Main Article

Ger J Exerc Sport Res
https://doi.org/10.1007/s12662-025-01050-z
Received: 22 December 2024
Accepted: 14 May 2025

Phil Ljubic[1] · Benedikt Broda[1] · Sabine Windmann[2] · Chris Englert[1]

[1] Department of Sports Sciences, Goethe University Frankfurt, Frankfurt/Main, Germany
[2] Department of Psychology, Goethe University Frankfurt, Frankfurt, Germany

Measuring habit strength in physical activity: Re-evaluation of the German Self-Report Habit Index and validation of the German version of the Self-Report Behavioural Automaticity Index

Introduction

Regular physical activity (PA) benefits physical and mental health. For example, PA reduces the risk of developing cardiovascular diseases and certain cancers (Garcia et al., 2023). Furthermore, it mitigates the severity of depressive disorders (Imboden, Claussen, Seifritz, & Gerber, 2021). The World Health Organization (WHO) recommends that adults should engage in at least 150–300 min of moderate-intensity PA per week or 75–150 min of vigorous-intensity PA per week or a combination of the two, supplemented by two sessions of muscle-strengthening activities involving all major muscle groups (Bull et al., 2020). However, more than one-third of the global adult population fails to meet the WHO recommendations (Strain et al., 2024). Although seven out of ten individuals intend to engage in PA on a regular basis, only about half of those actually engage in PA (Feil, Fritsch, & Rhodes, 2023). This discrepancy is referred to as the PA intention-behaviour gap (IBG; Sheeran, 2002).

Over the years, several theories have been developed to explain and bridge the IBG (Pfeffer, Hamilton, Hannan, & Wegner, 2023). For instance, classical social cognitive theories suggest that explicit processes, such as forming expectations about the potential outcome of a certain behaviour (e.g. "physical activity helps me to get in better shape"), are key determinants of PA behaviour (Rhodes, McEwan, & Rebar, 2019). More recent dual-process theories not only focus on explicit, consciously controlled processes but also stress the importance of implicit processes such as affects or habits which are not or only partially under conscious control. From the perspective of dual-process theories in the PA domain (e.g. Brand & Ekkekakis, 2018; Strobach, Englert, Jekauc, & Pfeffer, 2020), habits are highly important in explaining behaviour and behaviour change. A meta-analysis of several health-related outcomes revealed that habits can explain roughly 20% of the variation in nutrition and PA related behaviours (Gardner, de Bruijn, & Lally, 2011). Similarly, the application of habit theory to other behaviours (e.g. skill acquisition) appears useful. Given the importance of habits in shaping behaviour, the research interest in habits across scientific disciplines is considerable and has increased in recent years (Watson, O'Callaghan, Perkes, Bradfield, & Turner, 2022).

Nonetheless, there is no consensus on the definition of habits (Gardner et al., 2024b). Gardner defines habit as a "cognitive process whereby a contextual cue automatically triggers an impulse to act, based on a cue-behaviour association learnt through consistent repetition in the presence of the cue" (Gardner, 2015, p. 280). They regard *habits* and *habitual behaviours* as two different processes where habitual behaviour is seen as a specific action instigated by the underlying habit (Gardner & Lally, 2023). In contrast, some define habit as a type of behaviour rather than a cognitive process, meaning that the origin of the habitual behaviour remains unspecified (e.g. Watson et al., 2022). Others define habit as a tendency to act in a specific manner in specific contexts, leaving the origin of this tendency unspecified (e.g. Ouellette & Wood, 1998).

The definition by Gardner (2015) has also raised criticism. De Houwer (2019) argues that the inclusion of additional causal elements (like learning processes) makes it even more difficult to establish whether a behaviour is habitual. Moreover, Orbell and Verplanken (2015) argue that including an impulse in the habit definition risks confusion with related concepts such as impulsivity, addiction, or

Published online: 26 June 2025

temptation. Hewig (2024) also questions the rationale behind the inclusion of an impulse, proposing that an impulse is already an internal response triggered by a cue-response association. While external responses refer to behaviour, internal responses refer to changes in for example thoughts, emotions, or memory. In sum, these researchers prefer to define habit as the stimulus-response association itself whereby internal responses are included.

