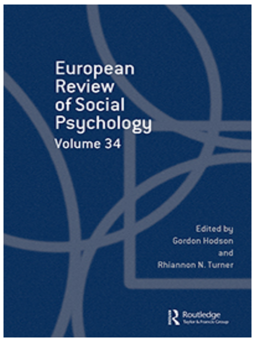

European Review of Social Psychology

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/pers20

Longitudinal tests of the theory of planned behaviour: A meta-analysis

Martin S. Hagger & Kyra Hamilton

EUROPEAN REVIEW OF SOCIAL PSYCHOLOGY
https://doi.org/10.1080/10463283.2023.2225897

Routledge
Taylor & Francis Group

OPEN ACCESS Check for updates

Longitudinal tests of the theory of planned behaviour: A meta-analysis

Martin S. Hagger [a,b,c,d] and Kyra Hamilton [b,c,d,e]

[a]Department of Psychological Sciences, University of California, Merced; [b]Health Sciences Research Institute, University of California, Merced; [c]Faculty of Sport and Health Sciences, University of Jyväskylä; [d]School of Applied Psychology, Griffith University; [e]Menzies Health Institute Queensland, Griffith University

ABSTRACT

In a meta-analysis of longitudinal analyses of the theory of planned behaviour, we tested a series of extended or *auxiliary* theory-consistent hypotheses: construct stability, theory predictions within and between occasions, consistency over time or stationarity in theory effects and reciprocal effects among constructs. We also tested the effects of moderators on theory effects: measurement lag, health behaviour type (protection, risk) and specific health behaviours (alcohol, dietary and physical activity). A systematic search identified 87 studies eligible for inclusion. Meta-analytic structural equation models supported construct stability and theory effects within and between occasions. Only the perceived behavioural control–intention effect exhibited stationarity. We found little evidence of reciprocal effects, and theory effects were small after accounting for reciprocal effects. We observed theory-consistent effects for the behaviour-type moderators, but no variation in model effects for the measurement lag moderators. Findings advance knowledge of the correlates of intentional behaviour and associated processes over time.

ARTICLE HISTORY Received 2 July 2022; Accepted 12 June 2023

KEYWORDS Social cognition theory; attitude theory; reciprocal effects; stability; panel design

A fundamental goal of social psychological research is to identify correlates of intentional, motivated action and to outline the mechanisms involved (Ajzen, 1988; DeCharms & Muir, 1978; Geen, 1995; Weiner, 1990). The development and application of social cognition theories has featured prominently in research seeking to address this goal. Common to these theories is the assumption that individuals are reasoned, rational decision makers whose intentions to perform

CONTACT Martin S. Hagger mhagger@ucmerced.edu Social and Health Psychology Applied Behavioral Research for Prevention and Promotion (SHARPP) Lab, Department of Psychological Sciences, and Health Sciences Research Institute, University of California, 5200 N. Lake Rd, Merced, CA 95343, USA
Supplemental data for this article can be accessed online at https://doi.org/10.1080/10463283.2023.2225897

behaviours in future are based on the available information regarding the target behaviour and evaluative prior knowledge derived from past experience represented in their beliefs (Ajzen & Driver, 1991; Conner & Norman, 2015; Sherman et al., 1989). The theories specify relations between belief-based constructs (e.g., attitudes and normative beliefs), dispositions to act (e.g., motives and intentions) and behavioural responses. Prominent among these theories is the theory of planned behaviour (Ajzen, 1991), which identifies intentions, a motivational construct, as the most proximal correlate of motivated action. Intentions mediate the effects of a series of belief-based constructs concerning the utility, normative concerns and personal capacity with respect to a target behaviour on its future performance. The theory has been applied extensively to predict behaviour across a broad range of behaviours, contexts and populations and meta-analytic syntheses of research applying the theory have largely corroborated its predictions (e.g., Armitage & Conner, 2001a; McEachan et al., 2011).

However, such syntheses have tended to focus on research adopting the received prospective study design in which measures of the belief-based constructs and intentions with respect to a target behaviour are set as predictors of a subsequent measure of the behaviour taken on a later occasion (Ajzen, 1991). Some studies have adopted longitudinal designs to test the theory in which some or all of the constructs are measured on two or more occasions. Such designs not only permit the testing of the original theory predictions but also offer the opportunity to test a number of key *auxiliary* hypotheses implied by the theory. These auxiliary hypotheses are additional predictions implied by the original theory and aimed at addressing the boundary conditions that place limits on the theory or extending its predictions. In the case of the theory of planned behaviour, these hypotheses include testing the degree of stability in its constructs over time, the extent to which its predictions hold over time, known as *stationarity*, and the extent to which theory effects conform to the theory-stipulated unidirectional effects, or exhibit mutual effects on each other, known as *reciprocal* effects.

Prior studies adopting longitudinal designs have not tended to systematically test these auxiliary hypotheses. In the current study, we aimed to address this evidence gap by utilising synthesised data from studies reporting longitudinal tests of the theory to test these auxiliary hypotheses. The value of this research is that it will contribute robust evidence to the extant literature on the extent to which predictions of the theory hold over time when accounting for construct stability and provide a robust large-sample test of the auxiliary hypotheses including the extent to which theory constructs change with time, that is, their stability (e.g., do what extent do attitudes, subjective norms and perceived behavioural control vary over time?), the directionality or reciprocity in the proposed effects between its constructs (e.g., do attitudes predict intentions over time, is the effect in the opposing direction or are both present?) and whether the predicted theory effects hold over time (e.g., to what extent is the intention-behaviour effect consistent or *stationary* over time?). Our analysis will also permit tests of effects of key moderators on theory effects over time.

The theory of planned behaviour

The theory of planned behaviour (Ajzen, 1991) is considered prototypical of social cognition approaches adopted to predict social behaviour, and it has been widely applied (Armitage & Conner, 2001b; Conner & Sparks, 2015; Hagger, 2019). Intention is a focal construct and summarises the strength of an individual's motivation to perform a target behaviour and how much effort they are prepared to invest in pursuing it. Intention is proposed as a function of three sets of beliefs: attitudes, subjective norms and perceived behavioural control. Attitudes reflect estimates of the utility of performing the behaviour underpinned by beliefs that the behaviour will result in certain outcomes and the value attached to those outcomes. Subjective norms reflect perceived social influence on the performance of the behaviour underpinned by beliefs concerning significant others' influence and motivation to comply with these influences. Perceived behavioural control reflects estimates of personal capacity to perform the behaviour underpinned by beliefs in the degree of perceived control over the behaviour and the strength of those beliefs. These belief-based constructs are considered integral to intention formation and are proposed to mediate their effects on behaviour. Perceived behavioural control is proposed to have multiple functions in the theory: as a moderator of the effect of intention on behaviour and as a moderator of the effects of attitudes and subjective norms on intentions (Ajzen, 1991; Lawton et al., 2009). However, researchers have tended to treat perceived behavioural control as a predictor of both intention and behaviour, with a direct effect on behaviour representing instances where perceived and actual control converge (Ajzen, 2002b; Hagger et al., 2022).

The theory has been applied extensively to the prediction of behaviour across multiple behaviours, in diverse populations and in numerous contexts (Ajzen, 2011; Conner & Sparks, 2015; Hagger, 2019). Although the predictions of the theory have been supported experimentally (e.g., Ajzen & Madden, 1986; Chatzisarantis & Hagger, 2005; Sniehotta, 2009), the vast majority of research testing theory predictions have adopted prospective correlational designs in which the belief-based social cognition constructs and intentions are measured on an initial occasion with a follow-up measure of behaviour taken on a subsequent occasion. This has resulted in a burgeoning research literature comprising hundreds of studies (Hagger, 2019).

Findings of these studies have been effectively summarised in many meta-analyses, utilising the synthesised data to test theory predictions including unique effects of direct measures study constructs on intentions, intentions on behaviour and the mediation of the effects of the constructs on behaviour by intentions (e.g., Albarracín et al., 2001; Armitage & Conner, 2001b; Cooke et al., 2016; Hagger et al., 2002, 2016; Hamilton, van Dongen, et al., 2020; McEachan et al., 2011; Rich et al., 2015; Sheeran & Taylor, 1999). These analyses have largely supported the predicted pattern of effects among the theory constructs, and the mediating effect of intentions, with small-to-medium-sized effects, albeit with considerable between-study heterogeneity. The perceived behavioural control moderating effects have also been tested meta-analytically and supported moderation of the intention–behaviour relationship (Hagger et al., 2022). In addition, meta-analyses have provided qualified support for the effects of moderators, such as type of behaviour and temporal lag between theory and behavioural measures, on theory effects (e.g., Hamilton, van Dongen, et al., 2020; McEachan et al., 2011). The theory has also shown utility in guiding interventions aimed at promoting behaviour change using persuasive communications (Ajzen & Schmidt, 2020).

Limitations and issues

The enduring attraction of the theory of planned behaviour lies in its elegant parsimony and in the broad generalisability of its predictions. The theory features a narrow set of "core" constructs, specifies clear directional predictions among them characterised as "risky" in that they are likely to be wrong if the theory is wrong (Trafimow, 2009) and has received broad support for its predicted effects and efficacy in accounting for variance in behaviour (Armitage & Conner, 2001b; McEachan et al., 2011). These virtues and the extensive empirical support notwithstanding, the basic predictions of the theory have been questioned over decades of research from early in its inception (e.g., Bentler & Speckart, 1981; Liska, 1984; Liska et al., 1984) to relatively recent applications (e.g., Hobbis & Sutton, 2005; Ogden, 2003; Sniehotta et al., 2014). Prominent criticisms include the relatively "static" approach taken by the theory in behavioural prediction in that it does not explicitly account for change in its constructs, and the proposed effects among them, over time (e.g., McAuley, 1992); the lack of consideration of potential for reciprocal effects among theory constructs, a criticism largely directed at its "recursive causal structure" (Liska, 1984, p. 67); and the lack of explicit account for past behaviour and its inclusion as standard in theory tests to confirm its sufficiency (see Hagger et al., 2016). Many of the criticisms concern boundary conditions or auxiliary assumptions not explicitly expressed in the original conceptualisation of the theory (Hagger, Gucciardi, et al., 2017; Trafimow, 2012) but are also directed towards researchers' fixation with prospective study designs to test its predictions.

Next, we summarise these criticisms and their relevance to providing better tests of the predictive validity of the theory and outline how researchers have adopted longitudinal study designs with relevant analytic methods to resolve them. Importantly, we highlight the potential of a meta-analytic synthesis of existing longitudinal studies to provide a robust test of a set of auxiliary hypotheses that address the limiting boundary conditions of the theory. Acknowledging that the subsequent discussion makes reference to a number of technical terms, we have made an accompanying glossary available to facilitate comprehension (see Appendix A, online supplemental materials).

Modelling construct change

One of the main concerns regarding the theory is that it does not explicitly account for change in its constructs over time (Bentler & Speckart, 1981; Liska, 1984; Reinecke et al., 1996). In the original conceptualisation of the theory, Ajzen (1991) recognised that the strength of theory predictions would likely be maximised if measures of its belief-based constructs were taken in close temporal proximity with measures of behaviour. It was highlighted that the belief-based constructs reflect individuals' judgements with respect to future behaviour performance at the moment of measurement, which tend not to be temporally invariant-like stable, trait-like constructs (e.g., personality), but are, instead, subject to change with the advent of new information that leads individuals to modify their beliefs. For example, an individual may report a strong intention to donate blood at a future blood drive or to vote in a forthcoming election. However, they may later receive reports from a friend who experienced adverse health effects after donating at that particular drive or hear that their preferred party candidate has dropped out of the election – information that would likely alter their utility, social influence or capacity beliefs with respect to performing the behaviour. The longer the time gap between measures of these constructs, the greater the likelihood that additional information will come to light and affect construct change and reduce their stability. Such change may also increase the likelihood that the constructs will be less effective in predicting behaviour. Of course, new information may alternatively bolster individuals' beliefs so as to maintain them and strengthen the prediction of behaviour. For example, an individual may receive reports of positive donation experiences or receive a leaflet through the mail from an election candidate whose policies they support.[1]

[1]It should be noted that the type of change referred to here relates to the *temporal stability* of constructs and their liability to change over time due to variation in extraneous sources of information. This should be distinguished from the extent to which the constructs are subject to change through, for example, the introduction of manipulations or techniques designed to affect subsequent change in these constructs, namely their *pliability*.

At this juncture, it is also important to highlight the role of the *pliability* of constructs in the theory and how that relates to their stability. While stability refers to the degree of consistency in a given construct, it is distinct from pliability, which represents the propensity of a given construct to change as a consequence of exposure to persuasive communication or the introduction of information that leads individuals to revise their beliefs (see Hamilton & Johnson, 2020). However, pliability and stability are related insofar as change in a given construct as a result of exposure to persuasive communication or new information after initial measures of the construct have been taken will be subsequently reflected in changes in the stability of the construct (see Bassili, 1996). Highly pliable constructs will be more subject to change and, therefore, are more liable to exhibit lower stability. In the context of the theory of planned behaviour, constructs such as attitudes have exhibited highly variable pliability, which has been linked to increased variability in their stability (Armitage & Conner, 2000; Conner & Sparks, 2002).

However, the theory does not explicitly outline how relations within and among the theory constructs change over time or specify means to model such temporal change. This limitation has been reinforced by researchers' tendency to adopt prospective study designs to test theory predictions. Partial resolution for this concern lies in longitudinal tests of the theory, in which theory constructs including behaviour are measured simultaneously on two or more occasions. Such designs allow researchers to model temporal change in the constructs by, for example, regressing each construct on itself over time. In such models, the prospective or follow-up measures of each construct are effectively controlled for variation, or change, in itself over time, known as *covariance stability* (Collins, 2006; Finkel, 1995; Hertzog & Nesselroade, 1987). Furthermore, by measuring within-occasion effects among constructs, the researcher is also able to evaluate whether proposed theory predictions vary over time while accounting for covariance stability in the constructs. In particular, it allows researchers to test the hypothesis as to whether theory effects remain relatively stable or unchanged over time, known as *stationarity* (Gollob & Reichardt, 1987) – an important auxiliary hypothesis test because otherwise the causal and stability effects could not be disambiguated (Rogosa, 1980). The design enables tests of the extent to which effects of attitude, subjective norm and perceived behavioural control on intention remain consistent, that is, exhibit stationarity, or vary across the two occasions. If relations among these constructs tend to wane over time, then it may be indicative that the model is *entropic* such that the effects among its constructs will eventually be extinguished with time (Hertzog & Nesselroade, 1987). This would be one potential consequence if stabilities in the constructs, namely, the regressions of the constructs on themselves over time, are imperfect, which is highly likely given the high possibility for at least some degree of change in these constructs due to the advent of new information, as outlined previously. However, entropy in model effects is only likely in the absence of extraneous variables or unmeasured constructs that affect relations among the constructs over time, which is improbable given the myriad sources of information that likely influence individuals' behaviour.

Indirect and reciprocal effects

A related question in longitudinal tests of the theory is the extent to which the belief-based social cognition constructs and intentions account for variance in intentions and behaviour over time. Ajzen's (1991) original specification of the theory highlighted that variables extraneous to the theory such as environmental factors and psychological traits should affect subsequent behaviour by influencing individuals' sets of beliefs (for examples see Conner & Abraham, 2001 McAnally & Hagger, 2023). As before, these variables act as sources of information that individuals explicitly or implicitly account for when estimating their beliefs and subsequent intentions to perform the behaviour in future. Extrapolating this premise, prior beliefs should also serve a similar informational function, and researchers have alluded to the presence of indirect effects in longitudinal models (Liska, 1984). As a consequence, effects of prior beliefs and intentions towards a given target behaviour should be informative for individuals in their formation of subsequent beliefs and intentions towards that behaviour. This can be tested in longitudinal models of theory predictions by specifying indirect effects of the theory constructs measured on an initial occasion on intentions and behaviour measured on a subsequent follow-up occasion mediated by measures of constructs themselves taken on the subsequent occasion and through measures of intentions and behaviour taken at the initial time point. Such tests may provide indication of the extent to which individuals' prior beliefs inform subsequent decision relative to belief estimates that are more proximal to the behaviour.

A further concern relating to the specification of the theory, and, in particular, the study designs utilised to test its predictions, is that it fails to account for potential for theory constructs to act as mutual effects or "causes" of each other, known as *reciprocal* effects.[2] Despite the intuitively appealing simplicity afforded by the proposed directional effects in the theory, numerous authors have criticised this "recursive causal structure" (Liska, 1984, p. 67), highlighting that the theory constructs are not orthogonal and often share considerable variance. Ajzen (1991) recognised this and explicitly

[2]Note that "cause" here refers to direction in effects rather than change through the experimental manipulation or extraneous influence on one theory construct and its concomitant effect on another, see Liska (1984).

stated, for example, that social cognition predictors of intention should covary, but did not elaborate on the consequences of that shared variance for individuals' subsequent decision-making. To account for this, models testing theory predictions, including those derived from large-scale meta-analyses (e.g., Hagger et al., 2002; Hamilton, van Dongen, et al., 2020; McEachan et al., 2011), specify covariances among the attitude, subjective norms and perceived behavioural control constructs, a practice which appropriately recognises the shared variance and facilitates model fit but is silent on the possible consequences of that shared variance.

An alternative conceptualisation of the shared variance between these constructs is to specify how they might serve as sources of information for each other over time. For example, individuals' beliefs in the utility of performing a future target behaviour may also inform their subsequent beliefs in how much control they have over the behaviour, and vice versa. Similarly, individuals' intention to perform the behaviour on one occasion may also inform their subsequent beliefs with respect to performing that behaviour in future. Such mutual "causes" are not explicitly accounted for in model specification and may, as a consequence, neglect to elucidate potential "dynamic interplay" between the theory constructs over time as a further mechanism by which theory constructs relate to intentions and behaviour. The adoption of longitudinal *panel* designs to test theory constructs may assist in addressing this limitation by incorporating cross-lagged or *reciprocal* effects among theory constructs.

