Current Psychology (2021) 40:1015–1030
https://doi.org/10.1007/s12144-018-0021-6

CrossMark

Trait self-control and self-discipline: Structure, validity, and invariance across national groups

Martin S. Hagger[1,2,3,4] · Chun-Qing Zhang[3] · Eva-Maria Kangro[5] · Francis Ries[6] · John C. K. Wang[7] · Brody Heritage[8] · Derwin K. C. Chan[1,9]

Published online: 15 October 2018

Abstract

The aim of the present study was to test the validity of the Brief Self-Control Scale (BSCS; Tangney, Baumeister and Boone 2004) including its dimensional structure based on competing one- and two-factor models, discriminant validity from the conceptually-related self-discipline construct, invariance across multiple samples from different national groups, and predictive validity with respect to health-related behaviors. Samples of undergraduate students (total $N = 1282$) from four national groups completed the brief self-control scale, the self-discipline scale from the NEO-PI-R, and self-report measures of binge drinking, exercise, and healthy eating. Confirmatory factor analytic models supported a two-factor structure of self-control encompassing restraint and non-impulsivity components. The model exhibited good fit in all samples and invariance of factor loadings in multi-sample analysis. The restraint and non-impulsivity components exhibited discriminant validity and were also distinct from self-discipline. Structural equation models revealed that non-impulsivity predicted binge drinking in three of the samples, and restraint predicted exercise in two samples, with no role for self-discipline. Results point to a multi-dimensional structure for trait self-control consistent with previous theory separating impulsive- and control-related components.

Keywords Self-control · Self-discipline · Self-regulation · Restraint · Impulsivity

Introduction

The construct of self-control has received considerable attention in the personality and social psychology literature and has been incorporated in multiple theories of motivation, volition, and action regulation (e.g., Carver 2005; Fishbach and Shah 2006; Gottfredson and Hirschi 1990; Hofmann et al. 2009; Kuhl 2000; Metcalfe and Mischel 1999; Wills et al. 2011). Self-control encompasses a wide range of responses including ability to exert control over, suppress, or inhibit thoughts, emotions, impulses, urges, temptations, and 'dominant responses', better performance regulation, and breaking habits and ingrained, well-learned responses (Baumeister and Heatherton 1996; Hofmann et al. 2009). Self-control has typically been conceptualized as a trait-like construct representing individuals' capacity to actively exert control

Electronic supplementary material The online version of this article (https://doi.org/10.1007/s12144-018-0021-6) contains supplementary material, which is available to authorized users.

✉ Martin S. Hagger
martin.hagger@curtin.edu.au

1 Health Psychology and Behavioural Medicine Research Group and Laboratory of Self-Regulation, School of Psychology, Faculty of Health Sciences, Curtin University, GPO Box U1987, Perth, WA 6845, Australia

2 Faculty of Sport and Health Sciences, University of Jyväskylä, Jyväskylä, Finland

3 Department of Sport and Physical Education, Hong Kong Baptist University, Hong Kong, China

4 Department of Psychology, Université Paris, Nanterre, Paris, France

5 Interim Gene Consultancy, Tallinn, Estonia

6 Departamento de Educación Física y Deporte, Facultad de Ciencias de la Educación, Universidad de Sevilla, Sevilla, Spain

7 Physical Education and Sports Science, National Institute of Education, Nanyang Technological University, Singapore, Singapore

8 School of Psychology and Exercise Science, Murdoch University, Perth, Australia

9 School of Public Health, the University of Hong Kong, Hong Kong, China

over impulsive responses (de Ridder et al. 2012; Metcalfe and Mischel 1999; Tangney et al. 2004). Theories of trait self-control highlight its self-organizing function; self-control in conceptualized as individuals' capacity to organize and structure long-term goals, recognize and predict costs and consequences of future actions, and monitor and detect shifts in attention and motivation away from goal-directed actions and rectify them (Gottfredson and Hirschi 1990; Inzlicht and Schmeichel 2012). Similarly, self-control has been identified as a core component of volition (Kuhl 1984, 2000). For example, Kuhl proposed that self-control is akin to self-discipline, and comprises a number of volitional components involved in actively inhibiting motives or impulses that detract from intentional action such as goal recollection, forgetfulness prevention, planning skill, impulse control, and initiating control. In effect, these theories outline strategies or competencies that individual may employ to manage alternative actions and pathways that may derail goal directed behavior.

Interest in self-control has been spurred by a burgeoning body of research that has positively linked self-control and associated constructs with adaptive outcomes in multiple domains. Good self-control is linked with better performance in school, university, and the workplace, better social functioning and cohesive relationships, less psychopathology and susceptibility to crime, delinquency, and drug abuse, and better physical and mental health (de Ridder et al. 2012; Hamilton et al. 2018; Tangney et al. 2004). Analogously, poor self-control is associated with poorer functioning and maladaptive outcomes.

Despite the proliferation of evidence demonstrating correlations between self-control and adoptive outcomes, there is considerable variability in the conceptualization and measurement of self-control. An ongoing debate in the scientific literature is whether self-control is unidimensional or comprises multiple domains, and this has been reflected in measures developed to tap self-control (de Vries and van Gelder 2013; Maloney et al. 2012; Tangney et al. 2004; Williams et al. 2007). Furthermore, there are a number of terms that have been applied to the domain of self-control and have often been used synonymously such as willpower, self-discipline, response inhibition, and impulse control. These issues present problems when attempting to ascertain the true nature of associations between self-control and key outcomes, and, by implication, the development of fit-for-purpose tests of the mechanisms by which self-control impacts behavior, and interventions or recommendations for practitioners.

An important endeavor in research on self-control is to ensure that measures exhibit adequate construct validity, internal consistency, discriminant validity from conceptually related but distinct constructs, and predictive and nomological validity, particularly relative to behavior. The purpose of the current study is to assess the validity of the Brief Self-Control Scale (BSCS; Tangney et al. 2004), a leading self-report measure of trait self-control. We will test the construct and factorial validity of the measure, its dimensional nature including uni- and multidimensional conceptualizations based on theories of self-control, its discriminant validity from the conceptually-related measure of self-discipline from the conscientiousness scale of the revised NEO personality inventory (NEO-PI-R; McCrae and Costa 2004), its generalizability across multiple samples from different national groups, and the predictive validity of the measure with respect to health-related behaviors. The research will add to the literature by demonstrating whether or not the measure exhibits adequate validity and is fit-for-purpose when it comes to assessing self-control in multiple samples and behavioral contexts.

Trait Self-Control Measurement

Self-control has typically been conceptualized as a generalized tendency to engage in conscious, deliberative control over actions and suppress impulsive, habitual, well-learned dominant responses that occur with little thought or conscious intervention. These conceptualizations are reflected in a number of theories and models of self-control. Metcalfe and Mischel's (1999) 'hot' and 'cool' pathways to action and Strack and Deutch's (2004) 'impulsive' and 'reflective' processes are two examples. The 'hot' or 'impulsive' components reflect emotive, spontaneous responses to stimuli driven by well-learned cue-response pairings with little conscious control. In contrast, the 'cool' or 'reflective' components reflect reasoned, deliberative processes that involve effortful, conscious control over actions. Individuals with high trait self-control tend to be more effective in enacting the 'cool' or 'reflective' pathway and, therefore, exert effective control over actions. Theories of self-control also suggest that individuals with high trait self-control are also more effective in structuring their environment so as to reduce the potential for derailing circumstances such as cues to impulsive behaviors or competing courses of action to interfere with goal-directed behaviors. For example, Gottfredson and Hirschi (1990) suggest that the mechanism by which these traits lead to more effective self-control is through better capacity to organize and structure long-term goals, and recognize and predict the benefits and costs of acting. Consistent with this proposal, research has indicated that individuals high in trait self-control ironically tend to exert less self-control than those low in trait self-control, suggesting that individuals with high self-control structure their goals and behaviors in such a way to reduce the use of self-control by avoiding temptations and relying on habitual enactment of goal-directed behaviors (Ent et al. 2015; Galla and Duckworth 2015). Importantly, trait self-control, like many personality and individual difference constructs, has been conceptualized as domain-general and, therefore, the benefits of good self-control and maladaptive consequences of poor self-control are likely to generalize across multiple contexts and behaviors. Effects of self-control are also expected to generalize across multiple populations and national groups.

Several prominent self-report measures of trait self-control have been developed. Some have conceptualized self-control as a unitary generalized construct (e.g., Forstmeier et al. 2011; Marcus 2003; Tangney et al. 2004), while others have developed multi-dimensional measures that identify specific facets that pertain to the overall self-control construct (e.g., Grasmick et al. 1993; Neal and Carey 2005). Although many measures align with different theoretical perspectives on self-control, there are frequent overlaps among item content, and correlations among the measures have suggested considerable shared variance. Recent analyses have indicated that even though many of the unidimensional self-control scales purport to contain items that capture the essence of a global self-control construct, factor analyses have indicated that separable underlying dimensions clearly exist (e.g., Allom et al. 2016; de Ridder et al. 2011; Maloney et al. 2012). The identification and isolation of the components of self-control may shed light on its conceptualization, how it may be operationalized in theory and empirical research and provide further evidence for the mechanisms and pathways by which it relates to behavioral outcomes.