Another question concerns the differences between habits and automatic responses. While some researchers see habit as a form of automaticity, there is also no consensus regarding the definition of automaticity (Moors & De Houwer, 2006). Moors and De Houwer (2006) concluded that (1) automaticity can be diagnosed by looking at the presence of automaticity features, (2) these features should be investigated separately, if possible, (3) these features should be construed as gradual, and (4) the degree to which a process is automatic should be evaluated relative to an adequate criterion. This conclusion raises the question of which automation features should be considered within the definition and measurement of habit. Bargh (1992) proposed that a core feature of automaticity is the autonomy of a response which "once started, did not require conscious guidance to run to completion" (Bargh, 1992, p. 186). Other features could be lack of awareness, control, intention as well as the efficiency of a response (Bargh, 1994). Considering Gardner's (2015) definition of habit, it seems relevant to specify which automatic features must be present in an observed association between cue and response for a habit to exist. Contrary to habits, other automatic processes such as involuntary muscle contractions are not the result of individually learned behaviour.

Over the years, several measures have been proposed to capture distinct aspects of the various habit definitions. Among these are the number of reinforcements of a behaviour (e.g. Hull, 1943), the frequency of a behaviour in the past (e.g. Bagozzi, 1981), the product of behavioural frequency and context stability (e.g. Wood, Quinn, & Kashy, 2002), the habit strength based on questionnaires (e.g. Mittal, 1988), and response-frequency measures (Verplanken, Aarts, van Knippenberg, & van Knippenberg, 1994). Based on the definition by Gardner (2015), the focus on frequency seems insufficient as it is the extent to which a respective response was triggered *automatically* that is relevant, not how often a behaviour has been performed. For instance, one might consider studying for a test. This behaviour can occur with considerable frequency but studying itself is inherently an explicitly controlled behaviour. Thus, more frequent studying will not be associated with more automatic studying while other non-varying aspects (studying routine, problem-solving strategies etc.) could undergo a habit formation.

The *Self-Report Habit Index* (SRHI; Verplanken & Orbell, 2003) is one of the most frequently used habit questionnaires (Gardner et al., 2011). Despite the inclusion of theoretical distinct features such as items regarding self-identity, the SRHI has been assumed to be unidimensional (e.g, Verplanken & Orbell, 2003). The SRHI consists of 12 items, each of which starts with the statement "[Behaviour] is something …" and is answered on a 5-, 7-, or 11-point Likert scale ranging from 1 (dis*agree*) to 5/7/11 (*agree*) with higher values representing a stronger habit. Recent studies in English (Morean et al., 2018), German (Opwis, Bartel, Salewski, & Schmidt, 2023), and Spanish (Reyes Fernández, Monge-Rojas, Solano López, & Cardemil, 2019) across different contexts found more complex structures of the SRHI's latent factors. For example, Opwis et al. (2023) and Reyes Fernández et al. (2019) found the questionnaire to represent the three dimensions *lack of awareness* (e.g. "I do automatically"), *lack of control* (e.g. "I would find hard not to do"), and *behavioural repetition* (e.g. "I do frequently").

The *Self-Report Behavioral Automaticity Index* (SRBAI; Gardner, Abraham, Lally, & de Bruijn, 2012) was later developed with the intention of aligning the scale more strongly with the introduced definition of habit. To that end, Gardner and colleagues (2012) filtered the SRHI items to retain only those pertaining to habitual actions as automatic behaviours. The SRBAI maintained the SRHI's item stem "[Behaviour] is something..." as well as the following selection of *lack of awareness*-scale items (1) "I do automatically", (2) "I do without having to consciously remember", (3) "I do without thinking" and (4) "I start doing before I realize I'm doing it". Thus, the SRBAI contains a subset of four items from the SRHI, dropping the ones regarding lack of control and behavioural repetition and keeping the ones regarding lack of awareness (except one item). As the scale cannot distinguish between habit and other forms of automaticity, the name of the scale targeted automaticity in general instead of habit specifically (Gardner et al., 2012). Nevertheless, the SRBAI is the most frequently applied scale to measure habit strength (Gardner et al., 2024b).

