that either process could increase the likelihood of behaviour that is consistent with the representation.

Indirect support for the role of behavioural simulation in the QBE is provided by demonstrations that ease of representation influences the effect. Levav and Fitzsimons (2006) argued that being asked to predict one's future behaviour is likely to lead to participants mentally representing the behaviour, and that participants subsequently reflect upon how easy or difficult they found it to represent the behaviour. Greater ease of representation is misinterpreted as an increased likelihood of the behaviour's occurrence that is then translated into an increase in actual behaviour. Relatedly, Song and Schwarz (2008) suggest that a fluent simulation of the behaviour at the time of questioning may lead to the

impression that the behaviour is easier to perform and thereby increase motivation to act. This ease of representation hypothesis suggests that the QBE should be attenuated for behaviours that are more difficult to represent. Consistent with this prediction, Levav and Fitzsimons (2006) reported that questions thought to promote ease of representation (e.g., asking participants with likely negative attitudes towards fatty food about their intentions to *avoid eating* fatty foods) increased the QBE. In a meta-analysis of QBE studies, Wood et al. (2016) tested whether the congruence between the question frame and likely attitude distribution moderated the QBE. The meta-analysis did reveal a significant negative effect of rated ease of representation on the QBE ($\beta = -0.11$, $p = .02$). However, this finding is not consistent with a behavioural simulation explanation of the QBE because greater rated ease of representation was associated with *smaller* effect sizes. Wood et al. (2016) observed the strongest QBEs when questions tapped self-predictions ($d_+ = 0.29$), followed by studies where the questions tapped self-predictions and intentions ($d_+ = 0.14$), and then studies where the questions tapped intentions only ($d_+ = 0.12$). To the extent that self-predictions are more likely to generate behavioural simulation, this finding could be interpreted as evidence supporting behavioural simulation as the mechanism underlying the QBE.

Research has also provided more direct support for the role of processing fluency in enhancing the accessibility or the perceived likelihood of the behaviour. Janiszewski and Chandon (2007) argued that the QBE prompts processing fluency effects in the form of transfer-appropriate processing. The proposal is that for those who have made predictions about their own behaviour, activation of the behavioural representation and processes involved in deciding whether to act will be facilitated because the same behavioural representation and processes are accessed at the moment of acting as when participants previously predicted their behaviour. Janiszewski and Chandon (2007) suggest that this increased processing fluency may be misinterpreted as an increased probability of the behaviour actually occurring (i.e., an inclination towards the behaviour) that serves to change subsequent behaviour. Consistent with this explanation, Janiszewski and Chandon (2007) reported a larger QBE when the correspondence between the intention and behaviour measures was greater. Wood et al. (2016) used the principle of correspondence (e.g., Azjen & Fishbein, 1977; Fishbein & Ajzen, 1975) to quantify the match between questions and behaviour (along the dimensions of target, action, context and time), in order to examine the effect of processing fluency on the QBE across studies. However, they found no significant effect of correspondence on the size of the QBE both across all studies ($\beta = -0.03$, $p = .43$) or for studies that used objective measures of behaviour ($k = 76$, $\beta = -0.05$, 95% $CI = -0.13, 0.03$, $p = .23$) that are

less likely to be biased by common method variance effects (see Conner, Warren, Close, & Sparks, 1999).

Wood et al. (2016) concluded that their meta-analysis provided little support for the behavioural simulation and processing fluency as general explanations of the QBE. We are not aware of more recent research that counters such a negative conclusion about this theoretical account of the QBE. However, further studies that manipulate and/or measure processing fluency are required before a definitive conclusion about this mechanism can be reached. In addition, exploration of more specific conditions under which such a QBE mechanism might operate may be warranted (e.g., exploring what aspects of behavioural representation are important in behavioural enactment).

Attitude accessibility

A second theoretical account of the QBE focuses on the impact of asking questions on the accessibility of attitudes towards that behaviour. In this attitude accessibility account of the QBE, it is assumed that asking individuals to report their behavioural intentions or to predict their behaviour activates the attitude underlying that behaviour and so makes it more accessible in memory. Consequently, this heightened accessibility of the relevant attitude increases the likelihood that individuals will act in a manner consistent with their attitude (Dholakia, 2010; Morwitz & Fitzsimons, 2004; Morwitz et al., 1993). There is some evidence supporting both of these links in the path from questioning to behaviour.

In relation to the impact of questioning on attitude accessibility, studies have demonstrated that participants who are asked to report their intentions or to predict their behaviour exhibit more accessible attitudes relative to those who are not asked (Chapman, 2001; Fitzsimons, Nunes, & Williams, 2007; Morwitz & Fitzsimons, 2004; Wood, Conner, Sandberg, Godin, & Sheeran, 2014). Evidence shows that accessible attitudes are associated with stronger attitude-behaviour relationships (Chen & Bargh, 1999; Fazio, Chen, McDonel, & Sherman, 1982; Fazio & Williams, 1986; for a meta-analysis, see Cooke & Sheeran, 2004). Wood et al. (2014) recently demonstrated that intention questions increased attitude accessibility and that attitude accessibility was related to behaviour within a single study. Moreover, attitude accessibility mediated the relationship between intention measurement and subsequent behaviour.

The attitude accessibility account of the QBE suggests that the effects of questioning should depend upon the valence of attitudes towards the behaviour (Fitzsimons & Morwitz, 1996; Godin et al., 2008; Morwitz et al., 1993). For instance, Morwitz and Fitzsimons (2004) observed that completing purchase intention questions increased the activation level of

preexisting brand attitudes. When the brand attitude was both highly accessible and positively valenced, participants were more likely to choose that brand, whereas when the activated attitude was both highly accessible and negatively valenced, participants were less likely to choose that brand (Morwitz & Fitzsimons, 2004). Other research shows that the valence of attitudes towards the behaviour moderates the QBE in line with an attitude accessibility account, such that participants reporting positive attitudes show a stronger QBE than do those with negative attitudes (Conner et al., 2011). Ayres et al. (2013) showed this effect experimentally, with the QBE being stronger when combined with a manipulation designed to increase positive attitudes. Other studies show that asking questions can decrease behavioural performance among participants with negative attitudes (e.g., Conner et al., 2011, Study 2).

Support for the attitude accessibility account of the QBE is by no means ubiquitous, however. Both Perkins, Smith, Sprott, Spangenberg and Knuff (2008) and Spangenberg et al. (2012) found no significant differences in attitude accessibility between participants who did or did not predict their own behaviour. In addition, a number of demonstrations of the QBE occurred under conditions not easily accounted for by attitude accessibility. For example, attitude accessibility does not provide a convincing explanation of the QBE for behaviours performed long after questioning, when increases in attitude accessibility prompted by questioning have presumably decayed (e.g., Godin et al., 2008 observed QBE effects on blood donation 6 and 12 months after questioning).

The meta-analysis of the QBE by Wood et al. (2016) assessed a number of predictions derived from the attitude accessibility account of the QBE and observed only limited support for the mediating role of attitude accessibility. There were too few studies reporting response latency measures of attitude accessibility to permit a quantitative synthesis. Therefore, an indirect measure of accessibility was created based on the likely valence of the attitude (using scores from independent raters) and the proportion of participants whose attitude was activated (based on the response rate to completing questions). This indirect measure of accessibility was found to be significantly related to the size of the QBE, supporting the attitude accessibility account of the QBE. However, this effect became non-significant in multivariate analyses controlling for other QBE moderators. Moreover, a number of other tests conducted by Wood et al. (2016) did not support the attitude accessibility mechanism. First, given that repeated expression of an attitude increases attitude accessibility (Fazio et al., 1982), it might be expected that the number of intention or self-prediction questions or even the total number of questions relating to behaviour influences the size of the QBE. However, neither variable was related to the size of the QBE ($p_s > .11$). Second, although direct experience with the behaviour

increases attitude accessibility (Fazio & Zanna, 1981), experience was unrelated to the size of the QBE. Overall the evidence for an attitude accessibility mechanism underlying the QBE is at best mixed. As for the behavioural simulation and processing fluency account of the QBE, exploration of the specific conditions under which an attitude accessibility mechanism might operate would be valuable (e.g., exploring the extent to which change in accessibility is maintained over time and whether this parallels persistence of the QBE).

Cognitive dissonance

A third theoretical account of the QBE focuses on cognitive dissonance (Festinger, 1957). Festinger (1957, p. 3) defined cognitive dissonance as “the existence of non-fitting relations among cognitions” where cognitions include “any knowledge, opinion or belief about the environment, about oneself, or about one’s behavior”. Cognitive dissonance is a tension state that motivates efforts to reduce dissonance. In relation to the QBE, cognitive dissonance accrues when an individual performs a behaviour that is inconsistent with a relevant standard of judgment, i.e., when people’s actions and their beliefs about how they should act are inconsistent (Stone & Cooper, 2001). Answering questions about a behaviour can increase the salience of social norms associated with the behaviour (a standard of judgment) and also any previous failures to behave in a manner that is consistent with such norms (discrepancies from standards). Perceived inconsistency between the two should generate cognitive dissonance. Cognitive dissonance should be reduced by subsequently acting in accordance with the social norms or standards (Aronson, 1992), resulting in a QBE. Stone and Cooper (2001), in their self-standards model of cognitive dissonance, noted that both normative (i.e., perceived norms) and personal (i.e., individual attitudes) standards can act as anchors for judgement. Personal goals or resolutions can also serve as standards in QBE studies (Dholakia, 2010).

Another way to reduce dissonance is to engage in downward comparisons. Consistent with the dissonance account of the QBE, participants asked to predict their own behaviour were more likely to engage in downward comparisons, presumably in order to reduce any dissonance generated by making self-predictions (Spangenberg et al., 2003). No published QBE study has directly measured cognitive dissonance and tested its potential mediation effect. However, various studies have explored moderator effects that might offer indirect evidence for the cognitive dissonance account of the QBE (Sprott et al., 2006a) but have only found mixed support (Spangenberg & Sprott, 2006; Spangenberg et al., 2003, 2012). For example, Spangenberg et al. (2003) reported that a self-affirmation manipulation

(known to reduce cognitive dissonance) attenuated the QBE. In contrast, Sprott et al. (2003) found preference for consistency (that increases susceptibility to cognitive dissonance; Cialdini, Trost & Newsom, 1995) did not moderate the QBE.

