Consistency Tendency and The Theory of Planned Behavior: A Randomized Controlled Crossover Trial in a Physical Activity Context

Derwin K. C. Chan[1,2,3], Andreas Stenling[4], Cleoputri Yusainy[5], Ziadatul Hikmiah[5], Andreas Ivarsson[6], Martin S. Hagger[7,8], Ryan E. Rhodes[9], Mark R. Beauchamp[10]

Faculty of Education and Human Development, The Education University of Hong Kong, Hong Kong, China[1]
School of Public Health, The University of Hong Kong, Hong Kong, China[2]
School of Psychology, Curtin University, Australia[3]
Department of Psychology, Umeå University, Umeå, Sweden[4]
Psychology Department, Faculty of Social and Political Sciences, Brawijaya University, Malang, Indonesia[5]
Center for research on Welfare, Health and Sport, Halmstad University, Halmstad, Sweden[6]
SHARPP Lab, Psychological Sciences, University of California, Merced, USA[7]
Faculty of Sport and Health Sciences, University of University of Jyväskylä[8]
Behavioural Medicine Laboratory, School of Exercise Science, Physical and Health Education, University of Victoria, Victoria, BC, Canada[9]
School of Kinesiology, University of British Columbia, Vancouver, Canada[10]

Correspondence concerning this article should be addressed to Dr Derwin Chan, Department of Early Childhood, Faculty of Education and Human Development, The Education University of Hong Kong, 10 Lo Ping Road, Tai Po, New Territories, Hong Kong. Email: derwin@eduhk.hk

Acknowledgments

This project is supported by (1) Seed Funding of Basic Research of The University of Hong Kong [104004966.092935.22400.301.01] awarded to Dr Derwin Chan and Prof Ryan Rhodes, and (2) World Class University Funding Program for International Research Publication of Brawijaya University, Indonesia awarded to Dr. Cleoputri Yusainy. The authors would like to thank Hanifah Ramadhani Sembiring, Atika Nurhidayah, Calvin Octavianus Anggono, Gitajiwa Sekarnusa Ganies, and Chairun Nissa Huwaida for their help in data collection.

Citation of this paper (please check for any updates)

Chan, D. K. C., Stenling, A., Yusainy, C., Hikmiah, Z., Ivarsson, A., Hagger, M. S., Rhodes, R. E., Beauchamp, M. R. (advanced online publication). Consistency tendency and the theory of planned behavior: A randomized controlled crossover trial in a physical activity context. *Psychology & Health*. doi: 10.1080/08870446.2019.1677904

Abstract

Objective: This study examined the effects of consistency tendency on the predictive power of the theory of planned behavior (TPB) in relation to physical activity behavior.

Methods: In this randomized controlled cross-over trial, we recruited 770 undergraduate students from Indonesia who were randomly assigned into two groups. Participants completed physical activity versions of TPB measures at T1 (baseline) and T2 (post 1 week), and the International Physical Activity Questionnaire at T3 (post 1 month). At T1 and T2, the TPB questions were either presented in ensemble-order (i.e., consistency tendency supressed) or alternate-order (i.e., consistency tendency facilitated).

Results: The parameter estimates of the model (CFI > .92, TLI > .90, SRMR < .08, RMSEA < .08) aligned with the tenets of TPB. As compared to ensemble-order, a TPB measured in alternate-order yielded stronger cross-sectional relationships, but this pattern did not appear in the prospective relationships in TPB (i.e., intention/ perceived behavioral control and behavior).

Conclusions: Consistency tendency inflated the factor correlations of cross-sectionally measured TPB variables, but the inflation was not observed in the prospective prediction of behavior. Health psychology questionnaires with items presented in ensemble order may represent a viable means of reducing the confounding effect of consistency tendency.

Keywords: Consistency motif; proximity effect; socratic effect; common method variance; response bias; general response tendency.

Consistency Tendency and The Theory of Planned Behavior: A Randomized Controlled Crossover Trial in Physical Activity Context

The theory of planned behavior (TPB) is one of the most widely-used health psychology frameworks in the explanation of individuals' behaviors in physical activity (Hagger, Chatzisarantis, & Biddle, 2002; Hausenblas, Carron, & Mack, 1997) as well as other health settings (Godin & Kok, 1996; McEachan, Conner, Taylor, & Lawton, 2011). The measurement of TPB constructs has predominantly relied on self-reported questionnaires in which the findings could be confounded to some extent by response bias, common method effects, and other methodological artifacts (Budd, 1987; Budd & Spencer, 1986; Courneya, Conner, & Rhodes, 2006; Rhodes, Matheson, & Blanchard, 2006; Rhodes, Matheson, Blanchard, & Blacklock, 2008; Rhodes, Matheson, & Mark, 2010). In particular, *common method bias* is an inherently unobservable phenomenon that captures the variance attributed to measurement method rather than the variance of the theoretical construct (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003; Podsakoff, MacKenzie, & Podsakoff, 2012; Richardson, Simmering, & Sturman, 2009). It is, therefore, undesirable if common method variance induces systematic errors that not only confound the validity of measurement of TPB variables, but also the explanatory power of the theory in relation to behavior change (Budd, 1987; Budd & Spencer, 1986; Courneya et al., 2006; Rhodes et al., 2006; Rhodes et al., 2010). In this study we sought to examine the effects of a common method bias, namely, consistency tendency (Chan, Ivarsson, et al., 2015), on the parameter estimates of the TPB in the prediction of physical activity.

Theoretical Explanation of Consistency Tendency

Consistency tendency (also known as Socratic effect) is the tendency that consecutive items in a questionnaire are answered in consistent way, and so it is speculated that this common method bias may induce artificial covariance between predictors and criterion variables (Salancik, 1984; Salancik & Pfeffer, 1977). Researchers have been aware of this method effect and have offered critical insights into how consistency tendency could confound research findings of questionnaire-based research (Chan, Ivarsson, et al., 2015; Cronbach, 1946; Doty & Glick, 1998; Podsakoff et al., 2012; Salancik, 1984). For decades there has been debate over the extent to which factor correlations are caused by the similarity or commonality of the form of measurement rather than the actual between-factor relationships (Cronbach, 1946; Doty & Glick, 1998; Lindell & Whitney, 2001; Warren & Halpern-Manners, 2012; Weijters, Geuens, & Schillewaert, 2009).

Cronbach (1946, 1950) first identified the tendency, known as *response sets*, of individuals to respond to items in a questionnaire according to the test format. The theoretical foundation of response sets is derived from the psychology literature that explains how individuals unconsciously establish frames or points of reference when making judgments or selections in a questionnaire, and the established referencing frames tend to be retained for subsequent assessments (Sherif & Cantril, 1945, 1946). According to response sets theory (Cronbach, 1946, 1950), response sets can cause a person to “give different responses to test items than he [or she] would when the same content is presented in a different form” (Cronbach, 1946, p. 146). Indeed, it may also make an individual offer similar responses to test items of different content “when similar situations are presented” (Cronbach, 1950, pp. 14-15), then “response set scores are significantly correlated” because of response sets rather than the true relationships between the factors. More recent literature refers to this specific type of response set as consistency tendency (Salancik, 1984; Salancik & Pfeffer, 1977), consistency motif (Podsakoff et al., 2003; Podsakoff & Organ, 1986), or Socratic effect (Henninger & Wyer, 1976; Wyer, 1974).