There is conceptual and empirical precedence in research applying such approaches. For example, previous studies examined reciprocal effects among attitudes and behaviour (Bentler & Speckart, 1981; Liska et al., 1984), and recent studies have provided more elaborate tests of reciprocal relations among multiple constructs of the theory (e.g., Eggers et al., 2015; Hagger, Chatzisarantis, Biddle, & Orbell, 2001; Marsh et al., 2006; Niepel et al., 2018). These studies provide some support for the proposed reciprocal relations among theory constructs, such as effects between attitudes and behaviour, a finding consistent with theory and research identifying prior experience as an influence on subsequent beliefs (e.g., Bem, 1972; Kroesen et al., 2017), as well as between perceived behavioural control and attitudes and subjective norms and attitudes. These effects suggest that individuals' beliefs regarding their performance of the behaviour in the future are informed, in part, by their prior beliefs. Longitudinal studies that model these effects add to the knowledge of the processes by which individuals' beliefs serve to inform their subsequent belief estimates and demonstrate that study designs restricted to testing directional predictions of the theory likely mask the presence of these "dynamic" processes. Alongside this, cross-lagged effects may also be informative in tests of stationarity of theory effects in longitudinal-design studies, given that reciprocal effects, that is constructs mutually causing each other, may maintain relations over time in the absence of perfect stability in the constituent constructs (Hertzog & Nesselroade, 1987). Adopting longitudinal models that specify reciprocal effects among theory constructs, therefore, has the potential to further elucidate the processes by which beliefs and intentions relate to behaviour over time.

Effects of past behaviour

A final issue highlighted in critiques of the theory relates to the effects of past behaviour. Although some social cognition theories have formally specified a role for past behaviour in behavioural prediction (e.g., Bagozzi, 1992; Triandis, 1977), the theory of planned behaviour does not make such predictions explicit. As Ajzen (1991, 2002b) points out, past behaviour is not a psychological construct, and its effects, therefore, are relatively uninformative when included as an additional predictor in tests of the theory. Nevertheless, Ajzen suggested that the inclusion of past behaviour as a predictor of subsequent behaviour in prospective studies of the theory provides a test of its *sufficiency*. If theory predictions hold when past behaviour is included, then it is considered sufficient, that is, its constructs and their effects are fit for purpose in accounting for unique variance in behaviour. By contrast, tests of theory predictions that exclude past behaviour will likely return biased estimates of model effects (Hagger et al., 2018; Liska, 1984). Research reporting tests of the theory that simultaneously control for past behaviour effects has indicated that theory predictions hold, supporting the sufficiency hypothesis (e.g., Albarracín et al., 2001; Hagger et al., 2002, 2016; Hamilton, van Dongen, et al., 2020; McEachan et al., 2011).

However, researchers have observed attenuation in the effect sizes among the theory constructs when past behaviour is included, likely attributable to variance shared between the belief-based constructs and intentions and behaviour that is also shared with past behaviour. As past behaviour is not a psychological construct, these effects model the influence of unmeasured variables or constructs that inform the decision-making processes, such as individual differences, implicit beliefs or habits (Ajzen, 1991, 2002b; Chatzisarantis et al., 2007; Hagger et al., 2023; Hamilton et al., 2023) and may also reflect inadequacies of the measures of the belief-based constructs. Furthermore, past behaviour–behaviour effects are expected, to some degree, to be indirect, mediated by the belief-based constructs and intentions as a reflection of the informational function of past behaviour on future decision-making (Ajzen, 2002b). This is also consistent with research applying panel designs to test reciprocal effects between behaviour and constructs, such as attitudes, alluded to previously – behaviour serves to inform beliefs just as beliefs account for variance in behaviour (Bentler & Speckart, 1979; Liska et al., 1984).

Despite the importance of the inclusion of past behaviour in providing evidence for theory sufficiency, relatively few studies formally account for past behaviour effects (Hagger et al., 2016, 2018). Theory tests adopting longitudinal designs, however, include the effects of past behaviour by design. They also offer an advance on studies using prospective designs by enabling examination of past behaviour effects while simultaneously accounting for the stability of other theory constructs, permitting tests of indirect effects of prior social cognition constructs on subsequent intentions and behaviour (Hertzog & Nesselroade, 1987; Liska, 1984). Furthermore, such models allow the specification of reciprocal effects between theory constructs and behaviour, which test the extent to which past behaviour informs subsequent decision-making and behaviour alongside the proposed directional effects from the theory (Hagger, Chatzisarantis, Biddle, & Orbell, 2001; Niepel et al., 2018). Specifically, longitudinal models provide the opportunity to estimate indirect effects of social cognition constructs from the theory taken at an initial occasion on intentions measured at a follow-up occasion mediated by intentions and behaviour. Such models provide a more comprehensive perspective on the processes by which the belief-based constructs from the theory account for intentions and behaviour over time, particularly the extent to which prior beliefs and past behaviour inform subsequent decision-making.

The present study

Despite the enduring appeal afforded by its inherent simplicity and demonstrable efficacy in explaining variance in intentions and behaviour, the theory of planned behaviour has notable conceptual and empirical limitations. For example, the theory does not explicitly account for: (a) the stability in, or temporal change in, its constructs; (b) simultaneous effects of construct stability and prior behaviour on theory relations; (c) stationarity in its predictions; (d) indirect effects of prior theory constructs on subsequent intentions over time and (e) reciprocal relations among its constructs over time. Furthermore, tests of the theory typically do not account for past behaviour effects and, in particular, seldom adopt designs to examine past behaviour effects while simultaneously accounting for construct stability. Longitudinal tests of the theory offer some resolution through testing key auxiliary hypotheses that address these limitations. They do so by measuring study constructs and estimating theory-stipulated relations among them over time, which afford opportunity for stability, stationarity and reciprocity tests, while simultaneously accounting for past behaviour. While studies adopting longitudinal designs have provided initial, albeit qualified, support for these auxiliary hypotheses (e.g., Eggers et al., 2015; Hagger, Chatzisarantis, Biddle, & Orbell, 2001; Marsh et al., 2006; Niepel et al., 2018; Reinecke et al., 1996), we aimed to provide further, more robust, support in the current study by quantitatively synthesising data from tests of the theory using longitudinal designs using meta-analysis. We expected such a synthesis would yield precise estimates of the size and variability in the proposed effects of the theory over time by leveraging the larger sample size afforded by the meta-analysis. Specifically, we aimed to locate studies reporting full or partial longitudinal tests of theory predictions on two or more occasions, extract effect size data for relations among the theory constructs over time, compute averaged sample-weighted correlations among theory constructs across studies using random effects meta-analysis and use the matrix of averaged correlations to test the hypotheses of longitudinal models of the theory using meta-analytic structural equation modelling. Our analysis was pre-registered on the Prospero database of systematic reviews: https://www.crd.york.ac.uk/prospero/display_record.php?RecordID=112780

Proposed longitudinal and panel models

We aimed to test the proposed auxiliary hypotheses of the theory in two models: (a) a longitudinal model in which we specified covariance stability effects of the theory variables, theory effects within each occasion to test stationarity and effects of past behaviour on theory constructs and (b) a panel model in which we specified cross-lagged effects among theory variables while controlling for stability in constructs. We tested the hypothesised longitudinal (e.g., stability, stationarity and effects over time) and reciprocal effects of the theory in separate models to minimise model complexity and promote ease of interpretation.

Proposed effects in our longitudinal model of the theory are summarised in Figure 1. We predicted non-zero averaged stability effects of each of the social cognition constructs, intentions and behaviours on themselves over time, known as *autoregressive* effects. In addition, we expected non-zero averaged direct effects of the social cognition constructs in the theory (attitudes, subjective norms and perceived behavioural control) on intentions and of behaviour on each of the theory variables, within the initial (T1) and follow-up (T2) measurement occasions. We specified non-zero averaged direct effects of intentions and perceived behavioural control at T1 on behaviour at T2. Taken together, the model was expected to yield estimates of stability in theory constructs over time, test standard theory hypotheses controlled for past behaviour within each time point, test for stationarity in theory effects through comparison of effects across occasions and provide estimates of the effects of past behaviour on theory constructs within each occasion. We also expected that there would be non-zero averaged indirect effects of social cognition

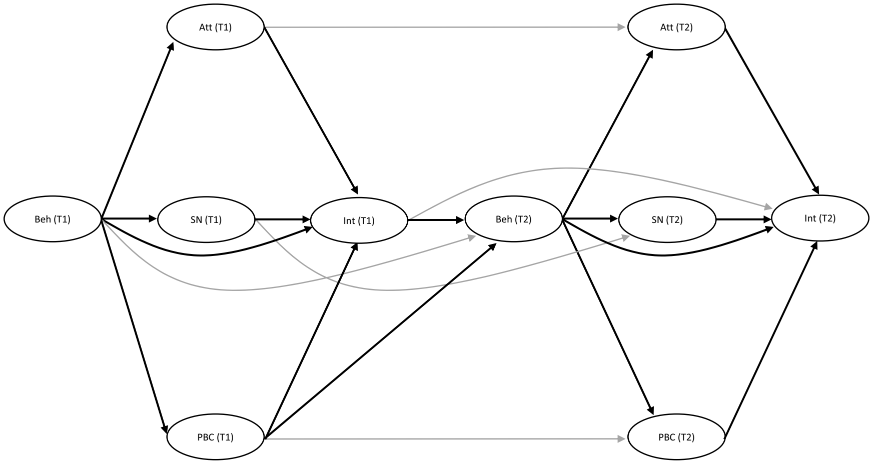

Figure 1. Proposed longitudinal model of the theory of planned behaviour. **Note**. T1 = Initial data collection occasion; T2 = Follow-up data collection occasion; Att = Attitude; SN = Subjective norm; PBC = Perceived Behavioural control; Int = Intention; Beh = Behaviour. Solid lines represent core hypothesised effects in the theory including past behaviour effects. Greyscale lines represent time-lagged stability coefficients or "autoregressions". Covariances between the error variance terms of the attitude, subjective norm and perceived behavioural control constructs within T1 and T2 omitted for clarity.

constructs at T1 on behaviour at T2, in keeping with typical predictions of the theory, and of the constructs at T1 on T2 intentions through intentions at T1, and social cognition constructs and behaviour at T2. As with all models based on the theory, error variances of the endogenous social cognition predictor constructs within each time point were set to covary (Ajzen & Driver, 1992; Hagger et al., 2002).

Proposed effects in our panel model of the theory are summarised in Figure 2. In the model, we specified autoregressive stability effects, expected to be non-zero and cross-lagged effects among social cognition constructs from the theory and intentions and behaviour across T1 and T2. The cross-lagged effects tested the hypothesised reciprocal relations among theory constructs while simultaneously controlling for construct stability over time. While we estimated these effects as free parameters, prior research has not provided equivocal support for systematic cross-lagged effects among theory constructs. As a consequence, we did not specify hypotheses regarding cross-lagged effects and considered these analyses exploratory. Consistent with analyses of panel model designs, the exogenous predictors at T1, and the error variance terms of the endogenous dependent variables at T2, were set to covary.

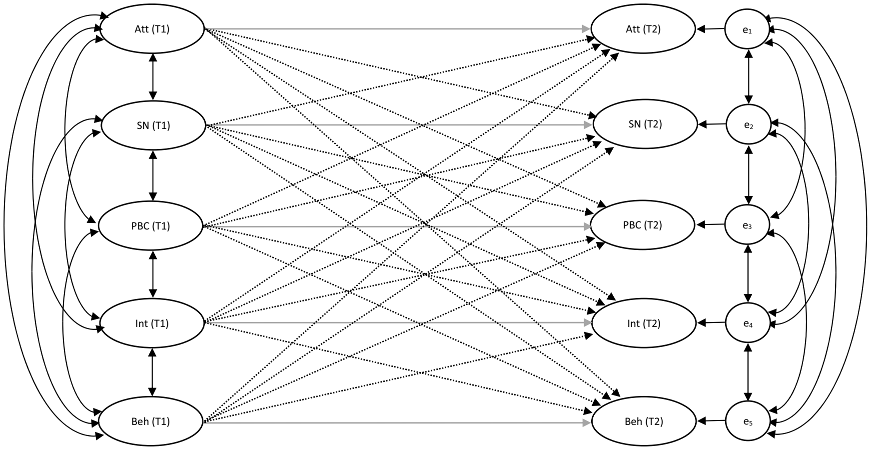

Figure 2. Proposed cross-lagged panel model of the theory of planned behaviour. **Note**. T1 = Initial data collection occasion; T2 = Follow-up data collection occasion; Att = Attitude; SN = Subjective norm; PBC = Perceived Behavioural control; Int = Intention; Beh = Behaviour; e = Error variance of dependent variable. Solid lines represent direct effects and covariances. Greyscale lines represent time-lagged stability coefficients or "autoregressions". Broken lines represent cross-lagged effects.

Moderators

Although the general patterns of prediction theory of planned behaviour effects are likely to be invariant across contexts, conditions and behaviours, the size of theory effects is likely to vary according to extraneous conditions or moderator variables. Meta-analyses of research on the theory have afforded researchers the opportunity to systematically test for variation in theory effects in groups of studies classified according to levels of key moderator variables (e.g., Hamilton, van Dongen, et al., 2020; McEachan et al., 2011). Several moderators have been examined, the most prominent of which have been the lag in measurement between measures of the theory constructs and the target behaviour (e.g., Hamilton, van Dongen, et al., 2020; McEachan et al., 2011) and the type of behaviour targeted – including conceptually driven behavioural categories such as health-promoting and health risk behaviours as well as specific behaviours that fall into each of these categories (e.g., McEachan et al., 2011). We aimed to test the effects of these moderators in our meta-analysis of longitudinal tests of the theory by grouping studies according to the levels of the proposed moderator, estimating our proposed model in each group and formally comparing model effects across the groups. Next, we outline the conceptual basis for each of these moderators.

Measurement lag

A prominent candidate moderator of the theory of planned behaviour effects studied in prior meta-analyses is the lag in time between measurement occasions. Meta-analytic tests of the lag between measures of theory constructs and behaviour have revealed smaller effect sizes with increased lag (McEachan et al., 2011). This is because longer periods between measures increase the possibility that new information will come to light that affects individuals' beliefs and, therefore, the relevance of those beliefs to explaining subsequent behaviour. The current analysis allowed us to extend the effects of this moderator to the stability of the theory constructs over time as well as between-occasion relations between study constructs and intention and behaviour. We therefore estimated our proposed model in groups of studies classified as having proximal (less than or equal to 4 weeks) and distal (greater than 4 weeks) lag in measurement of theory constructs (Hamilton, van Dongen, et al., 2020; McEachan et al., 2011). We expected smaller stability coefficients, within-occasion effects among theory constructs and intentions and direct and indirect effects of social cognition constructs at T1 on intentions and behaviour at T2, in the model estimated in studies adopting a distal measurement lag compared to the model estimated in studies adopting a proximal lag.

Behaviour type

Another key moderator of theory effects in prior meta-analyses has been the type of target behaviour. This moderator has been tested in a number of forms depending on data availability and conceptual basis. McEachan et al. (2011) distinguished between behaviours that conferred protection from illness or prevented onset of chronic disease and those that presented a health risk. The protection behaviour category was further subdivided into specific behaviours (abstaining from or quitting drugs, physical activity, safer sex and dietary behaviours), while the health risk behaviours encompassed risk-taking behaviours in a single category (e.g., speeding, drinking alcohol, smoking and drug use). Their analysis revealed larger effects of perceived behavioural control on intentions in studies on protection behaviours like physical activity and dietary behaviours and larger effects of subjective norms on intention in studies on health risk behaviours. They also found larger intention–behaviour relations in studies on physical activity behaviour compared to those on safer sex and abstinence behaviours. In the present study, we aimed to replicate this analysis by comparing theory effects in groups of studies that targeted health protection behaviours and those that targeted health risk behaviours, based on McEachan et al.'s coding scheme. Based on their findings, we predicted that individuals would tend to base their intentions on perceived control for protective behaviours and base their intentions on normative influences for risk behaviours. In addition, pending data availability was also aimed to compare effects in studies targeting specific, frequently studied behaviours from the included studies such as alcohol behaviour, dietary behaviour and physical activity.

Covariates

We also controlled the effects in our model for several key demographic variables: sample type (student vs. non-student, or clinical vs. non-clinical, participants), age (predominantly younger vs. predominately older samples) and sample sex (predominantly female vs. predominantly male). We also included ratings of the quality of the included studies, assessed using a validated multidimensional study quality assessment tool (Protogerou & Hagger, 2020), as a further covariate.

Method

Search strategy and study selection

We conducted a search of online databases (Web of Science, Scopus, PsycARTICLES and PubMed) for research items reporting full or partial longitudinal tests of the predictions of the theory of planned behaviour. The search was not restricted by date or language. We also searched for "fugitive literature" (Rosenthal, 1994) by contacting prominent authors in the field and by consulting the reference list of previous reviews and meta-analyses of the theory (e.g., Ajzen, 2011; Albarracín et al., 2001; Armitage & Conner, 2001b; Conner & Sparks, 2015; Cooke et al., 2016; Hagger, 2019; Hagger et al., 2002, 2016; Hamilton, van Dongen, et al., 2020; McEachan et al., 2011; Rich et al., 2015; Sheeran & Taylor, 1999). We adopted a three-stage screening procedure to assess the eligibility of the articles identified in the searches for eligibility for inclusion in our analysis. Specifically, the title, abstract and full text of the identified articles were screened against inclusion criteria by two researchers with training in systematic review methods. Consistency in decisions among researchers during the screening procedure was verified through double-screening of a subset ($k = 500$) of the articles screened, with good agreement between the researchers (97.40%, AC1/AC2 agreement statistic = .924, $p < .001$; Gwet, 2008). Inconsistencies were discussed and resolved through consensus, and the screening protocol and decision procedures were updated accordingly. Full search strings used in the database search and a flowchart outlining the search and inclusion and exclusion procedures are presented in Appendix B (online supplemental materials).

Inclusion Criteria

To be eligible for inclusion in our analysis, articles had to report at least one quantitative effect size between a measure of a construct from the theory of planned behaviour (attitudes, subjective norm and perceived behavioural control), or an equivalent measure, and either a measure of behavioural intention or a measure of behaviour, with both measures taken one more than one occasion – that is, measures of theory constructs, intention and behaviour had to be measured at least twice, once on an initial data collection occasion (T1) and on a subsequent follow-up occasion (T2) some time later or on further occasions. Studies measuring theory constructs over time but did not include a measure of intention or behaviour were excluded. Studies adopting case study, *n*-of−1, or qualitative designs were also excluded. The study protocol was registered in advance on the PROSPERO International Register of Systematic Reviews (see https://www.crd.york.ac.uk/PROSPERO/display_record.asp?ID=CRD42018112780).