Wood et al. (2016) tested the cognitive dissonance account of the QBE by rating studies for degree of dissonance likely created. Three ratings were used: the likely degree of discomfort (i.e., cognitive dissonance) experienced by participants at the time of prediction if their past behaviour was not consistent with the normative or personal standards conveyed by their self-predictions or intentions; the likely degree of discomfort participants would experience if their future behaviour was not consistent with their predictions/intentions at the time of prediction; and the likely degree of discomfort participants would experience if their future behaviour was not consistent with their predictions/intentions at the moment of enacting the behaviour. These three ratings were found to be consistent across raters ($ICC = .66-.78$) and to form a reliable scale ($alpha = .89$). However, rated cognitive dissonance was not significantly related to the size of the QBE ($k = 116$, $\beta = -0.05$, $95\% CI = -0.12, 0.01$, $p = .12$) in univariate analyses. Wood et al. (2016) also tested whether answering a greater number of questions about future behaviour (which might be expected to increase cognitive dissonance) increased the QBE. However, neither the number of intention or self-prediction questions nor the total number of questions relating to behaviour was related to the size of the QBE ($ps > .11$). In sum, there is only modest evidence to date that dissonance has an important role in explaining the QBE.

New evidence in relation to the cognitive dissonance explanation of the QBE

Here, we present two new lines of evidence concerning the dissonance explanation of the QBE. The first is a reanalysis of Wood et al. (2016) that focuses on how dissonance may moderate the impact of other drivers of the QBE. The second is a new study that offers a comparative test of the attitude accessibility and dissonance accounts of the QBE. We reanalysed Wood et al.'s data using a backwards elimination procedure for multivariate regression analysis (see Smit, Verdurmen, Monshouwer, & Smit, 2008; Steffgen, Recchia, & Viechtbauer, 2013) to see if the forced entry procedure that was originally deployed may have missed the possible role of rated dissonance. Findings revealed four significant moderators that explained approximately one-fifth of the between-study variance in effect sizes for QBE studies (adjusted $R^2 = 19.22\%$). Consistent with expectations, rated dissonance emerged as a significant predictor ($\beta = -0.08$, $p = .010$), as did social desirability ($\beta = 0.18$, $p < .001$), provision of an incentive ($\beta = 0.20$,

$p = .002$) and sample type ($\beta = 0.17, p = .004$)—three variables that were also significant in the univariate analyses reported by Wood et al. (2016).

It is notable that the beta for dissonance is *negative* in this analysis. One possible reason is that rated dissonance is a suppressor variable (e.g., Maassen & Bakker, 2001; Paulhus, Robins, Trzesniewski, & Tracy, 2004). However, contrary to this interpretation, rated dissonance was not significantly associated with the other predictors ($|r| \leq 0.18, ps > .05$), and did not appreciably enhance their ability to predict effect sizes (mean change in standardised $\beta = 0.01$). Another possibility is that rated dissonance interacts with the other predictors. Consistent with this idea, we observed two significant interactions.

First, the dissonance \times social desirability factor interaction proved reliable ($\beta = -0.084, 95\% CI = -0.17, -0.003, p = .042$). Simple slopes analyses (see Johnson, Low, & MacDonald, 2014) indicated that social desirability was associated with effect sizes when dissonance was low ($\beta = 0.211, 95\% CI = 0.10, 0.32, p < .001$) but not when dissonance was high ($\beta = 0.079, 95\% CI = -0.003, 0.16, p = .06$). Thus, the social desirability of the behaviour appears to be an important guide to acting on self-predictions/intentions when inaction arouses little dissonance. The second interaction indicated that dissonance altered the relationship between behavioural difficulty and magnitude of the QBE in a step function ($\beta = -0.50, 95\% CI = -0.80, -0.19, p = .001$; see Baron & Kenny, 1986, p. 1176). At extremely high levels of rated dissonance (a rating of 5 on the original 5-point scales), greater difficulty was associated with a *larger* QBE ($\beta = 0.11, 95\% CI = 0.01, 0.21, p = .04$). When rated dissonance was moderate or low (ratings = 1–4), however, behavioural difficulty was associated with a *smaller* effect of intention/prediction questions on behaviour ($\beta = -0.06, 95\% CI = -0.11, -0.01, p = .02$). Thus, behavioural difficulty generally attenuates the QBE, except when participants can anticipate very considerable discomfort from failing to act. Under these circumstances, the more difficult is the behaviour, the more likely participants are to act.

This interaction speaks to a long-standing anomaly in research concerning the role of perceived or task difficulty in behavioural performance. In research on goals and task performance, a positive relationship is observed such that people perform better on difficult goals (e.g., Locke, 1968; Locke, Shaw, Saari, & Latham, 1981). The idea is that difficult goals are more challenging than easy goals and engender greater effort mobilisation, and hence better performance. Research on behavioural prediction (e.g., Ajzen, 1991), on the other hand, finds that greater perceived difficulty is associated with weaker intentions and reduced performance of behaviour (see, e.g., Armitage & Conner, 2001; McEachan, Conner, Taylor, & Lawton, 2011; for meta-analyses). The present findings are consistent with both lines of research. When modest or little dissonance would accrue from failing to

act on intentions/predictions, then behavioural difficulty has a significant, negative relationship with effect sizes. However, when extremely high levels of discomfort are anticipated, then difficulty has a positive, linear relationship with the QBE. Thus, extreme dissonance serves to transform behavioural difficulty into motivation. Asking people their intentions or self-predictions regarding a behaviour makes subsequent action especially likely when the behaviour is both difficult to perform and extreme dissonance would accrue from non-performance. Thus, we observe some indirect support for the dissonance explanation of the QBE. But rather than a direct effect, dissonance appears to be influential in determining how social desirability and behavioural difficulty shape the impact of answering questions on behaviour change.

The second new line of evidence comes from a study that pitted the attitude accessibility and cognitive dissonance explanations of the QBE against one another by exploring the effects of incentives for completing questionnaires on the size of the effect on behaviour (Conner et al., [in preparation](#)). The attitude accessibility mechanism of the QBE would predict that incentives enhance the QBE for participants with positive attitudes but attenuate it for participants with negative attitudes. In contrast, the cognitive dissonance explanation of the QBE suggests that incentives for questionnaire completion could provide a sufficient justification for action (Festinger & Carlsmith, 1959) that attenuate the QBE. In the original Festinger and Carlsmith (1959) study, the effects of dissonance on attitude change were observed only when there was insufficient justification for action (i.e., participants were paid only \$1). When the action could be justified based on sufficient incentive (a \$20 payment), then dissonance effects were not observed. Thus, the standard QBE condition, where a questionnaire is sent with no incentive represents the insufficient justification condition. In contrast, sending a questionnaire with an incentive to complete and return it represents the sufficient justification condition. While in the former condition dissonance might be experienced after completing and returning a questionnaire and this dissonance might generate a QBE, in the latter condition dissonance might not be experienced after completing and returning a questionnaire and so no QBE would be generated. In addition, in the latter condition, the incentive might make more respondents with less positive views of the behaviour complete and return the questionnaire.

Conner et al. ([in preparation](#)) tested these predictions in two QBE studies on screening behaviour. In both studies there were three conditions: a control condition with an unrelated questionnaire, an experimental condition where participants were sent a questionnaire on the target behaviour and an experimental condition where participants were sent a questionnaire on the target behaviour plus an incentive to complete and return the

questionnaire. Study 1 focused on bowel screening kit return with the incentive being £5 for completing and returning the questionnaire. Among those who completed and returned questionnaires, there were no differences on mean cognitions (intention, attitude, etc.) about the behaviour between the condition where the questionnaire about bowel screening was or was not sent with an incentive. This suggests that the incentive did not lead to different samples completing and returning the questionnaire (e.g., more of those who were less positive about the behaviour completing and returning the questionnaire when an incentive was offered). However, rates of the behaviour (i.e., returning the screening kit) were significantly higher in the no incentive condition (97.8% return rate) compared with either the incentive condition (94.3%) or the control condition (94.5%). In addition, across the two experimental conditions there was no significant interaction between level of cognition and condition (standard vs. incentive QBE conditions) on screening rates. Study 2 showed a very similar pattern of results in relation to cervical screening attendance when using a different incentive to promote questionnaire return (social approval based on a request for help; see Garner, 2005). The pattern of findings offer better support for the cognitive dissonance than the attitude accessibility account of the QBE.

Although the present review offers new evidence concerning the dissonance explanation of the QBE, further studies are clearly required. Additional tests that directly pit one mechanism against another and further tests of incentives (Forster et al., 2014) would be desirable. However, progress is most likely to accrue from laboratory studies that deploy non-reactive measures of both attitude accessibility (Wood et al., 2014) and cognitive dissonance (e.g., physiological measures; Harmon-Jones, Brehm, Greenberg, Simon & Nelson, 1996). Studies that manipulate attitude accessibility, cognitive dissonance or behavioural fluency might also provide useful insights into the relative importance of these different mechanisms under varying conditions. Researchers need to be open to the possibility that different factors drive the QBE under different conditions and that more than one mechanism may underlie the QBE (Dholakia, 2010; Spangenberg et al., 2016).

Methodological factors underlying the QBE: A meta-analysis

This section reports a new meta-analysis of QBE studies. Details of the protocol for this systematic review were registered on PROSPERO and can be accessed at: www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42014006595. The current meta-analysis extends previous meta-analyses (Rodrigues et al., 2015; Spangenberg et al., 2016; Wood et al., 2016) in three ways. First, the present meta-analysis includes between 12 and 30 more papers (Table 2) than

previous meta-analyses (Rodrigues et al., 2015; Spangenberg et al., 2016; Wood et al., 2016). Second, the present meta-analysis was less restrictive than previous ones in terms of behaviours, designs and questions examined. It also includes a broad range of consumer, health and prosocial behaviours and a range of questions used to elicit the QBE including satisfaction, past behaviour, attitudes, intentions and self-predictions. Third, compared to our previous meta-analysis (Wood et al., 2016), the present meta-analysis considers ten new moderators (commitment, correspondence of question and behaviour, specific behaviour, type of health behaviour, normative behaviour, directedness of behaviour, baseline measure taken, research design, analysis method, risk of bias) and four similar moderators but with a more refined set of categories (sample type, study setting, question type, delivery method). See on-line materials Appendix D for coding and findings of the seven moderators similar to those explored by Wood et al. (2016). The moderators cover aspects of the sample, intervention, outcomes and methodology and offer an analysis of the impact of methodological bias on the magnitude of the QBE within this wide range of studies. As noted earlier, Table 2 summarises the differences between the different meta-analyses of the QBE.

Methodology of meta-analysis

Inclusion criteria

To be included in the review, studies had to meet all of the following criteria: (a) at least one group of participants were questioned on cognitions and/or behaviour before follow-up, (b) at least one group of participants were not questioned on cognitions and/or behaviour before follow-up and (c) there was a measure of behaviour at follow-up in both treatment and comparison groups.