A German version of the SRBAI (SRBAI-G) has been applied in previous studies (e.g. Jekauc et al., 2024; Labudek et al., 2023; Pfeffer & Strobach, 2018), but its psychometric properties have remained unclear. Only the German SRHI (SRHI-G) has been evaluated by Thurn, Finne, Brandes, and Bucksch (2014) in the PA domain, and more recently by Opwis et al. (2023) in the context of caffeine consumption and smartphone/tablet use. These studies (1) differ from each other in terms of their implications for measurement practice and (2) did not validate the SRBAI-G.

The aim of the present research is therefore to re-evaluate the use of the SRHI-G, to validate the SRBAI-G in the context of PA, and to discuss further enhancements for the measurement of habit. To this end, we conducted two preregistered studies (https://osf.io/vxntg; https://osf.io/ja8p7) following the best practice recommendations for questionnaire validations by Boateng et al. (2018). Data, code, and supplementary material are available in an OSF's repository (https://osf.io/4u9hs). In the first study, we evaluated the psychometric properties of the SRHI-G and SRBAI-G in the context of vigorous PA. In the second study, we compared the factorial

structure of the SRBAI-G across the three PA contexts walking, moderate PA, and vigorous PA. As the scale offers the opportunity to adopt the item stem, we investigated whether variations in context and in type (fixed vs. individualised) would affect its psychometric properties.

Study 1

The aim of Study 1[1] was to investigate the factorial validity of the SRHI-G. In accordance with Opwis et al. (2023), it was assumed that, in terms of model-fit, a three-factorial solution with one factor being similar to the SRBAI-G would be acceptable and superior to a single-factor solution or to a solution with an additional global factor (explaining all indicators of the SRHI-G). Furthermore, the convergent validity of the SRBAI-G was investigated assuming a significant positive correlation between the SRBAI-G and the automaticity subscale of the Indicators of Habit Strength in Physical Activity (IHSPA; Grove, Zillich, & Medic, 2014). Additionally, a significant positive correlation was expected for the test-retest-reliability. We did not expect any significant correlation between the SRBAI-G and the rating of perceived exertion (RPE; Borg, 1982), as perceived effort is a relevant construct for explaining PA in general (e.g. Cheval & Boisgontier, 2021) but should not be meaningful associated with a habit measure according to habit theory.

Methods

Participants and procedures

An a-priori power analysis (Jorgensen, Pornprasertmanit, Schoemann, & Rosseel, 2012; RMSEA = 0.06, df = 39, 1 – β = 0.95, α = 0.05) plus 10% for expected data loss resulted in a target sample size of N = 308[2]. Following approval from the local ethics committee, we advertised the study on social media as well as at the Goethe University Frankfurt and invited participants to complete an online survey via SoSci Survey. We recruited N = 384 individuals (≥ 18 years) who were proficient in the German language and participated in vigorous PA. In line with the preregistered criteria, we excluded participants: (1) with a relative speed index greater than 2 (10 cases), (2) who failed the awareness checks (47 cases), (3) who reported giving non-serious responses (0 cases), and (4) participants with no responses in the habit questionnaire section (60 cases). After this exclusion, n = 267 data sets remained for the analyses. Out of these participants, n = 93 participants answered the follow-up questionnaire. Data of the individuals were matched through pseudonymisation. Participants were 27.32 ± 9.62 years old (n = 127 female). A total of 45.3% of respondents reported having a university degree. In addition, the sample indicated to spent 302.73 ± 215.28 min/week exercising at vigorous intensity.