Characteristics of included studies

Our search strategy identified 7,187 articles meeting inclusion criteria following screening and after removal of duplicates. Some articles reported data from multiple samples, which yielded additional independent samples for inclusion ($k = 12$), while a few studies used the same dataset ($k = 10$). The final sample comprised 87 independent samples with a total sample size of 23,149. A list of included articles is provided in Appendix C (online supplemental materials), and a full summary of the characteristics of studies included in the analysis is presented in Appendix D (online supplemental materials).[3] A diverse range of target behaviours was represented in the sample. Most of the studies focused on health-related behaviours (e.g., physical activity, dietary behaviours, health-care appointment attendance, smoking cessation, parent-for-child behaviours, blood donation and personal health-care behaviours such as dental flossing and sleep hygiene; $k = 70$) and a substantive minority targeting a diverse set of behaviours outside the health domain (e.g., entrepreneurialism, learning behaviours, knowledge sharing, job searching, technology use, political behaviour and pro-environmental behaviours; $k = 17$). Across the included studies, alcohol ($k = 15$), dietary ($k = 10$) and physical activity ($k = 21$) behaviours were the most frequently targeted. Most studies adopted longitudinal correlational designs, but some adopted experimental or intervention designs ($k = 11$). However, the latter studies either applied experimental manipulations or intervention

[3]A spreadsheet providing full details of studies including full sample demographics, detailed description of constructs measured and target behaviours, operationalisation of the behaviour, and moderator coding is provided online: https://osf.io/xfjq7/

techniques that did not target constructs from the theory of planned behaviour or reported single-group designs adopting methods manipulating or affecting change in more than one theory construct simultaneously. We therefore used the control group data in each case to minimise potential bias, and as a consequence, included studies were exclusively treated as correlational.

Effect size data extraction and classification of constructs

Data extraction

Available effect size, variability and sample data for relations among measures of the theory of planned behaviour constructs, intention and/or behaviour across measurement occasions were extracted from the included studies. Given all included studies were treated as correlational in design, the zero-order correlation coefficient was identified as the appropriate effect size metric for use in analyses. Where studies did not report zero-order correlations among the variables of interest, we used appropriate conversion formulae to compute a correlation coefficient using available effect size data, such as differences in means, tests of difference (e.g., *t*-tests, *F*-ratios) or *p*-values (Borenstein et al., 2009). For example, if a study reported mean differences in a measure of a theory of planned behaviour construct across a dichotomised intention or behaviour measure, we were able to compute a correlation by standardising the difference in the construct means at each level of the behaviour and converting to a correlation coefficient. In cases where studies did not have sufficient data to compute effect sizes, reported correlations that were corrected for measurement error such as those using latent variable analyses, or did not report zero-order effect sizes such as those using multiple regression or multivariate analyses, we contacted study authors to request the required data.

Construct classification

We also applied a classification procedure to ensure that construct measures adopted in the included studies were adequately aligned with definitions and measurement characteristics of the theory of planned behaviour constructs. Our procedure matched the content of the items used in the study measures with Ajzen's (1991) definitions of theory constructs, consistent with published procedures to classify social cognition constructs (Conner, 2016; McMillan & Conner, 2007; Protogerou et al., 2018). The vast majority of the included studies reported following Ajzen's (2002a) published guidelines for developing measures of theory constructs, or cited an equivalent source consistent with these guidelines, although some studies adopted non-standard measures. Using this procedure, we were able to match study measures to the attitudes, subjective norms and perceived behavioural control constructs. Where studies exclusively adopted indirect, expectancy-value measures of the theory constructs, we treated the indirect measures as their equivalent direct measures – measures of behavioural, normative and control beliefs were classified as measures of attitude, subjective norms and perceived behavioural control, respectively. Measures of intention largely followed guidelines, and measures of other dispositions to act such as behavioural willingness or protection motivation were also classified as measures of intention. Finally, measures of behaviour predominantly comprised study participants' self-reports of the frequency of performing the behaviour, although a small minority of studies adopted non-self-report behavioural measures such as the use of gym attendance records or direct observation for physical activity or expired carbon monoxide for smoking or used performance or direct outcome measures as a proxy for behaviour such as GPA in maths for studying behaviour or medical attendance records for health-care appointment compliance.

Moderator and covariate coding

We extracted sufficient data to classify studies according to levels of our candidate moderator variables: measurement lag, health behaviour type and specific health behaviours.[4] In addition, we also extracted data on sample and study characteristics, which we used to classify studies into a series of variables used as covariates in our analyses: sample age, sample gender distribution, sample type (student vs. non-student; clinical vs. non-clinical) and study quality. Moderator and covariate coding for each study is summarised in the study characteristic table (Appendix D, online supplemental materials).[5] It is also important to note that our moderator analyses were not pre-registered.

[4]Other moderators were coded including likelihood of the target behaviour to be formed as a habit and the complexity of the target behaviour (see Hagger et al., 2023), but as these analyses were not directly germane to our longitudinal analysis, we have instead reported them in the online supplemental materials (see Appendix I). Other moderator variables considered were behaviour measure type (self-report vs. non-self-report) and other specific behaviours. However, studies adopting non-self-report measures of behaviour, or targeted other specific behaviours (e.g., smoking, entrepreneurialism, safe-sex behaviour, sleep-related behaviours, and learning behaviours), numbered very few ($k < 10$), or had empty cells in their pooled correlation matrix of study constructs, or both, precluding model estimation.

[5]A spreadsheet providing full details of study characteristics and moderator coding is available online along with analysis scripts and output for inter-rater agreement analyses where relevant:https://osf.io/xfjq7/

Moderator variables

Measurement lag

Given that the measurement of theory constructs and intention and/or behaviour was an inclusion criterion for studies in the current analysis, the lag period between the measurement of study constructs and outcomes over time was considered an important moderator of effects in our model. Most studies adopted a two-wave design with social cognition constructs and/or intention or past or concurrent behaviour measured on the first occasion and the same variables measured on the second subsequent occasion. We extracted the between-occasion time gap in each sample and classified studies as either having a proximal or distal measurement lag. Studies with a lag period of 4 weeks or fewer were assigned to a proximal moderator category ($k = 30$), and studies with a lag period of more than 4 weeks were assigned to the distal category ($k = 58$), based on meta-analyses of previous test of the theory (Hagger et al., 2018; Hamilton, van Dongen, et al., 2020; McEachan et al., 2011). A small minority of studies included multiple data collection occasions and reported data for multiple follow-up measures within the same study, which were accommodated using a multi-level analytic design. For these studies, we coded data from each occasion as proximal or distal in measurement lag, accordingly.[6]

Behaviour type

We subdivided the included studies in our analysis into conceptually based categories of behaviour adopting coding schemes from prior meta-analyses (Hamilton, van Dongen, et al., 2020; McEachan et al., 2011). Specifically, we coded studies targeting behaviours aimed at promoting health (e.g., dietary behaviours and physical activity) and preventing ill health (e.g., blood donation and parent-for-child sunscreen use) into a "health promoting behaviours" category and studies targeting behaviours that posed a risk to health or would likely to lead to maladaptive health consequences (e.g., alcohol consumption and smoking) into a "health risk behaviours" category.[7] Two researchers independently coded studies into these moderator categories, with perfect agreement (Gwet's (2008) AC1/AC2 agreement statistic = 1.00). Consistent with McEachan et al.'s analysis, there were sufficient studies in our sample targeting groups of specific behaviours to code moderator

[6]One study reported collecting data on multiple occasions, some with a measurement lag of four weeks or fewer, and others with a lag of greater than four weeks (Wanberg et al., 2005). Data from this study were therefore included in both the proximal and distal groups of the measurement lag moderator.

[7]We also coded an additional behaviour type moderator variable, categorising studies into those that targeted one-off and those that targeted repeated behaviours. This moderator almost exactly mirrored the likelihood of habit formation moderator coded for our ancillary analysis (see Appendix I) and, given the strong conceptual basis for the latter, superseded this moderator. The coding for the one-off vs. repeated behaviour moderator variable is included in the data file available online: https://osf.io/xfjq7/

variables comprising studies targeting two types of health promoting behaviour, dietary behaviours (e.g., fruit and vegetable consumption, sugar consumption, following a low-fat diet) and physical activity and one health risk behaviour, alcohol-related behaviours (e.g., binge drinking and low-risk drinking).[8]

Covariates

Age

Study samples were classified as older-aged samples ($k = 10$) if the reported average age of the sample was 40 years or older with a standard deviation below 15, or, in instances where average age was not reported, the majority of the sample was aged 40 years or older or had an age range with lower limit greater than 40 years. By contrast, samples were classified as younger-aged samples ($k = 60$) if the average sample age was younger than 40 years with a standard deviation below 15, the majority of the sample was younger than 40 years or had an age range with an upper limit less than 40 years. Studies that comprised a substantial range of ages and fell outside these criteria were classified as "mixed"-aged samples ($k = 17$). The three-category age variable was included as a covariate in tests of our proposed model, and we designated the younger age sample category as the reference group.

Gender

As few studies we conducted exclusively on male or female samples, we classified studies into studies with majority female (≥75% female; $k = 23$), majority male (≤25% female; $k = 4$) or balanced gender (>25% female and <75% female; $k = 60$) sample profiles. The resulting variable was included as a covariate in our analysis with the balanced gender profile category as the reference group.

Sample type

Many of the studies included in the current analysis were conducted on student samples. Given concerns over the representativeness of research findings conducted on student samples (e.g., Henrich et al., 2010), we included student status as a covariate in our analyses. We therefore classified studies into those conducted exclusively on student participants, or a combination of student and non-student participants ($k = 50$), and those conducted on non-student participants ($k = 37$). In addition, some studies were conducted on samples in clinical conditions or with clinically diagnosed

[8]One study reported within-study effect sizes for alcohol consumption, dietary behaviour, and physical activity (Norman et al., 2018). Data from this study were therefore represented in all three categories of our specific behaviour moderator variable.

conditions, which may affect individuals' beliefs relative to samples in non-clinical contexts or without reported clinical conditions. So, we classified studies according to whether they were conducted on samples in clinical contexts such as in a hospital or rehabilitation clinic ($k = 12$) or in non-clinical contexts ($k = 75$). The dichotomous sample-type variables were included as covariates in our model tests.

Study quality

We assessed the quality of the included studies using the 20-item Quality of Survey Studies in Psychology (Q-SSP) checklist (Protogerou & Hagger, 2020), which comprises items assessing study quality in four key areas: study rationale and justification ($n = 3$), participant recruitment ($n = 4$), data treatment and interpretation ($n = 10$) and ethical procedures ($n = 3$). Studies meeting the specific quality criterion of each item were assigned a score of 1 and those not meeting the quality standard, or provided insufficient information to evaluate the criterion, were assigned a score of zero. Scores for each criterion were summed to provide a total quality score out of 20. Each study was scored using a checklist by one researcher with training in the assessment of study quality using the checklist. To corroborate quality assessment, a subset of the sample of studies ($k = 20$) was independently scored by an additional trained researcher. Inter-rater agreement analyses indicated good absolute (average = 91.75%) and statistical (average AC1/AC2 = .874, $ps < .001$) agreement in the researchers' scores on the checklist items and an acceptable intra-class correlation (R) between the researchers' total scores on the checklist ($R = .893$, 95% CI = [.768, .951], $p < .001$). Identified inconsistencies were due to variation in interpretation and application of the quality criteria across the researchers. These were resolved through discussion, and assessment procedures were subsequently re-calibrated and applied to the quality assessment of the entire sample.[9] The total quality score derived from the checklist was used as a continuous covariate in our analyses.

Data analysis

Our goal was to estimate the pattern of effects among theory of planned behaviour constructs over time, as well as estimate the stability, stationarity and reciprocal effects among theory constructs using quantitatively synthesised data from the included samples identified in our searches. To do so, we conducted a two-stage analysis combining multivariate multi-level meta-analysis and meta-analytic structural equation modelling (MASEM) using

[9]Data files and analysis code and output for the inter-rater agreement analyses are available online: https://osf.io/xfjq7/

the procedures outlined by Cheung (2015a, b) and Wilson et al. (2016). In the first stage, we estimated pooled sample-weighted zero-order correlations among the constructs from the theory of planned behaviour, intention and behaviour from the data extracted from the included studies using multivariate multi-level meta-analysis. In the second stage, we used the pooled correlation matrix from the first stage as input for MASEM analyses testing the predictions of our proposed longitudinal model of the theory of planned behaviour and the cross-lagged effects among theory constructs in our proposed panel model.

Multivariate multi-level meta-analysis

Many of the included studies reported multiple within-sample effect sizes (e.g., multiple measures of social cognition constructs, intention or behaviour or multiple follow-up occasions). Including these multiple measures in our analysis would violate the assumption of independence of effect sizes as appropriate in traditional meta-analytic procedures. The multi-level meta-analytic approach is a recommended means to account for this within-study dependency. In the first stage of our analysis, therefore, we produced a pooled correlation matrix with its associated sampling covariance matrix among theory of planned behaviour constructs, intentions and behaviour using multivariate multi-level meta-analysis. The procedure also allowed us to produce a further pooled correlation matrix adjusted for our proposed covariates (sample age, gender, study design and study quality) using the weighted regression procedure suggested by Wilson et al. The analysis produces estimates for each correlation among the study variables with standard errors and 95% confidence intervals, unadjusted and adjusted for covariates. In addition, this analytic approach is designed to handle synthesised data from studies that only contribute a few effect sizes to the correlation matrix and yields precise pooled point and variability estimates for each effect size based on data sets with these kinds of missing data patterns. The analysis also yields estimates of variance attributable to the level 2 (between-study effects) and level 3 (multiple within-study effects) variance components. The proportion of each variance component relative to the overall variance is estimated Cheung's (2014) formula. Heterogeneity in correlations is estimated using Cochran's (1952) Q statistic, and the I^2 statistic provides an estimate of the overall variability in a set of studies not attributable to the variance components corrected for in the analysis. A statistically significant Q-value and I^2 value exceeding 25% are considered indicative of non-trivial heterogeneity.

Meta-analytic structural equation models

In the second stage of the analysis, we fitted our two models to the pooled sample-weighted correlation matrices and associated sampling covariance matrices derived from the multivariate multi-level meta-analysis from the first stage. Our first model specified the proposed effects of our longitudinal model of the theory of planned behaviour (see Figure 1). Our second model adopted a panel design to test cross-lagged effects among theory constructs (see Figure 2). Fit of the proposed models with the data from the first stage was evaluated using multiple recommended indices for goodness-of-fit: the comparative fit index (CFI), the Tucker-Lewis index (TLI), the standardised root mean square of the residuals and the root mean error of approximation (RMSEA). Values for the CFI and TLI that approach or exceed .95, a SRMSR value of less than .08 and a RMSEA value of .07 or less indicate an acceptable fit of the model with the data (Hu & Bentler, 1999; Steiger, 2007). The analysis produces standardised parameter estimates for model effects with accompanying Wald confidence intervals. Estimates were considered non-zero if the lower bound of the confidence intervals about the estimate did not encompass zero. Adoption of standardised estimates enabled comparison of effect sizes within and across models according to suggested rules of thumb (Cohen, 1992). However, it should be noted that evaluation and comparison of effect sizes is more difficult for indirect effects given they are products of standardised estimates. Each model was estimated using the unadjusted and covariate-adjusted correlation matrices. Comparisons of the models using the adjusted and unadjusted matrices were made using Akaike's Information Criterion (AIC) and the Bayesian Information Criterion, with lower values representing the most parsimonious model. Missing data are imputed using the full information maximum likelihood estimation method. The analyses were implemented using the metafor (Viechtbauer, 2010) and metaSEM (Cheung, 2015b) packages in R.

Moderator analyses

Effects of candidate moderator variables on the proposed effects in our longitudinal model of the theory of planned behaviour were tested by estimating the model separately in groups of studies at each level of the moderator, unadjusted and adjusted for covariates. As before, models were estimated using the multi-level MASEM approach, and model fit in each moderator group was evaluated using the same multiple goodness-of-fit criteria. Differences in the standardised parameter estimates for model effects across moderator groups were assessed using the 95% confidence intervals about the parameter estimate differences across models (Schenker & Gentleman, 2001). Confidence intervals excluding zero signalled a statistically significant difference in the parameter estimates across moderator groups with a formal test provided using Welch's *t*-test.

Bias assessment

Effects of selective reporting bias in relations between constructs in our proposed model were assessed using a series of recommended bias-correction methods (Carter et al., 2019). We used tests based on the "funnel" plot of the effect size from each included study against an estimate of its precision (e.g., the inverse standard error). The extent to which values in the plot deviate from the expected "funnel" shape under conditions of no bias provides an indication of small study bias, often attributed to selective reporting or "publication" bias. Three tests based on the "funnel" plot were used: Begg and Mazumdar's (1994) rank correlation test, Duval and Tweedie's "trim and fill" analysis and Egger et al. (1997) regression test and two variants thereof. A statistically significant rank correlation test based on Kendall's tau (τ), a large number of excluded and imputed studies from the "trim and fill" analysis and a significant estimate (z-test) assessing whether the intercept of Egger et al.'s regression model is different from zero were used as indicators of non-trivial selective reporting bias. The "trim and fill" analysis also produces a "corrected" value for the correlation after imputation. We also estimated two variants of Egger et al.'s original regression test: the precision effect test (PET) and the precision effect estimate with standard error (PEESE) (Stanley & Doucouliagos, 2014), each using different precision estimates. Both estimate the presence or absence of bias and yield a bias-corrected effect size estimate. As research has suggested that the PET may underestimate the true effect size under conditions of a non-zero effect, we used Stanley and Doucouliagos' conditional procedure: where the PET estimate is statistically significant, implying a non-zero effect, the PEESE estimate is taken, while in the absence of a statistically significant PET estimate, the PET estimate is taken. All analyses were implemented using the metafor package in R.

