Exploring The Question Behaviour Effect and its Impact on Exercise Behaviour during the COVID -19 Pandemic, and Cognitive Dissonance as a Mediator

---

DANIEL LEEDS

INTRODUCTION **3**

Background 3

The Question Behaviour Effect Across Domains 5

Mechanisms underlying the QBE 6

Physical Activity During the Covid-19 Pandemic 10

The Present Study 12

METHOD **13**

Participants 13

Design 14

Materials 14

Procedure 17

RESULTS **19**

Effect of QBE condition on exercise scores 20

Effect of QBE Condition on Cognitive Dissonance Ratings 20

DISCUSSION **22**

Aims and Findings 22

Limitations and Future Directions 24

Conclusion 27

REFERENCES **27**

Exploring The Question Behaviour Effect and its Impact on Exercise Behaviour during the COVID -19 Pandemic, and Cognitive Dissonance as a Mediator

Introduction

Background

The Question Behaviour Effect (QBE) can be described as the phenomenon by which asking questions influences performance of a focal behaviour (Wood et al., 2016). Research into this phenomenon can be traced back to 1980 with a study by Sherman (1980) first documenting its impact on socially normative behaviour and labelling the effect as the self-erasing nature of errors of prediction. Sherman asked participants to predict their behaviour towards socially desirable and undesirable behaviours, such as volunteering for charity or singing the national anthem over the telephone. Results showed that participants who had given predictions overestimated their probability of behaving in a socially normative manner but were also more likely to perform the target behaviour in a socially normative way, compared to those who made no prediction. Thus, the errors of prediction were self-erased.

Since Sherman's original study research has developed through multiple avenues and several terms have developed to describe the phenomenon. Most prominently, the mere measurement effect (Morwitz et al., 1993), and self-prophesy effect (Spangenberg et al., 2003). The QBE label can be considered a common descriptor of such terms as the mere measurement effect (Morwitz et al., 1993), self-prophecy effect (Spangenberg et al., 2003), self-generated validity (Sandberg & Conner, 2011), and self-erasing errors of prediction (Sherman, 1980). The QBE label is the current preference in part because it encompasses all branches of research into the effect. Although research has taken diverse paths, commonality

exists in that the act of asking questions regarding a target behaviour has been documented as having a small but significant effect on subsequent performance of that behaviour (Wilding et al., 2016). These questions have most often asked participants to make self-predictions (under self-prophesy label) or indicate their intentions (under mere measurement label). Indeed, Meta-analyses have indicated that the type of question employed is a key factor in the size of effect seen, with prediction and intention questions yielding largest effects (Wilding et al., 2016; Wood et al., 2016).

The impact of the QBE has been documented across a wide range of behaviour, most notably the domains of prosocial, consumer, and health behaviour. However, sufficient evidence has not been found to enable consensus on the mechanisms underlying the QBE. To date, the most well supported theoretical explanations of the QBE have focused on accessibility of attitudes, and cognitive dissonance when exploring the potential factors underpinning the QBE. The attitude accessibility explanation of the QBE suggests that asking questions increases the salience of attitudes towards the target behaviour, heightening the chance that participants will act in line with their pre-existing attitude. The cognitive dissonance explanation of the QBE suggests that asking questions highlights inconsistencies between participants behaviour and beliefs, this in turn creates dissonance which motivates behaviour in line with beliefs. Importantly, as branches of QBE research remained distinct from one another for so long (Sprott et al., 2006) the various aspects of QBE studies have often coalesced around the label of research in a particular way e.g. mere measurements studies using intention questions, for consumer behaviour, with attitude accessibility as the preferred proposed mechanism, and self-prophesy studies using prediction questions for prosocial behaviour with cognitive dissonance as the preferred proposed mechanism. Considering this shortfall in crossover, it is then not unexpected that Meta-analyses have called attention to the fact that very little research has directly tested the proposed

explanations of the QBE against one another, or across domains (Wilding et al., 2016; Wood et al., 2016). Moreover, meta-analyses have been unable to find support for either attitude accessibility and cognitive dissonance explanations of the QBE and have highlighted the need for direct measures to be employed to better elucidate the factors underpinning the QBE (Wilding et al., 2016; Wood et al., 2016). Nevertheless, the QBE has been demonstrated to have great potential in promoting change across a number of important areas of behaviour.

The Question Behaviour Effect Across Domains

Studies investigating the efficacy of harnessing the QBE have predominantly fallen into one of three domains, namely prosocial, consumer, and health behaviour. Substantial heterogeneity exists between the behaviours targeted. Therefore, it is of particular interest that evidence of the QBE can be found across these diverse areas of behaviour.

Prosocial behaviour refers to wide range of actions that can be characterised as a voluntary behaviour providing a benefit for others or society as a whole, such as cooperating, helping, sharing, and complying with requests (Xiao et al., 2019; Penner et al., 2005). Prosocial behaviour has great importance for social welfare and responsibility, and for these reasons is widely studied across disciplines (Thielmann et al., 2020). Within this domain the QBE has shown efficacy as a behavioural change strategy for a variety of behaviours, such as voter participation (Greenwald et al., 1987), blood donation (Godin, et al., 2008), and academic honesty (Spangenberg & Obermiller, 1996). Research within this domain has demonstrated that questioning can both increase performance of socially desirable behaviour (Cioffi & Garner, 1998; Godin et al., 2008), and also reduce performance of socially undesirable behaviour (Spangenberg & Obermiller, 1996). However, it should be noted that replications have sometimes found contrasting results (Smith et al., 2003).