Measures

After giving informed consent, we collected sociodemographic data and information about PA. Next, the habit-related scales were presented in a randomised order. At the end of the survey, participants could leave a comment and we asked them to indicate whether they had answered the questionnaire seriously. Additionally, participants were given the opportunity to leave their email address, so they were able to participate in the retest survey that took place one week later. The email addresses were stored separately from the other data to secure anonymity.

[1] In the preregistration we planned to use the *5-item* lack of awareness subscale of the SRHI-G. However, the main analyses in this article were performed with the more frequently used *4-item* SRBAI-G which comprises one item less (Item 10, see **Table 2**). The analyses and results for the lack of awareness subscales are similar and can be found in the repository (see supplemental material 3).

[2] Erratum: The model with the fewest df was the bifactor model with df = 36 instead of df = 39. This would have led to a required N = 295.

Ger J Exerc Sport Res https://doi.org/10.1007/s12662-025-01050-z

P. Ljubic · B. Broda · S. Windmann · C. Englert

Measuring habit strength in physical activity: Re-evaluation of the German Self-Report Habit Index and validation of the German version of the Self-Report Behavioural Automaticity Index

Abstract

Introduction. The Self-Report Behavioural Automaticity Index (SRBAI) features automaticity-related items of the Self-Report Habit Index (SRHI) and is the most widely used instrument to assess habit strength. The aim of the present investigation was to validate the German SRBAI (SRBAI-G) in the context of physical activity (PA) and to re-evaluate the use of the German Self-Report Habit Index (SRHI-G).

Methods. Two online surveys with convenience samples of young adults (N_1 = 267, N_2 = 334) were conducted to evaluate the reliability (internal consistency and test-retest reliability) and validity (factorial, convergent and discriminant) of habit strength measures. Study 1 took all items of the SRHI-G to re-evaluate its presumed factorial structure in the context of vigorous PA. Study 2 examined the SRBAI-G in three PA contexts. In both studies, the frequency and duration of PA were assessed through self-report.

Results. The SRBAI-G proved to be a reliable and valid instrument for measuring behavioural automaticity in the context of PA. For the SRHI-G, a three-factorial model was shown to be superior to a single-factor solution but not acceptable regarding the cut-off values set forth by Hu and Bentler (1999).

Discussion. In comparison to the SRHI-G, the SRBAI-G is more closely aligned with current habit definitions and meets conventional psychometric quality benchmarks in the context of PA. More research is needed to clarify whether automaticity assessed by available self-report measures can generally tackle changes in habit formation or disruption.