In addition, we applied a panel of selective bias tests based on selection methods including Hedges (1984) original selection model (Iyengar & Greenhouse, 1988; Vevea & Hedges, 1995) and recent variants, known as the p-curve (Simonsohn et al., 2014) and p-uniform* (van Aert & van Assen, 2018) tests. The selection model approach requires a researcher to specify a "data model", which provides a description of how the data are generated, and a hypothetical selection model, which models conditions of bias, such as publication of only statistically significant effects. We employed a three-parameter selection model (Carter et al., 2019; McShane et al., 2016), which yields a corrected estimate of the effect size and a likelihood ratio (χ^2) test testing the difference of selection from the standard meta-analytic model, with a non-significant value indicative of an absence of bias. The p-curve and p-uniform* tests assume that distributions of p-values in studies should conform to a characteristic distribution under conditions of no bias.

Specifically, the *p*-curve should exhibit significant right-skewness and non-significant estimates of "flatness" under conditions of non-bias. The *p*-uniform test provides a corrected estimate of the averaged effect size and an estimate of the true study variance (τ^2) along with a likelihood-ratio test, with a non-significant value indicating the absence of bias. The three-parameter selection model, *p*-curve and *p*-uniform* analyses were implemented using the weightr (Coburn & Vevea, 2019), dmetar (Harrer et al., 2019) and puniform (van Aert, 2020) packages, respectively, in R.

Given that these bias detection techniques have not been implemented in multi-level meta-analytic models, we applied the bias correction methods for each correlation separately using conventional random effects meta-analysis using a maximum likelihood estimation method. Aggregation of multiple effect sizes from the same study was, therefore, necessary to implement these analyses. Aggregation was implemented using Hunter and Schmidt's (2015) formula assuming a 0.50 correlation between the effect sizes within studies using the MAc package (Del Re & Hoyt, 2018) in R.

Results

Zero-order correlations

Estimates of the averaged sample-weighted zero-order correlations (r^+) among the theory of planned behaviour constructs, intention and behaviour at both measurement occasions (T1 and T2) from the multivariate multi-level meta-analysis adjusted for covariates are presented in Table 1 along with variability estimates and 95% confidence intervals.[10] Each correlation coefficient was non-zero based on 95% confidence intervals corroborated by formal tests of difference. Heterogeneity statistics for each model are presented in Table 2. Values for the I^2 and Q statistics indicated substantive heterogeneity in the correlations in both the unadjusted and covariate-adjusted models. Comparison of the I^2 statistics and level 2 and level 3 variance estimates for the unadjusted and unadjusted models indicated no notable changes in heterogeneity as a consequence of adjusting for covariates. However, the generally larger observed averaged correlations in the adjusted model indicated that exclusion of covariates tended to attenuate the correlations among constructs, so we interpret results of the covariate-adjusted models.

Consistent with prior research and expectations from the theory, we found non-zero averaged correlations between the theory of planned behaviour constructs (attitudes, subjective norms and perceived behavioural control), and intentions within each measurement occasion, with medium

[10]Correlations from the multi-level multivariate meta-analysis model unadjusted for covariates are presented in the online supplemental materials (see Table E1, Appendix E).

Table 1. Zero-Order averaged bias-corrected correlations among theory constructs, intention and behaviour from the multi-level multivariate meta-analysis of longitudinal tests of the theory of planned behaviour.

Effect	r^{+a}	SE	95% CI LL	95% CI UL	Effect	r^{+a}	SE	95% CI LL	95% CI UL
Beh (T1)-Int(T1)	.575	.022	.531	.618	Att (T1)-PBC (T2)	.491	.020	.452	.529
Beh (T1)-Att (T1)	.437	.023	.391	.483	SN (T1)-PBC (T1)	.403	.022	.361	.446
Beh (T1)-SN (T1)	.401	.022	.357	.444	SN (T1)-Beh (T2)	.378	.019	.340	.416
Beh (T1)-PBC (T1)	.458	.024	.412	.505	SN (T1)-Int (T2)	.457	.018	.421	.493
Beh (T1)-Beh (T2)	.738	.021	.698	.779	SN (T1)-Att (T2)	.349	.020	.309	.388
Beh (T1)-Int (T2)	.520	.022	.478	.563	SN (T1)-SN (T2)	.625	.018	.589	.661
Beh (T1)-Att (T2)	.388	.022	.346	.430	SN (T1)-PBC (T2)	.318	.021	.278	.359
Beh (T1)-SN (T2)	.348	.021	.307	.388	PBC (T1)-Beh (T2)	.421	.021	.381	.462
Beh (T1)-PBC (T2)	.393	.022	.351	.436	PBC (T1)-Int (T2)	.478	.021	.438	.518
Int (T1)-Att (T1)	.618	.021	.577	.660	PBC (T1)-Att (T2)	.491	.020	.453	.530
Int (T1)-SN (T1)	.550	.020	.510	.590	PBC (T1)-SN (T2)	.341	.021	.300	.382
Int (T1)-PBC (T1)	.586	.024	.538	.633	PBC (T1)-PBC (T2)	.662	.019	.625	.700
Int (T1)-Beh (T2)	.520	.019	.483	.558	Beh (T2)-Int (T2)	.618	.022	.575	.660
Int (T1)-Int (T2)	.688	.018	.653	.724	Beh (T2)-Att (T2)	.455	.024	.409	.501
Int (T1)-Att (T2)	.479	.020	.439	.518	Beh (T2)-SN (T2)	.448	.023	.404	.493
Int (T1)-SN (T2)	.470	.019	.433	.506	Beh (T2)-PBC (T2)	.483	.024	.436	.531
Int (T1)-PBC (T2)	.449	.021	.409	.489	Int (T2)-Att (T2)	.599	.022	.557	.642
Att (T1)-SN (T1)	.466	.021	.426	.507	Int (T2)-SN (T2)	.587	.021	.546	.628
Att (T1)-PBC (T1)	.537	.021	.497	.578	Int (T2)-PBC (T2)	.590	.022	.546	.634
Att (T1)-Beh (T2)	.399	.020	.359	.438	Att (T2)-SN (T2)	.474	.022	.431	.518
Att (T1)-Int (T2)	.483	.020	.444	.521	Att (T2)-PBC (T2)	.552	.021	.510	.594
Att (T1)-Att (T2)	.657	.019	.620	.693	SN (T2)-PBC (T2)	.435	.023	.390	.480
Att (T1)-SN (T2)	.368	.020	.328	.408					

Note. Correlations reported are from the model adjusted for the following covariates: age, gender, sample type (student vs. non-student), sample type (clinical vs. non-clinical) and study quality. [a]All coefficients are statistically significant ($p < .001$). r^+ = Zero-order correlation corrected for sampling error; 95% CI = 95% confidence interval of r^+; LL = Lower limit of 95% confidence interval; UL = Upper limit of 95% confidence interval; SE = Standard error; T1 = First measurement occasion; T2 = Second measurement occasion; Att = Attitude; SN = Subjective norm; PBC = Perceived behavioural control; Int = Intention; Beh = Behaviour.

effect sizes (r^+ range = .550 to .618, $ps < .001$). Similarly, we found non-zero correlations between these constructs measured at T1 and intentions measured at T2 with small-to-medium effect sizes (r^+ range = .457 to .483, $ps < .001$). The intention–behaviour averaged correlations within (T1: $r^+ = .575$, $p < .001$; T2: $r^+ = .618$, $p < .001$) and across ($r^+ = .520$, $p < .001$) measurement occasions were also non-zero and medium in size. We note that the absolute magnitude of these correlations is consistent with meta-analytically derived correlations reported in prior studies (e.g., Hamilton, van Dongen, et al., 2020; McEachan et al., 2011). We also found non-zero correlations between behaviour and all of the social cognition constructs of the theory within (r^+ range = .401 to .483, $ps < .001$) and across (r^+ range = .348 to .421, $ps < .001$) measurement occasions, with small-to-medium effect sizes. The average stability correlations for each study variable on itself over the two measurement occasions were all non-zero and medium-to-large in size (r^+ range

Table 2. Heterogeneity statistics from the multi-level multivariate meta-analytic models of longitudinal tests of the theory of planned behaviour for the full sample and moderator groups.

Model	L2 σ^2	L3 σ^2	Q^a	df	I^2	L2 var	L3 var
Full sample model	0.018	0.015	43147.135***	4264	92.365	50.148	42.217
	0.017	0.015	43008.484***	4264	92.321	50.031	42.290
Moderator: Measurement lag							
Proximal	0.011	0.015	5549.331***	871	86.208	35.929	50.280
	0.012	0.015	7537.214***	871	86.913	39.463	47.450
Distal	0.017	0.014	37067.473***	3348	92.663	49.879	42.785
	0.015	0.014	36633.495***	3348	92.286	47.438	44.847
Moderator: Health behaviour type							
Health protection behaviours	0.016	0.016	31617.341***	3327	92.593	46.803	45.791
	0.014	0.016	31436.193***	3327	92.283	44.580	47.703
Health risk behaviours	0.012	0.008	6210.431***	892	84.360	50.686	33.674
	0.010	0.008	6473.240***	892	82.976	46.321	36.654
Moderator: Specific behaviours							
Alcohol behaviours	0.014	0.008	6404.574***	893	86.091	54.896	31.294
	0.010	0.008	7083.624***	893	83.309	45.741	37.568
Dietary behaviours	0.007	0.008	1280.705***	267	76.258	33.972	42.286
	0.001	0.008	984.652***	267	67.079	8.517	58.562
Physical activity	0.008	0.010	6474.806***	1131	83.056	35.896	47.160
	0.005	0.010	5525.933***	1131	80.893	27.688	53.205

Note. Values printed on upper line are for models unadjusted for covariates, values printed on lower line are for models adjusted for the following covariates: age, gender, sample type (student vs. non-student), sample type (clinical vs. non-clinical) and study quality. L2 = Level 2 variance component of multi-level model (variance between effect sizes within studies); L3 = Level 3 variance component of the multi-level meta-analytic model (variance between studies); σ^2 = Estimate of "true" variability in the effect; Q = Cochran's Q test; df = Degrees of freedom for Q; I^2 = Higgins and Thompson's (2002) I^2 statistic; L2 var. = Percentage of total variability attributable to variability between effect sizes within studies (level 2); L3 var. = Percentage of total variability attributable to variability between studies (level 3).
***$p < .001$ **$p < .01$ *$p < .05$.

= .625 to .738, *p*s < .001). The stability correlations excepted, averaged correlations between study variables within each measurement occasion (e.g., the correlation between intention and attitude at T1, $r^+ = .618$, $p < .001$ and between intention and attitude at T2, $r^+ = .599$, $p < .001$), were all larger than the same correlations across time points, i.e., across T1 and T2 (e.g., the correlation between attitude at T1 and intention at T2, $r^+ = .483$, $p < .001$ and between attitude at T2 and intention at T1, $r^+ = .479$, $p < .001$), confirmed by formal tests of difference (*t*s >2.916, *p*s < .004). We also found no differences in the cross-lagged correlations among the theory of planned behaviour constructs and intentions (r^+ range = .449 to .483, *p*s < .001), confirmed by formal tests of difference (*t*s <0.994, *p*s > .320). Considerable consistency was noted, therefore, in the averaged zero-order correlations across studies and across measurement occasions.

Meta-analytic structural equation models

We tested the predictions of our longitudinal model of theory of planned behaviour (Figure 1) and our panel model specifying cross-lagged effects (Figure 2) by fitting each proposed model to the averaged sample-weighted correlation matrices derived from the multivariate multi-level meta-analysis and the associated sampling variance–covariance matrices. As with the multivariate multi-level meta-analysis models, we estimated covariate-unadjusted and covariate-adjusted models. Goodness-of-fit, variability and homogeneity statistics for the models are presented in Table 3. The models exhibited acceptable model fit according to the multiple criteria adopted, and AIC and CAIC values indicated that adjustment for covariates led to only very small differences in model fit. However, as with the zero-order correlations from the multivariate multi-level meta-analytic model, we focus on results from the covariate-adjusted models.[11]

Longitudinal model

Stability. Standardised parameter estimates from the MASEM analysis of our longitudinal model of the theory of planned behaviour are summarised in Table 4.[12] Regression of each of the theory of planned behaviour constructs and behaviour measures on itself over the two measurement occasions within our model, effectively tests of covariance stability or *autoregressive* effects, revealed non-zero effects for all variables with small-to-medium-sized effects.

Theory effects and stationarity. Consistent with predictions of the theory of planned behaviour, we observed non-zero averaged effects of attitude, subjective norms and perceived behavioural control on intentions at the T1 and T2 measurement occasions, with small effect sizes. Tests of difference in these effects across time points indicated that the effects of attitude and subjective norms on intentions were larger at T1, but there were no differences in the estimates for perceived behavioural control on intention. Only the perceived behavioural control–intention effect, therefore, exhibited stationarity. In addition, we observed a non-zero averaged effect of intention at T1 on behaviour at T2, consistent with typical prospective tests of the theory, with a small effect size. Consistent with standard tests of the theory of planned behaviour, we observed non-zero indirect effects of attitude, subjective norms and perceived behavioural control at T1 on behaviour at T2

[11]Parameter estimates and variability statistics for all of the meta-analytic structural equation models unadjusted for covariates are presented in the online supplemental materials (see Appendices F, G, and K).

[12]Full results of the longitudinal model are presented in Table F1, Appendix F (online supplemental materials).

Table 3. Fit indexes for multi-level meta-analytic structural equation models of longitudinal model of the theory of planned behaviour for the full sample and moderator analyses.

			Fit indexes										
			Goodness-of-fit							RMSEA 95% CI		Model selection criteria	
Model	N	k	χ^2	df	p	CFI	TLI	SRMR	RMSEA	LL	UL	AIC	BIC
Full sample MASEM model[a]	23149	87	131.475	18	<.001	0.963	0.908	.032	.017	.014	.019	95.475	−49.420
	23149	87	113.274	18	<.001	0.973	0.933	.023	.015	.013	.018	77.274	−67.621
Moderator: Measurement lag													
Proximal	4892	30	19.708	18	.350	0.998	0.996	.023	.004	.000	.014	−16.293	−133.209
	4892	30	20.077	18	.329	0.998	0.994	.023	.005	.000	.014	−15.923	−132.840
Distal	19160	58	127.002	18	<.001	0.954	0.885	.038	.018	.015	.021	91.002	−50.489
	19160	58	102.233	18	<.001	0.971	0.928	.024	.016	.013	.019	66.232	−75.258
Moderator: Health behaviour type													
Health protection behaviours	19624	74	106.792	18	<.001	0.968	0.920	.031	.016	.013	.020	70.792	−71.130
	19624	74	91.165	18	<.001	0.978	0.946	.023	.014	.012	.017	55.165	−86.756
Health risk behaviours	4209	14	69.656	18	<.001	0.927	0.817	.049	.026	.020	.033	33.656	−80.553
	4209	14	70.357	18	<.001	0.928	0.819	.042	.026	.020	.033	34.357	−79.853
Moderator: Specific behaviour													
Alcohol behaviours	4387	15	65.350	18	<.001	0.934	0.834	.061	.025	.018	.031	29.350	−85.605
	4387	15	69.263	18	<.001	0.930	0.825	.040	.026	.019	.032	33.263	−81.693
Dietary behaviours	2424	10	14.185	18	.717	1.000	1.021	.028	.000	.000	.014	−21.815	−126.092
	2424	10	13.937	18	.733	1.000	1.009	.027	.000	.000	.014	−22.063	−126.340
Physical activity	5907	21	63.104	18	<.001	0.967	0.922	.035	.021	.015	.026	27.104	−93.206
	5907	21	62.504	18	<.001	0.972	0.931	.031	.021	.015	.026	26.504	−93.806

Note. Values printed on upper line are for models unadjusted for covariates, values printed on lower line are for models adjusted for the following covariates: age, gender, sample type (student vs. non-student), sample type (clinical vs. non-clinical) and study quality. [a]Fit indices presented are for the proposed longitudinal model of the theory of planned behaviour using meta-analytic structural equation modelling for the full sample, fit indexes for the proposed panel model of the theory of planned behaviour are not presented, it was fully a saturated model with zero degrees of freedom so model fit was perfect. N = Total sample size across studies contributing to model; k = Number of studies contributing to estimated model; χ^2 = Model goodness-of-fit chi-square relative to independence (totally free) model; df = Degrees of freedom; CFI = Comparative fit index; TLI = Tucker-Lewis Index; SRMR = Standardised root mean square residual; RMSEA = Root mean square error of approximation; RMSEA 95% CI = 95% confidence intervals of RMSEA; LL = Lower limit of the RMSEA 95% confidence interval; UL = Upper limit of the RMSEA 95% confidence interval; AIC = Akaike's information criterion; BIC = Bayesian information criterion; MASEM = Meta-analytic structural equation modelling.

Table 4. Standardised parameter estimates from the meta-analytic structural equation modelling analysis of the proposed longitudinal model of the theory of planned behaviour.

		95% CI					95% CI		
Effect	β	LB	UB	*p*	Effect	β	LB	UB	*p*
Direct effects					Autoregressions (continued)				
Beh (T1)→Att (T1)	.429	.393	.464	<.001	Int (T1)→Int (T2)	.323	.281	.365	<.001
Beh (T1)→SN (T1)	.392	.357	.426	<.001					
Beh (T1)→PBC (T1)	.443	.403	.482	<.001	Indirect effects				
Beh (T1)→Int (T1)	.271	.227	.314	<.001	Att (T1)→Int (T1)→Beh (T2)	.046	.028	.064	<.001
Att (T1)→Int (T1)	.264	.210	.318	<.001	SN (T1)→Int (T1)→Beh (T2)	.046	.030	.063	<.001
SN (T1)→Int (T1)	.267	.229	.304	<.001	PBC (T1)→Int (T1)→Beh (T2)	.107	.063	.151	<.001
PBC (T1)→Int (T1)	.187	.128	.245	<.001	Att (T1)→Att (T2)→Int (T2)	.084	.052	.116	<.001
PBC (T1)→Beh (T2)	.075	.024	.125	.004	SN (T1)→SN (T2)→Int (T2)	.105	.082	.129	<.001
Int (T1)→Beh (T2)	.174	.120	.229	<.001	PBC (T1)→PBC (T2)→Int (T2)	.089	.057	.122	<.001
Beh (T2)→Att (T2)	.234	.196	.271	<.001	Att (T1)→Int (T2)[a]	.015	.009	.022	<.001
Beh (T2)→Sn (T2)	.254	.223	.286	<.001	SN (T1)→Int (T2)[b]	.015	.009	.022	<.001
Beh (T2)→PBC (T2)	.233	.193	.274	<.001	PBC (T1)→Int (T2)[c]	.036	.020	.051	<.001
Beh (T2)→Int (T2)	.215	.165	.265	<.001					
Att (T2)→Int (T2)	.144	.090	.198	<.001	Sum of indirect effects				
SN (T2)→Int (T2)	.193	.152	.234	<.001	Att (T1)→Int (T2)[d]	.099	.067	.131	<.001
PBC (T2)→Int (T2)	.153	.099	.208	<.001	SN (T1)→Int (T2)[e]	.121	.098	.144	<.001

(*Continued*)

Table 4. (Continued).