Recently, research has examined the dimensionality of the brief version of the self-control scale, a measure that has been frequently used to assess self-control in the extant literature (e.g. Lindner et al. 2015). Maloney et al. (2012) proposed a two-dimensional structure with one factor comprising items that reflected disciplined control over responses and actions, termed *restraint*, and another factor that reflected the tendency to be spontaneous or to act on the basis of intuition, heuristics, and well-learned cue-response tendencies, termed *impulsivity*. While both capacity for restraint and impulsive tendency are both defining characteristics of self-control, researchers using the Tangney et al. (2004) BSCS recognized differential associations with conceptually related constructs. From a theoretical perspective, the restraint and impulsivity conceptualization is consistent with the 'hot' vs. 'cool' distinction, which suggests affective- and cognitive-mediated pathways to action, such that good self-control is dependent on the extent to which the cognitive restraint system is able to 'put the brakes on' affectively-driven impulsive tendencies (Carver 2005; Hofmann et al. 2009). Maloney et al. (2012) found support for their proposed two factor structure, that made a clear distinction between items reflecting the *restraint* and *impulsivity* components. Furthermore, they found differential prediction of the scales with impulsivity predicting poor workplace practices and restraint predicting emotional exhaustion. This pattern of results is consistent with Metcalfe and Mischel's (1999) distinction and those suggested by others.

In a similar approach, de Ridder et al. (2011) conducted an analysis to identify inhibitory and initiatory dimensions of trait self-control based on the brief self-control scale. They surmised that individuals have tendencies to exert self-control for two different kinds of behavioral response: those that require inhibiting automated response tendencies, or *inhibition* and those that required active and deliberative engagement in behaviors or *initiation*. Again, with reference to Metcalfe and Mischel's (1999) 'hot' and 'cool' distinction indicating that inhibiting responses is almost always to service some sort of long-term, delayed goal which incurs a short term cost, not only in terms of delaying immediate gratification, but also in investing effort in behaviors that will assist in reaching that goal. Following a systematic classification of items from the brief self-control scale, de Ridder et al. (2012) found support for a distinct factor structure and also demonstrated that the inhibitory factor was more strongly related to maladaptive, undesirable health-related outcomes i.e. behaviors closely linked to impulse control that require desistence for improved outcomes (e.g., smoking cigarettes and binge drinking), while the initiatory factor was more strongly associated with adaptive, desirable behaviors i.e. behaviors in which engagement is necessary for better outcomes (e.g., physical activity and studying). It is important to note that while there seems to be common theoretical underpinning and conceptual bases for the restraint or inhibition and impulsivity or initiation components from Maloney et al.'s (2012) and de Ridder et al.'s (2011) analyses, and considerable overlap in the items identified to delineate the two components, they were not identical in terms of the exact item make up. This means that the two factor structures should not be considered equivalent, and which conceptualization most effectively captures the underlying structure of self-control has yet to be resolved.

A related issue for the trait self-control scale is the problems associated with redundancy across existing scales that may be tapping the same construct. This presents considerable challenges for researchers seeking to identify a valid and reliable measure of self-control that will be fit-for-purpose in assessing self-control. An imperative, therefore, is to establish the extent to which the measure of self-control exhibits discriminant validity from other measures that tap closely-related constructs. A prominent candidate likely to be closely associated with trait self-control is self-discipline, a sub-facet of the conscientiousness scale from NEO-PI-R (McCrae and Costa 2004). Self-discipline is defined as individual's capacity to actively work toward long-term goals and to resist temptations. Unsurprisingly, self-discipline shares many of the defining characteristics of self-control as captured by the BSCS (Tangney et al. 2004), particularly the restraint or inhibitory components outlined in Maloney et al.'s and de Ridder et al.'s analyses. Similarity can also be observed at the item level. For example, items 9 ("Pleasure and fun sometimes keep me from getting work done") and 11 ("I am able to work effectively toward long-term goals") from the BSCS bear close resemblance with items 3 ("I find it difficult to get down to work") and 10 ("I tend to carry out my plans"), respectively, from the self-discipline scale of the NEO-PI-R. This raises concerns as to whether the potential overlap self-discipline and components of self-control represent an example a 'jangle' fallacy (Block 1995; Hagger 2014). That is, constructs with the same underlying content labelled differently. Such phenomena present problems for researchers: the introduction of redundancy impedes scientific progress by causing conceptual confusion. There is, therefore, a need for resolution in terms of the measures of self-control and self-discipline which have ostensibly similar content but have typically been tapped with different scales and referred to using different terminology (Hagger and Hamilton 2018). Examining the discriminant validity of a leading measure of self-control, such as the brief self-control scale, and the facet of self-discipline will attempt to identify the level of redundancy and, if substantial overlap exists, may help restore some parsimony to the terminology and measurement of these constructs.

The purpose of the present study is to examine the construct, discriminant, and predictive validity of the brief self-control and self-discipline scales in multiple samples. Specifically, the research aims to identify the dimensional structure of the brief self-control scale, testing the unidimensional model proposed by Tangney et al., as well as alternative two-dimensional models comprising restraint and impulsivity dimensions (Maloney et al. 2012) and inhibition and initiation dimensions (de Ridder et al. 2011). In addition, the discriminant validity of the unidimensional and multi-dimensional scales will also be tested alongside the self-discipline facet from the conscientiousness scale of the NEO-PI-R. Finally, we will test the predictive validity of the scales in accounting for variance in three self-reported health-related behaviors likely related to self-control: binge drinking, exercise, and healthy eating. We hypothesize that the Maloney et al. (2012) and de Ridder et al. (2011) two-factor solutions for the BSCS will be superior to the one-dimensional model. We also expect to identify the factor structure that exhibits optimal fit in multiple samples from different national groups, and for the structure to exhibit invariance across samples. In addition, while we expected components of the BSCS to correlate significantly with the self-discipline scale, we predicted that the scales would exhibit discriminant validity. Finally, we expect the initiation or non-impulsivity components of the two-factor self-control models to be positively related to adaptive health-related outcomes (exercise and healthy eating) and negatively related to health behaviors for which disengagement is more adaptive (binge drinking). In contrast we expected the opposite pattern of effects for the inhibition or restraint components with respect to these behaviors.

Method

Participants

Participants were first-year university students from Universities in Estonia ($N = 347$, M age = 28.40, SD = 7.95, 123 males and 224 females), Luxembourg ($N = 207$, M age = 22.34, SD = 2.16, 70 males and 137 females), Spain ($N = 291$, M age = 22.34, SD = 3.41, 106 males and 185 females), and the United Kingdom ($N = 437$, M age = 20.80, SD = 2.55, 79 males, 337 females and 21 not reported). Students were majoring in psychology and were recruited at the request of their instructors during university class time.

Measures

Self-control and self-discipline constructs were measured using the BSCS developed by Tangney et al. (2004) and the items from the self-discipline subscale of the consciousness domain of the NEO-PI-R available from the International Personality Item Pool (IPIP 2017). Estonian, French (Luxembourg), and Spanish versions of the scales were developed using standardized back-translation techniques (Bracken and Barona 1991). An initial translation was vetted by two independent and proficient bilingual translators who translated the questionnaires back into English. We then compared the back-translated versions with the original English version for errors, biases, and incongruences. These were removed in further back-translations by the translators in an iterative process repeated until the versions were semantically identical.

Trait Self-Control The BSCS comprises 13 items (e.g., "I am good at resisting temptation") with responses made on five-point scales (1 = *not at all* and 5 = *very much*). The full scale is presented in Appendix A (supplemental materials).

Self-Discipline The self-discipline scale comprises 10 items (e.g., "II start tasks right away") with responses made on five-point scales (1 = *strongly disagree* and 5 = *strongly agree*). The full scale is presented in Appendix B (supplemental materials).