Keywords

Automaticity · Habit · Physical activity · Questionnaire · SRBAI · SRHI

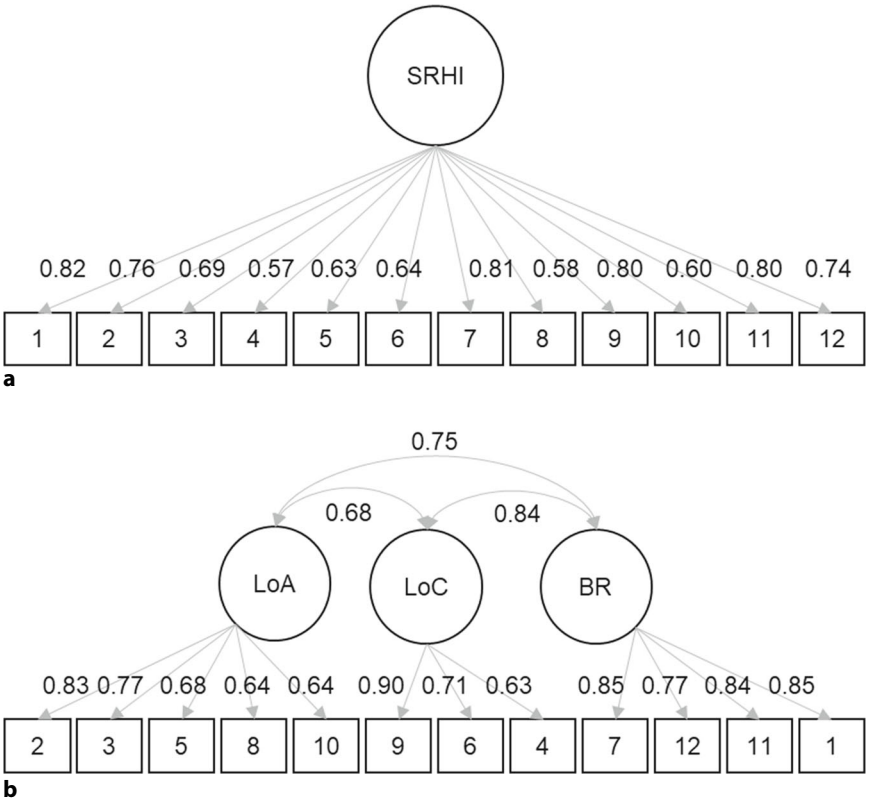

Fig. 1 ▲ Factor loadings and correlations between factors for the **a** single- and **b** three-factor CFA of the SRHI-G related to the main vigorous PA of the participants in Study 1. *Note.* The full description of each item is shown in **Table 2**. *SRHI* Self-Report Habit Index, *LoA* Lack of Awareness subscale, *LoC* Lack of Control subscale, *BR* Behavioural repetition subscale

Some items were included for exploratory purposes and are reported as part of the survey documentation (German) in the OSF repository (see supplemental material 4).

Sociodemographic data. We assessed age, gender, and educational level to characterise the sample.

Vigorous PA. Participants were requested to report the specific vigorous PA they typically spent the most time with (main vigorous PA). Vigorous PAs were described as activities that require a lot of physical exertion, where you breathe much harder than normal (with examples such as running, football, badminton, etc.). Descriptions and items were based on the German short form of the *International Physical Activity Questionnaire* (IPAQ; Craig et al., 2003). For instance, we assessed on how many days in a typical week they engaged with this activity on average, and for how long.

SRHI. We adapted the item stem of the 12-item SRHI-G (Thurn et al., 2014; Verplanken & Orbell, 2003) so it would refer to each individual's main vigorous activity ("[Main vigorous PA] is something …"). In line with the answer categories used for the SRHI-G (Thurn et al., 2014), each item was answered on a 5-point Likert-type scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). All items and internal consistencies are reported in the results section. We calculated mean scores for the SRHI-G and the SRBAI-G items with higher scores indicating stronger habits. Only this measure was presented for a second time one week later.

Convergent Validity Measure: IHSPA. The IHSPA is a four dimensional 17-item questionnaire validated by Grove et al. (2014). Only the four-item automaticity subscale (e.g. "I exercise without having to think about it") was used for our investigation with the mean value being the convergent validity measure. The scale was back-and-forth translated in collaboration with native speakers and a think aloud pilot study was conducted to assess the perception of the items of the SRHI-G and the IHSPA. In advance to any analysis with the scale, we checked the factorial validity and internal consistency of the subscale. Each item was answered on a 6-point Likert-type scale ranging from 1 (*not true for me*) to 6 (*very true for me*), meaning that a higher value indicates stronger automaticity. A CFA confirmed the factorial validity of the automaticity subscale ($\chi^2(2) = 2.70$, $p = 0.259$, RMSEA = 0.031 [0.000, 0.122], CFI = 1.00, TLI = 1.00) and the internal consistency was acceptable (McDonald's $\omega = 0.79$ [0.74, 0.83]).

Discriminant Validity Measure: RPE. Participants were asked to report their typical RPE (Borg, 1982) experienced during the reported vigorous PA on a 10-point Likert-type scale from 1 (*no exertion at all*) to 10 (*maximum [compared to the maximum exertion you have ever experienced]*).