The origin of consistency tendency is believed to come from individuals’ awareness of their own responses to questions (Budd, 1987; Budd & Spencer, 1986; Salancik, 1984;

Salancik & Pfeffer, 1977). This awareness may lead to unconsciously answering consecutive questions consistently because responses themselves include salient information that could constrain participants' subsequent responses in a test battery (Salancik, 1984; Salancik & Pfeffer, 1977). It is somewhat consonant with the mere measurement effect evidenced in the recent psychology literature (Godin, Belanger-Gravel, Amireault, Vohl, & Perusse, 2011; Mankarious & Kothe, 2015; Morwitz & Fitzsimons, 2004) that explains how the act of asking questions about a person's intention can itself change behavior. The mere measurement effect has been attributed to the increased saliency and accessibility of the thoughts, memories, attitudes, and response choices associated with the behavior after responding to items concerning a person's intention (Mankarious & Kothe, 2015; Morwitz & Fitzsimons, 2004). Similarly, recent research on the longitudinal effect of panel conditioning suggests that measuring attributes (e.g., attitudes) could change participants' responses toward the attributes in follow-up measures (Sturgis, Allum, & Brunton-Smith, 2009; Warren & Halpern-Manners, 2012), so it might imply that the after-effect of item responses could operate at the survey level and affect participants' subsequent responses within the questionnaire (Duan, Alegria, Canino, McGuire, & Takeuchi, 2007). In general, the literature appears to support the proposition that participants' responses in the measure of attitude and behavior could be affected by their prior response pattern (Godin et al., 2011; Mankarious & Kothe, 2015; Morwitz & Fitzsimons, 2004; Sturgis et al., 2009; Warren & Halpern-Manners, 2012).

Consistency Tendency and Response Order

Researchers have proposed several methods for controlling the effects of consistency tendency, such as placing correlational markers, factor analysis, and counter-balancing the item-order (Lindell & Whitney, 2001; Podsakoff et al., 2003; Richardson et al., 2009). For instance, physical separation can reduce the consistency tendency by minimizing the saliency of contextual cues that might exist in the previous measurement, and more importantly, it impairs respondents' ability to use information or answers from previous questions to inform their responses (Podsakoff et al., 2003). Podsakoff and colleagues (2003) suggested that temporal or proximal separation could be done by intentionally placing a time-lag or physical gap between the measurements of different variables. In this case, response bias could be reduced because (1) short term memory is less likely to hold the information from previous responses; (2) the perceived relevance of previous items would be diminished; and (3) individuals might have impaired ability or motivation to use their prior responses for completing subsequent items.

In support of the proposition about the temporal or proximal effects of response sets, Hui and Triandis (1985) consistently found in three studies (i.e., with different instruments) that the inter-item correlations were stronger between adjacent items than the items that were further apart. In a similar regard, a qualitative investigation of the proximity effect model (Weijters et al., 2009) showed that when respondents answered two items that were close to each other, they tended to perceive them as being related or that the second item was redundant or repetitive, so they were likely to carry the same beliefs to answer the items. Such a behavioral pattern occurs because representation of the previous item in highly-accessible short-term memory gives the respondent the perception that the next question was similar or alike even if they were not identical. Similarly, the meta-analysis of common method variance by Doty and Glick (1998) found that time-lag was a significant source of common method bias that negatively correlated with the within-study factor correlations.

As a complement to this evidence, there have been attempts in previous research to examine the effect of item-ordering within questionnaires (Duan et al., 2007; Holbrook, Krosnick, Moore, & Tourangeau, 2007; Hui & Triandis, 1985; McClendon, 1991; Stone & Gueutal, 1984; Weijters et al., 2009). However, these studies did not test how the item-order moderated the strength of relationships between theoretically linked factors. The only exception was the work by Budd and colleagues (Budd, 1987; Budd & Spencer, 1986) concerning response bias in relation to the theory of reasoned action (Ajzen & Fishbein, 1980), which represented the preceding theory of the TPB. It was found that the inter-factor correlations of the theory of reasoned action (e.g., attitude, subjective norm, and intention) were stronger when the items were presented in a non-random-(item)order (i.e., an item of one factor was followed by an item of another factor) than that in a random-(item)order (Budd, 1987; Budd & Spencer, 1986). It was argued that when the items were presented in a non-random logical order, the format of the questionnaire artificially made respondents aware of the factor relationships, which further brought about a response consistency and inflated the inter-factor correlations reported with respect to the theory of reasoned action (Budd, 1987; Budd & Spencer, 1986). However, these studies were limited by having a quasi-experiment with two-group comparison design, in which the authors did not take inter-item distance into account within the questionnaire design and interpretation of study findings. Therefore, the findings of Budd and colleagues (Budd, 1987; Budd & Spencer, 1986) were unable to reveal if a consistency tendency was initiated by the inter-item distance of the questionnaire variables.

As far as we know, there has only been one recent investigation that formally and experimentally examined the effects of consistency tendency on factor correlations by manipulating the inter-item distance in a questionnaire (Chan, Ivarsson, et al., 2015). The study employed a two-wave two-arm randomized control cross-over design to examine if the parameter estimates in the motivational process model (i.e., autonomy support → autonomous motivation → intention) of self-determination theory (Deci & Ryan, 1985), were moderated by the inter-item distance. In line with the consistency tendency effect, the parameter estimates were stronger when items were arranged in an *alternate-order* than in *ensemble-*

order. Although the differences were not statistically significant, potentially due to the small sample size and data attrition, it was argued that the effect of consistency tendency in an *alternate-order* would be more salient (due to shortened inter-item distance between factors) when compared to an *ensemble-order* (due to lengthened inter-item distance between factors). Nevertheless, the study was limited because there were no inclusion of behavioral assessment and prospective follow-up, so it was unclear how consistency tendency could moderate the prospective relationship between intention and behavior. More importantly, the study only applied self-determination theory (Deci & Ryan, 1985) in the context of sport injury prevention, and so the effects of consistency tendency on the measures of other classic health psychology theories in relation to health behavior measures are currently unknown.

In sum, research over the years has attempted to examine response bias due to questionnaire format and structure (Podsakoff et al., 2003; Podsakoff et al., 2012; Richardson et al., 2009). Indeed, it has been increasingly recognized that one's propensity to display a response tendency could substantively (i.e., in non-trivial ways) interfere with the findings derived from health psychology measures (Chan, Ivarsson, et al., 2015). Previous studies have not formally examined 'method variance' by using inter-item distance to manipulate consistency tendency, and tested how this cognitive bias may affect the predictive power of the TPB (Ajzen, 1985).

Present Study

To address these research gaps, the present study examined the effects of consistency tendency on the pathways of the TPB in the physical activity context, using a 3-wave randomized cross-over trial with 1-month prospective follow-up of behavior. According to the tenets of the TPB (Ajzen, 1985) and its research findings in physical activity settings (Hagger et al., 2002; Hausenblas et al., 1997), the social cognitive variables of attitude, subjective norm, perceived behavioral control are positive predictors of intention which in turn predict future engagement of physical activity. Perceived behavioral control (PBC) is also positioned as a direct predictor of behavior (i.e., unmediated by intention). Based on previous literature about consistency tendency (Henninger & Wyer, 1976; Salancik, 1984; Salancik & Pfeffer, 1977; Wyer, 1974), and the findings related to inter-item separation (Hui & Triandis, 1985; Weijters et al., 2009), we hypothesized that positive pathways of TPB would be stronger in the questionnaire with alternate-order (consistency tendency facilitated) as compared to that in an ensemble-order (consistency tendency suppressed). We further hypothesized that the inflations would only apply to the cross-sectional relationships in this study but not the prospective relationships that were assessed in relation to the TPB.[1] A time-lag or temporal gap was expected to reduce common method variance (vis-à-vis the saliency of previous measurement) on subsequent assessments (Podsakoff et al., 2003).