Health-Related Behaviors Participants completed a series of two-item measures of their recent engagement in three health related behaviors relevant to the undergraduate student population: frequency of binge drinking, exercise, and eating a healthy diet. Participants self-reported how frequently they exceeded guideline limits of alcohol in the previous four weeks on two items (e.g., "In the course of the past four weeks, how often have you engaged in binge drinking (i.e., consumed over the levels of alcohol intake given above in a single 'session')?" The items were preceded by the definition of binge drinking: "Binge drinking is considered drinking 10 units of alcohol (equivalent to 5 'pints' (approx. half-liter glasses) of normal strength beer or 10 spirits or liqueur 'shots' or measures) for men or 7 units of alcohol (equivalent to 3½ pints (approx. half-liter glasses) or 7 spirits or liqueur 'shots' or measures) for women in any *single* 'session'". Self-reported exercise behavior during leisure-time was measured using an adapted version of Godin and Shepherd's (1985) Leisure-Time Exercise Questionnaire (LTEQ). Participants rated their four-week behavioral frequency on two items (e.g., "In the course of the past two weeks, how often have you participated in vigorous physical activities for more than 20 minutes at a time?") using six-point Likert scales with scale endpoints *never* (1) and *everyday* (6). Participants rated the frequency with which they had watched their diet for health reasons in the previous week on two items (e.g. "In the course of the past four weeks, how often have you watched your diet at mealtimes and when snacking?") using six-point Likert scales with scale endpoints *never* (1) and *everyday* (6). All three behavioral measures have been used to indicate latent measures of exercise, following a healthy diet, and binge drinking with high factor loadings and average variance extracted in previous studies providing support for their construct validity and internal consistency (Arnautovska et al. 2017; Hagger and Chatzisarantis 2005; Hagger et al. 2012).

Procedure

Clearance from the Institutional Review Boards of each university was obtained prior to data collection. Data were collected during university seminars and lectures with participants being asked to participate by their instructor. Prior to data collection, participants were informed that they were being asked to participate in a study on personality and asked to complete an informed consent form. Measures were administered in sealed envelopes and participants were asked to complete all measures including demographic variables. They were also informed that responses were unique to individuals and asked not to confer with other students while completing the measures.

Data Analysis[1]

Confirmatory factor analysis (CFA) was conducted to test the factorial validity of Tangney et al.'s (2004) one-factor model of the brief self-control scale, Maloney et al.'s (2012) and de Ridder et al.'s (2011) two-factor models of the same scale, and the one-factor model of the self-discipline scale. Models were estimated using the Mplus 7.31 (Muthén and Muthén 2015) software using a maximum likelihood estimation method and robust standard errors (Satorra and Bentler 1988). Goodness of fit of the models with the data was tested using multiple fit indices, including the scaled comparative fit index (CFI), root mean square error of approximation (RMSEA) and its 90% confidence interval, and standardized root mean square residual (SRMSR). The model fit was considered acceptable if the CFI exceeded .90, SRMSR was equal to or below .05, and the RMSEA was equal to or below .08 with narrow 90% confidence intervals (Hu and Bentler 1999). We also examined the solution estimates of these models including the factor loadings, average variance extracted, and Cronbach alpha and composite reliability statistics. Based on these statistics we identified the model for the BSCS that was most optimal in representing the data across the four samples for use in subsequent analyses.

Measurement invariance of the selected model for the BSCS and one-factor model of the self-discipline scale across national samples was tested using multi-group CFA. Three levels of measurement invariance were examined by progressively constraining the parameter estimates of the models to be equal cross the groups, in order to demonstrate configural invariance (no equality constraints), metric invariance (factor loadings constrained to equality), and strong invariance (factor loadings and intercepts constrained to equality) (Byrne et al. 1989). Full measurement invariance was supported when the fit of the restricted metric and strong invariance models did not differ substantially from the configural model, marked by a change in the value of CFI by less than .01 (Cheung and Rensvold 2002). Partial metric invariance was demonstrated when change in CFI was less than .01 after removing equality constraints with the highest modification indices.

Discriminant validity of subscales from the selected two-factor model of BSCS and the self-discipline factor was assessed by estimating three-factor CFA models in each sample. Discriminant validity of the factors was supported if the 95% confidence interval of the correlation between the factors did not encompass unity and if removal of a constraint fixing the factor correlation to unity resulted in a significant change in model fit according to the Wald test.

We examined the predictive validity of the self-control and self-discipline scales using structural equation modelling. Constructs from the selected two-factor model of the BSCS and the self-discipline construct were set as predictors of self-reported binge drinking, exercise, and healthy eating. Specifically, latent factors representing the subscales of adequately fitting two-factor model of the brief-trait self-control scale and the self-control factors were set as predictors of each health-related behavior. Adequacy of the models in accounting for the data was evaluated using the same criteria used to evaluate the CFA models. Structural parameters with their associated confidence intervals were used to evaluate the relative contribution of each factor in the prediction of each health behavior.

[1] Data files, analysis scripts and output used in our data analyses are provided on the Open Science Framework Project for this study: https://osf.io/r36jt/

Results

Preliminary Analyses

Distributional properties of the data set from each national sample were examined prior to data analysis. Although there are no established cutoff values on the acceptable percentage of missing data, rates of missing data should be kept to a minimum (e.g., 5% or less; Dong and Peng 2013). There were no missing data points in the datasets from Estonia and Luxemburg, only one missing data point in the Spanish dataset, and no systematic pattern of missing the in the UK dataset (missing data = 2.42%). Missing data was imputed using full-information maximum likelihood estimation in Mplus. Skewness and kurtosis values were within acceptable cutoff values for items from the brief self-control and self-discipline scales indicating few instances of departures from normality.

Factorial Validity

Goodness-of-fit of the one- and two-factor CFA models of self-control for the full sample and each national sample are presented in Table 1. Solution estimates for the full sample and each individual sample are presented in Tables 2 and 3, respectively. The Maloney et al. two-factor model consistently yielded acceptable fit with the data in the full sample and each of the four national samples based on the multiple criteria for goodness-of-fit (CFI range = .92 to .97; RMSEA range = .034 to .070; SRMSR range = .036 to .082). By comparison, indices for the Tangney et al. (CFI range = .71 to .84; RMSEA range = .079 to .099; SRMSR range = .058 to .090) and de Ridder et al. (CFI range = .74 to .94; RMSEA range = .051 to .103; SRMSR range = .043 to .094) models fell below acceptable values in the full sample and most of the national samples. Examination of the solution estimates revealed at least two factor loadings at or below .40 for the Tangney et al. one-factor model in each sample. Factor loadings for Maloney et al. and de Ridder et al. two-factor models were within acceptable range in most cases, but on the low side in a few cases, with a few falling outside this range particularly for the de Ridder model (range = .34 to .87). Reliability and AVE estimates were acceptable in most samples, but fell below acceptable levels for the restraint scale for the Maloney et al. model and the initiation scale for the de Ridder et al. model.

Goodness-of-fit statistics and solution estimates for the one-factor model of self-discipline are presented in Tables 1 and 4, respectively. The model exhibited satisfactory goodness-of-fit indices in the full sample and all national samples (CFI range = .86 to .95; RMSEA range = .066 to .094; SRMSR range = .036 to .075) once the error variances for some items were correlated. This indicated some redundancy in the error variance across items that was not accounted for by the latent factor. Examination of the solution estimates revealed that overall, items 2, 4, and 10 did not perform well in terms of their relative contribution to the overall factor, meaning that the self-discipline factor was generally defined by a smaller subset of items.

Measurement Invariance

As goodness of fit statistics fell below acceptable levels for the Tangney et al. and de Ridder et al. models for the brief self-control scale, we restricted our invariance tests to the Maloney et al. two-factor model. Results of the invariance analyses are provided in Table 5. While configural invariance for the model was established across all four national samples, full metric invariance could only be confirmed across the Estonia and Luxembourg samples. We did, however, find partial metric invariance across other pairs of national samples, indicating that while constraining the majority of factor loadings to equality led to few misspecifications in model comparisons, a select few were non-invariant. Specifically, factor loadings for items 1 ("I am good at resisting temptation") and 2 ("I have a hard time breaking bad habits") were set to be freely estimated (i.e., not constrained to be invariant) across the Estonia and Spain samples, and across the UK and Spanish samples, and factor loadings for items 7 ("I wish I had more self-discipline"), 8 ("People would say that I have iron self-discipline"), and 12 ("Sometimes I can't stop myself from doing something, even if I know it is wrong") were set to be freely estimated across the Estonia and UK samples, across the Luxembourg and Spanish samples, and across the Luxembourg and UK samples. The lack of invariance for these parameters notwithstanding, we found that the majority of factor loadings were equivalent across the four samples.

We also tested the measurement invariance of the one-factor self-discipline scale across samples. Results of the invariance analyses are provided in Table 6. Analyses provided support for configural invariance in all tests, with the exception of the analysis comparing the Luxembourg and Spanish samples, which exhibited substantial misspecification in the configural model. We found support for metric invariance in all samples comparisons, with the exception of the analysis for the Estonia and Spanish samples. Fit of the metric invariance models was substantially improved when item 3 ("I find it difficult to get down to work") was set to be non-invariant for the analysis comparing the model in the Estonia and Luxembourg samples, and marginally improved item 2 ("I get my chores done right away") was set to be non-invariant for the analyses comparing the model in the Luxembourg and UK samples and the Spanish and UK samples. These items had the largest modification indexes. However, in no case did we find support for strong invariance in any of the analyses. Overall, results provide general support for metric invariance for the self-discipline scale across samples.