Effect	β	95% CI LB	95% CI UB	*p*	Effect	β	95% CI LB	95% CI UB	*p*
Autoregressions					PBC (T1)→Int (T2)[f]	.125	.092	.158	<.001
Beh (T1)→Beh (T2)	.592	.545	.639	<.001	Int (T1)→Int (T2)[g]	.058	.038	.078	<.001
Att (T1)→Att (T2)	.582	.547	.616	<.001					
SN (T1)→SN (T2)	.547	.515	.578	<.001	Total effects				
PBC (T1)→PBC (T2)	.584	.546	.621	<.001	Int (T1)→Int (T2)[h]	.381	.341	.421	<.001

Note. Model parameters are adjusted for the following covariates: age, gender, sample type (student vs. non-student), sample type (clinical vs. non-clinical) and study quality. Error covariances among independent variables omitted for brevity. [a]Indirect effect of attitude (T1) on intention (T2) mediated by intention (T1), attitude (T2) and behaviour (T2); [b]Indirect effect of subjective norm (T1) on intention (T2) mediated by intention (T1), subjective norm (T2) and behaviour (T2); [c]Indirect effect of PBC (T1) on intention (T2) mediated by intention (T1), PBC (T2) and behaviour (T2); [d]Sum of indirect effects of attitude (T1) on intention (T2) including all effects through attitude at T1, intention (T1) and behaviour (T2); [e]Sum of indirect effects of subjective norm (T1) on intention (T2) including all effects through subjective norm at T1, intention (T1) and behaviour (T2); [f]Sum of indirect effects of PBC (T1) on intention (T2) including all effects through PBC at T1, intention (T1) and behaviour (T2); [g]Sum of indirect effects of intention (T1) on intention (T2) through each social cognition construct (T2) and behaviour (T2); [h]Total effect of intention (T1) on intention (T2) through social cognition constructs (T2) and behaviour (T2) excluding the direct effect of intention (T1) on intention (T2); β = Standardised parameter estimate; 95% CI = 95% confidence interval of standardised parameter estimate; LB = Lower bound of 95% CI; UB = Upper bound of 95% CI; T1 = First measurement occasion; T2 = Second measurement occasion; Att = Attitude; SN = Subjective norm; PBC = Perceived behavioural control; Int = Intention; Beh = Behaviour.

through intentions at T1. Finally, we identified non-zero effects of behaviour, effectively measures of past behaviour, on the theory of planned behaviour variables and intention at the T1 and T2 measurement occasions, with small-to-medium effect sizes.

Indirect effects. We also predicted that our longitudinal model would provide evidence for the extent to which prior social cognition constructs measured at T1 accounted for variance in subsequent intentions at T2. Indirect effects indicated that attitudes, subjective norms and perceived behavioural control measured at T1 at all had unique effects on intentions at T2 mediated by intentions at T1 alone, intentions at T1 and social cognition constructs at T2 and intentions at T1 and behaviour at T2, leading to non-zero sums of indirect effects in all cases. Finally, there were non-zero sums of indirect effects of intention at T1 on intention at T2 through the social cognition constructs and behaviour at T2. However, when including the stability of intention effects across T1 and T2, the total effect was substantially larger, indicating that, unsurprisingly, a substantive proportion of the effect of intention on itself over time was accounted for by its stability.

Cross-lagged panel model

Results of our analysis testing the cross-lagged panel model of the theory of planned behaviour are summarised in Table 5.[13] The analysis revealed effects of subjective norms and perceived behavioural control at T1 on intentions at T2 and of intentions at T1 on behaviour at T2, all with small effect sizes, consistent with theory predictions. However, the effect of attitude at T1 on intentions at T2 was no different from zero, a finding that is inconsistent with theory. Examination of the cross-lagged effects revealed relatively few consistent reciprocal effects among the social cognition constructs. Specifically, we only found reciprocal relations between attitude and perceived behavioural control, and between subjective norms and intentions, and we observed no differences in the sizes for these effects.[14] Consistent with the findings of our longitudinal model, we also found non-zero stability effects for theory constructs, intention and behaviour over time, and within-occasion covariances among social cognition constructs and intentions were all larger at T1 than their corresponding error covariances at T2.

Moderator analyses

We tested for moderation of effects in our longitudinal model by estimating the model in groups of studies defined by levels of the measurement lag (proximal, distal), health behaviour type (protection behaviours and risk

[13]Full results of the panel model are presented in Table G1, Appendix G (online supplemental materials).
[14]Results of these difference tests are available online: https://osf.io/xfjq7/

Table 5. Standardised parameter estimates from the meta-analytic structural equation modelling analysis of the cross-lagged panel model of the theory of planned behaviour.

Effect	β	95% CI LB	95% CI UB	*p*	Effect	β	95% CI LB	95% CI UB	*p*
Direct effects/autoregressions					Direct effects/autoregressions (cont'd)				
Att (T1)→Att (T2)	.527	.473	.582	<.001	Beh (T1)→PBC (T2)	.081	.019	.143	.011
Att (T1)→SN (T2)	.007	−.055	.069	.825	Beh (T1)→Int (T2)	.160	.105	.216	<.001
Att (T1)→PBC (T2)	.181	.121	.240	<.001	Covariances/error covariances[a]				
Att (T1)→Int (T2)	.037	−.021	.096	.214	Att (T1)↔SN (T1)	.466	.426	.507	<.001
Att (T1)→Beh (T2)	.016	−.053	.085	.644	Att (T1)↔PBC (T1)	.537	.497	.578	<.001
SN (T1)→SN (T2)	.516	.473	.558	<.001	Att (T1)↔Int (T1)	.618	.577	.660	<.001
SN (T1)→Att (T2)	−.001	−.052	.050	.960	Att (T1)↔Beh (T1)	.437	.391	.483	<.001
SN (T1)→PBC (T2)	−.004	−.058	.051	.900	SN (T1)↔PBC (T1)	.403	.361	.446	<.001
SN (T1)→Int (T2)	.079	.037	.120	<.001	SN (T1)↔Int (T1)	.550	.510	.590	<.001
SN (T1)→Beh (T2)	.046	−.007	.098	.086	SN (T1)↔Beh (T1)	.401	.357	.444	<.001
PBC (T1)→PBC (T2)	.545	.491	.599	<.001	PBC (T1)↔Int (T1)	.586	.538	.633	<.001
PBC (T1)→Att (T2)	.167	.114	.220	<.001	PBC (T1)↔Beh (T1)	.458	.412	.505	<.001
PBC (T1)→SN (T2)	.026	−.034	.087	.397	Int (T1)↔Beh (T1)	.574	.531	.618	<.001
PBC (T1)→Int (T2)	.064	.004	.125	.037	Att (T2)↔SN (T2)	.192	.153	.232	<.001
PBC (T1)→Beh (T2)	.049	−.018	.115	.149	Att (T2)↔PBC (T2)	.148	.110	.186	<.001
Int (T1)→Int (T2)	.492	.427	.557	<.001	Att (T2)↔Int (T2)	.218	.180	.257	<.001
Int (T1)→Att (T2)	.013	−.071	.098	.758	Att (T2)↔Beh (T2)	.114	.068	.159	<.001
Int (T1)→SN (T2)	.140	.069	.211	<.001	SN (T2)↔PBC (T2)	.169	.127	.212	<.001
int (T1)→PBC (T2)	−.027	−.120	.066	.575	SN (T2)↔Int (T2)	.215	.181	.250	<.001
Int (T1)→Beh (T2)	.089	.005	.173	.037	SN (T2)↔Beh (T2)	.133	.091	.175	<.001
Beh (T1)→Beh (T2)	.639	.583	.696	<.001	PBC (T2)↔Int (T2)	.220	.179	.262	<.001
Beh (T1)→Att (T2)	.074	.013	.134	.017	PBC (T2)↔Beh (T2)	.137	.090	.185	<.001
Beh (T1)→SN (T2)	.046	−.009	.100	.101	Int (T2)↔Beh (T2)	.172	.131	.212	<.001

Note. Model parameters are adjusted for the following covariates: age, gender, sample type (student vs. non-student), sample type (clinical vs. non-clinical) and study quality.
[a]Parameter estimates reported for T1 are covariances among exogenous predictor variables, and estimates reported for T2 are covariances among error terms of the endogenous dependent variables; β = Standardised parameter estimate; 95% CI = 95% confidence interval of standardised parameter estimate; LB = Lower bound of 95% CI; UB = Upper bound of 95% CI; T1 = First measurement occasion; T2 = Second measurement occasion; Att = Attitude; SN = Subjective norm; PBC = Perceived behavioural control; Int = Intention; Beh = Behaviour.

behaviours) and specific behaviour (alcohol behaviours, dietary behaviours and physical activity). Models in each moderator group exhibited adequate fit with the data based on multiple criteria (Table 3). Standardised parameter estimates and comparisons across moderator groups are presented in Tables H1 to H5 (Appendix H, online supplemental materials).

Measurement lag. We found there were very few differences in model effects across moderator groups for the measurement lag moderator. These findings suggest that the stability, within-measurement occasion and between-occasion direct and indirect effects of the longitudinal model were similar in size and pattern regardless of measurement lag across T1 and T2.

Health behaviour type. Comparing model effects across health protection and health-risk behaviours revealed larger effects of attitudes on intentions, and smaller effects of subjective norms on intentions, within each occasion (T1 and T2), with the same pattern of effects observed for the indirect effects and sums of indirect effects in studies targeting risk behaviours compared to those targeting protection behaviours. Effects of prior behaviour on subjective norms, perceived behavioural control, and intentions at T1, and the autoregressions for all constructs, were also larger in studies targeting protection behaviours. [15]

Specific health behaviours. Estimating our model in groups of studies targeting alcohol, dietary and physical activity behaviours revealed relatively few differences in effect sizes. However, we report some notable differences. Within-occasion effects of attitude on intention (T1 and T2) were larger studies targeting alcohol behaviours compared to the studies targeting physical activity behaviours, while the within-occasion effect of perceived behavioural control on intention at T2 was larger in studies targeting physical activity compared to those targeting alcohol and dietary behaviours. In addition, effects of prior behaviour within both occasions (T1 and T2) on intention on both occasions were smaller in studies targeting alcohol behaviours compared to those targeting dietary behaviours and physical activity, while the within-occasion effects of prior behaviour on attitude and perceived behavioural control at T2 were smaller in studies targeting dietary behaviours compared to those targeting alcohol behaviours and physical activity. In addition, the within-occasion effect of behaviour on subjective norm at T1 was smaller, and the within-occasion effects of behaviour on

[15]As ancillary analyses, we also examined effects of three other moderator variables on model effects: likelihood of habit formation (high, low), behavioural complexity (high, low), and type of behaviour (health behaviour, non-health behaviour. The latter moderator was pre-registered. Full descriptions, rationale, results, and discussion of these moderator analyses appear in Appendix I of the online supplemental materials.

attitude and perceived behavioural control at T2 were larger, in studies targeting physical activity compared to those targeting dietary behaviours. The sums of indirect effects of perceived behavioural control at T1 on intention at T2 tended to be larger and effect of attitude at T1 on intention at T2 smaller, in studies targeting physical activity compared to studies targeting alcohol and dietary behaviours, although the effects for studies targeting dietary behaviours fell short of conventional criteria for statistical significance.

Assessment of bias

We applied a panel of recommended tests to assess the degree of selective reporting bias in each of the correlations included in the current meta-analysis. Results are summarised in Table J1, Appendix J (online supplemental materials). Tests based on funnel plots revealed generally non-significant rank correlation tests, low numbers of imputed effects in the trim-and-fill analysis, non-significant values for Egger et al.'s regression test and non-significant effects of precision estimates in the PET and PEESE versions of the regression tests with corrected effect size estimates that did not differ substantially from the uncorrected estimate. Similarly, tests based on selection models revealed significant right-skew and significant flatness tests for the *p*-curve analysis, non-significant bias for the *p*-uniform test with corrected effect size estimates that did not deviate from the uncorrected estimate and non-significant values for the difference between the data and selection models in the three-parameter selection model. Taken together, results of these tests revealed scant evidence for the presence of bias in correlations among theory constructs in this sample of studies.

Discussion

The purpose of the current meta-analysis was to test a series of key auxiliary hypotheses in longitudinal tests of the theory of planned behaviour. We tested our hypotheses in longitudinal and cross-lagged panel design models fitted to data from studies reporting longitudinal tests of the theory using multi-level meta-analytic structural equation modelling. Specifically, we tested the following hypotheses: (a) the stability hypothesis, in which constructs of theory were expected to exhibit non-zero stability coefficients across initial (T1) and follow-up (T2) measurement occasions; (b) the hypothesised within-measurement occasion predictions of the theory, that is, proposed effects of the social cognition constructs of attitudes, subjective norms and perceived behavioural control on intentions at both T1 and T2 while simultaneously controlling for stability and effects of past behaviour, the latter of which tested theory sufficiency; (c) the stationarity hypothesis, that is, the extent to which effects of the social cognition constructs on intentions within each measurement occasion are invariant across occasions; (d) the hypothesised prospective predictions of the theory, that is, effects of the social cognition constructs and intention at the first measurement occasion on behaviour at the follow-up measurement occasion, again controlling for stability and past behaviour effects and (e) the hypothesised reciprocal effects among theory constructs across measurement occasions by specifying cross-lagged effects consistent in a panel design. Finally, assuming a non-trivial residual heterogeneity in relations among the longitudinal model of the theory of planned behaviour across studies, we tested for differences in proposed theory effects in our longitudinal model across groups of studies defined by levels of key moderator variables: measurement lag, health behaviour type (health-promoting and health-risk behaviours) and specific health behaviours (alcohol behaviours, dietary behaviours and physical activity).

Results of our longitudinal model indicated non-zero stability effects for theory of planned behaviour constructs across measurement occasions, with medium-sized effects, supporting our predictions that model effects would exhibit non-trivial stability. Theory predictions, that is, effects of attitudes, subjective norms and perceived behavioural control on intentions, within each measurement occasion were also supported, with non-zero effects reported at the follow-up data collection occasion controlled for the covariance stability of the theory constructs and past behaviour effects. Effect sizes for the attitude–intention and subjective norm–intention effects within time points were smaller at T2 compared to these effects at T1, suggesting tha we should reject the stationarity hypothesis for these effects over time. The effect of intention at T1 on behaviour at T2 was supported, consistent with theory predictions. Analysis of indirect effects revealed non-zero sums of indirect effects of the social cognition constructs and intention at T1 on intentions and behaviour at T2, although it should be noted that effect sizes in all cases were small, with substantive variance in intentions and behaviour at T2 remaining unexplained. Examination of cross-lagged effects in our panel model of the theory of planned behaviour yielded little evidence for reciprocal effects of the theory constructs across measurement occasions, with the exception of reciprocal effects between perceived behavioural control and attitudes and between subjective norms and intentions.

Finally, our moderator analysis for measurement lag revealed few differences in model effects across groups of studies with proximal and distal measurement lag. However, our analyses of the behaviour-type moderator variables revealed notable differences in model effects. Specifically, we found direct and indirect effects of attitudes on intentions that were larger, and effects of subjective norms on intentions that were smaller, in studies targeting health risk behaviours relative to those targeting health protection behaviours. In addition, effects of perceived behavioural control on intentions tended to be larger in studies targeting physical activity relative to studies targeting alcohol and dietary behaviours, while effects of attitude on intention tended to be larger in studies targeting alcohol behaviours relative to those targeting physical activity and dietary behaviours.

Interpreting theory of planned behaviour effects over time

Findings of the current research provide evidence in support of several of our auxiliary hypotheses of theory of planned behaviour effects and further elucidate the processes by which the theory constructs relate to behaviour. Foremost among these findings is support for the temporal stability of the theory constructs. Providing estimates of the degree of construct stability is important given research indicating that stability of constructs such as attitudes and intentions, which effectively captures construct “strength”, moderates the effects of these constructs on intention and behaviour, respectively (e.g., Cooke & Sheeran, 2004; Doll & Ajzen, 1992; Sheeran & Abraham, 2003). Current findings across the extant research provide generalised guidance to researchers on the expected size and variability in the stability expected to be observed in these constructs over time, which would, in turn, partially determine their efficacy in predicting theory-relevant outcomes such as intentions and behaviour. A further notable finding was the greater stability the belief-based theory constructs relative to intentions. This may be because intentions are more subject to the advent of new information and other variables beyond these belief-based determinants and are, therefore, more liable to change. This signals the imperative of identifying additional predictors of intention in augmented or integrated models based on the theory.

In addition, our findings also provided support for theory predictions within and across measurement occasions. Consistent with the theory, intentions were shown to be a function of the belief-based constructs at the initial and follow-up measurement occasions, and intentions at the initial occasion predicted behaviour and mediated the effects of the social cognition constructs. These findings corroborate previous meta-analytic research of prospective studies that test theory predictions using meta-analytically synthesised data (e.g., Albarracín et al., 2001; Hagger et al., 2016; Hamilton, van Dongen, et al., 2020; McEachan et al., 2011) and extend them by demonstrating that the effects hold when simultaneously accounting for temporal change in theory constructs and past behaviour effects. Testing theory constructs while accounting for construct change extends the predictive validity of the theory given that such constructs are subject to change due to the advent of additional information in the interim between measurement occasions that may lead individuals to modify their beliefs.