Within the domain of consumer behaviour, research has demonstrated equally promising results (Chandon et al., 2004; Fitzsimons & Morwitz, 1996; Janiszewski & Chandon, 2007). This is particularly significant when taking in to account the high variability often found in consumer behaviour towards brands (Fan et al., 2012). In an example of this, Chandon et al. (2004) showed that consumers who were asked intention-based questions were more likely to make repeat purchases, took less time to make those purchases, and were more profitable customers than those not asked.

Within the domain of health behaviour, change promoting interventions are identified as key to reducing the premature morbidity and mortality associated with the noncommunicative diseases that account for a large percentage of deaths globally (Boyle et al., 2003; Johnson et al., 2010; Van Camp, 2014; World Health Organization [WHO], 2020) . Within this domain a large number of studies have evidenced the potential of QBE interventions for a range of health promoting behaviours, including health screening (Sandberg & Conner, 2009), vaccination uptake (Conner et al., 2011), and exercise (Wilding et al., 2019). QBE interventions demonstrating efficacy in the health behaviour domain may be considered of particular relevance when taking into account the resources required to implement such programs (Wilding et al., 2016; Wilding et al., 2019; Wood et al., 2016) and the impact complexities around delivery quality currently have on intervention outcomes (Bellg et al., 2004; Hatfield et al., 2020; Lambert et al., 2017).

Mechanisms underlying the QBE

Within the literature to date several mechanisms have been suggested to underpin the QBE, the two most empirically supported and dominant explanations are attitude accessibly, and cognitive dissonance (Spangenberg et al., 2012).

Attitude Accessibility

The attitude accessibility account of the QBE proposes that questions which ask participants to predict or indicate their intentions towards enacting a target behaviour activates their existing attitude towards that behaviour and in turn makes it more accessible in memory. This increased accessibility of the pre-existing attitude then increases or decreases the likelihood of participants performance of the target behaviour, in line with the associated attitude (Dholakia, 2009; Morwitz & Fitzsimons, 2004; Morwitz et al., 1993). The link between increased attitude accessibility and changes in behaviour associated with those attitudes is well established in literature (Chen & Bargh, 1999; Fazio et al., 1984; Fazio & Williams, 1986). The influence of questioning on attitude accessibility is also well supported, particularly by studies within the mere-measurement branch of QBE research (Dholakia, 2009). Several studies have shown that participants who are asked to predict their behaviour or provide an indication of their intentions towards a behaviour display more accessible attitudes compared to participants not asked (Chapman, 2001; Morwitz & Fitzsimons, 2004; Wood et al., 2014). Studies demonstrating that asking questions can increase attitude accessibility have also provided evidence that attitude accessibility acts as a mediator between intention measurement and behaviour (Wood et al., 2014). Within the attitude accessibility account, it is suggested that the effect witnessed depends on the direction and power of attitudes towards the target behaviour. Evidence of this can be seen in Fitzsimons and Morwitz (2004) study on consumer brand choices, in which it is demonstrated that intention measurement increased accessibility of pre-existing attitudes, and in turn increased purchasing incidence when consumers most salient attitudes towards a brand were positive, and decreased purchasing incidence when consumers most salient attitudes of a brand were negative. Additionally, research has also indicated that attitude accessibility acts as a moderator of effect, in such a way that stronger QBE is seen when participants hold positive attitudes towards a behaviour rather than negative (Ayres et al., 2013; Conner et al., 2011).

However, while evidence in support of the attitude accessibility account of the QBE has been provided in the literature, a number of studies have found conflicting evidence to suggest that attitude accessibility is not affected by questioning (Perkins et al., 2006; Spangenberg et al., 2012). Further to this, it is well established that memory accessibility reduces over time (Feldman & Lynch, 1988) and so it is difficult to attribute credit to attitude accessibility for effects seen in studies over longer time periods where increased any accessibility caused by questioning would likely have diminished (Godin et al., 2008).

Cognitive Dissonance

Cognitive dissonance has been described as a psychological state in which a person's cognitions (beliefs, attitudes, and behaviours) are conflicting (Festinger, 1957; **Lawrence & Festinger, 1962**). The experience of cognitive dissonance is aversive (Elliot & Devine, 1994; Stone & Cooper, 2001) and provides motivation for efforts to reduce dissonance (Aronson et al., 1991; Stone, & Fernandez (2008). A state of cognitive dissonance can be aroused nomothetically through misalignment with normative standards, or ideographically through misalignment with personal standards (Dholakia, 2009; Stone & Cooper, 2001). The cognitive dissonance account of the QBE proposes that questions which ask participants to predict or indicate their intentions towards a particular behaviour serve to induce cognitions of socially normative behaviour while concurrently reminding participants of their own failings in relation to those norms. When this simultaneous awareness highlights inconsistencies between a person's behaviour and beliefs, cognitive dissonance is aroused. In turn, the aversive experience of cognitive dissonance provides motivation to enact behaviour adherent to the relevant standard of judgment in order to reduce dissonance (Aronson, 1992). Thus, creating a QBE. In support of this, research has shown that participants predictions of future behaviour (Spangenberg & Greenwald, 1999) and also recall of past behaviour (Spangenberg et al, 2012) are biased in a normative direction, indicating their understanding