Table 1 Goodness-of-Fit Statistics of the Proposed One-Factor and Two-Factor Confirmatory Factor Analytic Models of the Brief Self-Control and Self-Discipline Scales in the Full Sample and Four National Samples

Model	χ^2R	*df*	CFI	RMSEA	RMSEA 90% CI	SRMR
Full sample (N = 1282)						
Tangney	660.92	65	.78	.085	.079	.090
Maloney	140.153	19	.92	.071	.060	.082
de Ridder	352.381	34	.84	.085	.078	.094
One-factor self-discipline model	402.04	35	.89	.090	.083	.098
Modified self-discipline model	211.87	32	.95	.066	.058	.075
Three-factor model	742.59	129	.89	.061	.057	.065
Estonia (N = 347)						
Tangney	206.59	65	.84	.079	.067, .092	.058
Maloney	49.84	19	.94	.068	.046, .092	.043
de Ridder	64.23	34	.94	.051	.031, .069	.043
One-factor self-discipline model	154.25	35	.89	.099	.083, .115	.051
Modified self-discipline model[b]	99.02	34	.94	.074	.057, .091	.045
Three-factor model	349.95	131	.89	.069	.061, .078	.057
Luxembourg (N = 207)						
Tangney	152.77	65	.80	.081	.064, .097	.065
Maloney	24.92	19	.97	.039	.000, .076	.039
de Ridder	77.01	34	.87	.078	.055, .101	.059
One-factor self-discipline model	110.09	35	.85	.102	.081, .124	.072
Modified self-discipline model	85.36	34	.90	.085	.063, .108	.058
Three-factor model	307.09	134	.80	.079	.067, .091	.170
Spanish (N = 291)						
Tangney	202.89	65	.71	.085	.072, .099	.067
Maloney	25.39	19	.97	.034	.000, .065	.036
de Ridder	91.58	34	.82	.076	.058, .095	.060
One-factor self-discipline model	216.69	35	.69	.134	.117, .151	.090
Modified self-discipline model[a]	113.73	32	.86	.094	.075, .113	.069
Three-factor model	293.25	129	.83	.066	.056, .076	.069
UK (N = 437)						
Tangney	341.38	65	.71	.099	.088, .109	.073
Maloney	59.35	19	.92	.070	.050, .090	.042
de Ridder	192.85	34	.74	.103	.089, .118	.069
One-factor self-discipline model	144.44	35	.91	.085	.071, .099	.052
Modified self-discipline model	118.43	34	.93	.075	.061, .090	.051
Three-factor model	284.09	131	.93	.052	.043, .060	.052

Note. χ^2 R = Robust chi-square statistic; *df* = Degrees of freedom for chi-square statistic; CFI = Comparative fit index; RMSEA = Root mean square error of approximation; CI = Confidence intervals; SRMSR = Standardized root mean square of residuals. Tangney = Tangney et al.'s (2004) one-factor model of the brief self-control scale; Maloney = Maloney et al.'s (2012) two-factor model of the brief self-control scale; de Ridder = de Ridder et al.'s (2011) two-factor model of the brief self-control scale; Three-factor model = Three factor model comprising Maloney et al.'s two factor model of the BSCS and the one-factor self-discipline model. In Maloney et al.'s (2012) two-factor model, the restraint factor comprised items 1, 2, 7, 8 from the brief self-control scale, and the non-impulsivity factor comprised items 5, 9, 12, 13. In de Ridder et al.'s (2011) two-factor model, the inhibitory self-control factor comprised items 1, 2, 5, 6, 9, 12, and the initiatory self-control factor comprised items 3, 10, 11, 13; [a] items 1 and 3, items 2 and 10, and items 6 and 8, were set to be correlated; [b] items 6 and 8 were set to be correlated; [c] items 2 and 10 were set to be correlated

Table 2 Means, Standard Deviations, Reliability Coefficients, Average Variance Extracted, and Standardized Factor Loadings of the Proposed One- and Two-Factor Confirmatory Factor Analytic Models of the BSCS for the Full Sample

λ	Self-control	Maloney		de Ridder	
		Restraint	Non-Impulsivity	Inhibitory self-control	Initiatory self-control
SC1	.52	.56	-	.50	-
SC2	.54	.57	-	.55	-
SC3	.56	-	-	-	.59
SC4	.43	-	-	-	-
SC5	.58	-	.62	.63	-
SC6	.43	-	-	.46	-
SC7	.55	.63	-	-	-
SC8	.52	.55	-	-	-
SC9	.46	-	.39	.46	-
SC10	.45	-	-	-	.51
SC11	.30	-	-	-	.34
SC12	.57	-	.70	.61	-
SC13	.40	-	.44	-	.40
Mean	3.28	3.01	3.33	3.45	3.25
SD	.62	.80	.79	.70	.72
α	.80	.66	.61	.70	.52
ρ	.81	.67	.62	.70	.52
AVE	.63	.27	.24	.35	.17

Note. Self-control = Tangney et al.'s (2004) one-factor model of self-control; Maloney = Maloney et al.'s (2012) two-factor model of the brief self-control scale; de Ridder = de Ridder et al.'s (2011) two-factor model of the brief self-control scale; [a] Factor loading not statistically significant. α = Cronbach alpha coefficient; ρ = Composite reliability coefficient; AVE = Average variance extracted

Discriminant Validity

We tested discriminant validity of the restraint, non-impulsivity, and self-discipline factors by computing latent factor correlations in three-factor CFA models for the full sample and each national sample. Fit statistics for the three-factor model are presented in Table 1. The models generally exhibited sub-optimal fit with the data with misspecifications largely attributable to the poor performance of some items. Discriminant validity statistics are presented in Table 7. Although latent factor correlations among the constructs were large and statistically significant, confidence intervals for each correlation did not encompass unity and the Wald test was statistically significant in all cases (*p*s < .001) providing support for discriminant validity.[2]

[2] For comparison, we provide correlations among all variables from the present study using composite (averaged) scales in Appendix C (supplemental materials). Correlations among the restraint, non-impulsivity, and self-discipline constructs using composite scales were substantially smaller (attenuated) than the correlations among the latent factors for the same variables (Table 7). This illustrates the effect of measurement error in attenuating correlations among scales constructs and the value of using latent constructs. Although correlations among the latent constructs were large, this did not alter our conclusions regarding discriminant validity of the constructs.

Predictive Validity

We examined the predictive validity of the self-control constructs from the Maloney et al. two-factor model (restraint and non-impulsivity factors) and self-discipline by simultaneously regressing scores for the three health behaviors (binge drinking, exercise, and healthy eating) on the self-control and self-discipline constructs in a series of structural equation models for each sample. Goodness-of-fit statistics of the models and standardized parameter estimates for the proposed effects are provided in Table 8. The most consistent effect was for the non-impulsivity component of self-control on binge drinking which was large, negative, and statistically significant in the Estonia ($\beta = -.61$, $p < .001$), Spanish ($\beta = -.60$, $p < .05$), and UK ($\beta = -.78$, $p < .001$) samples. The effect of the restraint component of self-control on exercise was also significant in the Luxembourg ($\beta = .43$, $p < .05$) and UK ($\beta = .24$, $p < .05$) samples. Finally, self-discipline significantly predicted exercise ($\beta = -.22$, $p < .05$) and binge drinking ($\beta = .32$, $p < .05$) in the Luxembourg and UK samples, respectively. However, neither effect was in the expected direction. Examination of the correlation matrices suggest that the latter effects are likely to be suppressor effects as the zero-order correlation between

Table 3 Means, Standard Deviations, Reliability Coefficients, Average Variance Extracted, and Standardized Factor Loadings of the Proposed One- and Two-Factor Confirmatory Factor Analytic Models of the BSCS in the Four National Samples