Data analyses

Data analyses were conducted in R (scripts are available in the repository). At first, we pre-processed the data, checked assumptions and analysed the missing data. The assumption of multivariate normality was violated for both the SRHI-G and the SRBAI-G. Additionally, multivariate outliers were identified for each scale. We therefore used the mean- and variance-adjusted weighted least squares (WLSMV) estimator to conduct the CFA (Brauer, Ranger, & Ziegler, 2023) and conducted sensitivity analyses without the outliers (which led to comparable results, see supplemental material 2). Visual inspection confirmed linear relationships between all calculated correlations and no univariate outliers. Missing data analysis showed that after applying the exclusion criteria there were no missing data left.

To investigate the factorial structure of the SRHI, we conducted confirmatory factor analyses (CFA) with tau-congeneric models, as analyses with the SRBAI-G data showed tau-congeneric model to be superior to other measurement models (see supplemental material 1). Interpretations were based

Table 1 Preregistered criteria to evaluate factor loadings, convergent and discriminant validity, and test-retest reliability

Factor loadings	Acceptable	Moderate	High	Very High
λ	> 0.50–0.60	> 0.60–0.70	> 0.70–0.80	> 0.80
Convergent validity	Low	Moderate	High	Very High
r_p	< 0.20	≥ 0.20–0.40	> 0.40–0.60	> 0.60
Discriminant validity	Unacceptable	Low	Moderate	High
$\|r_p\|$	> 0.60	≥ 0.40–0.60	> 0.20–0.40	≤ 0.20
Test-Retest Reliability	Low	Moderate	High	Excellent
r_p	< 0.30	≥ 0.30–0.60	> 0.60–0.80	> 0.80

We determined the criteria based on the results of previous studies conducted with self-report habit scales (Gardner et al., 2012; Thurn et al., 2014)

Table 2 Item characteristics of the SRHI-G in Study 1

Items (English/German) **[Main vigorous activity] is something, …/ [Anstrengende Hauptaktivität] ist für mich etwas, …**	***M***	***SD***	**Skew**	**Diff**	**Discr**
1. that I do frequently./das ich häufig tue.	4.51	0.67	−1.25	0.90	0.62
2. I do automatically./das ich automatisch tue.[a]	3.40	1.19	−0.46	0.68	0.68
3. I do without having to consciously remember./das ich tue, ohne mich bewusst daran erinnern zu müssen.[a]	3.72	1.16	−0.58	0.74	0.64
4. that makes me feel weird if I do not do it./bei dem ich mich komisch fühle, wenn ich es nicht tue.	3.69	1.24	−0.87	0.74	0.46
5. I do without thinking./das ich tue, ohne nachzudenken.[a]	3.35	1.19	−0.32	0.67	0.51
6. that would require effort not to do it./das mich Anstrengung kosten würde, es nicht zu tun.	3.20	1.19	−0.15	0.64	0.55
7. that belongs to my weekly routine./das zu meiner wöchentlichen Routine gehört.	4.69	0.59	−2.19	0.94	0.55
8. I start doing before I realize I'm doing it./mit dem ich anfange, ohne zu bemerken, dass ich es tue.[a]	2.24	1.19	0.74	0.45	0.50
9. I would find hard not to do./das ich schwer fände, es nicht zu tun.	4.00	1.07	−0.91	0.80	0.67
10. I have no need to think about doing. /über das ich nicht nachdenken muss, um es zu tun.[b]	3.45	1.22	−0.41	0.69	0.52
11. that is typically "me"./das typisch für mich ist.	4.13	0.98	−0.84	0.83	0.66
12. that I have been doing for a long time./das ich seit langer Zeit mache.	4.51	0.84	−1.92	0.90	0.51