Search terms

PsycINFO 1806-February 2015, MEDLINE 1946-February 2015 and EMBASE 1946-February 2015 were searched using OVID for articles published between 1980 (when the first study of the QBE was published: Sherman, 1980) and February 2015 (see on-line materials Appendix A for search terms). To supplement the database searches, the reference lists of identified studies were examined along with those of recent reviews (Dholakia, 2010; Rodrigues et al., 2015; Sprott et al., 2003). Contact was made with the first author of each of the included studies to identify additional studies including yet to be published studies. The titles and abstracts were screened by the lead author and also independently screened by two further authors; discrepancies across coders were discussed and agreement on coding reached. All full text screening was carried out independently by two authors.

Moderators

Where studies had multiple experimental conditions compared against a single control, we selected the experimental condition that was most similar to the control condition. The first author extracted the data for all studies and 10% of studies selected at random were coded by a co-author. Interrater reliability was perfect for all extracted moderators ($k = 1.0$ indicating perfect agreement). The key-dependent variable identified by the authors was used in the meta-analysis and input into Comprehensive Meta-Analysis (Borenstein, Hedges, Higgins, & Rothstein, 2005) and Stata (StataCorp, 2013). Four groups of moderators were assessed: population, intervention, outcome and methodology.

Population. *Sample type* was coded into: (a) students (university), (b) medical patients, (c) schoolchildren (adolescents/school pupils), (d) workers (recruited from specific workplaces) or (e) other (not in other categories). Foot and Sanford (2004) noted that students may be more likely to complete questionnaires honestly and rationally. This may enhance the QBE in student samples. *Study setting* was coded into (a) education, (b) medical, (c) community, (d) laboratory or (e) online in order to assess if there was an influence of setting on QBE. The controlled conditions of laboratory settings are expected to increase engagement with questions. *Commitment* levels were coded based on level of contact with the experimenter and coded as (a) low (little contact) or (b) high (moderate or high contact).

Intervention. *Question type* was coded as (a) prediction, (b) intention only, (c) intention combined with other cognitions, (d) satisfaction or (e) behaviour. *Correspondence of question and behaviour* was coded into (a) question: cognition or (b) question: behaviour. The *number of items* was also coded.

Outcomes. *Specific behaviour* was coded into (a) flossing, (b) health assessment, (c) risky driving, (d) drug use, (e) physical activity, (f) purchasing, (g) vaccination, (h) blood donation, (i) screening, (j) condom use, (k) voting or (l) alcohol consumption. *Type of health behaviour* was coded into (a) approach where performing the behaviour is healthy (desirable) or (b) avoid where not performing the behaviour is healthy (undesirable). *Normative behaviour* was coded as (a) normative when behaviour encouraged by most others (e.g., eating healthily), (b) unclear or (c) non-normative when behaviours would generally be discouraged by most others (e.g., smoking). *Directedness of behaviour* was coded in terms of (a) self-directed when the behaviour is one that is performed primarily for the interest of the individual performing it (e.g., healthy eating), (b) other directed when it is

performed for another person, (c) both directed where it is performed for both (e.g., voting) or (d) unclear directed when directedness of behaviour is unclear.

Methodology. *Baseline measure* was coded into (a) not assessed, or (b) assessed. *Delivery method* was coded as (a) face-to-face, (b) mailed, (c) telephone, (d) PC/Internet or (e) other/unclear. *Research design* was coded as (a) Randomised Controlled Trial (RCT), (b) non-RCT or (c) Solomon group design. *Analysis* was coded into (a) per protocol if only participants completing measures were analysed, or (b) intention to treat if all participants were analysed. For example, in studies where participants are mailed questionnaires, the QBE can be assessed in the overall sample (intention-to-treat) or just among participants who completed and returned questionnaires (per protocol). Studies were also coded by *time interval* between questioning and measurement of cognitions or behavioural DVs. *Return rate* at final follow-up was assessed as the percentage reported in the paper. *Risk of bias* used the Cochrane Collaboration's measure to assess the risk of bias and covered sequence allocation, allocation concealment, blinding, incomplete outcome reporting, selective outcome reporting and other bias. Risk of bias was coded into (a) low or (b) unclear/high. When any of the six categories of risk of bias (sequence generation, allocation concealment, blinding, incomplete outcome reporting data, selective outcome reporting, other sources of bias) was rated as unclear or high risk of bias, the overall study was categorised as falling in the unclear/high risk of bias category. Each study was only rated as low risk of bias, if all six risk of bias categories were rated as low risk.

Analysis

Comprehensive meta-analysis software (Borenstein et al., 2005) was used to calculate effect sizes and for subgroup analyses. Stata (StataCorp, 2013) was used to carry out meta-regression analyses. Hedges g and 95% confidence intervals were calculated for each study based on a random effects model.

Meta-analysis findings

The 65 papers that met the inclusion criteria reported 94 tests of the QBE ($N = 116,087$; see on-line materials Appendices B and C for list of studies and coding of studies). Overall random effects based on 94 tests showed a small but significant QBE ($g = 0.14$, 95% $CI = 0.11, 0.18$, $p < .001$). This is similar in magnitude to that reported in previous meta-analyses of the QBE (Rodrigues et al., 2015: Cohen's $d = 0.09$, 95% $CI = 0.04, 0.13$; Spangenberg et al., 2016: Cohen's $d = 0.28$, 95% $CI = 0.24, 0.32$; Wood et al., 2016: $d = 0.24$, 95% $CI = 0.18, 0.30$; current meta-analysis: Cohen's $d = 0.15$, 95%

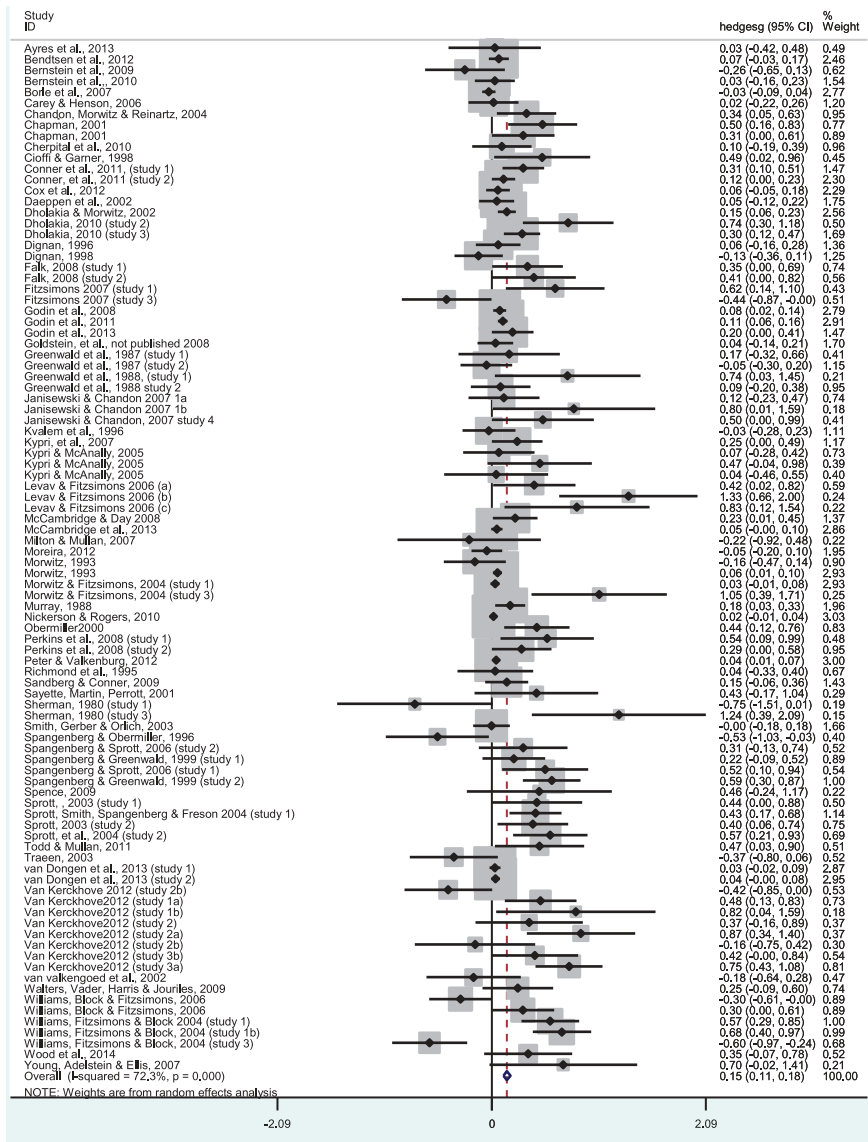


Figure 1. Study effect sizes from random effects meta-analysis. Diamonds represent Hedges' g effect size, horizontal lines indicate 95% confidence intervals and shaded sections indicate study percentage weight.

CI = 0.11, 0.19, $p < .001$). There was moderate-to-high heterogeneity among study effect sizes ($I^2 = 72.9\%$, $Q = 343.12$, $p < .001$). **Figure 1** shows the forest plot depicting the effect size for each study. Funnel plots showed that the effect sizes were not symmetrical: more studies with larger effect sizes had larger standard errors. Egger's regression revealed significant

asymmetry ($p < .001$), suggesting the findings were susceptible to publication bias. Trim and fill analysis (Taylor & Tweedie, 1998) estimated that there were 20 missing studies and inclusion of such studies would produce a smaller, but still significant QBE ($g = 0.08$, 95% $CI = 0.04, 0.12$).

The heterogeneity in findings supported the exploration of moderation effects. Only significant moderators are reported in any detail below, with non-significant moderator findings noted. Table 3 reports the subgroup moderator analyses and pairwise comparisons between categories based on mixed effects analysis (Borenstein, Hedges, Higgins, & Rothstein, 2009).