of the norm associated with the target behaviour. Multiple dissonance reduction strategies are well established in literature (McGrath, 2017). Indeed, support for the cognitive dissonance explanation of the QBE has been inferred from the rate at which participants engage in dissonance reduction strategies. For example, a study by Spangenberg et al. (2003) demonstrated that participants asked to make self-predictions subsequently made downward comparisons at higher rates than those who had not made any prediction. While a later study provided evidence demonstrating that participants asked to make predictions exhibit a downward reporting bias of others behaviour (Spangenberg et al., 2012). However, support for the cognitive dissonance explanation of the QBE is far from universal (Wood et al, 2016). Additionally, despite direct measures of dissonance being established in research for some time (Croyle & Cooper, 1983; Elkin & Leippe, 1986; Elliot & Devine, 1994; Martinie et al., 2013; de Vries et al., 2015) it is still considered difficult to measure (Wilding et al., 2019). Further, Meta-analyses show that QBE studies have largely not employed direct measures (implicit or explicit) of cognitive dissonance (Wood et al., 2016), and have overwhelmingly focused on behaviour as the focal outcome, with the role of dissonance inferred from moderator effects (Wilding et al., 2019; Wood et al., 2016). Nevertheless, recent studies have indicated towards cognitive dissonance as a key mechanism underlying the QBE (Wilding et al., 2019).

The degree to which a person experiences cognitive dissonance has been shown to vary (Lavergne and Pelletier, 2016; Leippe & Eisenstadt, 1999; McGrath, 2017), and dissonance reduction is proportional to the dissonance being experienced (Festinger, 1957). Wilding et al. (2019) investigated this granular element of dissonance in relation to behavioural outcomes. Based on the concept that if the cognitive dissonance explanation of QBE is correct, and also that dissonance is granular, then manipulating questions to magnify dissonance should yield a larger QBE. Findings of Wilding et al. (2019) study showed that

the level of dissonance generated by questioning can be magnified by enhancing cognitions of normative standards in relation to personal behaviour. Moreover, that higher levels of dissonance at the time of questioning translate into a larger QBE. Importantly, the design of this study enabled a more direct assessment of cognitive dissonance as a mechanism and highlighted the need for further exploration of how questions can be used to optimise outcomes.

In short, from a dissonance perspective there are three requirements which need to be met in order to elicit a dissonance driven QBE. first, the individual needs to hold normative beliefs in regard to the target behaviour (I should behave this way). Secondly, the individual needs to possess a self-concept which directs the belief that they are type of person who would behave in line with the normative belief. Thirdly, the individual must have not behaved in line with those standards (Aronson & Carlsmith, 1962; Sprott et al, 2003). If all three of these conditions are met, then a self-prediction task directly challenges the individual's idea of self (self-concept) and motivates them to adapt their behaviour (Aronson, 1992; Dholakia, 2009; Stone et al., 1997). Deploying this theory to engineer questions which highlight discrepancies between normative standards and behaviour can enhance the level of dissonance experienced, and has proven to be beneficial in achieving larger effects on behavioural outcomes (Wilding et al., 2019).

Physical Activity During the Covid-19 Pandemic

Physical activity can be defined as any bodily movement which uses skeletal muscle and expends energy, often measured in Metabolic Equivalent of Task (MET) values (Caspersen et al., 1985). National Health Service [NHS] (2019) guidelines state that adults should perform at least 150 minutes of moderate physical activity, and 75 minutes of vigorous physical activity each week, while reducing sedentary time. Physical inactivity is a primary cause of chronic disease (Booth et al., 2012; Boyle et al., 2003; Johnson et al., 2010;

Van Camp, 2014; WHO, 2021). Moreover, a curvilinear relationship exists between physical activity and health outcomes, meaning that substantial health benefits are observed from relatively small amounts of physical activity (Van Camp, 2014; Warburton, & Bredin, 2017).

Significant life events are known to be a key determinant in changes to the amount of physical activity people undertake, with changes in employment status, residence, physical status, relationships, and family structure all listed as having an impact on physical activity levels (Allender & Hutchinson, 2008). In January 2020 an outbreak of Novel strain of Severe Acute Respiratory Syndrome Coronavirus 2 (SARS-CoV2, Covid-19) was identified in Wuhan, China. By March 2020 the WHO had declared a global pandemic (Tarique et al., 2021), on the 23rd of March 2020 the UK government announced social distancing measures in response to the COVID-19 pandemic (Prime Minister's Office, 2020). On the 26th of March these "lockdown" measures legally came into force. Lockdown restrictions permitted UK residents to leave home for essential shopping, medical needs, travelling to and from work where absolutely necessary, or for one form of exercise per day. Under all other circumstances people were required to stay at home and reduce interpersonal contact with anyone outside of their home. These restrictions significantly altered the way in which people were able to approach most aspects of daily life, including physical activity (Mutz, 2021). Studies of physical activity during the COVID-19 pandemic from several countries have reported that inactivity increased among certain groups, particularly the young adult group (Guo et al., 2021; Stanton et al., 2020; Visser et al., 2020). Importantly, COVID-19 related stress was shown to significantly affect energy levels and in turn reduce physical activity output (Giessing et al., 2021). In line with research on the effects of physical activity on mental health and wellbeing (Chekroud et al., 2018), people who report negative changes in physical activity experience higher rates of depression, anxiety, and stress (O'Connor et al.,