Methods

Participants and Study Protocol

Upon the ethical approval from the Institutional Review Board of the first author's institution [UW 18-523] and local ethics committee of the third author, we recruited 770 undergraduate students of social and political science from Indonesia (mean age = 19.12, *SD* = 1.28, range = 16 to 30; 34.8% male; response rate = 81.99%). Based on the criteria of the International Physical Activity Questionnaire research committee (IPAQ; 2005), participant who had high, moderate, or low level of physical activity[2] were respectively 11.4%, 30.1%,

---

[1]In the current study, cross-sectional relations were assessed between attitude, subjective norm, PBC, and intention; whereas prospective relations were measured between PBC, intention, and physical activity.

[2]The criteria of high physical activity: (a) vigorous-intensity activity on at least 3 days achieving a minimum total physical activity of at least 1500 MET-minutes/week; or (b) 7 or more days of any combination of walking, moderate-intensity, or vigorous-intensity activities achieving a minimum total physical activity of at least 3000 MET-minutes/week. The criteria of moderate physical activity: (a) 3 or more days of vigorous-intensity activity of at least 20 minutes per day; or (b) 5 or more days of moderate-intensity activity and/or walking of at least 30 minutes per day; or (c) 5 or more days of any combination of walking, moderate-

and 58.4%. A large proportion (70.3%) participated in various leisure-time physical activity (e.g., jogging, running, swimming, soccer, badminton, volleyball, fitness training, cycling).

Before the start of the study, participants signed informed consent forms (i.e., to acknowledge their understanding of voluntary participation, rights of withdrawal, and confidentiality of data). They were randomly assigned into either Group 1 (n = 384) and Group 2 (n = 386) by computer balloting, and were blinded from the group allocation. Following a 3-wave cross-over design in this study, Group 1 completed the physical activity version of TPB measures in alternate-order and ensemble order, respectively at T1 (baseline) and T2 (post 1 week), and Group 2 did so in the opposite sequence (T1 = ensemble order, T2 = alternate-order). Both groups completed the follow-up questionnaire of physical activity at T3 (post 1 month). The questionnaire was translated into Indonesian with the translate-back-translate procedure of Hambleton (2005) in a team of two Indonesian-English bilingual speakers and a psychologist with extensive research experience in applying TPB cross-culturally.

Measures

TPB Variables. We used the physical activity version (Chan, Yang, et al., 2015) of standard TPB measures (Ajzen, 2002) to assess attitude, subjective norm, PBC, and intention of engaging in adequate physical activity. Clear definition of “adequate physical activity for adults” was based on the public health guidelines recommended by the World Health Organization (2010). This target was provided at the top of the TPB scale for instruction. To make the item-order changeable without affecting the presentation of the scale, we made sure the items could be read independently by presenting items in full form (i.e., without shared item stem), using a 7-point Likert scale for each item. The data derived from TPB measures

---

intensity, or vigorous intensity activities achieving a minimum total physical activity of at least 600 MET-minutes/week. Individuals who did not meet the criteria for moderate or high were considered to have low levels of physical activity.

displayed sound reliability, as reflected by Omega coefficients [ω] (Mcdonald, 1970) ranging from .728 to .919. Internal consistency values were in general higher when the items were presented in ensemble order ($M\omega = 0.85$) than alternate-order ($M\omega = 0.80$). See Appendix A for all the items and anchors used with the TPB measures.

Physical Activity. We used the short-form of the International Physical Activity Questionnaire (IPAQ; Craig et al., 2003) to assess the physical activity levels at T3. IT has been shown that the short-form of IPAQ produces reliable responses that are reflective to objectively measured physical activity (Ekelund et al., 2006). Participants reported their physical activity in terms of "how many days in a week" and "how much time per day" they spent on respectively vigorous and moderate intensity physical activity, and walking. We converted the scorings of IPAQ into a single indicator of metabolic equivalent of task (MET)-minutes/week according to the guideline of IPAQ research committee (2005).

Manipulation of Consistency Tendency

The primary manipulation of this study was item-order of the study variables presented in the questionnaire. Each item was displayed in a standardized format with consistent fonts in which the question was placed at the top, and a seven-point Likert scale and its anchors were independently presented below (see Figure 1 as an example). The order of presenting the questionnaire items followed one of two different sequences. In ensemble-order, all the items of one construct were presented sequentially before the next construct; whereas in alternate-order, we inter-mixed the items such that one item designed to measure a construct was followed by a separate item designed to assess another construct. As some of the TPB variables (i.e., intention and subjective norm) had fewer items, we placed two 7-point Likert scale items related respectively to the frequency and effort of "doing physical activity in the last 2 weeks" to make the inter-mixing of items more balanced across the TPB variables in the alternate-order questionnaire, and these two items were also presented at the end of the ensemble order questionnaire. The average inter-item distance (i.e., average distance between the items of distinct constructs) in ensemble-order questionnaire (7.34 item-units, *SD* = 1.77; *t*(30) = 1.00, $p < .001$, Cohen's *d* = .52) was significantly higher than that of the alternate-order questionnaire (i.e., 2.25 item-units, *SD* = 1.56).

INSERT FIGURE 1 ABOUT HERE

Statistical Analysis

We examined the effects of item-order between groups and within groups over time on the parameter estimates in the hypothesized model by employing structural equation modeling with Mplus version 8.1 (Muthén & Muthén, 1998-2017) using the robust maximum likelihood estimator (MLR). Missing data were handled by the full information maximum likelihood (FIML) estimator, which provided unbiased estimates under the assumption of the data missing at random (Enders, 2010). We relied on goodness-of-fit indices to evaluate model fit, given the known oversensitivity of the chi-square test of exact fit to sample size and minor model misspecifications (Marsh, Hau, & Grayson, 2005). Conventional fit indices were used to evaluate the model fit in the confirmatory factor analyses, such as the comparative fit index (CFI), Tucker-Lewis Index (TLI), the root mean square error of approximation (RMSEA), and standardized root mean square residual (SRMR). Traditional cut-off criteria with CFI and TLI values of .90 and RMSEA and SRMR values of .08 were used to indicate acceptable fit (Marsh, 2007). We computed the statistical power of each model we ran using MacCallum, Browne, and Sugawara (1996)'s algorithm. A statistical power of .80 or above was expected to support the adequacy of the sample size in each model (MacCallum et al., 1996).