Population moderation effects

Subgroup analyses showed significant heterogeneity between studies using different *sample types* ($Q = 22.16$, $p = .002$). The majority of studies ($k = 46$) used student samples and the QBE was found to be largest in this group ($g = 0.27$, 95% $CI = 0.18, 0.35$, $p < .01$). A much smaller proportion of studies investigated employee, healthcare and school pupil samples and the QBE was found to be smaller in these three groups. Pairwise comparisons (Table 3) indicated a significantly larger effect size in studies using student samples compared to healthcare patients ($Q = 11.47$, $p = .003$), school pupil samples ($Q = 6.77$, $p = .03$), specific employee samples ($Q = 6.50$, $p = .04$) and samples that did not fit into one of these categories ($Q = 15.61$, $p < .001$). No other differences between pairs of categories were significant. *Study setting* was also a significant moderator of effect sizes ($Q = 25.06$, $p < .001$). Laboratory based QBE studies produced the largest overall effect on cognitions or behaviour ($g = 0.33$, 95% $CI = 0.19, 0.47$, $p < .001$, $k = 19$). Studies using a laboratory setting produced a significantly greater effect size than those observed in medical ($Q = 11.24$, $p < .001$), community ($Q = 11.68$, $p = .001$) and online ($Q = 12.26$, $p < .001$) settings (Table 3). No other differences between categories were significant. There were no significant effects of *commitment* ($Q = 1.28$) on effect sizes (Table 3).

Intervention moderation effects

The *question type* used varied between studies: prediction ($k = 24$), intention ($k = 26$) and behaviour ($k = 23$) were the most frequently used questions; smaller proportions of studies assessed intentions combined with other cognitions ($k = 16$) or satisfaction measures ($k = 5$). Subgroup analysis indicated that the questions used significantly influenced the size of the QBE ($Q = 17.96$, $p < .001$). Prediction ($g = 0.25$, 95% $CI = 0.13, 0.37$, $p < .01$, $k = 24$) and intention ($g = 0.23$, 95% $CI = 0.14, 0.31$, $p < .001$, $k = 26$) questions produced the largest QBE and were not significantly different from one another ($Q = 1.17$, $p = .28$). Studies tapping satisfaction ($g = 0.16$, 95% $CI = 0.002, 0.32$, $p = .004$, $k = 5$), intention combined with other measures ($g = 0.08$, 95% $CI = 0.02, 0.13$, $p = .004$, $k = 16$) and

Table 3. Subgroup analysis of potential QBE moderators.

Moderator	Category	<i>k</i>	<i>g</i>	<i>CI</i>	<i>p</i> ³	<i>Q</i>	<i>p</i>	
<u>Population</u>								
<u>Sample type</u>	Students	46	.27 ^b	18, .35	.00	22.16	.002	
	Medical patients	9	.08 ^a	.01, .15	.02			
	Schoolchildren	6	.12 ^a	-.04, .29	.13			
	Workers	2	.11 ^a	.003, .23	.04			
	Other	31	.09 ^a	.05, .13	.00			
<u>Study setting</u>	Education	18	.16 ^{ab}	.006, .32	.04	25.06	.001	
	Medical	10	.06 ^a	-.01, .13	.10			
	Community	26	.07 ^a	.04, .11	.00			
	Laboratory	19	.33 ^b	.19, .47	.00			
	Online	6	.05 ^a	-.01, .11	.10			
<u>Commitment</u>	Low	79	.15 ^a	.11, .19	.00	1.28	.25	
	High	15	.08 ^a	-.01, .18	.00			
<u>Intervention</u>								
<u>Question type</u>	Prediction	24	.25 ^b	.13, .37	.00	17.96	.001	
	Intention only	26	.23 ^{ab}	.14, .31	.00			
	Combined intention	16	.08 ^a	.02, .13	.004			
	Satisfaction	5	.16 ^{ab}	.002, .32	.04			
	Behaviour	23	.07 ^a	.01, .13	.01			
<u>Correspondence of question and behaviour</u>	Cognition-Behaviour ¹	57	.16 ^b	.11, .20	.01	9.14	.003	
	Behaviour-Behaviour ¹	26	.07 ^a	.02, .11	.02			
<u>Outcomes</u>								
<u>Specific behaviour</u>	Flossing	3	.61 ^a	.39, .83	.00	83.14	.00	
	Health assessment	5	.41 ^{ab}	.26, .56	.00			
	Risky driving	2	.37 ^{ab}	.11, .63	.005			
	Drug use	2	.28 ^{bcd}	.04, .52	.02			
	Physical activity	4	.22 ^{bc}	.07, .38	.004			
	Purchasing	14	.18 ^{ce}	.09, .27	.00			
	Vaccination	2	.08 ^{cf}	.009, .17	.03			
	Blood donation	5	.06 ^{df}	.02, .10	.002			
	Screening	4	.01 ^{def}	-.12, .15	.80			
	Condom use	2	-.15 ^{ef}	-.47, .17	.36			
	Voting	8	.06 ^{cf}	-.03, .16	.19			
	Alcohol consumption	12	-.05 ^h	-.11, .007	.09			
	<u>Type of health behaviour</u>	Avoid (unhealthy)	15	-.07 ^a	-.15, .00	.051	5.37	.02
		Approach (healthy)	34	.17 ^b	.10, .23	.00		
<u>Normative behaviour</u>	Normative	45	.12 ^a	.08, .16	.00	3.26	.20	
	Unclear	22	.07 ^a	.007, .13	.03			
<u>Directedness of behaviour</u>	Non-normative	12	.08 ^a	.11, .28	.38			
	Self-directed	38	.14 ^a	.06, .21	.00	2.07	.36	
	Other-directed	20	.07 ^a	.02, .12	.003			
	Both directed	20	.13 ^a	.07, .19	.00			
Unclear directed	1	.02 ^a	-.09, .03	.40				
<u>Methodology</u>								
<u>Baseline measure</u>	Not assessed	81	.16 ^b	.12, .20	.00	14.86	.001	
	Assessed	13	.05 ^a	.01, .09	.007			
<u>Delivery method</u>	Face-to-face	42	.23 ^a	.14, .33	.00	20.84	.001	
	Mail	8	.06 ^b	.03, .09	.00			
	Phone	10	.11 ^{ab}	.03, .20	.00			
	PC/Internet	15	.10 ^{ab}	.03, .17	.00			
	Other/unclear	9	.25 ^{ab}	.01, .49	.04			
	Unclear	1	.02 ^a	-.09, .03	.40			
<u>Research design</u>	Randomised Controlled Trials	22	.07 ^a	.04, .11	.00	17.29	.001	
	Non-Randomised Controlled Trials	64	.20 ^b	.15, .25	.00			
	Solomon four group	8	.02 ^a	-.13, .17	.77			
<u>Analysis</u>	Per protocol	7	.42 ^a	.26, .58	.001	15.94	.001	
	Intention to treat	7	.08 ^b	.04, .13	.001			
<u>Risk of bias</u>	Low	27	.07 ^a	.04, .10	.001	14.88	.001	
	Unclear/High	67	.20 ^b	.14, .25	.001			

¹ Question first, dependent variable second: Cog: cognition, Beh: behaviour. ^{ab} Pairwise comparisons within specified moderator (effect sizes not sharing a postscript are significantly different based on fixed effects test, $p < .05$). ³ p -values marked as .00 indicate $p \leq .001$.

behaviour ($g = 0.07$, 95% $CI = 0.01, 0.13$, $p = .01$, $k = 23$) reported smaller QBEs. Pairwise comparisons indicated that studies using prediction items reported significantly higher effect sizes than those using intention items combined with other cognitions items ($Q = 6.77$, $p = .009$) and those assessing behaviour ($Q = 7.41$, $p = .006$). No other differences were significant (Table 3).

Subgroup analyses were performed on the *correspondence* of question and behaviour, i.e., whether studies measured cognitions or behaviour as the QBE intervention. Studies using cognition measures as the intervention ($g = 0.16$, 95% $CI = 0.11, 0.20$, $p = .01$, $k = 57$) reported significantly stronger QBE than studies measuring behaviour at both intervention and follow-up ($g = 0.07$, 95% $CI = 0.02, 0.11$, $p = .02$, $k = 26$), $Q = 9.14$, $p = .003$. The *number of items* did not significantly predict effect size ($\beta = -0.001$, 95% $CI = -0.003, 0.005$, $p = .16$, $k = 77$, see on-line materials Table D2).

Outcome moderation effects

In relation to *specific behaviours*, subgroup analyses showed significant heterogeneity between studies on different behaviours ($Q = 83.14$, $p < .001$). The QBE was found to have a medium-to-large effect on flossing ($g = 0.61$, 95% $CI = 0.39, 0.83$, $p < .001$, $k = 3$), a small-to-medium effect in health assessment ($g = 0.41$, 95% $CI = 0.26, 0.56$, $p < .001$, $k = 5$), risky driving ($g = 0.37$, 95% $CI = 0.11, 0.63$, $p = .005$, $k = 5$), drug use ($g = 0.28$, 95% $CI = 0.04, 0.52$, $p = .018$, $k = 2$), physical activity ($g = 0.22$, 95% $CI = 0.07, 0.38$, $p = .004$, $k = 4$), and purchasing ($g = 0.18$, 95% $CI = 0.09, 0.27$, $p = .001$, $k = 14$). Small effects were found in vaccination ($g = 0.08$, 95% $CI = 0.009, 0.17$, $p = .03$, $k = 2$) and blood donation ($g = 0.06$, 95% $CI = 0.02, 0.10$, $p = .002$, $k = 5$). Non-significant effects were found for screening ($g = 0.01$, 95% $CI = -0.12, 0.15$, $p = .80$, $k = 4$), condom use ($g = -0.15$, 95% $CI = -0.47, 0.17$, $p = .36$, $k = 2$), voting ($g = 0.06$, 95% $CI = -0.03, 0.16$, $p = .19$, $k = 8$) and alcohol consumption ($g = -0.05$, 95% $CI = -0.11, 0.007$, $p = .09$, $k = 12$). Table 3 shows that a number of these specific behaviours had significantly different QBEs, e.g., flossing had a higher QBE than all other specific behaviours except health assessment and risky driving. For type of *health behaviour*, effect sizes were significantly stronger among “approach” compared to “avoid” health behaviours ($Q = 5.37$, $p = .02$). The QBE was associated with performing significantly more approach behaviours ($p < .001$) and non-significantly less avoid behaviours ($p = .05$).

Q-tests revealed no significant differences for *normative behaviour* ($Q = 3.26$, $p = .20$), and *directedness of behaviour* ($Q = 2.07$, $p = .36$); examination of differences between pairs of effect sizes also revealed no significant differences (Table 3).