Our findings demonstrate that while individuals' beliefs do exhibit change – their stability coefficients are imperfect – they still account for unique variance in intentions and behaviour at follow-up when accounting for that change. These findings are consistent with primary studies that have controlled for temporal change in theory constructs over time (e.g., Eggers et al., 2015; Hagger, Chatzisarantis, Biddle, & Orbell, 2001; Niepel et al., 2018). In the original conceptualisation of the theory, Ajzen (1991) anticipated that individuals' beliefs regarding performing a future behaviour would change over time and therefore posited that the theory would be more effective in predicting behaviour when its constructs were measured in proximity to the measure of behaviour (see also St Quinton & Trafimow, 2022). However, the theory is not explicit in how temporal change may be accounted for in tests of the theory. Longitudinal models such as the one tested here present a solution by modelling one form of change, covariance stability, over time. Our findings based on the synthesis of multiple studies provide robust evidence that theory predictions hold over time when accounting for stability.

We were also able to examine the extent to which the social cognition constructs from the theory at the initial measurement occasion indirectly accounted for variance in intentions at the follow-up occasion through intentions at the initial occasion and follow-up measures of the constructs and behaviour. Our findings supported these proposed effects with non-zero indirect effects of the constructs and intentions on follow-up intentions. This pattern of effects provides further elaboration of the decision-making process not captured in typical prospective tests of the theory. Although Ajzen (1991) was not explicit on *when* attitudes, subjective norms and perceived behavioural control lead to intention formation, the underlying assumption of the theory is one of the information processings, in keeping with the social cognition approach. When individuals are prompted to report their intentions to perform a given future behaviour, therefore, they draw from various sources of information, summarised in their belief estimates. The current research illustrates the extent to which individuals' proximal (i.e., within-occasion) and distal (i.e., prior occasion) belief estimates of their beliefs remain informationally salient to the formation of intentions over time. Thus, the relative change in the theory constructs over time notwithstanding, individuals' earlier estimates of the utility, normative support and their capacity with respect to performing the behaviour in future still have resonance in informing their intentions and are, therefore, somewhat enduring over time.

Alongside testing theory predictions within and across measurement occasions while controlling for stability, our analysis also afforded the opportunity to formally test whether effects sizes among theory constructs were consistent over time, that is, a test of the *stationarity* in theory effects.

While we expected consistency in the effects, we found that the sizes of two of the effects, those of attitude and subjective norms on intention, were non-invariant, with smaller effect sizes observed at follow-up. Only the effect of perceived behavioural control on intention, therefore, exhibited stationarity across measurement occasions. These findings suggest that although the expected pattern of prediction is consistent over time, effect sizes for two of the key constructs appear to wane.

Where stabilities of constructs in longitudinal models are imperfect and the model reflects a "closed" system unaffected by other extraneous factors, such findings may imply that the model is entropic – without stability in its constituent constructs to maintain the within-time effects among them, model effects will asymptotically approach the null over time (Hertzog & Nesselroade, 1987). However, such hypothetically rigid conditions seldom apply in models estimated in behavioural data such as these because there are likely other unmeasured extraneous variables that predict intentions and behaviour that may serve to maintain these effects over time, so they are not fully extinguished. Examples of such extraneous factors are environmental influences (e.g., Godin et al., 2010), other beliefs not encompassed by the theory (e.g., Rivis et al., 2009) and constructs that represent non-conscious processes such as implicit cognition (e.g., Chevance et al., 2017; Hagger et al., 2019; Keatley et al., 2012) or habit (e.g., Hagger et al., 2023; Hamilton et al., 2023; Hamilton, Gibbs, et al., 2020). These extraneous variables may also be a reason for the lack of observed stationarity in theory effects. The relatively low stability in intentions relative to other belief-based constructs coinciding with reductions in attitude–intention and subjective norm–intention relations suggests divergence in these constructs and intentions over time. Beliefs may, therefore, become less salient as a basis for individuals' intentions over time, while other extraneous factors become more relevant. There is also possible that cross-lagged effects among constructs may serve to bolster within-occasion relations among constructs insofar as they account for some of the variance shared between the variables involved in the within-occasion theory effects that would otherwise be accounted for by their stability over time.[16] This is an issue we focus on next in our discussion of cross-lagged effects.

[16]In a closed-system cross-lagged model with perfect stabilities, variance shared between model constructs within occasions reflected in, for example, a correlation or direct effect between them, would be unchanged over time, and the model would exhibit perfect stationarity in these effects. However, where stabilities are imperfect, within occasion effects will decline in value over time – such an isolated stability model is, therefore, entropic. Cross-lagged effects between the constructs involved in the within-occasion effects, however, can serve to maintain stationarity in these effects to the extent that the cross-lags account for the shortfall in shared variance attributed to the imperfect stabilities. For further details, the reader is directed to the lucid treatment of this topic by Hertzog and Nesselroade (1987).

Our analysis also presented the opportunity to test for reciprocal effects among theory constructs, intentions and behaviour across measurement occasions. The model indicated support for directional theory effects over time, although, contrary to hypotheses, the effect of attitudes on intention was no different from zero. Furthermore, we only identified reciprocal relations between attitudes and perceived behavioural control and between subjective norms and intentions.

Focusing on the attitude-perceived behavioural control reciprocal relationship, although prior research has supported conceptual and empirical distinctions between these constructs (Trafimow & Duran, 1998), and they demonstrate predictive validity in that they account for unique variance in intentions in prospective tests of the theory, they usually share substantive variance and are often correlated in tests of the theory (e.g., Ajzen & Driver, 1992; Hagger et al., 2002; McEachan et al., 2011). Our study extends these findings suggesting that individuals' attitude and perceived behavioural control estimates serve to inform each other and are inextricably linked, and it is control perceptions that ultimately account for unique variance in behaviour. A possible reason for this reciprocal effect may be overlap or congruence in the sets of beliefs purported to underpin these constructs. For example, individuals who view a behaviour as having utility in producing valued outcomes are also more likely to perceive fewer barriers to performing it in future and similarly cite favourable evaluations of their capacity to perform it. This is consistent with theories suggesting that individuals overall strive for the main consistency in their beliefs (e.g., Guadagno & Cialdini, 2010; Kroesen et al., 2017); hence, favourable behavioural evaluations with respect to a behaviour are likely to align with estimates of capacity to perform it in future. The subjective norm–intention reciprocal relationship illustrates that forming an intention to perform a target behaviour previously will inform individuals' subsequent estimates of the level of normative support for performing the behaviour in future. This is a specific instance that serves to illustrate the generalised prediction that prior decisions serve to inform subsequent beliefs (e.g., Marsh et al., 2006). However, current findings do not provide unequivocal support for this generalised premise.

Returning to the issue of entropy, cross-lagged effects may contribute to maintaining an effect in a panel design over time and, therefore, its stationarity. However, the lack of reciprocal effects between perceived behavioural control and intentions suggests that we can rule out the possibility that reciprocity in these constructs contributes substantively to the observed stationarity in this effect. Ultimately, therefore, our panel model may provide an indication of stationarity in this effect across two measurement occasions, but does not enable us to definitively confirm or reject it, or provide information on why, for two reasons. First, we did not test an elaborate model of theory predictions over time. The model is, in effect, a "first order" longitudinal design, i.e., one with only two measurement occasions. More elaborate models tested on synthesised data from tests of the theory across multiple measurement occasions (i.e., more than two) may be more informative as to the extent of consistency in theory effects over time. Second, as alluded to in our previous discussion, the longitudinal tests of the theory included in the current analysis were collected in the field rather than in controlled conditions such as those of a laboratory. As a consequence, there was minimal control for potential extraneous factors that may alter model effects. It is therefore feasible that although effects among constructs at the follow-up occasions were reduced in size across measurement occasions in the current model, it does not confirm that they will decrease further on subsequent occasions.

It is also important to note that theory effects in both the longitudinal and cross-lagged models accounted for effects of past behaviour and its stability across occasions. Although it has been argued that inclusion of past behaviour effects in tests of social cognition theories is not informative in that past behaviour is not a psychological construct, its inclusion can nevertheless serve to provide important information on theory processes. Specifically, Ajzen (1991) suggested that support for theory effects when accounting for past behaviour provides confirmation of theory sufficiency. This has been confirmed in multiple meta-analytic tests of the theory using prospective designs (e.g., Albarracín et al., 2001; Hagger et al., 2002, 2016; Hamilton, van Dongen, et al., 2020). Our longitudinal model, which included effects of past behaviour by design, provided further confirmation of this hypothesis and extends it by demonstrating that theory effects hold while accounting for the stability of theory constructs and past behaviour across occasions. Our model, therefore, supports theory sufficiency when accounting for behaviour change.

A possible interpretation of past behaviour effects in our longitudinal test of the theory is that they indirectly represent effects of extraneous, unmeasured constructs on theory constructs. These unmeasured constructs are unlikely to be mediated by theory constructs or intentions because they effectively capture the reasoned processes that precede behaviour. Candidate unmeasured constructs likely include those that represent implicit, non-conscious processes such as habits or implicit cognition. Research integrating these constructs as additional predictors in theory tests corroborates this perspective in that they predict behaviour independent of the social cognition constructs and mediate past behaviour effects (e.g., Hagger, Trost, et al., 2017; Hamilton et al., 2017; Phipps et al., 2021). While our current test of the theory is unable to definitively support this interpretation, it nevertheless signals the need for systematic research testing the theory while simultaneously accounting for the effects of these additional constructs on behaviour and their capacity to mediate past behaviour.

Effects of moderators

Measurement lag

We expected model stability constructs, and, possibly, effects of measures of constructs taken at the initial measurement occasion on measures of intentions and behaviour at follow-up, would be smaller in groups of studies adopting a longer measurement lag. This prediction was based on our auxiliary hypothesis that theory constructs will likely be more effective in predicting behaviour when measured in proximity to the behaviour because there is less opportunity for new information to come to light that may alter individuals' beliefs and reduce their predictive validity (Ajzen, 1991). Such a prediction has been supported in prior meta-analytic research reporting attenuation in model effects as lag increases (Hagger et al., 2002; McEachan et al., 2011). However, the fact that we found little evidence for these predictions suggests that stability effects tended to be consistent regardless of measurement lag. This observed invariance in construct stability by measurement lag meant that the indirect effects of theory constructs on intentions over time would also be unlikely to vary because they operate through stability effects. The observed similarity in stability effects across proximal and distal measurement lag constructs may have been due to researchers' fastidious adoption of measures with strong temporal measurement correspondence in the included studies. Another possibility is that there were insufficient studies in the distal lag moderator group that provided very long-term predictions – for example, relatively few studies reported a lag greater than 12 months ($k = 12$). Taken together, the current findings did not provide support for the moderation of longitudinal effects in the theory by measurement lag. However, the current findings should be augmented by future large-sample longitudinal studies applying the theory which systematically vary in measurement lag.

Behaviour type

We anticipated that the pattern of effects among constructs in our longitudinal model of the theory would likely vary across the type of behaviour targeted, including moderator variables distinguishing between health protection and health-risk behaviours as well as specific health behaviours. Focusing first on model effects for studies targeting health protection and health risk behaviours, our analysis identified a more prominent role for attitudes on intentions within each occasion and over time, and a lesser role for subjective norms, for health-risk behaviours. This pattern was also corroborated in our analysis of specific health behaviours. For example, we observed larger effects of attitudes on intentions in studies targeting alcohol behaviours, a risk behaviour, compared to studies targeting specific health-promoting behavioural dietary behaviours and physical activity. This pattern of effects contrasts with those identified elsewhere. For example, effects of subjective norms on behaviour were larger in studies targeting risk behaviours in McEachan et al. (2011) analysis, although it should be noted that effects are not directly comparable given we estimated unique effects within a full test of the model, while McEachan et al. confined their analysis to zero-order correlations.

A possible interpretation of the larger attitude effects may be the preponderance of studies in the current sample targeting behaviours likely to be highly rewarding, such as smoking and excessive patterns of alcohol consumption associated with hedonic motives (e.g., binge drinking). This interpretation is inferred from research indicating that affective attitude effects are important predictors of intentions and behaviour for more impulsive behaviours like alcohol consumption (Conner et al., 2015; Lawton et al., 2009). Furthermore, consistent with this premise, a further meta-analysis that made the distinction between instrumental and affective (experiential) attitudes on intentions found larger affective attitude–intention effects in studies targeting health-risk behaviours (McEachan et al., 2016), and however, recent research indicated that affective attitudes are also important to the prediction of physical activity intentions (Phipps et al., 2020). However, we could not provide definitive empirical support for this explanation in our analysis as few of the included studies made the distinction between affective and cognitive forms of attitude, precluding a separate analysis of these attitude components.

By contrast, our analysis indicated that perceived behavioural control seemed to be more relevant to intention formation in studies targeting physical activity compared to studies targeting the other two behaviours. This finding is consistent with the pattern of effects in McEachan et al. (2011) meta-analysis. Control-related factors, such as barriers and capacity constraints, therefore, seem be the most salient correlate of physical activity intentions compared to alcohol and dietary behaviours. To speculate, this may be because physical activity is less subject to rewarding contingencies and, as such, more likely to demand greater consideration of barriers and capacity prior to intention formation. By contrast, alcohol and dietary behaviours are inherently more rewarding, so anticipated affective outcomes captured by the attitude construct may be more relevant as a basis for intentions.

Limitations and avenues for future research

Our meta-analytic tests of a longitudinal model of the theory of planned behaviour advance knowledge by providing robust tests of construct stability and key theory hypotheses including stationarity and time-lagged predictions while accounting for temporal change in study variables and past behaviour. It also afforded us the opportunity to examine cross-lagged effects among theory constructs in our panel model, which provides better means to test the directional effects proposed in the theory effects, and the extent to which prior social cognition constructs and intentions towards a target behaviour informs subsequent decision-making with respect to performing the behaviour in future. However, current findings should be interpreted in the light of a number of limitations of the current analyses. These include limitations in the designs adopted in the included studies, and in the availability of data, that placed restrictions on our analyses and, therefore, the inferences that could be drawn from them.

A key limitation of the included studies is that they tended not to tap the systems of beliefs that underpin the direct measures of the social cognition constructs of the theory. Ajzen (1991) argues that direct measures of these constructs should effectively summarise their component beliefs, and, as a consequence, the vast majority of studies have tended to adopt direct measures (i.e., the global belief-based constructs of attitude, subjective norms and perceived behavioural control), in keeping with early prospective designs (e.g., Ajzen & Driver, 1992). However, research has indicated that direct measures are frequently associated only with a small handful of beliefs (Armitage & Conner, 1999; Hagger, Chatzisarantis, & Biddle, 2001), which suggests that either individuals tend to rely only on a very small subset of beliefs when making decisions, or the direct measures do not effectively account for these beliefs. The relatively dearth of studies adopting indirect measures of theory constructs has limited meta-analytic studies testing its predictions relying solely on the direct measures (c.f., Armitage & Conner, 2001b). This was the case in the present study, with relatively few studies ($k = 3$) measuring the beliefs, precluding an analysis of our longitudinal model of the theory that included beliefs alongside direct measures. Researchers adopting longitudinal designs, therefore, should consider measuring theory constructs using direct and indirect measures concurrently and use both measure types to test theory predictions longitudinally.

It is also important to acknowledge that although we included studies that adopted experimental or intervention designs, all included data were correlational. This is because identified studies adopting such designs and reporting sufficient longitudinal data adopted manipulations or intervention techniques that either targeted a construct other than those in the theory of planned behaviour or adopted single group designs with manipulations or techniques that targeted multiple theory constructs simultaneously. The experimental or intervention effects in both cases were, therefore, not equivalent to "zero-order" or unattenuated tests of theory effects. For this reason, we used control group data so that the included effect sizes were comparable across studies. As a consequence, we were unable to make comparisons between experimental or intervention designs with correlational designs as a moderator of theory effects. This is an important consideration for future research given that adopting correlational designs may inflate or attenuate effect sizes relative to those from experimental or intervention research due to common method variance (Podsakoff et al., 2003). We look to future longitudinal experimental or intervention research that adopts manipulations or techniques that target change in individual theory constructs or factorial designs that allow for isolation of the effects of such manipulations or techniques on their respective outcomes.

A further limitation of our analysis is that we did not test the hypothesised-moderating effects of perceived behavioural control in the theory, that is, the extent to which perceived behavioural control moderates the effects of intentions on behaviour and of attitudes and subjective norms on intention (e.g., Conner & McMillan, 1999; Schifter & Ajzen, 1985). Testing these hypotheses would require samples to report sufficient data to compute correlations between interaction terms and the theory constructs. None of the included studies did so, and few data sets were openly available for us to compute them ourselves, which is necessary to test these effects meta-analytically (Hagger et al., 2022). Further research testing whether the proposed interaction effects in longitudinal tests of the theory hold over time is needed, and researchers testing the theory are encouraged not only to test the interaction effects as standard practice but also to make their data freely available to permit ancillary analyses such as these.

It should also be noted that while the current longitudinal test of the theory is an advance on prospective designs in that it is more effective in capturing change in theory constructs over time, and in testing associated auxiliary hypotheses, it is still suboptimal as it does not account for rapid, frequent change in study constructs. As Ajzen (1991) intimated, individuals are likely to encounter and interpret a number of additional sources of information regarding their future performance of a target behaviour, information that may lead to fluctuations in the strength of the belief-based constructs that impact their intentions and behaviour, particularly in contexts where rapid change may occur (Randolph, 1981). Resolution may lie in the collection of theory measures on multiple occasions in the interim between actions, and the extent of the variability could be captured in time-series analyses of theory effects using such data, such as autoregressive integrated moving average (ARIMA) models, or latent growth curve models, of theory predictions. There are currently few precedents applying these kinds of research designs to testing theory approaches (c.f., Hanbury et al., 2009; Joensuu et al., 2013), but such approaches would be an important avenue for future research to elucidate the effects of fluctuations in beliefs over time impact intentions towards, and actual participation in, subsequent behaviour.