λ	Estonia (n = 347)					Luxembourg (n = 207)					Spanish (n = 291)					UK (n = 437)				
	SC	Maloney		de Ridder		SC	Maloney		de Ridder		SC	Maloney		de Ridder		SC	Maloney		de Ridder	
		Res	Non- Imp	Inhib	Init		Res	Non- Imp	Inhib	Init		Res	Non-Imp	Inhib	Init		Res	Non- Imp	Inhib	Init
SC1	.60	.61	-	.55	-	.41	.48	-	.41	-	.50	.53	-	.50	-	.51	.62	-	.44	-
SC2	.55	.52	-	.56	-	.46	.51	-	.47	-	.47	.64	-	.49	-	.34	.42	-	.29	-
SC3	.54	-	-	-	.64	.58	-	-	-	.60	.52	-	-	-	.55	.45	-	-	-	.44
SC4	.37	-	-	-	-	.40	-	-	-	-	.46	-	-	-	-	.41	-	-	-	-
SC5	.52	-	.57	.61	-	.53	-	.57	.57	-	.50	-	.53	.53	-	.49	-	.50	.51	-
SC6	.49	-	-	.57	-	.45	-	-	.44	-	.28	-	-	.30	-	.34	-	-	.34	-
SC7	.64	.73	-	-	-	.44	.54	-	-	-	.31	.32	-	-	-	.58	.71	-	-	-
SC8	.71	.80	-	-	-	.41	.44	-	-	-	.27	.20	-	-	-	.50	.58	-	-	-
SC9	.43	-	.48	.45	-	.64	-	.56	.66	-	.50	-	.44	.49	-	.61	-	.60	.65	-
SC10	.38	-	-	-	.45	.49	-	-	-	.57	.45	-	-	-	.49	.57	-	-	-	.63
SC11	.45	-	-	-	.51	.32	-	-	-	.36	.35	-	-	-	.34	.44	-	-	-	.44
SC12	.53	-	.66	.60	-	.65	-	.87	.70	-	.62	-	.77	.69	-	.49	-	.59	.53	-
SC13	.38	-	.40	-	.39	.46	-	.43	-	.49	.42	-	.49	-	.41	.48	-	.59	-	.49
Mean	3.13	2.73	3.11	3.13	3.13	3.40	3.25	3.35	3.36	3.56	3.63	3.42	3.79	3.81	3.57	3.10	2.84	3.17	3.00	3.35
SD	.61	.81	.74	.61	.71	.57	.72	.77	.68	.68	.56	.69	.73	.68	.69	.57	.75	.75	.64	.68
α	.82	.75	.60	.72	.55	.80	.56	.69	.72	.58	.75	.46	.62	.65	.47	.80	.67	.66	.65	.57
ρ	.82	.72	.61	.73	.57	.80	.56	.70	.72	.58	.75	.52	.66	.67	.50	.79	.65	.68	.62	.58
AVE	.69	.40	.23	.37	.20	.62	.19	.39	.37	.21	.52	.17	.26	.32	.21	.61	.28	.26	.27	.21

Note. SC = Tangney et al.'s (2004) one-factor model of self-control; Maloney = Maloney et al.'s (2012) two-factor model of the brief self-control scale; de Ridder = de Ridder et al.'s (2011) two-factor model of the brief self-control scale; Res = Restraint subscale; Non-Imp = Non-impulsivity subscale; Inhib = Inhibitory self-control; Init = Initiatory self-control subscale; [a] Factor loading not statistically significant. α = Cronbach alpha coefficient; ρ = Composite reliability coefficient; AVE = Average variance extracted

Table 4 Means, Standard Deviations, Reliability Coefficients, Average Variance Extracted, and Standardized Factor Loadings for the One-Factor Model for the SelfDiscipline Scale in the Full Sample and Four National Samples

National sample loading (λ)	*M*	*SD*	α	ρ	AVE	Standardized factor									
						Item 1	Item 2	Item 3	Item 4	Item 5	Item 6	Item 7	Item 8	Item 9	Item 10
Full sample (N = 1282)	3.20	.68	.85	.84	.71	.72	.46	.73	.41	.67	.63	.63	.55	.69	.35
Estonia (n = 347)	3.20	.69	.86	.86	.82	.76	.38	.72	.33	.69	.68	.71	.63	.69	.45
Luxembourg (n = 207)	3.18	.60	.81	.84	.66	.75	.21	.82	.36	.64	.58	.55	.58	.69	.08[a]
Spanish (n = 291)	3.47	.62	.80	.84	.54	.68	.36	.57	.50	.69	.49	.49	.45	.52	.35
UK (n = 437)	3.04	.68	.88	.87	.82	.75	.59	.80	.52	.67	.68	.62	.56	.72	.39

Note. [a] Factor loading not statistically significant. α = Cronbach alpha coefficient; ρ = Composite reliability coefficient; AVE = Average variance extracted

these factors and the respective behavior was not significant and negative in the Luxembourg and UK samples, respectively.

Discussion

The purpose of the present study was to examine the factor structure of the BSCS and the self-discipline scale from the NEO-PI-R in four national samples, test the invariance of the structure of both scales across the samples, and test the predictive validity of the scales in predicting health-related behaviors related to self-control. Specifically, we tested three candidate models that aimed to describe the underpinning structure of the BSCS using confirmatory factor analysis: the one-factor model originally proposed by Tangney et al. (2004), and the two-factor models proposed by Maloney et al. (2012) and de Ridder et al. (2011). Based on our evaluation of the effectiveness of the different models in accounting for scores from the brief self-control scale, we aimed to assess the discriminant validity of the most appropriate model of self-control and the self-discipline scale across the samples. Finally, pending support for discriminant validity, we tested the validity of the self-control factor or factors and self-discipline scale in predicting variance in binge drinking, exercise, and healthy eating using structural equation modelling.

Results revealed that the Maloney et al. model exhibited the most consistent goodness-of-fit statistics producing well-fitting models in the full sample and across the four samples. Neither the Tangney et al. one-factor model nor the two-factor de Ridder et al. model exhibited satisfactory goodness-of-fit in the full sample and national samples. These models were abandoned in favor of the two-factor Maloney model, which segregated the BSCS into restraint and non-impulsivity factors. The one-factor self-discipline model fit the data well in all four samples, although the solution estimates indicated low factor loadings for selected items. Invariance tests of the Maloney et al. two-factor model and the one-factor self-discipline model indicated support for partial metric invariance with factor loadings invariant across the four samples with a few exceptions. Discriminant validity tests based on the intercorrelations among the non-impulsivity, restraint, and self-discipline factors supported discriminant validity. Finally, structural equation models in which the self-control factors from the Maloney et al. model predicted binge drinking, exercise, and healthy eating indicated a prominent role for the non-impulsivity factor in predicting binge drinking in all but the Luxembourg sample, and restraint in predicting exercise behavior in the Luxembourg and UK samples.

Current analyses provide additional support for the multidimensionality of trait self-control based on Tangney et al.'s (2004) brief self-control scale. Our findings extend previous research by (i) providing further confirmation of the inadequacy of a one-factor model of self-control in multiple samples from different national groups; (ii) demonstrating the effectiveness of a two-factor model based on Maloney et al.'s (2012) original analysis in fitting data from multiple samples relative to the one-factor model and a competing two-factor model proposed by de Ridder et al. (2011); and (iii) providing evidence that the restraint and non-impulsivity factors from the Maloney et al. two-factor model achieve discriminant validity. A multi-dimensional conceptualization of trait self-control also fits well with contemporary and previous self-control theories. For example, Metcalfe and Mischel (1999) suggest that effective behavioral control is subject to restraint tendencies which may moderate or regulate the more impulsive, emotion-driven pathways to action consistent a 'hot' vs. 'cool' distinction in pathways to action. Interestingly, Tangney et al. conducted an exploratory factor analysis of their scale and differentiated between factors they termed self-discipline and others such as 'impulsivity' and 'work ethic', but focused on the overall scores of the scale due to finding substantive correlations among the factors. However, we argue that despite the significant inter-correlations among the factors, the distinction is valid as criterion for discriminant validity was satisfied in our current analysis. Although the factors are not

Table 5 Measurement Invariance and Overall Fit Indexes for Maloney et al.'s (2012) Two-Factor Measurement Model of the BSCS in the Four National Samples

Model	χ^2R	*df*	CFI	ΔCFI	RMSEA	RMSEA 90% CI
Estonia - Luxembourg						
M1: Configural invariance	74.25	38	.95	-	.059	.039, .078
M2: Metric invariance	77.54	44	.96	-.004	.052	.032, .071
M3: Strong invariance	380.63	52	.56	-.398	.151	.137, .165
Estonia - Spanish						
M1: Configural invariance	75.37	38	.95	-	.056	.037, .074
M2: Metric invariance	104.74	44	.92	-.032	.066	.050, .082
M2P1: Partial metric invariance[a]	89.65	43	.94	-.013	.058	.041, .075
M2P2: Partial metric invariance[b]	77.83	42	.95	.002	.052	.033, .069
M3: Strong invariance	560.11	50	.31	-.638	.179	.166, .192
Estonia - UK						
M1: Configural invariance	109.44	38	.93	-	.069	.054, .085
M2: Metric invariance	128.21	44	.92	-.012	.070	.056, .084
M2P1: Partial metric invariance[c]	121.68	43	.92	-.007	.068	.054, .083
M3: Strong invariance	564.13	51	.50	-.426	.160	.148, .172
Luxembourg - Spanish						
M1: Configural invariance	50.30	38	.97	-	.036	.000, .061
M2: Metric invariance	62.42	44	.96	-.014	.041	.011, .063
M2P1: Partial metric invariance[d]	58.38	43	.97	-.007	.038	.000, .061
M3: Strong invariance	194.13	51	.67	-.286	.106	.091, .122
Luxembourg - UK						
M1: Configural invariance	84.52	38	.93	-	.062	.044, .079
M2: Metric invariance	101.68	44	.92	-.016	.064	.048, .080
M2P1: Partial metric invariance[e]	91.81	43	.93	-.003	.059	.043, .076
M3: Strong invariance	231.80	51	.74	-.187	.105	.091, .119
Spanish - UK						
M1: Configural invariance	85.87	38	.93	-	.059	.042, .075
M2: Metric invariance	121.58	44	.89	-.042	.070	.055, .084
M2P1: Partial metric invariance[a]	101.77	43	.92	-.015	.061	.046, .077
M2P2: Partial metric invariance[b]	94.75	42	.93	-.006	.059	.043, .075
M3: Strong invariance	413.38	50	.49	-.439	.141	.129, .154