Initially, we ran single-group structural equation models to examine goodness of fit of the measurement models at each time point within each group. To examine the effects of group assignment, invariance tests were conducted between Group 1 and Group 2 at T1 and

T2. To examine the effects of time, invariance tests were performed between T1 and T2 for each group. We initially built a baseline model with no constrains to examine configural invariance across groups and time, and then progressively constrained the parameter estimates to be equal between the groups or across time to test invariance of the measurement model. We followed the recommendations from Little (2013) and constrained the factor loadings and then indicator intercepts to examine metric invariance and scalar invariance, respectively. We adopted Chen's (2007) criteria of ΔCFI < .01 and ΔRMSEA < .015 as an indication of measurement invariance. Differences in the strength of the factor correlations between T1 and T2 within groups and between groups at T1 and at T2 were evaluated using the Wald test of parameter equalities (Buse, 1982). Between-group differences of predictors T1 and T2 in the predictive ability of physical activity at T3 was also evaluated using the Wald test of parameter equalities. The significance level was set to 0.05.

Results

The responses of the study variables were within the normal range, skewness and kurtosis values were between -1 and 1 for all variables except for physical activity that had a skewness value of 2.20 and a kurtosis value of 6.77. Of the initial 770 respondents, 694 responded at T2, and 421 provided physical activity data at T3. Descriptive statistics are provided in Table 1. As displayed in Table 2, the model fit indices indicted that the measurement models had an adequate fit to the data within each group at each time point. The results also showed that the model was invariant within groups over time and between groups at each time point, as indicated by ΔCFI and ΔRMSEA of less than .01 and .015, respectively, in models with equality constraints imposed on the factor loadings and intercepts over time or across groups. Statistical power of the models ranged from .95 to .99, indicating adequacy of the sample size required for the complexity of the model.

INSERT TABLE 1 AND 2 ABOUT HERE

Within-group differences in the magnitude of the factor correlations were examined for each group separately. A consistent pattern of larger cross-sectional factor correlations was observed in both groups when the items were presented in alternate-order compared to ensemble order. Between-group comparisons at T1 indicated a similar pattern when compared to the within-group comparisons (i.e., larger factor correlations in the group with alternate-order). Although a similar pattern was observed between groups at T2, the magnitude of the differences in the factor correlations were smaller. Finally, for prospective relationships, multi-group SEM showed that differences in the predictive ability of the TPB variables at T1 and T2 on physical activity at T3 were not statistically significant. Table 3 shows the full results of the Wald test of parameter equalities and Table 4 displays the parameter estimates of TPB in both groups and timepoints.[3]

INSERT TABLE 3 AND 3 ABOUT HERE

Discussion

This study was the first randomized controlled experiment that examined the effect of consistency tendency on the predictive power of the TPB in the physical activity domain. A consistent pattern of results in supportive of a consistency tendency effect was observed at both T1 and T2. Particularly, the cross-sectional relationships assessed within the context of the TPB were consistently inflated when the TPB measures were structured in alternate-order. This is likely explained by the shortened inter-item distance which induced the effects of consistency tendency. However, we did not observe significant differences in the strength of all prospective relationships between the two item-orders which supported our hypothesis that

[3]We examined an alternative TPB model that included baseline physical activity as a covariate. This alternative model reflects, and accounts for, established relationships between previous behavior and downstream psychosocial constructs subsumed within the TPB (Hagger et al., 2002; Rhodes & Courneya, 2003) when examining the predictive utility of the model. Parameter estimates and Wald test of parameter equalities showed a pattern of results that is directly consistent with the primary TPB model which did not include baseline physical activity as a covariate (see Appendix B for details).

consistency tendency did not moderate the prospective relationships that we examined in this study. However, with a cross-over design, we managed to mirror the effects of T1 in T2 when the other group received the opposite treatment, so the effects of consistency tendency on cross-sectional relationships of TPB appeared to be highly robust.

These interesting findings were in line with the preliminary findings of Chan and colleagues (2015), and the literature regarding common method bias (Podsakoff et al., 2003; Podsakoff et al., 2012), consistency tendency (Salancik, 1984; Salancik & Pfeffer, 1977), and the Socratic effect (Henninger & Wyer, 1976; Wyer, 1974). The results may also explain why the inter-factor correlation of the social cognitive variables were inflated when they were aligned in a logical manner (Budd, 1987; Budd & Spencer, 1986). Our findings have important implications for questionnaire-based research in health psychology, as it was demonstrated that the covariance between two positively correlated variables (e.g., variable A and B) could be artificially inflated when the items of both factors were presented in alternate-order, and where inter-item distance (i.e., physical separation) was reduced. While temporal separation has often been used in various fields of psychology for reducing common method variance in questionnaire-based research (e.g., Chan & Hagger, 2012; Chan, Hagger, & Spray, 2011; Hagger, Chatzisarantis, Barkoukis, Wang, & Baranowski, 2005; Hagger, Chatzisarantis, Culverhouse, & Biddle, 2003), our findings suggest that researchers should *also* pay particular attention to the confounding effects of the proximity of the items *within* a given questionnaire.

How to Minimize Consistency Tendency?

To reduce the effect of consistency tendency, it is important to make sure that participants do not assume that consecutive items or study variables in a health psychology questionnaire are related or repetitive, and avoid relying on the previous responses to answer the next questions (Podsakoff et al., 2003; Weijters et al., 2009). The use of temporal or proximal separation may be an effective solution for reducing the consistency tendency. Other methods such as placing a time gap, section break, highlighting the distinct or common features of the items (e.g., item stem), might also help as they may reduce participants' cognitive burden of questionnaire completion (Chan, Ivarsson, et al., 2015; Krosnick, 1991; Podsakoff & Organ, 1986; Weijters et al., 2009). However, researchers should be careful when randomizing item-order because the effect of consistency tendency is unlikely to be neutralized but rather factors could correlate more highly with their neighboring items or factors than those located further apart (Hui & Triandis, 1985; Podsakoff et al., 2003), leading to heightened random measurement error in the questionnaire (Budd, 1987; Budd & Spencer, 1986).

Study Limitations and Unexplored Issues of Consistency Tendency

Balanced against the unique perspective provided by the present study, limitations should also be highlighted and addressed in future work. First, although our hypothesis was supported by non-significant between- and within-subject differences in prospectively measured relationships, the results could plausibly be due to the fact that both intention and PBC failed to significantly correlate with physical activity, and the effects of the predictions, in some group/ timepoints, were smaller than the meta-analytic findings derived from TPB research in relation to the explanation of health behaviors (Hagger & Chatzisarantis, 2009b; Hagger et al., 2002; McEachan et al., 2011). This could conceivably be explained by the manipulation of participants' consistency tendency or inter-item distance that interfered with the prospective prediction. It might also be plausible that the target behavior in TPB (i.e., adequate physical activity) and the measurement of physical activity in IPAQ (i.e., time spent in various level of physical activity in the last seven days) were somewhat incompatible. Future tests should explore if the compatibility of the behavioral components within a questionnaire is related to factor correlations or consistency tendency. Despite being statistically non-significant, the magnitude of the prospective intention → physical activity association appeared to be larger when the items were presented in an ensemble order (β_{T1}= 0.170, β_{T2} = 0.169) compared to the alternate-order (β_{T1}= -0.070, β_{T2} = 0.028), and this pattern was consistent across both groups/timepoints. Indeed, for the relationship between PBC and behavior, it was generally not significant except when PBC was measured in alternate-order in T1. In fact, PBC forms a direct link to behavior because it is proposed to be a proxy measure of control over certain behaviors that demand heavily on resources, opportunity, and barriers (Ajzen & Madden, 1986; Hagger et al., 2002). The inconsistent finding of this direct link could be due to the possibility that the mediation of intention may attenuate the relationship between PBC and behavior (Ajzen, 1985), or the effect of consistency tendency in this study. Further studies should resolve this muddle by investigating if the consistency tendency effect only affects the predictive power of certain types of TPB variables or type of questionnaire format, and the introduction of alternative measures or criterion variables of physical activity (e.g., objective measurement by accelerometers or pedometers, implicit association test) or testing the consistency tendency effect in another health contexts may also be warranted (Chan, Ivarsson, et al., 2015).