McDonald's ω of the SRHI-G was ω = 0.87 [0.85, 0.89]
Skew Skewness of the respective distribution of item responses. ***Diff*** Item difficulty. ***Discr*** Item discriminability
[a] Items of the SRBAI. McDonald's ω of the SRBAI-G was ω = 0.78 [0.74, 0.82]
[b] After pilot studies with the German questionnaire we decided to change the verb in item 10 from ***tun*** to ***beginnen*** because it was perceived as identical to item 5

on the following indices: The root mean square error of approximation (RMSEA) indicates a misfit of the model and takes the parsimony of a model into account. In contrast, the standardized root mean square residual (SRMR) is an absolute measure of model fit. Additionally, the comparative fit index (CFI) provides information about improvement of the specified model over the baseline model with penalties for non-parsimony, and the Tucker-Lewis Index (TLI) does the same without these penalties. To evaluate model-fit, we used the cut-off values of Hu and Bentler (1999) as reference (RMSEA < 0.06, SRMR < 0.08, CFI/TLI > 0.95). Given the results in the investigation of Thurn et al. (2014), we evaluated the factor loadings using the criteria shown in **Table 1** (anything below these thresholds was rated as inadequate for inclusion in the SRBAI-G). To evaluate the convergent and discriminant validity, we calculated correlation coefficients with the unweighted means of the scales (see **Table 1**). Internal consistencies were evaluated using McDonald's ω where values greater than ω > 0.70 were seen as acceptable. Confidence intervals (CI) are reported on a 95% level [lower, upper].

Results

Item characteristics

The item characteristics for the SRHI-G are shown in **Table 2** and inter-item correlations are shown in **Table 3**. The internal consistencies for the SRHI-G (ω = 0.87 [0.85, 0.89]) and for the SRBAI-G (ω = 0.78 [0.74, 0.82]) were acceptable.

Factorial validity

Both, the robust fit indices of the CFA for the single-factor ($X^2(54) = 256.83$, $p < 0.001$, RMSEA = 0.15 [0.13, 0.17], SRMR = 0.08, CFI = 0.82, TLI = 0.78) and the three-factor model ($X^2(51) = 100.45$, $p < 0.001$, RMSEA = 0.10 [0.08, 0.13], SRMR = 0.05, CFI = 0.92, TLI = 0.90) exceed all pre-set cut-off criteria. The corresponding factor loadings are displayed in **Fig. 1**. Standard errors and robust test statistics for the bifactor model were not calculated as the information matrix could not be inverted. A likelihood ratio test with non-robust estimates showed that the three-factor model with the factors *lack of awareness*, *lack of control*, and *behavioural repetition* was superior to the single-factor model ($X^2(3) = 73.58$, $p < 0.001$). Sensitivity analyses with a maximum likelihood estimator showed that the model fit indices were better if the violation of assumptions was not considered (see supplemental material 2).

Convergent validity

We calculated the Pearson correlation coefficient with the IHSPA automaticity subscale and the SRBAI-G. There was a significant positive correlation between

Table 3 Inter-item correlation for the items of the SRHI-G in Study 1

	1	2	3	4	5	6	7	8	9	10	11
Item 1	–	–	–	–	–	–	–	–	–	–	–
Item 2[a]	0.47	–	–	–	–	–	–	–	–	–	–
Item 3[a]	0.46	0.52	–	–	–	–	–	–	–	–	–
Item 4	0.36	0.35	0.37	–	–	–	–	–	–	–	–
Item 5[a]	0.32	0.56	0.42	0.22	–	–	–	–	–	–	–
Item 6	0.44	0.34	0.37	0.39	0.27	–	–	–	–	–	–
Item 7	0.62	0.43	0.41	0.34	0.31	0.36	–	–	–	–	–
Item 8[a]	0.20	0.49	0.33	0.25	0.41	0.36	0.21	–	–	–	–
Item 9	0.50	0.42	0.43	0.50	0.25	0.57	0.48	0.32	–	–	–
Item 10	0.37	0.43	0.47	0.27	0.46	0.25	0.27	0.33	0.28	–	–
Item 11	0.57	0.45	0.47	0.41	0.33	0.43	0.50	0.30	0.63	0.33	–
Item 12	0.53	0.41	0.34	0.34	0.27	0.34	0.35	0.27	0.49	0.35	0.57

Spearman correlation coefficients are shown. All correlations are statistically significant ($p < 0.001$)
[a] Items of the SRBAI

the SRBAI-G and the IHSPA automaticity subscale ($r_p = 0.64$ [0.57, 0.71], $p < 0.001$) demonstrating very high convergent validity.