Methodology moderation effects

The majority of studies did not take a *baseline measure* of behaviour ($k = 81$) and these studies ($g = 0.16$, 95% $CI = 0.12, 0.20$, $p < .001$) had larger effects than studies that did measure behaviour at baseline ($g = 0.05$, 95% $CI = 0.01, 0.09$, $p = .007$, $k = 13$), ($Q = 14.86$, $p = .001$). Subgroup analysis showed that *delivery method* of questioning had a significant effect on the QBE ($Q = 20.84$, $p = .001$). The majority of studies ($k = 42$) used a face-to-face delivery method and this produced the largest QBE ($g = 0.23$, 95% $CI = 0.14, 0.33$, $p < .001$). Smaller effects were found when questions were administered by telephone ($g = 0.11$, 95% $CI = 0.03, 0.20$, $p < .001$, $k = 10$), PC/Internet ($g = 0.10$, 95% $CI = 0.03, 0.17$, $p < .001$, $k = 15$) and mail ($g = 0.06$, 95% $CI = 0.03, 0.09$, $p < .001$, $k = 8$). Pairwise comparisons found that studies using face-to-face delivery produced significantly greater effect size than those using a mailed delivery method ($Q = 6.94$, $p = .008$). No other pairwise comparisons were significant.

Studies were compared based on the *research design* used. Non-RCTs produced the greatest effect size ($g = 0.20$, 95% $CI = 0.15, 0.25$, $p < .001$, $k = 64$), followed by studies using a RCT design ($g = 0.07$, 95% $CI = 0.04, 0.11$, $p < .001$, $k = 22$), with the smallest effects observed in studies using a Solomon group design ($g = 0.02$, 95% $CI = -0.13, 0.17$, $p = .77$, $k = 8$). There was significant heterogeneity between studies based on the study design used ($Q = 17.29$, $p < .001$). Pairwise comparisons showed a significant difference between studies using RCT design compared with non-RCT design ($Q = 16.05$, $p < .001$) and non-RCT design compared to studies using a Solomon group design ($Q = 4.81$, $p = .03$); no significant differences were found between studies using RCT design and Solomon group design ($Q = 0.43$, $p = .51$). The present meta-analysis also compared effect sizes in studies that provided both per protocol or intention-to-treat *analysis*. Seven studies reported both intention to treat and per protocol analyses. Meta-analysing just the per protocol analysis results produced a significantly larger QBE ($g = .42$, 95% $CI = .26, .58$, $p < .001$, $k = 7$) compared to just focusing on intention to treat analysis of the same studies ($g = 0.08$, 95% $CI = 0.04, 0.13$, $p < .001$, $k = 7$) and the difference between effect sizes based on these two subgroups was significant, $Q = 15.94$, $p < .001$.

The *time interval* between questioning and measurement of behavioural DVs varied between immediately (e.g., van Kerckhove et al., 2012) and 5 years (Murray, 1988). There was a negative relationship between time interval and the size of the QBE ($\beta = -0.001$, 95% $CI = -0.002, -0.0003$, $p = .01$; $k = 91$ see on-line materials Table D2), consistent with the idea that the QBE fades over time. *Return rates* in survey tests of the QBE ranged from 13% (Ayres et al., 2013) to 100% (Falk, 2010) and were unrelated to the QBE ($\beta = -0.0003$, 95% $CI = -0.002, 0.001$, $p = .78$, $k = 21$; see on-line materials Table D2).

In relation to *risk of bias* (Table 3), the majority of studies were rated as unclear or high risk of bias and among those studies the effect size was small ($g = 0.20$, 95% $CI = 0.14, 0.25$, $p < .001$, $k = 67$). For studies rated as a low risk of bias, using a random effect analysis yielded a lower overall effect size ($g = 0.07$, 95% $CI = 0.04, 0.11$, $p < .001$, $k = 27$). Heterogeneity between these two subgroups was significant ($Q = 14.88$, $p < .001$). Lower heterogeneity was also found among the low risk studies ($Q = 52.05$, $p < .001$, $I^2 = 51.96\%$, $p < .001$). Subgroup analysis also showed that all six categories of bias individually significantly impacted on effect size. The largest effects were found in studies biased on sequence allocation ($g = 0.16$, 95% $CI = 0.08, 0.02$), incomplete outcome reporting ($g = 0.14$, 95% $CI = 0.001, 0.12$), selective outcome reporting ($g = 0.14$, 95% $CI = 0.07, 0.22$) and allocation concealment ($g = 0.12$, 95% $CI = 0.05, 0.20$). Bias was also calculated as a continuous score from 0–6 across the 6 Cochrane risk of bias categories (sequence allocation, allocation concealment, blinding, incomplete outcome reporting, selective outcome, other bias; coded as 0: low risk or bias or 1: unclear/high risk). The same pattern of findings was found when using this continuous score in a meta-regression ($\beta = 0.02$, 95% $CI = 0.007, 0.01$, $p < .001$; $k = 94$), suggesting that bias was a significant predictor of QBE effect size, i.e., higher risk of bias was associated with a greater effect size.

Due to the different effect sizes produced when health behaviours were separated into approach (desirable) and avoid (undesirable) behaviours, risk of bias was reanalysed separately on these two categories of health behaviour. In conducting this analysis, a small non-significant negative effect size was observed in studies of avoid health behaviours with low risk of bias ($g = -0.03$, 95% $CI = -0.10, 0.04$) and a small non-significant negative effect was observed in studies at unclear/high risk of bias ($g = -0.15$, 95% $CI = -0.31, 0.01$). A significant positive effect was found in approach health behaviours in studies that were at low risk of bias ($g = 0.10$, 95% $CI = 0.04, 0.16$) although this was smaller than that observed in studies with a high risk of bias ($g = 0.22$, 95% $CI = 0.08, 0.36$).

Future directions for QBE research

The present meta-analysis confirms the results of previous QBE meta-analyses (Rodrigues et al., 2015; Spangenberg et al., 2016; Wood et al., 2016) in demonstrating that questions have a small but significant effect on behaviour. This was the case despite the present meta-analysis incorporating a larger range of studies, with broader inclusion criteria, than those used in previous reviews of the QBE (Table 2). As a result of the inclusion of a greater number of studies of the QBE ($k = 65$) and consideration of unique moderators ($N = 14$), the present review produced a number of

novel findings beyond those of previous reviews. A key novel finding is the different effects produced when health behaviours were separated into those that should be approached (desirable) or avoided (undesirable) to promote a healthy lifestyle. The present review found a small, negative and marginally significant ($p = .051$) effect in studies of undesirable health behaviours, but a small, positive and significant effect in desirable health behaviours. The present review also provides further insights into the types of QBE intervention that appear to produce the greatest effect. Questioning about cognitions produced a greater effect than questioning about behaviour, and questioning face-to-face produced a greater effect than when questions were mailed to participants. Finally, key novel findings in the present review relate to the type of study design and the influence of bias over findings. The approach taken for coding and analysing risk of bias in the present review was more rigorous than that employed by the only previous meta-analysis to consider bias (Rodrigues et al., 2015). Whereas Rodrigues et al. adopted a simple approach by grouping the high risk studies together and the low risk studies together, comparing them and detecting no difference in effect size, we take a much more refined approach and use a larger number of studies and detect, for the first time in any review, significant effects. Specifically, we also considered overall bias score (i.e., as a continuous variable predicting effect size) and find that greater bias is associated with larger effect sizes. From the present review, it is clear that the QBE is considerably smaller in studies using designs less likely to be at risk of bias.

The meta-analysis also suggests a number of directions for future research on the QBE. As noted earlier, we would argue that such research is likely to be most informative where it provides further insights into the mechanisms underlying the QBE. Consistent with the meta-analysis, we discuss these future directions in relation to the sample studied, intervention employed, behavioural outcomes examined and methodology employed.

Sample studied

Wood et al. (2016) reported the QBE to be significantly larger in student compared to non-student samples. The QBE was also found to be significantly larger in the present meta-analysis compared to medical patients, schoolchildren, workers and others. If students complete questionnaires in a more honest and rational manner (Foot & Sanford, 2004), then such greater and more careful engagement with the questions might be expected to lead to stronger QBEs via each of the mechanisms discussed earlier. Wood et al. (2016) reported the QBE to be significantly stronger in laboratory compared to field settings. In the current meta-analysis, QBEs were significantly stronger in laboratory compared to medical, community and

online settings but not compared to educational settings. Research in laboratory settings might be expected to be associated with greater and more careful engagement with the questions, although it is also associated with other factors. Future research might attempt to systematically tease apart the key influences on the magnitude of the QBE.

Neither the present meta-analysis nor that of Wood et al. (2016) revealed any significant effect of the sample's experience with the behaviour on the magnitude of the QBE. In both cases the QBE was non-significantly larger in the no experience group. Spangenberg et al. (2016) reported that a continuous measure of experience was significantly related to the magnitude of the QBE, with stronger effects where the behaviour was more novel for the sample. Future research could explore more systematically the impact of asking questions in matched samples that vary only in their degree of experience with a behaviour. For example, our work on blood donation suggests that less experience with the behaviour can attenuate the QBE: QBE was effective in changing donation in experienced (Godin et al., 2008) but not novice (Godin et al., 2010) blood donors. These contradictory findings point to the need to consider interactions of behavioural experience with other moderators of the QBE. In a recent study, Conner et al. (in preparation) reported that questions about exercising at the sports centre significantly increased sports centre use among those who had previously used the sports centre, whereas questions about exercising elsewhere significantly increased sports centre use among those who had not previously used the sports centre. The authors suggested that this finding is consistent with an attitude accessibility account of the QBE. The differences in the QBE observed may be attributable to the fact that variations in experience actually reflect differences in the valence of attitudes towards the behaviours. In the Conner et al. (in preparation) study, participants with more experience of sports centre use had more positive attitudes towards sports centre use, while those with less experienced of sports centre use had more positive attitudes towards exercising generally.

Intervention employed

In the present meta-analysis, we observed that questions tapping cognitions were associated with a larger QBE than those tapping past behaviour. This is in contrast to Rodrigues et al. (2015), who reported no significant difference between studies measuring past behaviour, cognitions or past behaviour and cognitions, although their comparisons were based on a smaller number of studies. More consistent with other reviews (Spangenberg et al., 2016; Wood et al., 2016), the present meta-analysis did indicate that prediction questions were associated with the largest QBE. As noted earlier, this finding can be seen as consistent with each of the behavioural simulation and processing

fluency, attitude accessibility and cognitive dissonance accounts of the QBE. It does suggest that self-prediction of behaviour questions should be used in studies attempting to maximise a QBE, but avoided in studies seeking to minimise the impact of a QBE. Senay, Albarracin and Noguchi (2010) make the distinction between an interrogative (i.e., “Will I?”) and declarative (i.e., “I will”) form in which intention questions can be phrased. They observed that the former led to a better performance in anagram solving (Experiments 1 and 2) and an increase in the intention to exercise (Experiments 3 and 4) in cross-sectional studies. Godin, Bélanger-Gravel, Vézina-Im, Amireault and Bilodeau (2012) showed that interrogative (vs. declarative) intention questions significantly increased self-reported physical activity in a sample of students. Conner, Sandberg, Jackson, Godin, and Sheeran (*in preparation*) reported interrogative (compared to declarative) intention questions to increase objectively assessed attendance for cervical screening. In contrast, Godin, Germain, Conner, Delage, and Sheeran (2014) found both declarative and interrogative intention questions to significantly increase blood donation rates, compared to a no-question control condition at 15 months, with no significant difference between the two QBE conditions. Conner (*in preparation*) reported that interrogative self-prediction questions produced stronger impacts on self-reported physical activity compared to interrogative intentions, declarative intentions, declarative self-predictions or unrelated questions. Wood et al. (2014) speculated that the use of interrogative questions may result in stronger effects on attitude accessibility. Future studies could usefully further test the use of interrogative vs. declarative forms of questions, preferably using strong designs with objective measures of behaviour. Such studies should also explore the impact of variations in questioning on attitude accessibility and other mechanisms assumed to underlie the QBE.