Note. χ^2 R = Robust chi-square; *df* = Degrees of freedom; CFI = Comparative fit index; RMSEA = Root mean square error of approximation; CI = Confidence interval. Factor loading of the item with highest modification index was set to be freely estimated, including: [a] Factor loading of the item with highest modification index (item 2) was set to be non-invariant; [b] Factor loadings of the items with highest modification indexes (items1 and 2) were set to be non-invariant; [c] Factor loading of the item with highest modification index (item 8) was set to be non-invariant; [d] Factor loading of the item with highest modification index (item 7) was set to be non-invariant; [e] Factor loading of the item with highest modification index (item 12) was set to be non-invariant

entirely orthogonal, substantial variance in each remains unexplained when examining the coefficients of determination for the intercorrelations. Aggregating responses to the BSCS may, therefore, mask or confound the effects of the separate components of self-control in research predicting important cognitive and behavioral outcomes relating to self-regulation. Our research also provides robust evidence for a two-factor structure by replicating it in multiple samples from different national groups. Given the invariance in factor structure across groups, we advocate differentiation of the restraint and non-impulsivity constructs in future research adopting the brief self-control model. This will provide better evaluation of the aspects of self-control most likely to account for variance in cognitive and behavioral outcomes and provide more comprehensive tests of self-regulatory processes underpinning action.

Tests of discriminant validity of the non-impulsivity and restraint factors from the Maloney et al. two-factor model with the self-discipline scale from the NEO-PI-R is also an important contribution of the current research. The presence of 'jangle' fallacies in social and personality psychology (Block

Table 6 Measurement Invariance and Overall Fit Indexes for the Measurement Model of the One-Factor Self-Discipline Scale in the Four National Samples

Model[a]	χ^2R	*df*	CFI	ΔCFI	RMSEA	RMSEA 90% CI
Estonia - Luxembourg						
M1: Configural invariance	118.53	58	.96	-	.061	.045, .077
M2: Metric invariance	152.51	68	.95	-.016	.067	.053, .081
M2P1: Partial metric invariance[b]	142.03	67	.95	-.010	.064	.049, .078
M3: Strong invariance	252.63	77	.89	-.076	.091	.078, .103
Estonia - Spanish						
M1: Configural invariance	140.497	49	.94	-	.077	.062, .091
M2: Metric invariance	164.00	59	.94	-.008	.075	.061, .088
M3: Strong invariance	281.93	69	.87	-.075	.098	.087, .110
Estonia - UK						
M1: Configural invariance	77.36	49	.99	-	.040	.023, .055
M2: Metric invariance	116.07	59	.98	-.012	.050	.036, .063
M3: Strong invariance	251.82	69	.92	-.066	.082	.071, .093
Luxembourg - Spanish						
M1: Configural invariance	180.02	57	.89	-	.093	.078, .109
M2: Metric invariance	199.60	67	.88	-.009	.089	.075, .104
M3: Strong invariance	275.79	77	.82	-.069	.102	.089, .115
Luxembourg - UK						
M1: Configural invariance	86.24	47	.98	-	.051	.034, .068
M2: Metric invariance	126.97	57	.96	-.017	.062	.047, .076
M2P1: Partial metric invariance[c]	100.51	56	.98	-.003	.050	.034, .065
M3: Strong invariance	253.77	66	.90	-.083	.094	.082, .106
Spanish - UK						
M1: Configural invariance	114.04	45	.96	-	.065	.050, .080
M2: Metric invariance	147.04	55	.95	-.012	.068	.055, .081
M2P1: Partial metric invariance[c]	137.10	54	.96	-.008	.065	.052, .079
M3: Strong invariance	314.35	64	.87	-.097	.104	.092, .115

Note. χ^2 R = Robust chi-square; *df* = Degrees of freedom; CFI = Comparative fit index; RMSEA = Root mean square error of approximation; CI = Confidence interval. [a] Correlations among error variances with highest modification indexes from the single-sample CFAs were included in each model; [b] Factor loading with highest modification index (item 3) set to be non-invariant across samples; [c] Factor loading with highest modification index (item 2) set to be non-invariant across samples

1995; Hagger 2014), that is, multiple constructs with similar content going by different terms, presents considerable problems for researchers seeking to identify a narrow, parsimonious set of factors that predict cognitive and behavioral responses. The terms self-control, self-discipline, and even conscientiousness, have been used interchangeably and, in doing so, researchers have implied considerable overlap or redundancy in the constructs at the conceptual level. For example, definitions of self-control as a capacity to inhibit impulses, responses, urges, habitual actions, and dominant responses appear also to overlap with the definition of self-discipline as the capacity to begin tasks and follow them through to completion despite boredom or distractions (Duckworth and Seligman 2005). In fact, Tangney et al. (2004), in their original development of their self-control scales, make explicit reference to self-discipline in their definition: “More generally, breaking habits, resisting temptation, and keeping good self-discipline all reflect the ability of the self to control itself, and we sought to build our scale around them” (p. 275).

The conceptual overlaps notwithstanding, our data indicates that despite sharing considerable variance, both the restraint and non-impulsivity factors were distinct from the self-discipline factor. Although the range of correlations among the self-control components and self-discipline were large in magnitude based on Cohen’s taxonomy of effect sizes, substantial variance in the two factors remains unexplained. Coupled with support for discriminant validity, current data provide little support for empirical overlap in the constructs suggested by conceptual similarly. Our research suggests, therefore, that multidimensional trait self-control and self-

Table 7 Latent Inter-Factor Correlations and Discriminant Validity Statistics for the Restraint, Impulsivity, and Self-Discipline Factors in the Four National Samples

	Res ↔ Non-Imp	Res ↔ SD	Non-Imp ↔ SD
Estonia (*n* = 347)			
Inter-factor correlation	.68	.64	.63
CI	.56, .80	.53, .75	.51, .76
Wald test	29.07***	38.87***	34.07***
Luxembourg (*n* = 207)			
Inter-factor correlation	.73	.52	.68
CI	.54, .92	.34, .71	.52, .89
Wald test	8.81**	19.10***	19.30***
Spanish (*n* = 291)			
Inter-factor correlation	.76	.63	.71
CI	.56, .96	.44, .82	.50, .92
Wald test	6.72**	15.69***	9.02**
UK (*n* = 437)			
Inter-factor correlation	.60	.54	.77
CI	.47, .72	.43, .65	.68, .85
Wald test	28.07***	56.81***	32.45***

Note. Res = Restraint; Non-Imp = Non-impulsivity; SD = Self-discipline; CI = 95% confidence intervals of latent factor correlations. Wald = Wald test constraining value of the latent-factor correlation to zero. * $p < .05$ ** $p < .01$ *** $p < .001$

discipline likely tap different aspects of self-control. Self-discipline may be a more ‘focused’ construct than self-control in that it focuses on goal-directed actions that lead to better self-regulations, consistent with its overarching trait of conscientiousness (Zimmerman and Kitsantas 2014). Although the BSCS, particularly the restraint component from the two-factor model, may make reference to working toward distal goals, it encompasses more than just a focus on “hard work” toward goals (Hagger and Hamilton 2018). Of course, while these possible conceptual distinctions and formal tests of factorial validity may point to distinctions between the constructs, the high correlations may present problems for predictive validity when the constructs are used to predict cognitive and behavioral outcomes.

An additional point worth noting is the low factor loadings for some of the items for the self-discipline scale. While the focus of the current study was on the discriminant and concurrent validity of the BSCS, our factor analyses also permitted an examination of the factor structure of the self-discipline scale from the NEO-PI-R. Our findings indicated that some of the items performed relatively poorly in indicating the latent self-discipline factor in all samples. Although this is not a problem for testing discriminant validity per se because the latent factor for self-discipline is largely indicated by the items with adequate loadings, it does suggest that the scale items may not perform well in capturing the essence of the construct. While a considerable body or research has supported the factor structure and integrity of the sub-facet scales from the NEO-PI-R (IPIP 2017), current results indicate that the scales may not perform optimally across samples and contexts, and points to the necessity of conducting rigorous factor analytic work prior to use of these scales.