2020; Stanton et al., 2020). Concerningly, suicidal ideation among young adults was shown to increasing during the UK's first lockdown (O'Connor et al., 2020).

Under these conditions, it is then of particular relevance that physical activity has been demonstrated to have beneficial effects for stress related physical and mental health conditions (Anderson & Shivakumar, 2013; Saxena et al., 2005), with research providing evidence that sufficient physical activity can buffer the effects of stress on health (Klaperski & Fuchs, 2021). Moreover, recent research has found that physical activity significantly affects the relationship between worry and stress during the COVID-19 pandemic (Green et al., 2021). In addition to the benefits of physical activity on mental health, sufficient physical activity is also shown to bolster immune health (Nieman et al., 2019), and reduce the risk of respiratory distress syndrome, listed as a major cause of death from COVID-19 (Lake, 2020). In spite of this, a significant percentage of the UK population do not meet the minimum standards of physical activity (NHS, 2017), and interventions targeting physical activity have been widely recommended for some time (European Union, 2008; WHO, 2004; WHO, 2013).

The Present Study

The aim of the present study was firstly to explore whether the QBE could be harnessed to increase exercise behaviour during the COVID-19 pandemic. The secondary aim of the present study was to explore the potential mediators of any QBE found, specifically whether cognitive dissonance mediated any increases in exercise activity prompted by the QBE. Based on the above findings and conceptualisations, it was firstly hypothesised that participants in a QBE prediction question condition would report significantly higher rates of exercise activity, compared to participants in the control condition. Further, it was hypothesised that participants in an enhanced dissonance QBE condition would report significantly higher rates of exercise activity, compared to participants in both the control and

QBE prediction only conditions. Secondly, in keeping with the secondary aim of the study to test for cognitive dissonance as a mediator of the QBE, it was hypothesised that participants in the QBE prediction condition would report greater levels of dissonance, compared with participants in the control condition. Further, it was hypothesised that participants in the enhanced dissonance QBE condition would report greater levels of dissonance than participants in both the standard QBE and control conditions.

Method

Participants

Participants were either psychology undergraduate students recruited via the Online Research Participation System (SONA), or were recruited from the general adult population by placing a recruitment advertisement on social media platforms, and online research participation forums. Two SONA system credits were used as incentive for participants to take part in the study. Only participants without pre-existing conditions which could be affected by exercise or spending time thinking about exercise were recruited for the study. Ethical approval was obtained from the University of Sheffield's Psychology Department Ethics Board. 152 participant responses were recorded for the first part in the study. However, a significant dropout occurred and only 76 responses were recorded for the follow up survey. From these 76 responses data were removed for participants who did not provide matching Identifying codes for both parts of the study (19), did not provide full informed consent (3), or who entered more than one submission in the first survey (5). Where participants had entered more than one submission to the second survey, only data from the first submission were retained and all others removed (12). Further to this, some participants data were also excluded as part of a 2 standard deviation trim to remove extreme outliers (4). The final sample consisted of 33 participants (5 male, 26 female, 1 prefer not to say, 1 non-conforming, mean age = 23.61, SD = 10.20, age range = 18 - 56). Participants were randomly

allocated to one of the three condition groups, with 13 in the control group, 11 in the QBE prediction group, and 9 in the QBE prediction + past group.

Design

The present study employed an experimental between participants' design. The independent variable for this study had three categorical levels, consisting of two experimental conditions and one control condition. Participants were randomly allocated to one of the three conditions and completed the study over two separate time points, spaced one week apart. The conditions were made up of a QBE prediction condition, in which participants were asked to predict their future exercise behaviour. A QBE prediction + past condition, in which participants were asked to predict their future exercise behaviour and also asked a second question regarding their past exercise behaviour. A control condition in which participants were not asked any questions regarding their exercise behaviour. All participants took part in their allocated condition at the first time point. The study had one continuous dependent variable; this was participants exercise behaviour in the seven days between the first and second time points of the study, this was measured using an adapted version of the Godin Shephard Leisure Time Physical Activity Questionnaire (GSLTPAQ) (Godin, 2011) at the second time point. Additionally, the study included two measures of cognitive dissonance, these were The Implicit Positive and Negative Affect Test (Quirin et al., 2009) (IPANAT) and the Dissonance thermometer (Elliot & Devine, 1994). Measures of dissonance were completed by all participants at the first time point, immediately after taking part in their allocated condition.

Materials

All of the materials used within the study were presented to participants via the online survey platform Qualtrics (https://www.qualtrics.com). Participants accessed the material through links provided within recruitment advertisements, SONA, and follow up emails.