A related issue in recent QBE studies is the impact of supplementing intention or self-prediction questions with other cognition questions. Little research has addressed the impact of supplemental questions about moral norms (Godin et al., 2014), positive self-image (Godin et al., 2014) and beneficence (Conner et al., *submitted*). More studies have examined supplemental questions such as attitudes, norms and perceived behavioural control from the Theory of Planned Behaviour (Ajzen, 1991) and anticipated regret. A number of studies have reported that asking Theory of Planned Behaviour questions is associated with a significant QBE (e.g., Conner et al., 2011; Godin et al., 2008). However, both the current meta-analysis (see on-line materials, Table D1) and that of Wood et al. (2016) indicated a larger QBE in studies not using Theory of Planned Behaviour items compared to those using such items, although the difference was not significant in either case and may be attributable to overlap with other moderators. Mankarious and Kothe (2015) explored the within-subjects effects of completing Theory of Planned Behaviour questions on changes in behaviour (i.e., no comparison to other

conditions or the usual randomisation to condition that typically occurs in most QBE studies). Across 66 studies, a small but non-significant negative effect on behaviour was observed. There was no effect in desirable behaviours ($d = 0.07$, 95% $CI = -0.009, 0.15$). However, significant *decreases* in behaviour were observed for socially undesirable behaviours such as binge drinking, risky driving and sugary snack consumption ($d = -0.28$, 95% $CI = -0.37, -0.18$). This suggests that QBE studies measuring Theory of Planned Behaviour components could be effective in reducing socially undesirable behaviours (see below for discussion of using the QBE for such behaviours), although in the present meta-analysis the QBE for undesirable behaviours was negative but only marginally significant, a pattern that was replicated when restricted to studies measuring Theory of Planned Behaviour components.

Testing the impact of adding anticipated regret questions to those tapping intentions (e.g., Godin et al., 2010, 2014) or those tapping components of the Theory of Planned Behaviour (Sandberg & Conner, 2009, 2011) has been the focus of a number of QBE studies. However, the reported effects have been mixed, with studies indicating that adding anticipated regret questions did (Sandberg & Conner, 2011) or did not (Godin et al., 2010; O'Carroll, Chambers, Brownlee, Libby, & Steele, 2015) increase the QBE. Wood et al. (2016) reported that QBE studies that included anticipated regret items had a significantly smaller effect size than studies that did not include such items. Sandberg and Conner (2011) showed that including anticipated regret questions to Theory of Planned Behaviour questions only increased the size of the QBE when they appeared before intention measures. Conner et al. (submitted) suggest that there may be complex effects of including anticipated regret in QBE studies linked to the nature of the behaviour and the sample. For example, it may be the case that among those with experience of the behaviour and a positive intention to perform the behaviour, anticipated regret plus (later) intention questions are an effective means to change behaviour by binding individuals to their intentions. In contrast, among those with less experience or more negative intentions about the behaviour, adding anticipated regret questions to intention questions may be counterproductive for behaviour change, and might even be associated with psychological reactance (Brehm, 1966). Future research might usefully explore the exact conditions under which adding anticipated regret questions increase or decrease the observed QBE.

Behavioural outcomes examined

The current meta-analysis indicated that there were significant differences in the magnitude of the QBE for different specific behaviours, with larger effects observed for behaviours such as flossing and health assessment and weaker

effects for condom use and alcohol consumption. Some of these differences may be attributable to the difficulty of the behaviour, with weaker QBEs being observed for more difficult behaviours (Wood et al., 2016; on-line materials Table D1). Future research might usefully further examine whether the QBE can be an effective means to change difficult behaviours. Our meta-analysis also indicated a weaker QBE for unhealthy (avoid, undesirable) compared with healthy (approach, desirable) behaviours. Whereas, the QBE was found to significantly increase desirable health behaviours, it was found to non-significantly ($p = .051$) reduce undesirable health behaviours. However, risk of bias appeared to have a greater influence over studies focusing on desirable health behaviours. Wood et al. (2016) similarly reported a significantly weaker QBE in undesirable behaviours (not significantly different from zero). The findings for the QBE on undesirable behaviours has been particularly mixed, with findings suggesting both decreases (Levav & Fitzsimons, 2006) and increases (Williams, Block, & Fitzsimons, 2006) in such behaviours. This has led to debate (Gollwitzer & Oettingen, 2008; Moore & Fitzsimons, 2008; Sherman, 2008) about the appropriateness of asking questions about undesirable behaviours, particularly in adolescent samples. It is suggested that asking about these behaviours may increase the likelihood that they are subsequently performed, particularly among adolescents with either positive or mixed attitudes towards such behaviours. In three laboratory studies, Wilding et al. (submitted) showed that questions tapping intentions to consume unhealthy snacks significantly increased objectively measured unhealthy snacking. However, in a survey study, Wilding, Conner, Lawton, Prestwich and Sheeran (in preparation) showed that Theory of Planned Behaviour questions about a range of both desirable and undesirable health behaviours resulted in significant increases in self-reports of the former but no effects on the latter 1 month later, compared with a condition with questions about unrelated behaviours. Consistent with early studies in the area (Sherman, 1980), reviews appear to indicate a small but significant effect of asking questions about approach or desirable behaviours. The QBE for undesirable behaviours appears to be much more mixed, with many studies observing no effect and a limited number of studies observing either an increase or decrease in such behaviours following questioning. Systematic exploration of the conditions under which questions prompt increases, decreases or no change in undesirable behaviours using studies with a low risk of bias would be particularly valuable in increasing our understanding.

Methodology employed in QBE studies

Although the meta-analysis of Spangenberg et al. (2016) reported that the QBE was significantly smaller for experimenter observed compared to self-reported behaviour, no significant differences were observed here or in

Wood et al. (2016). The present meta-analysis did observe a significantly stronger QBE in studies that did not take a baseline measure of behaviour and in studies that used face-to-face as opposed to mailed delivery of questions. Weaker QBEs were also observed in studies with longer time intervals, suggesting that the effect dissipates over time. In addition, several aspects of bias were related to the size of the QBE in the present meta-analysis. In particular, studies using stronger RCT designs and intention to treat analyses were associated with a significantly smaller QBE. Finally, those studies rated as being at low risk of bias reported a significantly smaller QBE compared with studies rated as having an unclear or high risk of bias.

Mankarious and Kothe (2015) have recently argued that the QBE could be produced purely as a result of the influence of demand effects. However, the fact that a number of well-designed studies have observed a QBE (e.g., Godin et al., 2008) even when demand effects unlikely provides strong evidence against this view. Mankarious and Kothe (2015) also argued that observed changes in behaviour may be due to self-selection bias in responding to the questionnaire. However, most QBE intervention studies do not make their aims clear to participants or inform participants that they are involved in an intervention to change their behaviour. Thus, it seems unlikely that self-selection bias is mainly responsible for QBE-based behaviour change. Future research that systematically explores the impacts of different sources of bias on the magnitude of the QBE (particularly for desirable vs. undesirable behaviours) would be useful in pinpointing the role of bias in this area.

Conclusions

This paper has provided a quantitative review of research on the QBE. Meta-analyses like the one presented here make it clear that asking questions about intentions and/or behaviour is associated with a small but significant effect on subsequent behaviour. This effect is smaller, but still significant, in studies with low risk of bias. The QBE extends across a number of different questions but is strongest for intention and self-prediction questions. The QBE is observable across a range of behaviours including health, consumer and voting behaviours. The majority of research has used the QBE to increase socially desirable behaviours, with only a modest focus on undesirable behaviours. Meta-analytic findings suggest that the QBE has either no significant effects on undesirable behaviours or reduces them, although individual studies (e.g., Wilding et al., [submitted](#)) show that questionnaires can also increase such behaviours. Further studies exploring the QBE for undesirable behaviours using low bias designs are warranted. A further area for research on the QBE is in relation to underlying mechanisms. As noted earlier, none of the

three main proposed theoretical explanations of the QBE (processing fluency, attitude accessibility and cognitive dissonance) has received strong support. Future research should pit different mechanisms against one another as rival explanations of findings, explore the factors that influence which mechanism may be operating, or test new mechanisms. In calling for further research, we also note the need for future studies to involve low risk of bias (e.g., well designed RCTs). Given the small overall effect size associated with the QBE, such studies will require large sample sizes to be appropriately powered. Nevertheless, the significant, albeit small-sized, changes in behaviour that are consistently observed to result from questioning suggest the need to gain a better understanding of the QBE in relation to efforts to both maximise and minimise its magnitude.

Acknowledgments

Preparation of this manuscript was supported in part by a grant from the UK Economic and Social Research Council (RES-062-23-2220) to Mark Conner, Cath Jackson, Paschal Sheeran and Gaston Godin (Understanding and using self-generated validity to promote behaviour change). Along with a UK Medical Research Council studentship awarded to Sarah Wilding supervised by Mark Conner, Rebecca Lawton and Andrew Prestwich (Using the question-behaviour effect to change health risk behaviours: Ref 1354363).

Funding

This work was supported by the Economic and Social Research Council: [Grant Number RES-062-23-2220].