A further important finding is the pattern of effects for the self-control and self-discipline factors in predicting behavioral outcomes. This is an important endeavor if researchers are to provide an evidence base of potentially modifiable factors that will serve as targets in behavior change interventions, such as intervention to promote increased participation in health behaviors (c.f., Hagger et al. 2018; Kok et al. 2016; Rich et al. 2015). Our findings indicate that the non-impulsivity component of self-control was negatively related to binge drinking behavior in three samples, while restraint was positively related to exercise in two of the samples. These results are consistent with our original hypotheses and recent theory that behaviors requiring active engagement and working toward attaining a distal goal (i.e., exercise) would be positively associated with restraint (Duckworth and Gross 2014; Hagger and Hamilton 2018). Similarly, we expected that suppressing cues and impulses to engage in a rewarding behavior (i.e., binge drinking) would be negatively associated with non-impulsivity. We also expected these factors to predict healthy eating, but neither were effective, perhaps indicating that eating behavior is more complex and may be accounted for by specific food-related cues and dietary restraint (Hofmann et al. 2007). The restraint aspect of trait self-control is consistent with Gottfredson and Hirschi’s (1990) hypothesis that individuals with good self-control are highly effective in recognizing the benefits and risks of their actions and the need to structure

Table 8 Goodness of Fit Statistics and Standardized Parameter Estimates of Structural Equation Models Predicting Health-Related Outcomes by Self-Control Dimensions from Maloney at al.'s Two-Factor Model and Self-Discipline

Sample and behavior	Model fit						Standardized parameter estimates (β)		
	$\chi^2_{S\text{-}B}$	*df*	CFI	RMSEA	RMSEA 90% CI	SRMSR	Res	Non- Imp	SD
Estonia (*n* = 347)									
Binge drinking	426.04	163	.87	.068	.060, .076	.057	.18	-.61***	.08
Exercise	378.81	163	.89	.062	.054, .070	.054	.14	.08	.23
Healthy eating	403.35	163	.89	.065	.057, .073	.056	-.05	.09	.10
Luxembourg (*n* = 207)									
Binge drinking	293.63	163	.87	.062	.051, .074	.064	-.10	-.17	.08
Exercise	284.33	163	.90	.063	.048, .071	.060	.43*	-.24	-.22*
Healthy eating	281.38	163	.90	.059	.047, .071	.061	.17	.05	.03
Spanish (*n* = 291)									
Binge drinking	342.42	161	.85	.062	.053, .071	.067	.26	-.60*	.09
Exercise	343.66	161	.88	.062	.053, .072	.067	.16	-.34	.25
Healthy eating	332.58	161	.88	.061	.051, .070	.065	.21	-.11	.17
UK (*n* = 437)									
Binge drinking	418.12	163	.92	.060	.053, .067	.054	.12	-.78***	.32*
Exercise	339.90	163	.93	.050	.042, .057	.044	.24**	-.06	.04
Healthy eating	318.26	163	.94	.047	.039, .054	.045	.04	-.03	.18

Note. χ^2 S-B = Robust Satorra-Bentler scaled chi-square statistic; *df* = Degrees of freedom for chi-square statistic; CFI = Comparative fit index; RMSEA = Root mean square error of approximation; CI = confidence interval; SRMSR = Standardized root mean square of residuals; Res = Restraint; Non-Imp = Non-impulsivity; SD = Self-discipline. $^{*}p < .05$ $^{**}p < .01$ $^{***}p < .01$

their environment accordingly to achieve long-term ends. Similarly, research in impulsivity and non-conscious pathways to action indicates that capacity for inhibiting impulsive tendencies and cues to well-learned behaviors that have previously been highly reinforced (e.g., binge drinking) is a major determinant of whether an individual will be more or less successful in regulating their behavior (e.g., Christiansen et al. 2012; Friese and Hofmann 2009).

Strengths and Limitations

The current research has a number of strengths. We adopted contemporary theory on trait self-control and personality to develop our hypotheses relating to the structure and validity of the BSCS and the self-discipline scale from the NEO-PI-R. We tested our hypotheses using fit-for-purpose analytic techniques that enabled us to specify hypothesized and competing model structures a priori and test them against our data. Our data was collected in multiple samples enabling us to test our hypotheses across multiple samples and national groups, and our analyses permitted formal comparisons of our hypothesized models across groups. It is, however, also important to acknowledge the limitations of the current research, which may constrain the generalizability of our findings and offer possible alternative interpretations. First, our data are correlational and cross-sectional and, therefore, do not enable us to infer the causal direction of our predictions beyond theory. For example, although we specified the self-control and self-discipline factors as predictors of health behaviors, the correlational data means that equally-plausible alternative models from an empirical perspective could be specified and would exhibit good fit with the data. Similar issues have been identified in research on other social psychological theories of intention and motivation (e.g., Hagger et al. 2016, 2018; Rich et al. 2015). Cross-lagged panel designs, in which measures of the self-control, self-discipline, and behavioral outcomes are collected across two time periods, would permit tests of the directionality or reciprocity of the proposed effects. Second, it is important to note that the factor loadings for some items from the Maloney et al. two-factor model self-control were sub-optimal. While two-factor structure may be optimal in terms of overall structural integrity and fit with the data, some items remain problematic and point to the need for further refinement and possible streamlining of items in future revisions. The poor performance of the items in some of the samples may have been due to participants' misunderstanding of some of the items. While the back-translation process indicated congruence in the translated and back-translated versions, it is still possible that there were semantic differences leading participants to respond to items differently across cultures. However, this interpretation remains speculative without further evidence. This issue may be resolved by conducting an additional study in which participants complete the translated scale using a 'think aloud' method (e.g., Darker and French 2009). This may capture how respondents interpret scale items and highlight any misunderstandings and remains an avenue for future research. Finally, the current student samples reflect a homogenous group which may not be representative of the general population, and replication in a general population should be considered in future.

Funding This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors. Martin S. Hagger's contribution was supported by a Finland Distinguished Professor (FiDiPro) fellowship from Business Finland, a Kennedy Y. H. Wong Distinguished Visiting Professorship from Hong Kong Baptist University, and a visiting Professorship from Université Paris Nanterre.

Compliance with Ethical Standards

Conflict of Interest On behalf of all authors, the corresponding author states that there is no conflict of interest.

References

Allom, V., Panetta, G., Mullan, B., & Hagger, M. S. (2016). Self-report and behavioural approaches to the measurement of self-control: Are we assessing the same construct? *Personality and Individual Differences, 90*, 137–142. https://doi.org/10.1016/j.paid.2015.10.051.

Arnautovska, U., Fleig, L., O'Callaghan, F., & Hamilton, K. (2017). A longitudinal investigation of older adults' physical activity: Testing an integrated dual-process model. *Psychology & Health, 32*, 166–185. https://doi.org/10.1080/08870446.2016.1250273.

Baumeister, R. F., & Heatherton, T. F. (1996). Self-regulation failure: An overview. *Psychological Inquiry, 7*, 1–15. https://doi.org/10.1207/s15327965pli0701_1.

Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin, 117*, 187–215. https://doi.org/10.1037/0033-2909.117.2.187.

Bracken, B. A., & Barona, A. (1991). State of the art procedures for translating, validating, and using psychoeducational tests in cross-cultural assessment. *School Psychology International, 12*, 119–132. https://doi.org/10.1177/0143034391121010.

Byrne, B. M., Shavelson, R. J., & Muthén, B. (1989). Testing for the equivalence of factor covariance and means structures: The issue of partial measurement invariance. *Psychological Bulletin, 105*, 456–466. https://doi.org/10.1037/0033-2909.105.3.456.

Carver, C. S. (2005). Impulse and constraint: Perspectives from personality psychology, convergence with theory in other areas, and potential for integration. *Personality and Social Psychology Review, 9*, 312–333. https://doi.org/10.1207/s15327957pspr0904_2.

Cheung, G. W., & Rensvold, R. B. (2002). Evaluating goodness-of-fit indexes for testing measurement invariance. *Structural Equation Modeling, 9*, 233–255. https://doi.org/10.1207/S15328007SEM0902_5.

Christiansen, P., Cole, J. C., & Field, M. (2012). Ego depletion increases ad-lib alcohol consumption: Investigating cognitive mediators and moderators. *Experimental and Clinical Psychopharmacology, 20*, 118–128. https://doi.org/10.1037/a0026623.

Darker, C. D., & French, D. P. (2009). What sense do people make of a theory of planned behaviour questionnaire? A think-aloud study. *Journal of Health Psychology, 14*, 861–871. https://doi.org/10.1177/1359105309340983.

de Ridder, D. T. D., de Boer, B. J., Lugtig, P., Bakker, A. B., & van Hooft, E. A. J. (2011). Not doing bad things is not equivalent to doing the right thing: Distinguishing between inhibitory and initiatory self-control. *Personality and Individual Differences, 50*, 1006–1011. https://doi.org/10.1016/j.paid.2011.01.015.

de Ridder, D. T. D., Lensvelt-Mulders, G., Finkenauer, C., Stok, F. M., & Baumeister, R. F. (2012). Taking stock of self-control: A meta-analysis of how trait self-control relates to a wide range of behaviors. *Personality and Social Psychology Review, 16*, 76–99. https://doi.org/10.1177/1088868311418749.

de Vries, R. E., & van Gelder, J.-L. (2013). Tales of two self-control scales: Relations with five-factor and HEXACO traits. *Personality and Individual Differences, 54*, 756–760. https://doi.org/10.1016/j.paid.2012.12.023.