The present study only looked at the relationships between positively linked factors, so we still know very little about how consistency tendency would affect the relationships between unrelated or negatively related constructs (see Hagger et al., 2007). The findings of Weijters and colleagues (2009) showed that proximity of items inflated the correlation between theoretically unrelated items, and reduced the correlation between a pair of reverse-scored, theoretically negatively-related items. In our study, the TPB variables were expected to be positively correlated (Ajzen, 1985), so we were unable to ascertain whether consistency tendency might moderate the correlations between factors that are expected to be negatively correlated or unrelated. Future studies may consider testing whether or not consistency tendency results in either (a) the effects of negative predictors on dependent constructs being less negative, or (b) unrelated constructs displaying (unfounded) correlations (Hagger et al., 2007), by including or integrating additional variables from other psychological frameworks. Similarly, some psychometric instruments contain inverted items that are intentionally designed to tap the opposite valence of the intended dimensions (Hagger et al., 2007; Weijters et al., 2009). Yet, the questionnaires used in the present study did not consist of inverted or negatively-worded items. Moreover, other method effects such as yes-saying and naysaying effects, recency and primacy effects, and scaling response options may take place in a health psychology questionnaire (Chan, Ivarsson, et al., 2015; Courneya et al., 2006; Rhodes et al., 2006; Rhodes et al., 2008; Rhodes et al., 2010). Therefore, it would be important for further studies to evaluate the extent to which other method effects could moderate the effect of consistency tendency on item or factor covariance (Richardson et al., 2009).

Finally, it is important to point out that we did not include a measurement point of TPB variables at T3 where both groups completed the TPB questions in the same item-order. Recent research in the health psychology has adopted cross-lagged panel designs to examine the causal relationship between longitudinally assessed psychological variables (Lindwall, Larsman, & Hagger, 2011; Marsh & Perry, 2005). It would be highly interesting to test the effects of consistency tendency on cross-lagged pathways in longitudinal studies. This may reveal how consistency tendency could moderate the change of TPB variables and behavior over time.

Conclusion

Self-reported questionnaires with an arbitrary item-order is one of the most commonly adopted research methods of psychological and behavioral assessment in the field of health psychology (Andersen, McCullagh, & Wilson, 2007; Godin et al., 2010; Hagger & Chatzisarantis, 2009a; Mankarious & Kothe, 2015), but the measurement responses could be confounded by common method bias. Our study demonstrated how the consistency tendency could moderate the strength of the pathways of the TPB in the prediction of physical activity. Inter-mixing the items may facilitate common method bias within TPB research, which results in artificially inflated cross-sectional factor correlations (i.e., attitude/ subjective norm/ PBC → intention). Consistency tendency does not appear to inflate prospective (i.e., intention → behavior) relationships in TPB. Health psychology researchers should be mindful of the confounding effect of consistency tendency when designing questionnaires for the assessment of psychological and behavioral variables. It is important to make sure that research participants do not assume that consecutive items in questionnaires are repetitive or related to previous items. From a measurement development perspective, the inclusion of physical gaps (e.g., section breaks, page breaks) and temporal separation between measures (e.g., short breaks, time gap) might be effective strategies to reduce common method bias induced by consistency tendency.

References

Ajzen, I. (1985). From intentions to actions: A theory of planned behavior. In J. Kuhl & J. Beckmann (Eds.), *From intentions to actions: A theory of planned behavior* (pp. 11-39). Berlin: Springer.

Ajzen, I. (2002, January, 2006). Constructing a TPB questionnaire: Conceptual and methodological considerations. Retrieved from University of Massachusetts Web site, http://people.umass.edu/aizen/tpb.html

Ajzen, I., & Fishbein, M. (1980). *Understanding attitudes and predicting social behavior*. Englewood Cliffs, NJ: Prentice-Hall.

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. Retrieved from <Go to ISI>://A1986F567300004

Andersen, M. B., McCullagh, P., & Wilson, G. J. (2007). But what do the numbers really tell us?: Arbitrary metrics and effect size reporting in sport psychology research. *Journal of Sport & Exercise Psychology, 29*(5), 664-672. Retrieved from <Go to ISI>://000250268600006

Budd, R. J. (1987). Response bias and the theory of reasoned action. *Social Cognition, 5*(2), 95-107. doi:10.1521/soco.1987.5.2.95

Budd, R. J., & Spencer, C. P. (1986). Lay theories of behavioral intention - a source of response bias in the theory of reasoned action. *British Journal of Social Psychology, 25*, 109-117. doi:10.1111/j.2044-8309.1986.tb00709.x

Chan, D. K. C., & Hagger, M. S. (2012). Trans-contextual development of motivation in sport injury prevention among elite athletes. *Journal of Sport and Exercise Psychology, 34*(5), 661-682.

Chan, D. K. C., Hagger, M. S., & Spray, C. M. (2011). Treatment motivation for rehabilitation after a sport injury: Application of the trans-contextual model. *Psychology of Sport and Exercise, 12*(2), 83-92. doi:10.1016/j.psychsport.2010.08.005

Chan, D. K. C., Ivarsson, A., Stenling, A., Yang, X. S., Chatzisarantis, N. L. D., & Hagger, M. S. (2015). Response-order effects in survey methods: A randomized controlled crossover study in the context of sport injury prevention. *Journal of Sport and Exercise Psychology, 37*(6), 666-673. doi:10.1123/jsep.2015-0045

Chan, D. K. C., Yang, S. X., Hamamura, T., Sultan, S., Xing, S., Chatzisarantis, N. L. D., & Hagger, M. S. (2015). In-lecture learning motivation predicts students' motivation, intention, and behaviour for after-lecture learning: Examining the trans-contextual model across universities from UK, China, and Pakistan. *Motivation and Emotion, 39*(6), 908-925. doi:10.1007/s11031-015-9506-x

Chen, F. F. (2007). Sensitivity of goodness of fit indexes to lack of measurement invariance. *Structural Equation Modeling, 14*(3), 464-504. doi:10.1080/10705510701301834

Courneya, K. S., Conner, M., & Rhodes, R. E. (2006). Effects of different measurement scales on the variability and predictive validity of the "two-component" model of the theory of planned behavior in the exercise domain. *Psychology & Health, 21*(5), 557-570. doi:10.1080/14768320500422857

Craig, C. L., Marshall, A. L., Sjostrom, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., . . . Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise, 35*(8), 1381-1395. doi:10.1249/01.Mss.0000078924.61453.Fb

Cronbach, L. J. (1946). Response sets and test validity. *Educational and Psychological Measurement, 6*(4), 475-493. doi:10.1177/001316444600600405

Cronbach, L. J. (1950). Further evidence on response sets and test design. *Educational and Psychological Measurement, 10*(1), 3-31. doi:10.1177/001316445001000101

Deci, E. L., & Ryan, R. M. (1985). *Intrinsic motivation and self-determination in human behavior*. New York: Plenum.