ORCID

Sarah Wilding  <http://orcid.org/0000-0002-7977-7132>

References

- Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, 50(2), 179–211. doi:10.1016/0749-5978(91)90020-T
- Ajzen, I., & Fishbein, M. (1977). Attitude–behavior relations: A theoretical analysis and review of empirical research. *Psychological Bulletin*, 84, 888–918. doi:10.1037/0033-2909.84.5.888
- Armitage, C. J., & Conner, M. (2001). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*, 40, 471–499. doi:10.1348/014466601164939
- Aronson, E. (1992). The return of the repressed: Dissonance theory makes a comeback. *Psychological Inquiry*, 3, 303–311. doi:10.1207/s15327965pli0304_1

- Ayres, K., Conner, M. T., Prestwich, A., Hurling, R., Cobain, M., Lawton, R., & O'Connor, D. (2013). Exploring the question-behaviour effect: Randomized controlled trial of motivational and question-behaviour interventions. *British Journal of Health Psychology*, 18, 31–44. doi:10.1111/j.2044-8287.2012.02075.x
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182. doi:10.1037/0022-3514.51.6.1173
- Bendtsen, P., McCambridge, J., Bendtsen, M., Karlsson, N., & Nilsen, P. (2012). Effectiveness of a proactive mail-based alcohol Internet intervention for university students: Dismantling the assessment and feedback components in a randomized controlled trial. *Journal of Medical Internet Research*, 14(5), e142. doi:10.2196/jmir.2062
- Bernstein, E., Edwards, E., Dorfman, D., Heeren, T., Bliss, C., & Bernstein, J. (2009). Screening and brief intervention to reduce marijuana use among youth and young adults in a pediatric emergency department. *Academic Emergency Medicine*, 16(11), 1174–1185. doi:10.1111/j.1553-2712.2009.00490.x
- Borenstein, M., Hedges, L., Higgins, J., & Rothstein, H. (2005). *Comprehensive meta-analysis version 2*. Englewood, NJ: Biostat.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
- Brehm, J. W. (1966). *A theory of psychological reactance*. New York, NY: Academic Press.
- Chandon, P., Morwitz, V. G., & Reinartz, W. J. (2004). The short- and long-term effects of measuring intent to repurchase. *Journal of Consumer Research*, 31, 566–572. doi:10.1086/425091
- Chapman, K. J. (2001). Measuring intent: There's nothing ?mere? about mere measurement effects. *Psychology & Marketing*, 18, 811–841. doi:10.1002/mar.1031
- Chen, M., & Bargh, J. A. (1999). Consequences of automatic evaluation: Immediate behavioral predispositions to approach or avoid the stimulus. *Personality and Social Psychology Bulletin*, 25(2), 215–224. doi:10.1177/0146167299025002007
- Cialdini, R. B., Trost, M. R., & Newsom, J. T. (1995). Preference for consistency: The development of a valid measure and the discovery of surprising behavioral implications. *Journal of Personality and Social Psychology*, 69(2), 318–328. doi:10.1037/0022-3514.69.2.318
- Cioffi, D., & Garner, R. (1998). The effect of response options on decisions and subsequent behavior: Sometimes inaction is better. *Personality and Social Psychology Bulletin*, 24(5), 463–472. doi:10.1177/0146167298245002
- Conner, M. (in preparation). *Tapping the question-behaviour effect using interrogative versus declarative intention or self-prediction questions*.
- Conner, M., Godin, G., Norman, P., & Sheeran, P. (2011). Using the question-behavior effect to promote disease prevention behaviors: Two randomized controlled trials. *Health Psychology*, 30, 300–309. doi:10.1037/a0023036
- Conner, M., Sandberg, T., Jackson, C., Godin, G., & Sheeran, P. (in preparation). *An eight arm randomized controlled trial of the question-behavior effect for cervical screening*.
- Conner, M., Sandberg, T., Netkiting, C., Wood, C., Jackson, C., Godin, G., & Sheeran, P. (submitted). *Varying cognitive targets and response rates to enhance*

- the question-behavior effect: An eight arm randomized controlled trial on influenza vaccination.*
- Conner, M., Sandberg, T., Wood, C., Jackson, C., Godin, G., & Sheeran, P. (in preparation). *A randomized controlled trial of the question-behavior effect: Moderating effects of past behavior.*
- Conner, M., Warren, R., Close, S., & Sparks, P. (1999). Alcohol consumption and the theory of planned behavior: An examination of the cognitive mediation of past behavior. *Journal of Applied Social Psychology*, 29, 1676–1704. doi:10.1111/j.1559-1816.1999.tb02046.x
- Conner, M., Wilding, S., Sandberg, T., Jackson, C., Wood, C., Godin, G., & Sheeran, P. (in preparation). *Testing the mechanisms underlying the question-behavior effect: Two experiments on cancer screening.*
- Cooke, R., & Sheeran, P. (2004). Moderation of cognition-intention and cognition-behaviour relations: A meta-analysis of properties of variables from the theory of planned behaviour. *British Journal of Social Psychology*, 43(2), 159–186. doi:10.1348/0144666041501688
- Dholakia, U. M. (2010). A critical review of question-behavior effect research. *Review of Marketing Research*, 7, 145–197. doi:10.1108/S1548-6435(2010)000007009
- Dholakia, U. M., & Morwitz, V. G. (2002). The scope and persistence of mere measurement effects: Evidence from a field study of customer satisfaction measurement. *Journal of Consumer Research*, 29(2), 159–167. doi:10.1086/341568
- Falk, B. (2010). Do drivers become less risk-prone after answering a questionnaire on risky driving behaviour? *Accident Analysis & Prevention*, 42(1), 235–244. doi:10.1016/j.aap.2009.08.003
- Fazio, R. H., Chen, J.-M., McDonel, E. C., & Sherman, S. J. (1982). Attitude accessibility, attitude-behavior consistency, and the strength of the object-evaluation association. *Journal of Experimental Social Psychology*, 18(4), 339–357. doi:10.1016/0022-1031(82)90058-0
- Fazio, R. H., & Williams, C. J. (1986). Attitude accessibility as a moderator of the attitude “perception and attitude “behavior relations: An investigation of the 1984 presidential election. *Journal of Personality and Social Psychology*, 51(3), 505–514. doi:10.1037/0022-3514.51.3.505
- Fazio, R. H., & Zanna, M. P. (1981). Direct experience and attitude-behavior consistency. *Advances in Experimental Social Psychology*, 14, 161–202. doi:10.1177/014616727800400109
- Festinger, L. (1957). *A theory of cognitive dissonance*. London: Tavistock Publications.
- Festinger, L., & Carlsmith, J. M. (1959). Cognitive consequences of forced compliance. *The Journal of Abnormal and Social Psychology*, 58, 203–210. doi:10.1037/h0041593
- Fishbein, M., & Ajzen, I. (1975). *Belief, attitude, intention, and behavior: An introduction to theory and research*. Reading, MA: Addison-Wesley.
- Fitzsimons, G. J., & Morwitz, V. G. (1996). The effect of measuring intent on brand-level purchase behavior. *Journal of Consumer Research*, 23, 1–11. doi:10.1086/jcr.1996.23.issue-1
- 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. doi:10.1086/513043

- Fitzsimons, G. J., & Williams, P. (2000). Asking question can change choice behavior: Does it do so automatically or effortfully? *Journal of Experimental Psychology: Applied*, 6(3), 195–206. doi:10.1037//1076-898X.6.3.195
- Foot, H., & Sanford, A. (2004). The use and abuse of student participants. *The Psychologist*, 17(5), 256–259.
- Forster, A. S., Burgess, C., McDermott, L., Wright, A. J., Dodhia, H., Conner, M. . . . Gulliford, M. C. (2014). Enhanced invitation methods to increase uptake of NHS health checks: Study protocol for a randomized controlled trial. *Trials*, 15, 342. doi:10.1186/1745-6215-15-342
- Garner, R. (2005). Post-It® note persuasion: A sticky influence. *Journal of Consumer Psychology*, 15, 230–237. doi:10.1207/s15327663jcp1503_8
- Godin, G., Bélanger-Gravel, A., Vézina-Im, L.-A., Amireault, S., & Bilodeau, A. (2012). Question-behaviour effect: A randomised controlled trial of asking intention in the interrogative or declarative form. *Psychology & Health*, 27, 1086–1099. doi:10.1080/08870446.2012.671617
- Godin, G., Germain, M., Conner, M., Delage, G., & Sheeran, P. (2014). Promoting the return of lapsed blood donors: A seven-arm randomized controlled trial of the question-behavior effect. *Health Psychology*, 33(7), 646–655. doi:10.1037/a0033505
- Godin, G., Sheeran, P., Conner, M., Delage, G., Germain, M., Bélanger-Gravel, A., & Naccache, H. (2010). Which survey questions change behavior? Randomized controlled trial of mere measurement interventions. *Health Psychology*, 29, 636–644. doi:10.1037/a0021131
- Godin, G., Sheeran, P., Conner, M., & Germain, M. (2008). Asking questions changes behavior: Mere measurement effects on frequency of blood donation. *Health Psychology*, 27, 179–184. doi:10.1037/0278-6133.27.2.179
- Gollwitzer, P. M., & Oettingen, G. (2008). The question-behavior effect from an action control perspective. *Journal of Consumer Psychology*, 18(2), 107–110. doi:10.1016/j.jcps.2008.01.004
- Greenwald, A. G., Carnot, C. G., Beach, R., & Young, B. (1987). Increasing voting behavior by asking people if they expect to vote. *Journal of Applied Psychology*, 72(2), 315–318. doi:10.1037/0021-9010.72.2.315
- Harmon-Jones, E., Brehm, J. W., Greenberg, J., Simon, L., & Nelson, D. E. (1996). Evidence that the production of aversive consequences is not necessary to create cognitive dissonance. *Journal of Personality and Social Psychology*, 70(1), 5–16. doi:10.1037/0022-3514.70.1.5
- Janiszewski, C., & Chandon, E. (2007). Transfer-appropriate processing, response fluency, and the mere measurement effect. *Journal of Marketing Research*, 44(2), 309–323. doi:10.1509/jmkr.44.2.309
- Johnson, B. T., Low, R. E., & MacDonald, H. V. (2014). Panning for the gold in health research: Incorporating studies' methodological quality in meta-analysis. *Psychology & Health*, 30(1), 135–152. doi:10.1080/08870446.2014.953533
- Kvalem, I. L., Sundet, J. M., Rivø, K. I., Eilertsen, D. E., & Bakketeig, L. S. (1996). The effect of sex education on adolescents' use of condoms: Applying the Solomon four-group design. *Health Education & Behavior*, 23(1), 34–47. doi:10.1177/109019819602300103
- Kypri, K., & McAnally, H. M. (2005). Randomized controlled trial of a web-based primary care intervention for multiple health risk behaviors. *Preventive Medicine*, 41(3–4), 761–766. doi:10.1016/j.ypmed.2005.07.010