Godin-Shephard Leisure-Time Physical Activity Questionnaire

The GSLTPAQ was developed by Godin and Shephard (1985) in order to classify physical activities and gain a measurement of adults' overall activity levels. Activities are classified based on their MET value and fall into one of three intensity subgroups, mild, moderate, and strenuous. The frequency and type (mild, moderate, strenuous) of activity performed each week by participants is used to calculate an overall weekly activity score by multiplying the number of times participants performed each type of exercise by its corresponding MET value (9 x frequency of strenuous, 5 x frequency of moderate, 3 x frequency of mild) and summing scores together. For the present study a modified version of the GSLTPAQ was used to measure participants exercise behaviour in the 7 days following the first time point of the study. The questionnaire was modified to remove exercises which would be inaccessible during the COVID-19 pandemic lockdown, and to add exercises which would be more accessible under lockdown restrictions. Exercises added were taken from the WHO (2021) and allocated an intensity subgroup based on their MET value. In the strenuous subgroup, squash, basketball, American football, cross country skiing, judo, skiing, and roller skating were removed, and home workouts and carrying or moving weight of >20kg were added. In the moderate subgroup, baseball, volleyball, badminton, and alpine skiing were removed, and home workouts and carrying or moving weight of >20kg were added. In the mild subgroup, archery, fishing from riverbank, bowling, horseshoes, golf, and snowmobiling were removed. The questionnaire was self-administered, with participants asked to respond to the following question for each subgroup of exercise activity; "During the last 7 days how many times on average did you do the following kinds of exercise for more than 15 minutes during your free time?".

Measures of Cognitive Dissonance

The Dissonance Thermometer. The Dissonance Thermometer developed by Elliot and Devine (1994) is a well-established means of assessing levels of cognitive dissonance (Bran & Vaidis, 2020) which calculates scores based on how much participants report feeling uncomfortable, uneasy, and bothered. The questionnaire for the current study contained 24 items representing dissonance relevant terms (e.g. bothered, uneasy, uncomfortable). As adapted from Lecrique (2007) asked participants to indicate how they were feeling "right now" on a 7-point Likert scale, where 1 = does not apply at all and 7 = applies very much. Overall scores were obtained by averaging participants' answers to the 'uncomfortable, uneasy and bothered' items, with higher scores indicating greater levels of cognitive dissonance.

The Implicit Positive and Negative Affect Test. The Implicit Positive and Negative Affect Test **(**Quirin et al., 2009) (IPANAT) implicitly measures affect through the process of affect misattribution (Payne et al., 2005; Quirin & Bode, 2014). It is intended that participants will indicate their affective state without being aware of the construct being measured. In order to achieve this, participants were shown an instruction intended to divert attention from the aim of affect measurement and direct attention towards the features (sounds) of words instead. Participants were asked to rate the degree to which each of six ambiguous non-words from an artificial language (SAFME, TALEP, BELNI, SUKOV, VIKES, TUNBA) expressed 6 different emotional adjectives (happy, helpless, energetic, tense, cheerful, inhibited) on a 1 (doesn't fit at all) to 4 (fits very well) scale. The six non-words were randomly presented to participants along with the mood adjectives to avoid order effects. Scores for positive and negative affect were computed in a two-step process. First, the average rating for each mood adjective is calculated for all six non-word judgements. Secondly, scores for positive and negative affect are calculated by averaging ratings for all positively valanced adjectives together (happy, energetic,

cheerful) and all negatively valanced adjectives together (helpless, tense, inhibited). Higher fit ratings for negatively valanced adjectives (and lower for positive) were scored as indicative of greater negative affect (and lower positive affect) and taken as a reflection of greater cognitive dissonance.

Procedure

The study was conducted via the online survey platform Qualtrics (https://www.Qualtrics.com). Participants completed the study across two time points, set 7 days apart. Participants were informed that the study was about exercise habits during the Covid-19 pandemic, and the time required to complete both parts of the study. For the first questionnaire, participants accessed the material through links provided within recruitment advertisements or SONA. For the second questionnaire, participants were provided with access via links in follow up emails. Participants were able to complete the constituent parts of each time point of the study at their own pace.

Time 1 Questionnaire.

At the onset of the first questionnaire participants were shown an information sheet for the study and asked to electronically indicate that they had read and fully understood all the information contained within the sheet, and that they met the eligibility criteria to take part. Following this, Participants were asked to generate a unique code to allow the matching of their data across both time points, and then asked to provide answers to three questions regarding their key demographics (age, ethnicity, gender). Participants were then randomly allocated to one of the three groups. Participants allocated to the QBE prediction group were informed that they would be presented with one question regarding their exercise habits in a yes/no format. Participants were also told that following this question they would be presented with words that describe types of feelings, and for each word shown they should indicate how well the word described how they were feeling at the time. These participants