Doty, D. H., & Glick, W. H. (1998). Common methods bias: Does common methods variance really bias results? *Organizational Research Methods, 1*(4), 374-406. doi:10.1177/109442819814002

Duan, N., Alegria, M., Canino, G., McGuire, T. G., & Takeuchi, D. (2007). Survey conditioning in self-reported mental health service use: Randomized comparison of alternative instrument formats. *Health Services Research, 42*(2), 890-907. doi:10.1111/j.1475-6773.2006.00618.x

Ekelund, U., Sepp, H., Brage, S., Becker, W., Jakes, R., Hennings, M., & Wareham, N. J. (2006). Criterion-related validity of the last 7-day, short form of the International Physical Activity Questionnaire in Swedish adults. *Public Health Nutrition, 9*(2), 258-265. doi:10.1079/Phn2005840

Enders, C. K. (2010). *Applied missing data analysis*. New York: Guilford.

Godin, G., Belanger-Gravel, A., Amireault, S., Vohl, M. C., & Perusse, L. (2011). The effect of mere-measurement of cognitions on physical activity behavior: a randomized controlled trial among overweight and obese individuals. *International Journal of Behavioral Nutrition and Physical Activity, 8*. doi:10.1186/1479-5868-8-2

Godin, G., & Kok, G. (1996). The theory of planned behavior: A review of its applications to health-related behaviors. *American Journal of Health Promotion, 11*(2), 87-98. Retrieved from <Go to ISI>://A1996VT74400002

Godin, G., Sheeran, P., Conner, M., Delage, G., Germain, M., Belanger-Gravel, A., & Naccache, H. (2010). Which survey questions change behavior? Randomized controlled trial of mere measurement interventions. *Health Psychology, 29*(6), 636-644. doi:Doi 10.1037/A0021131

Hagger, M. S., Asci, F. H., Lindwall, M., Hein, V., Mulazimoglu-Balli, O., Tarrant, M., . . . Sell, V. (2007). Cross-cultural validity and measurement invariance of the social physique anxiety scale in five European nations. *Scandinavian Journal of Medicine and Science in Sports, 17*(6), 703-719. doi:10.1111/J.1600-0838.2006.00615.X

Hagger, M. S., & Chatzisarantis, N. L. D. (2009a). Assumptions in research in sport and exercise psychology. *Psychology of Sport and Exercise, 10*(5), 511-519. doi:10.1016/j.psychsport.2009.01.004

Hagger, M. S., & Chatzisarantis, N. L. D. (2009b). Integrating the theory of planned behaviour and self-determination theory in health behaviour: A meta-analysis. *British Journal of Health Psychology, 14*, 275-302.

Hagger, M. S., Chatzisarantis, N. L. D., Barkoukis, V., Wang, C. K. J., & Baranowski, J. (2005). Perceived autonomy support in physical education and leisure-time physical activity: A cross-cultural evaluation of the trans-contextual model. *Journal of Educational Psychology, 97*(3), 376-390. doi:10.1037/0022-0663.97.3.376

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 & Exercise Psychology, 24*(1), 3-32. Retrieved from <Go to ISI>://000174213500001

Hagger, M. S., Chatzisarantis, N. L. D., Culverhouse, T., & Biddle, S. J. H. (2003). The processes by which perceived autonomy support in physical education promotes leisure-time physical activity intentions and behavior: A trans-contextual model. *Journal of Educational Psychology, 95*(4), 784-795. doi:10.1037/0022-0663.95.4.784

Hambleton, R. K. (2005). Issues, designs, and technical guidelines for adapting tests into multiple languages and cultures. In R. K. Hambleton, P. Merenda, & C. Spielberger (Eds.), *Adapting educational and psychological tests for cross-cultural assessment* (pp. 3-38). Mahwah, NJ: Lawrence Erlbaum.

Hausenblas, H. A., Carron, A. V., & Mack, D. E. (1997). Application of the theories of reasoned action and planned behavior to exercise behavior: A meta-analysis. *Journal of Sport & Exercise Psychology, 19*(1), 36-51. Retrieved from <Go to ISI>://A1997WL43100003

Henninger, M., & Wyer, R. S. (1976). The recognition and elimination of inconsistencies among syllogistically related beliefs: Some new light on the Socratic effect. *Journal of Personality and Social Psychology, 34*(4), 680-693. doi:10.1037/0022-3514.34.4.680

Holbrook, A. L., Krosnick, J. A., Moore, D., & Tourangeau, R. (2007). Response order effects in dichotomous categorical questions presented orally: The impact of question and respondent attributes. *Public Opinion Quarterly, 71*(3), 325-348. doi:10.1093/poq/nfm024

Hui, C. H., & Triandis, H. C. (1985). The instability of response sets. *Public Opinion Quarterly, 49*(2), 253-260.

IPAQ Research Committee. (2005). Guidelines for Data Processing and Analysis of the International Physical Activity Questionnaire (IPAQ) – Short and Long Forms. Retrieved from http://www.ipaq.ki.se/scoring.pdf

Krosnick, J. A. (1991). Response strategies for coping with the cognitive demands of attitude measures in surveys. *Applied Cognitive Psychology, 5*(3), 213-216. doi:10.1002/acp.2350050305

Lindell, M. K., & Whitney, D. J. (2001). Accounting for common method variance in cross-sectional research designs. *Journal of Applied Psychology, 86*(1), 114.

Lindwall, M., Larsman, P., & Hagger, M. S. (2011). The reciprocal relationship between physical activity and depression in older European adults: A prospective cross-lagged panel design using SHARE data. *Health Psychology, 30*(4), 453-462. doi:10.1037/A0023268

Little, T. D. (2013). *Longitudinal structural equation modeling*. NY: Guilford Press.

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods, 1*(2), 130-149. doi:10.1037/1082-989X.1.2.130

Mankarious, E., & Kothe, E. (2015). A meta-analysis of the effects of measuring theory of planned behaviour constructs on behaviour within prospective studies. *Health Psychology Review, 9*(2), 190-204. doi:10.1080/17437199.2014.927722

Marsh, H. W. (2007). Application of confirmatory factor analysis and structural equation modeling in sport/exercise psychology. In G. Tenenbaum & R. C. Eklund (Eds.), *Handbook of sport psychology* (3rd ed., pp. 774-798). Hoboken, NJ: Wiley.

Marsh, H. W., Hau, K. T., & Grayson, D. (2005). Goodness of fit evaluation in structural equation modeling. In A. Maydeu-Olivares & J. J. McArdle (Eds.), *Multivariate applications book series. Contemporary psychometrics: A festschrift for Roderick P. McDonald* (pp. 275-340). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.

Marsh, H. W., & Perry, C. (2005). Self-concept contributes to winning gold medals: Causal ordering of self-concept and elite swimming performance. *Journal of Sport & Exercise Psychology, 27*(1), 71-91. Retrieved from <Go to ISI>://000227461100005

McClendon, M. J. (1991). Acquiescence and recency response-order effects in interview surveys. *Sociological Methods and Research, 20*(1), 60-103. doi:10.1177/0049124191020001003

Mcdonald, R. P. (1970). Theoretical foundations of principal factor analysis, canonical factor analysis, and alpha factor analysis. *British Journal of Mathematical and Statistical Psychology, 23*(May), 1-21. doi:10.1111/j.2044-8317.1970.tb00432.x

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*, 97-144. doi:10.1080/17437199.2010.521684

Morwitz, V. G., & Fitzsimons, G. J. (2004). The mere-measurement effect: Why does measuring intentions change actual behavior? *Journal of Consumer Psychology, 14*(1-2), 64-74. doi:10.1207/S15327663jcp1401&2_8

Muthén, L. K., & Muthén, B. O. (1998-2017). *Mplus user's guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.