Dong, Y., & Peng, C.-Y. J. (2013). Principled missing data methods for researchers. *SpringerPlus, 2*, 222. https://doi.org/10.1186/2193-1801-2-222.

Duckworth, A. L., & Gross, J. J. (2014). Self-control and grit: Related but separable determinants of success. *Current Directions in Psychological Science, 23*, 319–325. https://doi.org/10.1177/0963721414541462.

Duckworth, A. L., & Seligman, M. E. P. (2005). Self-discipline outdoes IQ in predicting academic performance of adolescents. *Psychological Science, 16*, 939–944. https://doi.org/10.1111/j.1467-9280.2005.01641.x.

Ent, M. R., Baumeister, R. F., & Tice, D. M. (2015). Trait self-control and the avoidance of temptation. *Personality and Individual Differences, 74*, 12–15. https://doi.org/10.1016/j.paid.2014.09.031.

Fishbach, A., & Shah, J. Y. (2006). Self-control in action: Implicit dispositions toward goals and away from temptations. *Journal of Personality and Social Psychology, 90*, 820–832.

Forstmeier, S., Drobetz, R., & Maercker, A. (2011). The delay of gratification test for adults: Validating a behavioral measure of self-motivation in a sample of older people. *Motivation and Emotion, 35*, 118–134. https://doi.org/10.1007/s11031-011-9213-1.

Friese, M., & Hofmann, W. (2009). Control me or I will control you: Impulses, trait self-control, and the guidance of behavior. *Journal of Research in Personality, 43*, 795–805. https://doi.org/10.1016/j.jrp.2009.07.004.

Galla, B. M., & Duckworth, A. L. (2015). More than resisting temptation: Beneficial habits mediate the relationship between self-control and positive life outcomes. *Journal of Personality and Social Psychology, 109*, 508–525. https://doi.org/10.1037/pspp0000026.

Godin, G., & Shephard, R. J. (1985). A simple method to assess exercise behavior in the community. *Canadian Journal of Applied Sport Science, 10*, 141–146.

Gottfredson, M. R., & Hirschi, T. (1990). *A general theory of crime*. Stanford: Stanford University Press.

Grasmick, H. G., Tittle, C. R., Bursik Jr., R. J., & Arneklev, B. J. (1993). Testing the core empirical implications of Gottfredson and Hirschi's general theory of crime. *Journal of Research in Crime and Delinquency, 30*, 5–29.

Hagger, M. S. (2014). Avoiding the 'déjà-variable' phenomenon: Social psychology needs more guides to constructs. *Frontiers in Psychology, 5*, 52. https://doi.org/10.3389/fpsyg.2014.00052.

Hagger, M. S., & Chatzisarantis, N. L. D. (2005). First- and higher-order models of attitudes, normative influence, and perceived behavioural control in the theory of planned behaviour. *British Journal of Social Psychology, 44*, 513–535. https://doi.org/10.1348/014466604X16219.

Hagger, M. S., & Hamilton, K. (2018). Grit and self-discipline as predictors of effort and academic attainment. *British Journal of Educational Psychology*. https://doi.org/10.1111/bjep.12241.

Hagger, M. S., Lonsdale, A., Hein, V., Koka, A., Lintunen, T., Pasi, H. J., . . . Chatzisarantis, N. L. D. (2012). Predicting alcohol consumption and binge drinking in company employees: An application of planned behaviour and self-determination theories. *British Journal* *of Health Psychology, 17*, 379–407. https://doi.org/10.1111/j.2044-8287.2011.02043.x.

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., 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*. https://doi.org/10.1016/j.socscimed.2018.07.038.

Hamilton, K., Fleig, L., & Hagger, M. S. (2018). Being active in pregnancy: Theory-based factors associated with physical activity among pregnant women. *Women & Health*. https://doi.org/10.1080/03630242.2018.1452835.

Hofmann, W., Rauch, W., & Gawronski, B. (2007). And deplete us not into temptation: Automatic attitudes, dietary restraint, and self-regulatory resources as determinants of eating behavior. *Journal of Experimental Social Psychology, 43*, 497–504. https://doi.org/10.1016/j.jesp.2006.05.004.

Hofmann, W., Friese, M., & Strack, F. (2009). Impulse and self-control from a dual-systems perspective. *Perspectives on Psychological Science, 4*, 162–176. https://doi.org/10.1111/j.1745-6924.2009.01116.x.

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

Inzlicht, M., & Schmeichel, B. J. (2012). What is ego depletion? Towards a mechanistic revision of the resource model of self-control. *Perspectives on Psychological Science, 7*, 450–463.

IPIP. (2017). Self-discipline scale from the NEO-PI-R. Retrieved April 1, 2017, from http://ipip.ori.org/newNEOKey.htm#Self-Discipline

Kok, G., Gottlieb, N. H., Peters, G.-J. Y., Mullen, P. D., Parcel, G. S., Ruiter, R. A. C., . . . Bartholomew, L. K. (2016). A taxonomy of behavior change methods: An intervention mapping approach. *Health Psychology Review, 10*, 297–312. https://doi.org/10.1080/17437199.2015.1077155.

Kuhl, J. (1984). Volitional aspects of achievement motivation and learned helplessness: Toward a comprehensive theory of action control. In B. A. Maber (Ed.), *Progress in experiemental personality research* (pp. 99–171). New York: Academic Press.

Kuhl, J. (2000). A functional design approach to motivation and self regulation: The dynamics of personality systems interactions. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Self-regulation: Directions and challenges for future research* (pp. 111–169). New York: Academic Press.

Lindner, C., Nagy, G., & Retelsdorf, J. (2015). The dimensionality of the Brief Self-Control Scale—An evaluation of unidimensional and multidimensional applications. *Personality and Individual Differences, 86*, 465–473. https://doi.org/10.1016/j.paid.2015.07.006.

Maloney, P. W., Grawitch, M. J., & Barber, L. K. (2012). The multi-factor structure of the brief self-control scale: Discriminant validity of restraint and impulsivity. *Journal of Research in Personality, 46*, 111–115. https://doi.org/10.1016/j.jrp.2011.10.001.

Marcus, B. (2003). An empirical examination of the construct validity of two alternative self-control measures. *Educational and Psychological Measurement, 63*, 674–706.

McCrae, R. R., & Costa Jr., P. T. (2004). A contemplated revision of the NEO five-factor inventory. *Personality and Individual Differences, 36*, 587–596. https://doi.org/10.1016/S0191-8869(03)00118-1.

Metcalfe, J., & Mischel, W. (1999). A hot/cool-system analysis of delay of gratification: The dynamics of willpower. *Psychological Review, 106*, 3–19. https://doi.org/10.1037/0033-295X.106.1.3.

Muthén, L. K., & Muthén, B. O. (2015). *Mplus statistical software version 7.31*. Los Angeles: Muthén & Muthén.

Neal, D. J., & Carey, K. B. (2005). A follow-up psychometric analysis of the self-regulation questionnaire. *Psychology of Addictive Behaviors, 19*, 414–422. https://doi.org/10.1037/0893-164x.19.4.414.

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*, 673–688. https://doi.org/10.1007/s10865-015-9644-3.

Satorra, A., & Bentler, P. M. (1988). *Scaling corrections for statistics in covariance structure analysis*. Los Angeles: University of California at Los Angeles, Department of Psychology.

Strack, F., & Deutsch, R. (2004). Reflective and impulsive determinants of social behavior. *Personality and Social Psychology Review, 8*, 220–247. https://doi.org/10.1207/s15327957pspr0803_1.

Tangney, J. P., Baumeister, R. F., & Boone, A. L. (2004). High self-control predicts good adjustment, less pathology, better grades, and interpersonal success. *Journal of Personality, 72*, 271–324. https://doi.org/10.1111/j.0022-3506.2004.00263.x.

Williams, M. W. M., Fletcher, R. B., & Ronan, K. R. (2007). Investigating the theoretical construct and invariance of the self-control scale using confirmatory factor analysis. *Journal of Criminal Justice, 35*, 205–218. https://doi.org/10.1016/j.jcrimjus.2007.01.007.

Wills, T. A., Pokhrel, P., Morehouse, E., & Fenster, B. (2011). Behavioral and emotional regulation and adolescent substance use problems: A test of moderation effects in a dual-process model. *Psychology of Addictive Behaviors, 25*, 279–292. https://doi.org/10.1037/a0022870.

Zimmerman, B. J., & Kitsantas, A. (2014). Comparing students' self-discipline and self-regulation measures and their prediction of academic achievement. *Contemporary Educational Psychology, 39*, 145–155. https://doi.org/10.1016/j.cedpsych.2014.03.004.