(QBE prediction group) were then asked one question predicting their exercise behaviour. Specifically, they were asked; “Do you predict that you will exercise in the next week?: Yes/No”. Participants allocated to the QBE prediction + past group were informed that they would be presented with two questions regarding their exercise habits in a yes/no format. These participants were also told that following this question they would be presented with words that describe types of feelings and that for each word shown they should indicate how well the word described how they were feeling at the time. These participants (QBE prediction + past group) were then asked one question predicting their future behaviour and another question describing their past behaviour. Specifically, they were asked; “Do you predict that you will exercise in the next week? Yes/No”, and then “Have you exercised in the past week? Yes/No”. Participants allocated to the control group were directed to proceed to the next stage without being asked any questions. Upon completion of their allocated condition, all participants proceeded to the Dissonance Thermometer and were presented with instruction on how to complete the measure. Specifically, participants were told “Below are words that can describe different types of feelings. For each word, please indicate how much it describes how you are feeling right now by selecting a number on the scale. "1" means "does not apply at all" and "7" means "applies very much" to how you are feeling right now. Don't spend much time thinking about each word, just give a quick, gut-level response”. This measure was presented over two pages, each containing a 12 item 7-point Likert scale. Upon completion of the Dissonance Thermometer all participants proceeded to the IPANAT and were presented with instruction on how to complete this measure. Specifically, participants were told “You will be presented with words from an artificial language. They are intended to express various moods. In all languages there are words that help to express their meaning by the way they sound (for example, the word rattle almost sounds like something that rattles). In poetry and literature, this is known as onomatopoeia. Please rate how well each

artificial word expresses different moods, using a scale ranging from 1 (doesn't fit at all) to 4 (fits very well). For example: To what extent does the word FILNU convey each of the following moods: Happy, Helpless, Energetic, Tense, Cheerful, Inhibited. In making these ratings, let yourself be guided by your spontaneous feelings". All participants were then redirected to a separate page where they were asked to enter their email address and were presented with a debrief form which reminded them of the timing of the time 2 questionnaire.

Time 2 questionnaire.

At the start of the second questionnaire participants were asked to input the unique identifying code they had generated previously. Following this, participants were moved on to the modified GSLTPAQ. This measure was presented to participants as three separate questions, with each asking them to provide the number of times they had exercised over the previous seven days for an intensity subgroup. Participants were provided with a list of exercises for each particular intensity subgroup at the time of the corresponding question. Specifically, participants were presented with the question "during the last 7 days how many times on average did you do the following kinds of exercise for more than 15 minutes during your free time?". Upon completion all participants were redirected to a separate page which collected their email address in order to provide them with their allocated SONA credits, asked if they had felt suspicious of the study aims, and presented them with a debrief form.

Results

The final analysis was performed on a sample comprised of 33 participants (5 male, 26 female, 1 prefer not to say, 1 non-conforming, mean age = 23.61, SD = 10.20, age range = 18 - 56).

Effect of QBE condition on exercise scores

Normality checks indicated that the data for exercise scores was not normally distributed and would not meet the assumptions for parametric tests. Examination of histograms showed a positive skew for both the QBE prediction (skewness = 1.25) and also QBE prediction + past (skewness = 1.58) groups, while examination of Q-Q plots indicated that the data was not normally distributed for any of the groups. Additionally, box and whisker plots showed outliers for both QBE prediction (1) and QBE prediction + past (2) groups. Therefore, non-parametric tests were employed for the analysis of this data. As group sizes were not equal and histograms indicated variance in the spread of the data across groups (see figure 1) the assumptions for a Kruskal-Wallis test comparing medians were not met. Therefore, in order to examine the impact of QBE on exercise behaviour a Kruskal-Wallis test comparing mean ranks was performed on LSA scores across conditions. Results of the Kruskal-Wallis test showed no significant effect of QBE condition on the amount exercise performed by participants, $H(2) = 3.17$, $P = .205$, with a mean rank LSA score of 17.0 for control group, 13.6 for prediction group, and 21.3 for prediction + past group.

Effect of QBE Condition on Cognitive Dissonance Ratings

The Dissonance Thermometer

Normality checks indicated that the data for Dissonance Thermometer scores was not normally distributed and would not meet the assumptions for parametric tests. Examination of histograms showed bimodal distribution for both control and prediction groups, while examination of Q-Q plots indicated that the data was not normally distributed for any of the groups. Therefore, non-parametric tests were employed for the analysis of this data. As group sizes were not equal and histograms indicated variance in the spread of the data across groups the assumptions for a Kruskal-Wallis test comparing medians were not met. Therefore, in order to examine the impact of QBE condition on cognitive dissonance a Kruskal-Wallis test

comparing mean ranks was performed on Dissonance thermometer scores across conditions. Results of the Kruskal-Wallis test showed no significant effect of QBE condition on participants Dissonance Thermometer ratings, $H(2) = 4.08$, $P = .130$, with a mean rank score of 14.8 for control group, 21.8 for prediction group, and 14.3 for prediction + past group. These results suggest that QBE condition did not significantly affect participants level of cognitive dissonance.