Relatedly, typical means to analyse data from studies adopting standard panel designs to test theory predictions may not sufficiently capture intraindividual change in theory constructs over time. Specifically, standard analytic models, such as those used here, account only for temporal stability constructs, that is, the extent to which they vary or, rather, do not vary or remain stable over time, but do not account for intraindividual change, or lack thereof as might be assumed if the constructs were somewhat "trait like" in their nature. This is an unrealistic assumption for psychological constructs as most are expected to exhibit at least some degree of intraindividual stability. Given that standard analytic models used with panel data do not account for these within-person stabilities, the estimates of cross-lagged effects they yield are likely to be imprecise (Hamaker et al., 2015). Alternatives have been offered, such as the inclusion of random intercepts in analyses of panel designs (Hamaker et al., 2015; Luo et al., 2022; Orth et al., 2021). Such analytic designs seek to overcome the limitations of the estimates of cross-lagged effects provided in standard panel designs by introducing a random intercept, which effectively models the intraindividual change independent of the change modelled by the covariance stability in the panel design, akin to a multi-level approach which segregates at within- and between-person levels. Applying this approach to longitudinal models of the theory would be advantageous as it would illustrate the extent to which change, or lack thereof, in constructs such as attitudes and intentions over time could be attributed to intraindividual change and temporal stability. However, such an analytic approach was not possible in the current review because we did not have access to the original data sets. Furthermore, to date, we know of no study that has applied the random-intercepts analytic approach to longitudinal data using the theory of planned behaviour. We therefore advocate a comparison of effects from the model applied to the data from the current meta-analysis with theory of planned behaviour data from large-sample panel designs that adopt a random-intercepts analytic approach to shed light on the extent to which accounting for intraindividual change may affect estimates.

Finally, the designs of the included studies, and our analytic approach to testing the proposed longitudinal models in the current study, focused on modelling covariance stability in theory constructs, intentions and behaviour over time. While this provides valuable insight into ecologically valid change in constructs over time, it only focuses on one aspect of change. Other study methods may provide further insight into how these constructs change and shed light on the processes involved. For example, a synthesis of experimental and intervention research that targets change in all theory constructs over time across multiple follow-up occasions may provide further insight into how change in theory constructs impacts other constructs in the theory over time. Although meta-analyses have reported effects of manipulations of theory constructs on follow-up measures of behaviour (e.g., Sheeran et al., 2016; Steinmetz et al., 2016), few studies adopting these kinds of design report full study measures over time and, in particular, effects of the intervention on other theory constructs or their mediating effect on study outcomes. We look to future studies to adopt designs that manipulate theory constructs and simultaneously include full measures of study constructs on multiple occasions to provide further insight into effects of construct change by extraneous techniques within the theory over time.

Conclusion

While prior studies testing the theory of planned behaviour, including meta-analyses, have provided generalised, broad support for its predictions, the vast majority of studies have adopted the received prospective design (e.g., Albarracín et al., 2001; Armitage & Conner, 2001b; Hagger et al., 2002; McEachan et al., 2011). We aimed to extend this knowledge by conducting a meta-analytic synthesis of research applying longitudinal designs to test theory predictions, that is, studies measuring social cognition constructs, intentions and behaviour on an initial occasion and on one or more follow-up occasions. Our analyses afforded tests of several key auxiliary hypotheses of the theory in keeping with, or extrapolating from, those promulgated in its original form. Specifically, we addressed the following hypotheses: the stability of theory constructs and stationarity in its effects over time; effects of initially measured theory constructs on subsequent intentions over time and effects of past behaviour while accounting for stability; cross-lagged effects among theory constructs and the effects of key moderators of study effects in the longitudinal model. Meta-analytic structural equation models of longitudinal and panel design models provided robust support for the stability of theory effects over time, and the pattern of theory effects within and across time points, but did not provide strong support for stationarity in theory effects over time. We also found little evidence for cross-lagged effects among theory constructs, relations between attitudes and perceived behavioural control, and between subjective norms and intentions, excepted. Findings provide qualified support for some of the auxiliary hypotheses, but effect sizes were small, and some predicted effects did not hold, such as the directional effect of attitudes on intentions over time in our panel model. Moderator analyses revealed notable theory-consistent differences in theory effects over time according to levels of the health behaviour type and specific behaviours. For example, subjective norm–intention relations over time were larger in studies targeting health protection behaviours, while attitude–intention relations were larger in studies targeting health risk behaviours. Finally, theory effects did not vary according to lag in measurement of theory constructs and behaviour measures.

Our research has several important ramifications for the development of the theory and research on the social cognition correlates of goal-directed behaviour more broadly. Our findings indicate that the theory is minimally sufficient to explain variance in intentions and behaviour over time and does so once temporal change in theory constructs and past behaviour effects over time have been accounted for. From a practical perspective, our findings may add further to the growing evidence base of potentially modifiable targets that could be targeted for interventions that change behaviour. From this perspective, our analysis provides indication of the extent to which theory constructs change temporally in the absence of manipulations or techniques aimed at affecting a change. This temporal change component is an additional consideration that should be taken into account when designing and evaluating interventions based on the theory that target behaviour change through construct change (Ajzen, 1991). It should also be stressed that much variance in intentions and behaviour remains unexplained by theory, and past behaviour, a variable that is not informative of psychological processes, still exhibits substantive effects on intentions and behaviour over time. Going forwards, future researchers should consider developing studies that extend and modify the theory by integrating additional constructs and associated processes from other theories to account for additional variance in intentions and behaviour over time.

Acknowledgements

We thank Daniel J. Phipps for his assistance with data collection and collation.

Author Declarations

The authors received no funding for this research.

The authors report there are no competing interests to declare.

Data availability statement

Data files, data analysis scripts and output are available online: https://osf.io/xfjq7/

Disclosure statement

No potential conflict of interest was reported by the authors.

ORCID

Martin S. Hagger http://orcid.org/0000-0002-2685-1546

Kyra Hamilton http://orcid.org/0000-0001-9975-685X

References

Ajzen, I. (2002a, September 1). Constructing a TPB questionnaire: Conceptual and methodological considerations. Retrieved September 1, 2002, from http://people.umass.edu/~aizen/pdf/tpb.measurement.pdf

Ajzen, I. (1988). *Attitudes, personality, and behavior*. Dorsey Press.

Ajzen, I. (1991). The theory of planned behavior. *Organizational Behavior and Human Decision Processes*, *50*(2), 179–211. https://doi.org/10.1016/0749-59789190020-T.

Ajzen, I. (2002b). Residual effects of past on later behavior: Habituation and reasoned action perspectives. *Personality and Social Psychology Review*, *6*(2), 107–122. https://doi.org/10.1207/S15327957PSPR060202.

Ajzen, I. (2011). The theory of planned behaviour: Reactions and reflections. *Psychology & Health*, *26*(9), 1113–1127. https://doi.org/10.1080/08870446.2011.613995.

Ajzen, I., & Driver, B. E. (1991). Prediction of leisure participation from behavioral, normative, and control beliefs: An application of the theory of planned behavior. *Leisure Sciences*, *13*(3), 185–204. https://doi.org/10.1080/01490409109513137.

Ajzen, I., & Driver, B. L. (1992). Application of the theory of planned behavior to leisure choice. *Journal of Leisure Research*, *24*(3), 207–224. https://doi.org/10.1080/00222216.1992.11969889 .

Ajzen, I., & Madden, T. J. (1986). Prediction of goal directed behavior: Attitudes, intentions and perceived behavioral control. *Journal of Experimental Social Psychology*, *22*(5), 453–474. https://doi.org/10.1016/0022-1031(86)90045-4 .

Ajzen, I., & Schmidt, P. (2020). Changing behavior using the theory of planned behavior. In M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen, & T. Lintunen (Eds.), *The handbook of behavior change* (pp. 17–31). Cambridge University Press. https://doi.org/10.1017/97811086773180.002.

Albarracín, D., Johnson, B. T., Fishbein, M., & Muellerleile, P. A. (2001). Theories of reasoned action and planned behavior as models of condom use: A meta-analysis. *Psychological Bulletin*, *127*(1), 142–161. https://doi.org/10.1037/0033-2909.127.1.142 .

Armitage, C. J., & Conner, M. T. (1999). The theory of planned behavior: Assessment of predictive validity and ‘perceived control’. *British Journal of Social Psychology*, *38*(1), 35–54. https://doi.org/10.1348/014466699164022.

Armitage, C. J., & Conner, M. T. (2000). Attitudinal ambivalence: A test of three key hypotheses. *Personality and Social Psychology Bulletin*, *26*(11), 1421–1432. https://doi.org/10.1177/0146167200263009.

Armitage, C. J., & Conner, M. T. (2001a). Efficacy of the theory of planned behavior: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471–499. https://doi.org/10.1348/014466601164939/.

Armitage, C. J., & Conner, M. T. (2001b). Efficacy of the theory of planned behaviour: A meta-analytic review. *British Journal of Social Psychology*, *40*(4), 471–499. https://doi.org/10.1348/014466601164939.

Bagozzi, R. P. (1992). The self-regulation of attitudes, intentions and behavior. *Social Psychology Quarterly*, *55*(2), 178–204. https://doi.org/10.2307/2786945 .

Bassili, J. N. (1996). Meta-judgmental versus operative indices of psychological properties: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, *71*(4), 637–653. https://doi.org/10.1037/0022-3514.71.4.637.

Begg, C. B., & Mazumdar, M. (1994). Operating characteristics of a rank correlation test for publication bias. *Biometrics*, *50*(4), 1088–1101. https://doi.org/10.2307/2533446 .

Bem, D. J. (1972). Self-perception theory. In L. Berkowitz (Ed.), *Advances in Experimental Social Psychology* (Vol. 6, pp. 1–62). Academic Press. https://doi.org/10.1016/S0065-2601(08)60024-6.

Bentler, P. M., & Speckart, G. (1979). Models of attitude–behavior relations. *Psychological Review*, *86*(5), 452–464. https://doi.org/10.1037/0033-295X.86.5.452.

Bentler, P. M., & Speckart, G. (1981). Attitudes "cause" behaviors: A structural equation analysis. *Journal of Personality and Social Psychology*, *40*(2), 226–238. https://doi.org/10.1037/0022-3514.40.2.226.

Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Wiley. https://doi.org/10.1002/9780470743386.

Carter, E. C., Schonbrodt, F., Gervais, W., & Hilgard, J. (2019). Correcting for bias in psychology: A comparison of meta-analytic methods. *Advances in Methods and Practices in Psychological Science*, *2*(2), 115–144. https://doi.org/10.1177/2515245919847196.

Chatzisarantis, N. L. D., & Hagger, M. S. (2005). Effects of a brief intervention based on the theory of planned behavior on leisure time physical activity participation. *Journal of Sport and Exercise Psychology*, *27*(4), 470–487. https://doi.org/10.1123/jsep.27.4.470.

Chatzisarantis, N. L. D., Hagger, M. S., & Smith, B. (2007). Influences of perceived autonomy support on physical activity within the theory of planned behavior. *European Journal of Social Psychology*, *37*(5), 934–954. https://doi.org/10.1002/ejsp.407.

Cheung, M. W.-L. (2014). Modeling dependent effect sizes with three-level meta-analyses: A structural equation modeling approach. *Psychological Methods*, *19*(2), 211–229. https://doi.org/10.1037/a0032968.

Cheung, M. W.-L. (2015a). *Meta-analysis: A structural equation modeling approach*. Wiley.

Cheung, M. W.-L. (2015b). metaSEM: An R package for meta-analysis using structural equation modeling. *Frontiers in Psychology*, *5*, 1521. https://doi.org/10.3389/fpsyg.2014.01521.

Chevance, G., Caudroit, J., Romain, A. J., & Boiché, J. (2017). The adoption of physical activity and eating behaviors among persons with obesity and in the general population: The role of implicit attitudes within the theory of planned behavior. *Psychology, Health & Medicine*, *22*(3), 319–324. https://doi.org/10.1080/13548506.2016.1159705.

Coburn, K. M., & Vevea, J. L. (2019). Package 'weightr'. from https://vevealab.shinyapps.io/WeightFunctionModel/.

Cochran, W. G. (1952). The χ^2 test of goodness of fit. *Annals of Mathematical Statistics*, *23*(3), 315–345. https://doi.org/10.1214/aoms/1177729380.

Cohen, J. (1992). A power primer. *Psychological Bulletin*, *112*(1), 155–159. https://doi.org/10.1037/0033-2909.112.1.155.

Collins, L. M. (2006). Analysis of longitudinal data: The integration of theoretical model, temporal design, and statistical model. *Annual Review of Psychology*, *57*(1), 505–528. https://doi.org/10.1146/annurev.psych.57.102904.190146 .

Conner, M. T. (2016). Social cognitions in health behaviour. In Y. Benyamini, M. Johnston, & E. C. Karademas (Eds.), *Assessment in health psychology* (pp. 19–30). Hogrefe.

Conner, M. T., & Abraham C. (2001). Conscientiousness and the theory of planned behavior: Toward a more complete model of the antecedents of intentions and behavior. *Personality & Social Psychology Bulletin*, *27*(11), 1547–1561. 10.1177/01461672012711014

Conner, M. T., McEachan, R., Taylor, N., O'Hara, J., & Lawton, R. (2015). Role of affective attitudes and anticipated affective reactions in predicting health behaviors. *Health Psychology*, *34*(6), 642–652. https://doi.org/10.1037/hea0000143 .

Conner, M. T., & McMillan, B. (1999). Interaction effects in the theory of planned behaviour: Studying cannabis use. *British Journal of Social Psychology*, *38*(2), 195–222. https://doi.org/10.1348/014466699164121.

Conner, M. T., & Norman, P. (2015). *Predicting and changing health behaviour: Research and practice with social cognition models* (3rd ed.). Open University Press.

Conner, M. T., & Sparks, P. (2002). Ambivalence and attitudes. *European Review of Social Psychology*, *12*(1), 37–70. https://doi.org/10.1080/14792772143000012 .

Conner, M. T., & Sparks, P. (2015). The theory of planned behavior and reasoned action approach. In M. T. Conner & P. Norman (Eds.), *Predicting and changing health behaviour: Research and practice with social cognition models* (3rd ed., pp. 142–188). Open University Press.

Cooke, R., Dahdah, M., Norman, P., & French, D. P. (2016). How well does the theory of planned behaviour predict alcohol consumption? A systematic review and meta-analysis. *Health Psychology Review*, *10*(2), 148–167. https://doi.org/10.1080/17437199.2014.947547.

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. https://doi.org/10.1348/0144666041501688.

DeCharms, R., & Muir, M. S. (1978). Motivation: Social approaches. *Annual Review of Psychology*, *29*(1), 91–113. https://doi.org/10.1146/annurev.ps.29.020178.000515.

Del Re, A. C., & Hoyt, W. T. (2018). Package 'MAc': Meta-analysis with correlations. Retrieved November 1, 2018, from https://cran.r-project.org/web/packages/MAc/MAc.pdf .

Doll, J., & Ajzen, I. (1992). Accessibility and stability of predictors in the theory of planned behavior. *Journal of Personality and Social Psychology*, *63*(5), 754–765. https://doi.org/10.1037/0022-3514.63.5.754.

Egger, M., Smith, D. G., Schneider, M., & Minder, C. (1997). Bias in meta-analysis detected by a simple, graphical test. *BMJ*, *315*(7109), 629–634. https://doi.org/10.1136/bmj.315.7109.629.

Eggers, S. M., Taylor, M., Sathiparsad, R., Bos, A. E., & de Vries, H. (2015). Predicting safe sex: Assessment of autoregressive and cross-lagged effects within the theory of planned behavior. *Journal of Health Psychology*, *20*(11), 1397–1404. https://doi.org/10.1177/1359105313512354.

Finkel, S. E. (1995). *Causal analysis with panel data*. SAGE Publications, Inc. https://doi.org/10.4135/9781412983594.

Geen, R. G. (1995). *Human motivation: A social psychological approach*. Thomson Brooks/Cole Publishing Co.

Godin, G., Sheeran, P., Conner, M., Belanger-Gravel, A., Cecilia, M., Gallani, B. J., & Nolin, B. (2010). Social structure, social cognition, and physical activity: A test of four models. *British Journal of Health Psychology*, *15*(1), 79–95. https://doi.org/10.1348/135910709x429901.

Gollob, H. F., & Reichardt, C. S. (1987). Taking account of time lags in causal models. *Child Development*, *58*(1), 80–92. https://doi.org/10.2307/1130293 .

Guadagno, R. E., & Cialdini, R. B. (2010). Preference for consistency and social influence: A review of current research findings. *Social Influence*, *5*(3), 152–163. https://doi.org/10.1080/15534510903332378.

Gwet, K. L. (2008). Computing inter-rater reliability and its variance in the presence of high agreement. *British Journal of Mathematical and Statistical Psychology*, *61*(1), 29–48. https://doi.org/10.1348/000711006X126600.

Hagger, M. S. (2019). The reasoned action approach and the theories of reasoned action and planned behavior. In D. S. Dunn (Ed.), *Oxford bibliographies in psychology*. Oxford University Press. https://doi.org/10.1093/OBO/9780199828340-0240 .

Hagger, M. S., Chan, D. K. C., Protogerou, C., & Chatzisarantis, N. L. D. (2016). Using meta-analytic path analysis to test theoretical predictions in health behavior: An illustration based on meta-analyses of the theory of planned behavior. *Preventive Medicine*, *89*, 154–161. https://doi.org/10.1016/j.ypmed.2016.05.020 .

Hagger, M. S., Chatzisarantis, N. L. D., & Biddle, S. J. H. (2001). The influence of self-efficacy and past behaviour on the physical activity intentions of young people. *Journal of Sports Sciences*, *19*(9), 711–725. https://doi.org/10.1080/02640410152475847.

Hagger, M. S., Chatzisarantis, N. L. D., & Biddle, S. J. H. (2002). A meta-analytic review of the theories of reasoned action and planned behavior in physical activity: Predictive validity and the contribution of additional variables. *Journal of Sport and Exercise Psychology*, *24*(1), 3–32. https://doi.org/10.1123/jsep.24.1.3.

Hagger, M. S., Chatzisarantis, N. L. D., Biddle, S. J. H., & Orbell, S. (2001). Antecedents of children's physical activity intentions and behaviour: Predictive validity and longitudinal effects. *Psychology and Health*, *16*(4), 391–407. https://doi.org/10.1080/08870440108405515 .

Hagger, M. S., Cheung, M. W. L., Ajzen, I., & Hamilton, K. (2022). Perceived behavioral control moderating effects in the theory of planned behavior: A meta-analysis. *Health Psychology*, *41*(2), 155–167. https://doi.org/10.1037/hea0001153 .