- Levav, J., & Fitzsimons, G. J. (2006). When questions change behavior: The role of ease of representation. *Psychological Science*, *17*, 207–213. doi:10.1111/j.1467-9280.2006.01687.x
- Locke, E. A. (1968). Toward a theory of task motivation and incentives. *Organizational Behavior and Human Performance*, *3*(2), 157–189. doi:10.1016/0030-5073(68)90004-4
- Locke, E. A., Shaw, K. N., Saari, L. M., & Latham, G. P. (1981). Goal setting and task performance: 1969–1980. *Psychological Bulletin*, *90*, 125–152. doi:10.1037/0033-2909.90.1.125
- Maassen, G. H., & Bakker, A. B. (2001). Suppressor variables in path models: Definitions and interpretations. *Sociological Methods & Research*, *30*(2), 241–270. doi:10.1177/0049124101030002004
- Mankarious, E., & Kothe, E. (2015). A meta-analysis of the effects of measuring theory of planned behaviour constructs on behaviour within prospective studies. *Health Psychology Review*, *9*, 190–204. doi:10.1080/17437199.2014.927722
- McEachan, R. R. C., Conner, M., Taylor, N. J., & Lawton, R. J. (2011). Prospective prediction of health-related behaviours with the theory of planned behaviour: A meta-analysis. *Health Psychology Review*, *5*, 97–144. doi:10.1080/17437199.2010.521684
- Moore, S. G., & Fitzsimons, G. J. (2008). While parents might not want to, researchers really should ask questions about risky behaviors. *Journal of Consumer Psychology*, *18*(2), 111–115. doi:10.1016/j.jcps.2008.01.005
- Morwitz, V. G., & Fitzsimons, G. J. (2004). The mere-measurement effect: Why does measuring intentions change actual behavior. *Journal of Consumer Psychology*, *14*, 64–74. doi:10.1207/s15327663jcp1401&2_8
- Morwitz, V. G., Johnson, E., & Schmittlein, D. (1993). Does measuring intent change behavior? *Journal of Consumer Research*, *20*, 46–61. doi:10.1086/209332
- Murray, M. (1988). The Hawthorne effect in the measurement of adolescent smoking. *Journal of Epidemiology and Community Health*, *42*(3), 304–306. doi:10.1136/jech.42.3.304
- Nickerson, D. W., & Rogers, T. (2010). Do you have a voting plan?: Implementation intentions, voter turnout, and organic plan making. *Psychological Science*, *21*(2), 194–199. doi:10.1177/0956797609359326
- O’Carroll, R. E., Chambers, J. A., Brownlee, L., Libby, G., & Steele, R. J. C. (2015). Anticipated regret to increase uptake of colorectal cancer screening (ARTICS): A randomised controlled trial. *Social Science & Medicine*, *142*, 118–127. doi:10.1016/j.socscimed.2015.07.026
- Paulhus, D. L., Robins, R. W., Trzesniewski, K. H., & Tracy, J. L. (2004). Two replicable suppressor situations in personality research. *Multivariate Behavioral Research*, *39*(2), 303–328. doi:10.1207/s15327906mbr3902_7
- Perkins, A., Smith, R. J., Sprott, D. E., Spangenberg, E. R., & Knuff, D. C. (2008). Understanding the self-prophecy phenomenon. *Advances in Consumer Research - European Conference Proceedings*, *8*, 462–467.
- Rodrigues, A. M., O’Brien, N., French, D. P., Glidewell, L., & Sniehotta, F. F. (2015). The question-behavior effect: Genuine effect or spurious phenomenon? A systematic review of randomized controlled trials with meta-analyses. *Health Psychology*, *34*, 61–78. doi:10.1037/hea0000104
- Sandberg, T., & Conner, M. (2009). A mere measurement effect for anticipated regret: Impacts on cervical screening attendance. *British Journal of Social Psychology*, *48*, 221–236. doi:10.1348/014466608X347001

- Sandberg, T., & Conner, M. (2011). Using self-generated validity to promote exercise behaviour. *British Journal of Social Psychology*, 50(4), 769–783. doi:10.1111/j.2044-8309.2010.02004.x
- Senay, I., Albarracin, D., & Noguchi, K. (2010). Motivating goal-directed behavior through introspective self-talk: The role of the interrogative form of simple future tense. *Psychological Science*, 21, 499–504. doi:10.1177/0956797610364751
- Sherman, S. J. (1980). On the self-erasing nature of errors of prediction. *Journal of Personality and Social Psychology*, 39, 211–221. doi:10.1037/0022-3514.39.2.211
- Sherman, S. J. (2008). Should we ask our children about sex, drugs, and rock & roll? A different conclusion. *Journal of Consumer Psychology*, 18(2), 96–101. doi:10.1016/j.jcps.2008.01.003
- Smit, E., Verdurmen, J., Monshouwer, K., & Smit, F. (2008). Family interventions and their effect on adolescent alcohol use in general populations; a meta-analysis of randomized controlled trials. *Drug and Alcohol Dependence*, 97(3), 195–206. doi:10.1016/j.drugalcdep.2008.03.032
- Song, H., & Schwarz, N. (2008). If it's hard to read, it's hard to do: Processing fluency affects effort prediction and motivation. *Psychological Science*, 19, 986–988. doi:10.1111/psci.2008.19.issue-10
- Spangenberg, E. R., & Greenwald, A. G. (1999). Social influence by requesting self-prophecy. *Journal of Consumer Psychology*, 8(1), 61–89. doi:10.1207/s15327663jcp0801_03
- Spangenberg, E. R., Karelkas, I., Devezer, B., & Sprott, D. (2016). A meta-analytic synthesis of the question-behavior effect. *Journal of Consumer Psychology*, 26(3), 441–458. doi:10.1016/j.jcps.2015.12.004
- Spangenberg, E. R., & Sprott, D. E. (2006). Self-monitoring and susceptibility to the influence of self-prophecy. *Journal of Consumer Research*, 32, 550–556. doi:10.1086/500485
- Spangenberg, E. R., Sprott, D. E., Grohmann, B., & Smith, R. J. (2003). Mass-communicated prediction requests: Practical application and a cognitive dissonance explanation for self-prophecy. *Journal of Marketing*, 67, 47–62. doi:10.1509/jmkg.67.3.47.18659
- Spangenberg, E. R., Sprott, D. E., Knuff, D. C., Smith, R. J., Obermiller, C., & Greenwald, A. G. (2012). Process evidence for the question-behavior effect: Influencing socially normative behaviors. *Social Influence*, 7(3), 211–228. doi:10.1080/15534510.2012.694024
- Spence, J. C., Burgess, J., Rodgers, W., & Murray, T. (2009). Effect of pretesting on intentions and behaviour: A pedometer and walking intervention. *Psychology & Health*, 24, 777–789. doi:10.1080/08870440801989938
- Sprott, D. E., Smith, R. J., Spangenberg, E. R., & Freson, T. S. (2004). Specificity of prediction requests: Evidence for the differential effects of self-prophecy on commitment to a health assessment. *Journal of Applied Social Psychology*, 34, 1176–1190. doi:10.1111/j.1559-1816.2004.tb02002.x
- Sprott, D. E., Spangenberg, E. R., Block, L. G., Fitzsimons, G. J., Morwitz, V. G., & Williams, P. (2006a). The question-behavior effect: What we know and where we go from here. *Social Influence*, 1, 128–137. doi:10.1080/15534510600685409
- Sprott, D. E., Spangenberg, E. R., & Fisher, R. (2003). The importance of normative beliefs to the self-prophecy effect. *Journal of Applied Psychology*, 88, 423–431. doi:10.1037/0021-9010.88.3.423

- Sprott, D. E., Spangenberg, E. R., Knuff, D. C., & Devezer, B. (2006b). Self-prediction and patient health: Influencing health-related behaviors through self-prophecy. *Medical Science Monitor*, 12, RA85–RA91.
- StataCorp. (2013). *Stata statistical software: Release 13*. College Station, TX: StataCorp LP.
- Steffgen, G., Recchia, S., & Viechtbauer, W. (2013). The link between school climate and violence in school: A meta-analytic review. *Aggression and Violent Behavior*, 18(2), 300–309. doi:10.1016/j.avb.2012.12.001
- Stone, J., & Cooper, J. (2001). A self-standards model of cognitive dissonance. *Journal of Experimental Social Psychology*, 37, 228–243. doi:10.1006/jesp.2000.1446
- Taylor, S., & Tweedie, R. (1998). *Trim and fill: A simple funnel plot based method of testing and adjusting for publication bias in meta-analyses*. Fort Collins, CO: Colorado State University.
- Træen, B. (2003). Effect of an intervention to prevent unwanted pregnancy in adolescents. A randomized, prospective study from Nordland County, Norway, 1999–2001. *Journal of Community & Applied Social Psychology*, 13(3), 207–223. doi:10.1002/casp.712
- van Kerckhove, A., Geuens, M., & Vermeir, I. (2012). A motivational account of the question-behavior effect. *Journal of Consumer Research*, 39, 111–127. doi:10.1086/661936
- van Valkengoed, I. G., Morre, S. A., Meijer, C. J., van den Brule, A. J., & Boeke, A. J. P. (2002). Do questions on sexual behaviour and the method of sample collection affect participation in a screening programme for asymptomatic Chlamydia trachomatis infections in primary care? *International Journal of STD & AIDS*, 13(1), 36–38. doi:10.1258/0956462021924596
- Wilding, S., Conner, M., Lawton, R., Prestwich, A., Perugini, M., Mattavelli, S., & Wood, C. (submitted). *Questioning behavioral intentions increases both healthy and unhealthy snacking in three studies*.
- Wilding, S., Conner, M., Lawton, R., Prestwich, A., & Sheeran, P. (in preparation). *Question-behavior effects in health protection and risk behaviors in the same sample*.
- 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. doi:10.1080/15534510600630850
- Williams, P., Fitzsimons, G. J., & Block, L. G. (2004). When consumers do not recognize “benign” intention questions as persuasion attempts. *Journal of Consumer Research*, 31(3), 540–550. doi:10.1086/425088
- Wood, C., Conner, M., Miles, E., Sandberg, T., Taylor, N., Godin, G., & Sheeran, P. (2016). The impact of asking intention or self-prediction questions on subsequent behavior: A meta-analysis. *Personality and Social Psychology Review*, 20, 245–268. doi:10.1177/1088868315592334
- Wood, C., Conner, M., Sandberg, T., Godin, G., & Sheeran, P. (2014). Why does asking questions change health behaviours? The mediating role of attitude accessibility. *Psychology & Health*, 29, 390–404. doi:10.1080/08870446.2013.858343