IPANAT

Normality checks indicated that the data for IPANAT implicit positive affect (IPA) and implicit negative affect (INA) were not normally distributed and would not meet the assumptions for parametric tests. Examination of histograms showed bimodal distribution for both control and prediction groups in both IPA and INA data, while examination of Q-Q plots indicated that the data was not normally distributed for any of the groups across both IPA and INA. Additionally, examination of box and whisker plots for INA showed outliers (2) in the data for prediction + past group. Therefore, non-parametric tests were employed. As group sizes were not equal and histograms indicated variance in the spread of the data across groups the assumptions for a Kruskal-Wallis test comparing medians were not met. Therefore, in order to examine the impact of QBE on cognitive dissonance Kruskal-Wallis tests comparing mean ranks were performed on INA and IPA scores across conditions. Results of the Kruskal-Wallis test on IPA data showed no significant effect of QBE condition on IPA ratings, $H(2) = 3.91$, $P = .822$, with a mean rank score of 15.9 for control group, 17.2 for prediction group, and 18.4 for prediction + past group. Similarly, results of the Kruskal-Wallis test on INA data showed no significant effect of QBE condition on participants INA ratings, $H(2) = 0.11$, $P = .948$, with a mean rank score of 17.4 for control group, 16.2 for prediction group, and 17.3 for prediction + past group. These results suggest that QBE condition did not significantly affect participants level of cognitive dissonance.

Discussion

Aims and Findings

The aim of the present study was to explore whether the QBE could be harnessed to increase exercise behaviour during the COVID-19 pandemic, and to investigate whether cognitive dissonance mediated any increases in exercise activity prompted by the QBE. In line with the first aim of the study, it was examined whether asking participants to make self-predictions regarding their future exercise behaviour would cause a significant increase in subsequent exercise behaviour. Further, it was examined whether asking participants to make self-predictions regarding their future exercise behaviour and to also describe their previous exercise behaviour would generate a larger increase in exercise behaviour than a prediction question alone. In keeping with the secondary aim of the study to test for cognitive dissonance as a mediator of the QBE, it was examined to what extent, if any, participants who were asked these questions experienced significantly greater levels of cognitive dissonance than those not asked.

The present study found no significant difference between groups in exercise behaviour. The lack of difference between groups indicates that asking participants to provide self-predictions, or self-predictions plus descriptions of past behaviour, did not affect the amount of exercise they subsequently performed. Therefore, these prediction questions did not initiate a QBE. Additionally, this study found no significant difference between groups in rated cognitive dissonance. This finding indicates that in the moments just after being asked to make self-predictions or describe their past exercise behaviour, participants where not experiencing any greater level cognitive dissonance than if they were not asked. Although, the present studies focus on exercise behaviour and use of self-predictions was consistent with much prior research under the self-prophesy label and a cognitive dissonance

explanation of the QBE, these results are inconsistent with the majority of prior research on both counts.

The first finding, that asking participants to make self-predictions did not initiate a QBE on exercise scores is discordant with the first hypothesis as well as previous research. This suggests that in the case of exercise behaviour under the conditions of a lockdown, asking participants to make self-predictions did not initiate the mechanism which underlies the QBE as reliably as it has in previous research under standard conditions. Why the present study found no evidence of any QBE is difficult to determine. However, the results of this study clearly show that the prediction question used had no significant effect on behaviour or dissonance. The second finding, that there was no difference between groups for dissonance ratings, is discordant with the second hypothesis and is inconsistent with prior research supporting the cognitive dissonance explanation of the QBE (Spangenberg & Greenwald, 1999; Spangenberg et al, 2003; Wilding et al, 2019). However, taken together from a dissonance-based perspective it is possible to argue that finding no effect on either dissonance or behaviour is simply an indication that the criteria required to generate a QBE through a prediction question were not met, and that in the presence of no increase in dissonance, no QBE would be expected.

In light of the conditions participants undertook this study there is potential that the interplay between social desirability, cognitive dissonance and behavioural difficulty may have impacted results. It is expected that dissonance is evoked by the same process for behaviours which are of personal normative significance as for those which are socially normative (Stone & Cooper, 2001). Although the majority of research in this area has focused on established social norms (Dholakia, 2009), studies of QBE in relation to exercise behaviour have demonstrated the influence of normative beliefs on the magnitude of effect (Sprott et al., 2003). These normative beliefs are in large part made up of evaluations of what

is socially desirable (Sprott et al., 2003). It is of particular relevance to the present study that social desirability has been shown to be important in determining how people act on self-predictions when dissonance is low (Wilding et al., 2016). Specifically, because social desirability is instable (Haberecht et al., 2015), and is subject to situational specific effects due to changing context (Lönnqvist et al., 2007). The lockdown conditions of the COVID-19 pandemic presented a novel context, and whether participants beliefs regarding normative standards for this behaviour remained consistent with previous research is unknown. Therefore, the lack of impact made by self-predictions on participants behaviour and dissonance ratings within the present study may point towards a different set of normative standards of judgment (due to changes in what was socially desirable) being applied than in previous research. Additionally, behavioural difficulty has been shown to attenuate the QBE (Wilding et al, 2016). In particular, when rated dissonance is low, perceived behavioural difficulty is associated with prediction questions having a much smaller effect on behaviour (Wilding et al, 2016). Uniquely, participants within the present study were living under restrictions which limited access to a host of physical activities (Mutz, 2021) and are shown to reduce energy towards physical activity (Giessing et al., 2021). Therefore, it is plausible that the perceived behavioural difficulty associated with exercise was higher than is usual, and in conjunction with the lack of dissonance generated by the prediction question, may have attenuated the expected QBE.