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. doi:10.1037/0021-9101.88.5.879

Podsakoff, P. M., MacKenzie, S. B., & Podsakoff, N. P. (2012). Sources of method bias in social science research and recommendations on how to control it. . *Annual Review of Psychology, 63*, 539-569. doi:10.1146/annurev-psych-120710-100452

Podsakoff, P. M., & Organ, D. W. (1986). Self-reports in organizational research: Problems and prospects. *Journal of Management, 12*(4), 531-544. doi:10.1177/014920638601200408

Rhodes, R. E., & Courneya, K. S. (2003). Modelling the theory of planned behaviour and past behaviour. *Psychology, Health and Medicine, 8*, 57-69. doi:10.1080/1354850021000059269

Rhodes, R. E., Matheson, D. H., & Blanchard, C. M. (2006). Beyond scale correspondence: A comparison of continuous open scaling and fixed graded scaling when using social cognitive constructs in the exercise domain. *Measurement in Physical Education and Exercise Science, 10*(1), 13-39. doi:10.1207/s15327841mpee1001_2

Rhodes, R. E., Matheson, D. H., Blanchard, C. M., & Blacklock, R. E. (2008). Evaluating timeframe expectancies in physical activity social cognition: Are short- and long-term motives different? *Behavioral Medicine, 34*(3), 85-93. doi:10.3200/BMED.34.3.85-94

Rhodes, R. E., Matheson, D. H., & Mark, R. (2010). Evaluation of social cognitive scaling response options in the physical activity domain. *Measurement in Physical Education and Exercise Science, 14*(3), 137-150. doi:10.1080/1091367x.2010.495539

Richardson, H. A., Simmering, M. J., & Sturman, M. C. (2009). A tale of three perspectives examining post hoc statistical techniques for detection and correction of common method variance. *Organizational Research Methods, 12*(4), 762-800. doi:10.1177/1094428109332834

Salancik, G. R. (1984). On Priming, consistency, and order effects in job-attitude assessment: With a note on current research. *Journal of Management, 10*(2), 250-254. doi:10.1177/014920638401000209

Salancik, G. R., & Pfeffer, J. (1977). An examination of need-satisfaction models of job attitudes. *Administrative Science Quarterly, 22*(3), 427-456. doi:10.2307/2392182

Sherif, M., & Cantril, H. (1945). The psychology of'attitudes': Part I. *Psychological Review, 52*(6), 295-315. doi:10.1037/h0062252

Sherif, M., & Cantril, H. (1946). The psychology of'attitudes': Part II. *Psychological Review, 53*(1), 1-24. doi:10.1037/h0058561

Stone, E. F., & Gueutal, H. G. (1984). On the premature death of need-satisfaction models: An investigation of Salancik and Pfeffer's views on priming and consistency artifacts. *Journal of Management, 10*(2), 237-249. doi:10.1177/014920638401000208

Sturgis, P., Allum, N., & Brunton-Smith, I. (2009). Attitudes over time: The psychology of panel conditioning. In P. Lynn (Ed.), *Methodology of Longitudinal Surveys* (pp. 113-126). West Sussex, UK: Wiley.

Warren, J. R., & Halpern-Manners, A. (2012). Panel conditioning in longitudinal social science surveys. *Sociological Methods & Research, 41*(4), 491-534. doi:10.1177/0049124112460374

Weijters, B., Geuens, M., & Schillewaert, N. (2009). The proximity effect: The role of inter-item distance on reverse-item bias. *International Journal of Research in Marketing, 26*(1), 2-12. doi:10.1016/j.ijresmar.2008.09.003

World Health Organization. (2010). *Global recommendations on physical activity for health*. Retrieved from http://www.who.int/dietphysicalactivity/publications/9789241599979/en/

Wyer, R. S. (1974). Some implications of the “Socratic effect” for alternative models of cognitive consistency. *Journal of Personality, 42*(3), 399-419. doi:10.1111/1467-6494.ep8970049

1.

I intend to do adequate physical activity in the forthcoming month.						
Strongly Disagree			Neutral			Strongly Agree
1	2	3	4	5	6	7

Figure 1. An example item of intention for demonstration of the format of the questionnaire.

Table 1.

Correlation Matrix and Descriptive Statistics of the Study Variables.

		1	2	3	4	5	6	7	8	9
1. Intention T1		1	.676***	.260***	.671**	.770**	.619**	.454***	.597**	.220**
2. Attitude T1		.844**	1	.256**	.598**	.640**	.795**	.395**	.533**	.157*
3. Subjective Norm T1		.641**	.558**	1	.104	.282**	.209**	.663**	.134	.030
4. Perceived Behavioral Control T1		.791**	.825**	.604**	1	.607**	.570**	.301**	.734**	.174*
5. Intention T2		.674**	.492**	.349**	.519**	1	.855**	.613**	.844**	.176*
6. Attitude T2		.575**	.694**	.367**	.563**	.662**	1	.465**	.719**	.189**
7. Subjective Norm T2		.376**	.279**	.691**	.317**	.435**	.369**	1	.449**	.088
8. Perceived Behavioral Control T2		.562**	.545**	.314**	.698**	.711**	.638**	.389**	1	.210*
9. Physical Activity T3		.202**	.093	.124	.259**	.211*	.133*	.166*	.187**	1
Group 1	Mean	4.98	4.95	5.07	5.45	5.12	4.96	5.08	5.55	1522.75
	Standard Deviation	1.08	1.08	.92	1.00	1.07	1.17	.94	1.04	1701.82
	ω	.811	.814	.728	.737	.877	.919	.846	.821	
	Factor Loading	.776	.746	.721	.611	.845	.858	.791	.677	
	[range]	[.751-.814]	[.706-.767]	[.652-.838]	[.512-.714]	[.804-.892]	[.807-.887]	[.657-.909]	[.559-.802]	
Group 2	Mean	4.79	5.54	4.75	4.87	4.99	5.51	4.81	4.94	1490.94
	Standard Deviation	1.21	1.24	.97	1.07	1.08	1.09	0.98	1.04	1666.05
	ω	.871	.902	.819	.771	.827	.900	.796	.821	
	Factor Loading	.820	.835	.757	.635	.814	.834	.767	.667	
	[range]	[.812-.834]	[.803-.876]	[.619-.843]	[.395-.786]	[.741-.870]	[.769-.871]	[.621-.850]	[.444-.838]	

Notes. The zero-order correlations at lower and upper diagonal respectively showed the Pearson correlations of the study variables for Group 1 and Group 2.