Hagger, M. S., Gucciardi, D. F., & Chatzisarantis, N. L. D. (2017). On nomological validity and auxiliary assumptions: The importance of simultaneously testing effects in social cognitive theories applied to health behavior and some guidelines. *Frontiers in Psychology*, *8*, 1933. https://doi.org/10.3389/fpsyg.2017.01933 .

Hagger, M. S., Gucciardi, D. F., Turrell, A., & Hamilton, K. (2019). Self-control and health-related behavior: The role of implicit self-control, trait self-control, and lay beliefs in self-control. *British Journal of Health Psychology*, *24*(4), 764–786. https://doi.org/10.1111/bjhp.12378.

Hagger, M. S., Hamilton, K., Phipps, D. J., Protogerou, C., Zhang, C.-Q., Girelli, L., Mallia, L., & Lucidi, F. (2023). Effects of habit and intention on behavior:

Meta-analysis and test of key moderators. *Motivation Science*, *9*(2), 73–94. https://doi.org/10.1037/mot0000294.

Hagger, M. S., Polet, J., & Lintunen, T. (2018). The reasoned action approach applied to health behavior: Role of past behavior and test of some key moderators using meta-analytic structural equation modeling. *Social Science & Medicine*, *213*, 85–94. https://doi.org/10.1016/j.socscimed.2018.07.038.

Hagger, M. S., Trost, N., Keech, J., Chan, D. K. C., & Hamilton, K. (2017). Predicting sugar consumption: Application of an integrated dual-process, dual-phase model. *Appetite*, *116*, 147–156. https://doi.org/10.1016/j.appet.2017.04.032.

Hamaker, E. L., Kuiper, R. M., & Grasman, R. P. P. P. (2015). A critique of the cross-lagged panel model. *Psychological Methods*, *20*(1), 102–116. https://doi.org/10.1037/a0038889.

Hamilton, K., Gibbs, I., Keech, J. J., & Hagger, M. S. (2020). Reasoned and implicit processes in heavy episodic drinking: An integrated dual process model. *British Journal of Health Psychology*, *25*(1), 189–209. https://doi.org/10.1111/BJHP.12401.

Hamilton, K., & Johnson, B. T. (2020). Attitude and persuasive communication interventions. In M. S. Hagger, L. D. Cameron, K. Hamilton, N. Hankonen, & T. Lintunen (Eds.), *The handbook of behavior change* (pp. 445–460). Cambridge University Press. https://doi.org/10.1017/97811086773180.031.

Hamilton, K., Kirkpatrick, A., Rebar, A., & Hagger, M. S. (2017). Child sun safety: Application of an integrated behavior change model. *Health Psychology*, *36*(9), 916–926. https://doi.org/10.1037/hea0000533 .

Hamilton, K., Phipps, D. J., Loxton, N., Modecki, K. L., & Hagger, M. S. (2023). Reciprocal relations between past behavior, implicit beliefs, and habits: A cross-lagged panel design. *Journal of Health Psychology*, 13591053231164492. https://doi.org/10.1177/13591053231164492.

Hamilton, K., van Dongen, A., & Hagger, M. S. (2020). An extended theory of planned behavior for parent-for-child health behaviors: A meta-analysis. *Health Psychology*, *39*(10), 863–878. https://doi.org/10.1037/hea0000940.

Hanbury, A., Wallace, L., & Clark, M. (2009). Use of a time series design to test effectiveness of a theory-based intervention targeting adherence of health professionals to a clinical guideline. *British Journal of Health Psychology*, *14*(3), 505–518. https://doi.org/10.1348/135910708X369558 .

Harrer, M., Cuijpers, P., Furukawa, T. A., & Ebert, D. D. (2019). *Doing meta-analysis in R: A hands-on guide*. https://doi.org/10.5281/zenodo.2551803.

Hedges, L. V. (1984). Estimation of effect size under nonrandom sampling: The effects of censoring studies yielding statistically insignificant mean differences. *Journal of Educational and Behavioral Statistics*, *9*(1), 61–85. https://doi.org/10.3102/10769986009001061.

Henrich, J., Heine, S. J., & Norenzayan, A. (2010). The weirdest people in the world? *Behavioral and Brain Sciences*, *33*(2–3), 61–83. https://doi.org/10.1017/S0140525X0999152X.

Hertzog, C., & Nesselroade, J. R. (1987). Beyond autoregressive models: Some implications of the trait-state distinction for the structural modeling of developmental change. *Child Development*, *58*(1), 93–109. https://doi.org/10.2307/1130294 .

Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine*, *21*(11), 1539–1558. https://doi.org/10.1002/sim.1186.

Hobbis, I., & Sutton, S. (2005). Are techniques used in cognitive behaviour therapy applicable to behaviour change interventions based on the theory of planned behaviour? *Journal of Health Psychology*, *10*(1), 7–18. https://doi.org/10.1177/1359105305048549.

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling*, *6*(1), 1–55. https://doi.org/10.1080/10705519909540118 .

Hunter, J. E., & Schmidt, F. L. (2015). *Methods of meta-analysis: Correcting error and bias in research findings* (3rd ed.). Sage. https://doi.org/10.4135/9781483398105.

Iyengar, S., & Greenhouse, J. B. (1988). Selection models and the file drawer problem. *Statistical Science*, *3*(1), 109–117. https://doi.org/10.1214/ss/1177013012.

Joensuu, S., Viljamaa, A., Varamäki, E., Tornikoski, E., & Harry Matlay, P. (2013). Development of entrepreneurial intention in higher education and the effect of gender – a latent growth curve analysis. *Education + Training*, *55*(8/9), 781–803. https://doi.org/10.1108/ET-06-2013-0084.

Keatley, D. A., Clarke, D. D., & Hagger, M. S. (2012). Investigating the predictive validity of implicit and explicit measures of motivation on condom use, physical activity, and healthy eating. *Psychology and Health*, *27*(5), 550–569. https://doi.org/10.1080/08870446.2011.605451 .

Kroesen, M., Handy, S., & Chorus, C. (2017). Do attitudes cause behavior or vice versa? An alternative conceptualization of the attitude-behavior relationship in travel behavior modeling. *Transportation Research Part A: Policy & Practice*, *101*, 190–202. https://doi.org/10.1016/j.tra.2017.05.013.

Lawton, R., Conner, M. T., & McEachan, R. (2009). Desire or reason: Predicting health behaviors from affective and cognitive attitudes. *Health Psychology*, *28*(1), 56–65. https://doi.org/10.1037/a0013424.

Liska, A. E. (1984). A critical examination of the causal structure of the Fishbein/Ajzen attitude-behavior model. *Social Psychology Quarterly*, *47*(1), 61–74. https://doi.org/10.2307/3033889.

Liska, A. E., Felson, R. B., Chamlin, M., & Baccaglini, W. (1984). Estimating attitude-behavior reciprocal effects within a theoretical specification. *Social Psychology Quarterly*, *47*(1), 15–23. https://doi.org/10.2307/3033884 .

Luo, J., Zhang, B., Estabrook, R., Graham, E. K., Driver, C. C., Schalet, B. D., Turiano, N. A., Spiro Iii, A., & Mroczek, D. K. (2022). Personality and health: Disentangling their between-person and within-person relationship in three longitudinal studies. *Journal of Personality and Social Psychology*, *122*(3), 493–522. https://doi.org/10.1037/pspp0000399.

Marsh, H. W., Papaioannou, A., & Theodorakis, Y. (2006). Causal ordering of physical self-concept and exercise behavior: Reciprocal effects model and the influence of physical education teachers. *Health Psychology*, *25*(3), 316–328. https://doi.org/10.1037/0278-6133.25.3.316.

McAnally, K., & Hagger,M. S. (2023). Health literacy, social cognition constructs, and health behaviors and outcomes: A meta-analysis. *Health Psychology*, *42*(4), 213–234. 10.1037/hea0001266

McAuley, E. (1992). The role of efficacy cognitions in the prediction of exercise behavior in middle-aged adults. *Journal of Behavioral Medicine*, *15*(1), 65–88. https://doi.org/10.1007/bf00848378 .

McEachan, R. R. C., Conner, M. T., Taylor, N., & Lawton, R. J. (2011). Prospective prediction of health-related behaviors with the theory of planned behavior: A meta-analysis. *Health Psychology Review*, *5*(2), 97–144. https://doi.org/10.1080/17437199.2010.521684 .

McEachan, R. R. C., Taylor, N., Harrison, R., Lawton, R., Gardner, P., & Conner, M. T. (2016). Meta-analysis of the reasoned action approach (RAA) to understanding health behaviors. *Annals of Behavioral Medicine*, *50*(4), 592–612. https://doi.org/10.1007/s12160-016-9798-4 .

McMillan, B., & Conner, M. (2007). Health cognition assessment. In A. B. S. Ayers, C. McManus, S. Newman, K. Wallston, J. Weinman, & R. West (Eds.), *Cambridge handbook of psychology, health and medicine* (2nd ed., pp. 260–266). Cambridge University Press.

McShane, B. B., Böckenholt, U., & Hansen, K. T. (2016). Adjusting for publication bias in meta-analysis. *Perspectives on Psychological Science*, *11*(5), 730–749. https://doi.org/10.1177/1745691616662243.

Niepel, C., Burrus, J., Greiff, S., Lipnevich, A. A., Brenneman, M. W., & Roberts, R. D. (2018). Students' beliefs and attitudes toward mathematics across time: A longitudinal examination of the theory of planned behavior. *Learning and Individual Differences*, *63*, 24–33. https://doi.org/10.1016/j.lindif.2018.02.010.

Norman, P., Cameron, D., Epton, T., Webb, T. L., Harris, P. R., Millings, A., & Sheeran, P. (2018). A randomized controlled trial of a brief online intervention to reduce alcohol consumption in new university students: Combining self-affirmation, theory of planned behaviour messages, and implementation intentions. *British Journal of Health Psychology*, *23*(1), 108–127. https://doi.org/10.1111/bjhp.12277.

Ogden, J. (2003). Some problems with social cognition models: A pragmatic and conceptual basis. *Health Psychology*, *22*(4), 424–428. https://doi.org/10.1037/0278-6133.22.4.424.

Orth, U., Clark, D. A., Donnellan, M. B., & Robins, R. W. (2021). Testing prospective effects in longitudinal research: Comparing seven competing cross-lagged models. *Journal of Personality and Social Psychology*, *120*(4), 1013–1034. https://doi.org/10.1037/pspp0000358 .

Phipps, D. J., Hagger, M. S., & Hamilton, K. (2020). Predicting limiting 'free sugar' consumption using an integrated model of health behavior. *Appetite*, *150*, 104668. https://doi.org/10.1016/j.appet.2020.104668.

Phipps, D. J., Hannan, T. E., Rhodes, R. E., & Hamilton, K. (2021). A dual-process model of affective and instrumental attitudes in predicting physical activity. *Psychology of Sport and Exercise*, *54*, 101899. https://doi.org/10.1016/j.psychsport.2021.101899 .

Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, *88*(5), 879–903. https://doi.org/10.1037/0021-9010.88.5.879 .

Protogerou, C., & Hagger, M. S. (2020). A checklist to assess the quality of survey studies in psychology methods in psychology. *Methods in Psychology*, *3*, 100031. https://doi.org/10.1016/j.metip.2020.100031 .

Protogerou, C., Johnson, B. T., & Hagger, M. S. (2018). An integrated model of condom use in sub-Saharan African youth: A meta-analysis. *Health Psychology*, *37*(6), 586–602. https://doi.org/10.1037/hea0000604.

Randolph, W. A. (1981). Cross-lagged correlational analysis in dynamic settings. *Journal of Applied Psychology*, *66*(4), 431–436. https://doi.org/10.1037/0021-9010.66.4.431 .

Reinecke, J., Schmidt, P., & Ajzen, I. (1996). Application of the theory of planned behavior to adolescents' condom use: A panel study. *Journal of Applied Social Psychology*, *26*(9), 749–772. https://doi.org/10.1111/j.1559-1816.1996.tb01128.x .
Rich, A., Brandes, K., Mullan, B. A., & Hagger, M. S. (2015). Theory of planned behavior and adherence in chronic illness: A meta-analysis. *Journal of Behavioral Medicine*, *38*(4), 673–688. https://doi.org/10.1007/s10865-015-9644-3 .
Rivis, A., Sheeran, P., & Armitage, C. J. (2009). Expanding the affective and normative components of the theory of planned behavior: A meta-analysis of anticipated affect and moral norms. *Journal of Applied Social Psychology*, *39*(12), 2985–3019. https://doi.org/10.1111/j.1559-1816.2009.00558.x.
Rogosa, D. (1980). A critique of cross-lagged correlation. *Psychological Bulletin*, *88* (2), 245–258. https://doi.org/10.1037/0033-2909.88.2.245.
Rosenthal, M. C. (1994). The fugitive literature. In H. Cooper & L. V. Hedges (Eds.), *The handbook of research synthesis* (pp. 85–94). Russell Sage Foundation.
Schenker, N., & Gentleman, J. F. (2001). On judging the significance of differences by examining the overlap between confidence intervals. *The American Statistician*, *55* (3), 182–186. https://doi.org/10.1198/000313001317097960.
Schifter, D. E., & Ajzen, I. (1985). Intention, perceived control, and weight loss: An application of the theory of planned behavior. *Journal of Personality and Social Psychology*, *49*(3), 843–851. https://doi.org/10.1037/0022-3514.49.3.843.
Sheeran, P., & Abraham, C. (2003). Mediator of moderators: Temporal stability of intention and the intention-behavior relation. *Personality and Social Psychology Bulletin*, *29*(2), 205–215. https://doi.org/10.1177/0146167202239046.
Sheeran, P., Maki, A., Montanaro, E., Avishai-Yitshak, A., Bryan, A., Klein, W. M. P., Miles, E., & Rothman, A. J. (2016). The impact of changing attitudes, norms, and self-efficacy on health-related intentions and behavior: A meta-analysis. *Health Psychology*, *35*(11), 1178–1188. https://doi.org/10.1037/hea0000387.
Sheeran, P., & Taylor, S. (1999). Predicting intentions to use condoms: A meta-analysis and comparison of the theories of reasoned action and planned behavior. *Journal of Applied Social Psychology*, *29*(8), 1624–1675. https://doi.org/10.1111/j.1559-1816.1999.tb02045.x.
Sherman, S. J., Judd, C. M., & Park, B. (1989). Social cognition. *Annual Review of Psychology*, *40*(1), 281–326. https://doi.org/10.1146/annurev.ps.40.020189.001433.
Simonsohn, U., Nelson, L. D., & Simmons, J. P. (2014). P-curve and effect size: Correcting for publication bias using only significant results. *Perspectives on Psychological Science*, *9*(6), 666–681. https://doi.org/10.1177/1745691614553988.
Sniehotta, F. F. (2009). An experimental test of the Theory of Planned Behavior. *Applied Psychology: Health and Well-Being*, *1*(2), 257–270. https://doi.org/10.1111/j.1758-0854.2009.01013.x .
Sniehotta, F. F., Presseau, J., & Araújo-Soares, V. (2014). Time to retire the theory of planned behaviour. *Health Psychology Review*, *8*(1), 1–7. https://doi.org/10.1080/17437199.2013.869710 .
Stanley, T. D., & Doucouliagos, H. (2014). Meta-regression approximations to reduce publication selection bias. *Research Synthesis Methods*, *5*(1), 60–78. https://doi.org/10.1002/jrsm.1095.
Steiger, J. H. (2007). Understanding the limitations of global fit assessment in structural equation modeling. *Personality and Individual Differences*, *42*(5), 893–898. https://doi.org/10.1016/j.paid.2006.09.017 .
Steinmetz, H., Knappstein, M., Ajzen, I., Schmidt, P., & Kabst, R. (2016). How effective are behavior change interventions based on the Theory of planned

behavior? A three-level meta-analysis. *Zeitschrift Fur Psychologie-Journal of Psychology*, *224*(3), 216–233. https://doi.org/10.1027/2151-2604/a000255.

St Quinton, T., & Trafimow, D. (2022). The unappreciated relevance of auxiliary assumptions for evaluating theory-based interventions in health psychology. *Theory & Psychology*, *32*(6), 915–930. https://doi.org/10.1177/09593543221113263.

Trafimow, D. (2009). The theory of reasoned action: A case study of falsification in psychology. *Theory & Psychology*, *19*(4), 501–518. https://doi.org/10.1177/0959354309336319.

Trafimow, D. (2012). The role of auxiliary assumptions for the validity of manipulations and measures. *Theory & Psychology*, *22*(4), 486–498. https://doi.org/10.1177/0959354311429996.

Trafimow, D., & Duran, A. (1998). Some tests of the distinction between attitude and perceived behavioural control. *British Journal of Social Psychology*, *37*(1), 1–14. https://doi.org/10.1111/j.2044-8309.1998.tb01154.x .

Triandis, H. C. (1977). *Interpersonal behavior*. Brookes/Cole.

van Aert, R. C. M. (2020). Package 'puniform'. https://github.com/RobbievanAert/puniform

van Aert, R. C. M., & van Assen, M. A. L. M. (2018). Correcting for publication bias in a meta-analysis with the P-uniform* method. *MetaArXiv*. https://doi.org/10.31222/osf.io/zqjr9.

Vevea, J. L., & Hedges, L. V. (1995). A general linear model for estimating effect size in the presence of publication bias. *Psychometrika*, *60*(3), 419–435. https://doi.org/10.1007/BF02294384.

Viechtbauer, W. (2010). Conducting meta-analyses in R with the metafor package. *Journal of Statistical Software*, *36*(3), 1–48. https://doi.org/10.18637/jss.v036.i03.

Wanberg, C. R., Glomb, T. M., Song, Z., & Sorenson, S. (2005). Job-search persistence during unemployment: A 10-wave longitudinal study. *Journal of Applied Psychology*, *90*(3), 411–430. https://doi.org/10.1037/0021-9010.90.3.411.

Weiner, B. (1990). *Human motivation*. Erlbaum.

Wilson, S. J., Polanin, J. R., & Lipsey, M. W. (2016). Fitting meta-analytic structural equation models with complex datasets. *Research Synthesis Methods*, *7*(2), 121–139. https://doi.org/10.1002/jrsm.1199.