Limitations and Future Directions

When evaluating the findings from the present study, several limitations emerge that warrant reflection. Due to lockdown restrictions the present study was conducted entirely online, this impacted on both the method of delivery and the measures used within the study.

Sample Size and Power

Online studies are known to produce higher dropout rates than lab-based studies (Hoerger, 2010), it is theorised that this is in part due to a reduction in situational demands (Dandurand et al., 2008). Dropout rates among students can be reduced by offering course credit as an incentive (Dandurand et al., 2008), the present study offered such an incentive. Additionally, dropout rates can be reduced further by including contact information in surveys (Clifford & Jerit, 2015), the present study provided this contact information. Despite this, a substantial dropout rate occurred, from 152 responses in the first survey, down to 76 in the second. Partly as a result of this dropout rate, the final sample consisted of only 33 participants.

There are multiple issues which stem from small sample size which affect the interpretation of the results of the present study. First, there is a chance that low performing individuals were more likely to dropout, affecting the accuracy of dependent measures (Crump et al., 2013). Secondly, and most clearly, the possibility that a null result was obtained because of a lack of power due to small sample size. There is a risk of interpretive bias in downplaying non-significant findings as a result of small sample size or power (Hewitt et al., 2008; Kaptchuk, 2003). However, it is well established that the probability of finding effects which are true in a study with low power is itself low, meaning that low powered studies are known to give rise to false negatives more often than high powered studies (Ingre, 2013). The lower the power of a study, the higher the chance of finding false positives and false negatives becomes (Button et al., 2013; Chen et al., 2018, Cremers et al., 2017). Indeed, analysis of studies using similarly small sample sizes have demonstrated that effects can be over and also under-estimated (Cremers et al., 2017; Lorca-Puls et al., 2018). Moreover, it has been demonstrated that studies using small samples only achieve consistency in detecting large deviations from true population effect (Szucs & Ioannidis, 2017). Importantly, not only are small samples sizes associated with poor precision of

estimated effects (Ingre, 2013), but it is also unlikely that studies with such small sample sizes will attain statistical significance when effect sizes are correctly estimated (Lorca-Puls et al., 2018). It is therefore important to put the findings of this study in context with results of similar studies that possess greater statistical power, and also the reality of what can be feasibly be deduced from such a small sample size (33). The sample size and power of the present study clearly has a negative impact on the generalisability of the results. Future research might benefit from avoiding the uncertainty that small samples sizes and low power produce with recruitment strategies and study designs which reduce dropout rates and encourage better engagement.

Measures

The present study was performed via an online platform. Therefore, employing direct physiological tests of dissonance was not possible. The measures used were both established in literature as common means of measuring dissonance (Bran, & Vaidis, 2020; Quirin et al., 2009). However, questions remain over the accuracy of the Dissonance thermometer and it has been indicated that a floor effect likely exists with this measure (Bran, & Vaidis, 2020). Taking into account that both the IPANAT and the Dissonance Thermometer found no significant effect on dissonance ratings provides limited assurance of the accuracy of this finding. Moreover, as multi-modal measures were not employed determining the validity of both measures is difficult. Future studies may benefit from using multi-modal measures, including measuring dissonance directly with physiological arousal through skin conductance (Harmon-Jones et al., 1996) or heart rate (Mann et al., 1969) where this is an option.

A further limitation of the present study regards the reliance on self-report measures. The GPLTAQ required participants to self-report their exercise activity, validity for this method has been demonstrated prior research (Amireault, & Godin, 2015). However, there is an inherent risk of bias when using self-report measures (Althubaiti, 2016). In the case of the

present study, the risk of social desirability bias warrants consideration, particularly as self-reporting of exercise is subject to an over-reporting bias (Brenner, & DeLamater, 2014). Unfortunately, no other measures were used to allow comparison or to verify participants accounts of their behaviour, and so no internal or external validity was determined for this measure. Future studies might consider using more than one measure of exercise to enable a more comprehensive global analysis of this activity.

Conclusion

Overall, the literature on the QBE and the mechanisms which underlie are still developing. Despite its limitations, the present study expands on previous research by employing direct measures of cognitive dissonance. However, results were not consistent with those of previous research. The present study was unable find any evidence of the QBE or the mechanisms which underlie it. Asking participants to self-predict their behaviour neither affected their behaviour, or the level of cognitive dissonance they experienced. Findings suggest that in the context of a COVID-19 pandemic lockdown, using self-prediction questions is not effective at initiating a QBE on exercise behaviour. Findings also suggest that asking prediction questions regarding exercise behaviour during a lockdown is not effective at generating dissonance.

To conclude, this study extended prior research by employing direct measures of proposed mechanisms when investigating the QBE. However, no evidence of the QBE was found with the intervention used. Therefore, this study did not find any evidence in support for the QBE as an effective way of promoting exercise behaviour. In order to allow the QBE to be fully explored and interventions optimised, future studies should expand on existing knowledge by making use of direct measures to allow a more comprehensive analysis of the proposed mechanisms which underlie the phenomenon.