**p* < .05, ** *p* < .01.

Table 2

Model Fit Indices of the Confirmatory Factor Analyses

	$\chi 2$	*df*	*p*	RMSEA	CFI	TLI	SRMR	Statistical Power
				Group 1				
T1 Alternate	218.023	84	.000	.064 [.054, .075]	.919	.899	.052	.95
T2 Ensemble	241.712	84	.000	.074 [.063, .085]	.929	.911	.078	.97
Configural	690.056	362	.000	.049 [.043, .054]	.930	.915	.065	.99
Metric	710.886	373	.000	.049 [.043, .054]	.928	.916	.070	.99
Scalar	741.384	384	.000	.049 [.044, .055]	.923	.913	.070	.99
				Group 2				
T1 Ensemble	175.771	84	.000	.053 [.042, .064]	.957	.947	.062	.95
T2 Alternate	212.698	84	.000	.066 [.055, .077]	.933	.916	.060	.97
Configural	585.365	362	.000	.040 [.034, .046]	.956	.947	.062	.99
Metric	602.787	373	.000	.040 [.034, .046]	.955	.948	.067	.99
Scalar	622.001	384	.000	.040 [.034, .046]	.953	.947	.068	.99
				Between-Groups T1				
Configural	393.770	168	.000	.059 [.052, .067]	.941	.926	.057	.99
Metric	430.766	179	.000	.060 [.053, .068]	.934	.922	.072	.99
Scalar	455.100	190	.000	.060 [.053, .067]	.930	.923	.071	.99
				Between-Groups T2				
Configural	453.897	168	.000	.070 [.062, .078]	.931	.913	.069	.99
Metric	474.247	179	.000	.069 [.061, .077]	.928	.916	.076	.99
Scalar	497.752	190	.000	.068 [.061, .076]	.925	.918	.076	.99

Table 3
Within- and Between-Group Comparisons of Factor Correlations Using Wald Test of Parameter Equalities

	Within-Group Comparison							
	Group 1				Group 2			
	T1	T2	χ^2	*p*	T1	T2	χ^2	*p*
Intention ↔ Norm	.641*	.436*	9.833	.0017	.260*	.613*	28.845	.0000
Intention ↔ PBC	.791*	.711*	2.177	.1401	.671*	.844*	14.414	.0000
Intention ↔ Attitude	.843*	.661*	15.720	.0001	.676*	.855*	19.203	.0000
Norm ↔ PBC	.605*	.390*	11.476	.0007	.105	.449*	20.670	.0000
Norm ↔ Attitude	.558*	.369*	7.928	.0049	.256*	.465*	10.880	.0000
PBC ↔ Attitude	.825*	.638*	16.926	.0000	.598*	.719*	4.932	.0264
	Between-Group Comparison							
	T1				T2			
	Group 1	Group 2	χ^2	*p*	Group 1	Group 2	χ^2	*p*
Intention ↔ Norm	.674*	.227*	29.140	.0000	.437*	.590*	2.904	.0884
Intention ↔ PBC	.804*	.636*	6.319	.0119	.719*	.813*	1.618	.2034
Intention ↔ Attitude	.860*	.662*	17.825	.0000	.663*	.848*	12.492	.0004
Norm ↔ PBC	.623*	.075	36.616	.0000	.399*	.416*	.029	.8638
Norm ↔ Attitude	.590*	.224*	23.464	.0000	.369*	.441*	.671	.4128
PBC ↔ Attitude	.845*	.586*	15.556	.0000	.647*	.696*	.491	.4837

Note. Norm = subjective norm; PBC = perceived behavioural control.
*$p < .05$

Table 4

Between-Group Comparisons of the Effects of TPB Variables at T1 and T2 on PA Behavior at T3 Using Wald Test of Parameter Equalities (Standardized Coefficients are Presented)

	T1				T2			
	Group 1	Group 2	χ^2	p	Group 1	Group 2	χ^2	p
Norm → Intention	.206*	.116*	1.397	.2373	.140*	.201*	1.091	.2962
Attitude → Intention	.502*	.420*	.729	.3932	.313	.460*	2.364	.1242
PBC → Intention	.196*	.447*	3.700	.0544	.517*	.350*	.887	.3462
PBC → PA	.201	.038	.745	.3882	.116	.149	.018	.8941
Intention → PA	.041	.192	.826	.3633	.126	.065	.101	.7508
R^2 Intention	.666	.641			.686	.739		
R^2 PA	.054	.049			.052	.041		

Note. Norm = subjective norm; PBC = perceived behavioural control; PA = physical activity.
$*p < .05$

Appendix A

Item-content and item-order of the scale at T1 and T2

		Anchors	Item-order – Ensemble	Item-order – Alternate
	Intention			
1	I intend to do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	1	1
2	I will try to put great effort into doing adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	2	5
3	I plan to do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	3	9
	Subjective Norm			
4	Most people who are important to me think that I should do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	4	2
5	It is expected of me that I do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	5	6
6	The people in my life whose opinions I value would approve me to do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	6	10
	Perceived Behavioral Control			
7	It is possible for me to do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	7	3
8	If I want to I could do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	8	7
9	I have complete control over how to do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	9	11
10	It is completely down to me to decide to do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	10	14
11	It is easy for me to do adequate physical activity in the forthcoming month.	1 “Strongly Disagree” – 7 “Strongly Agree”	11	18
	Attitude			
12	For me doing adequate physical activity in the forthcoming month would be …	1 “Extremely Harmful” – 7 “Extremely Beneficial”	12	4
13	For me doing adequate physical activity in the forthcoming month would be …	1 “Extremely Unpleasant” – 7 “Extremely Pleasant”	13	8
14	For me doing adequate physical activity in the forthcoming month would be …	1 “Extremely Worthless” – 7 “Extremely Valuable”	14	12
15	For me doing adequate physical activity in the forthcoming month would be …	1 “Extremely Bad” – 7 “Extremely Good”	15	15
16	For me doing adequate physical activity in the forthcoming month would be …	1 “Extremely Unenjoyable” – “Extremely Enjoyable”	16	17
17	How frequent do you do adequate leisure-time physical activity in the last 2 weeks?	1 “Never” – 7 “Very Often”	17	13
18	How much effort do you put into doing adequate leisure-time physical activity in the last 2 weeks?	1 “Minimum Effort” – “Maximum Effort”	18	16

Notes. Item 16 was not included in the analyses because of weak factor loadings and inter-item correlations with the other attitude items.

Appendix B

Between-Group Comparisons of the Effects of TPB Variables at T1 and T2 on PA Behavior at T3 (While Controlling for PA Behavior at T1) Using Wald Test of Parameter Equalities (Standardized Coefficients are Presented)

	T1				T2			
	Group 1	Group 2	χ^2	*p*	Group 1	Group 2	χ^2	*p*
Norm → Intention	.207*	.116*	1.413	.2345	.140*	.201*	1.091	.2962
Attitude → Intention	.504*	.420*	.777	.3782	.311*	.459*	2.381	.1229
PBC → Intention	.193*	.448*	3.828	.0504	.521*	.353*	.900	.3427
PBC → PA T3	.198	-.105	3.583	.0583	.088	.033	.009	.9235
Intention → PA T3	-.054	.180	2.646	.1038	.054	.025	.105	.7464
PA T1 → PA T3	.507*	.556*	.134	.7147	.498*	.552*	.148	.7000
R^2 Intention	.667	.643			.688	.741		
R^2 PA T3	.300	.342			.289	.327		

Note. Norm = subjective norm; PBC = perceived behavioural control; PA = physical activity.
$^*p < .05$