Discriminant validity

There was a non-significant positive correlation between the SRBAI-G and the RPE for the related vigorous activity ($r_p = 0.10$ [−0.02, 22], $p = 0.089$). As the null hypothesis was of interest, we conducted a bootstrapped equivalence test (Lakens, 2017) which showed that the effect size of the correlation lies within the preregistered boundaries of $-0.20 < r < 0.20$ ($p = 0.039$), demonstrating high discriminant validity.

Test-retest reliability

There was a significant positive correlation between the SRBAI-G measured at T1 and at T2 ($n = 93$; 8.18 ± 3.48 days test-retest interval; $r_p = 0.68$ [0.55, 0.78], $p < 0.001$) demonstrating a high test-retest reliability.

Visualisations of the bivariate relationships are shown in **Fig. 2**.

Discussion

Study 1 aimed at re-evaluating the factorial structure of the SRHI-G and validating the SRBAI-G in the context of vigorous PA. The internal consistencies of both the SRHI-G and the SRBAI-G were acceptable and the CFAs confirmed the superiority of the three-factor against the single-factor model of the SRHI-G (Opwis et al., 2023; Reyes-Fernández et al., 2019). Unexpectedly, the model fit indices of the three-factor solution exceeded each pre-set cut-off criteria, meaning that the measurements cannot be explained adequately by the specified latent constructs. As expected, there was a significant positive relationship between the SRHI-G and the automaticity subscale of the IHSPA (i.e. convergent validity) and a near to zero association between the SRBAI-G and the RPE scale (i.e. discriminant validity). Additionally, the hypothesis regarding the test-retest reliability of the SRBAI-G can be accepted. From a quantitative perspective, the construct validity measures were similarly correlated with the SRHI-G as with the SRBAI-G, so these findings do not shed any light on the conceptual differences between these two scales. Future studies could therefore benefit from applying a mixed-methods approach. However, given the results of factorial analyses (Opwis et al., 2023; Reyes Fernández et al., 2019; Thurn et al., 2014), the use of the SRHI-G as a unidimensional measure is contraindicated. As the definition of habit underwent progress from the development of the SRHI (Verplanken and Orbell, 2003) to this day, measurements based on more precise definitions can also cover more precisely the constructs of interest such as lack of awareness for the SRBAI (Gardner et al., 2012).

Study 2

Given that the SRBAI-G met validity and validity criteria in Study 1, the aim of Study 2 was to evaluate the model fit of the SRBAI-G across PA contexts and types of application (fixed vs. individualised). It was assumed that the SRBAI-G models with a single-factor structure[3] meet the model fit recommendations set forth by Hu and Bentler (1999).

Method

Participants and procedures

An a-priori power analysis (Pornprasertmanit et al., 2012; factor loadings from Thurn et al. (2014), $1 - \beta = 0.80$, $\alpha = 0.05$) resulted in a target sample size of $N = 300$. We advertised the study on social media and at the Goethe University Frankfurt. Participants were invited to complete an online survey via SoSci Survey. We recruited $N = 356$ individuals (≥ 18 years) who were proficient in German. After excluding participants according to the preregistered criteria (relative speed index > 2, 14 cases; failed seriousness check, 6 cases), $n = 336$ data sets remained for statistical analyses. Participants were 27.25 ± 11.62 years old ($n = 164$ female, $n = 166$ male, $n = 4$ non-binary). A total of 41.9% of respondents reported having a university degree. In addition, the sample indicated to spent 249.45 ± 272.39 min/week with walking, 205.78 ± 213.16 min/week with moderate, and 301.01 ± 221.73 min/week with vigorous PA.

Measures

After giving informed consent, we collected sociodemographic data and information about the individual PA behaviour. Afterwards, we presented the habit scale related to the three PA con-

[3] Multifactor models are not identified when only four indicators are available and consequently no comparison was conducted as preregistered.

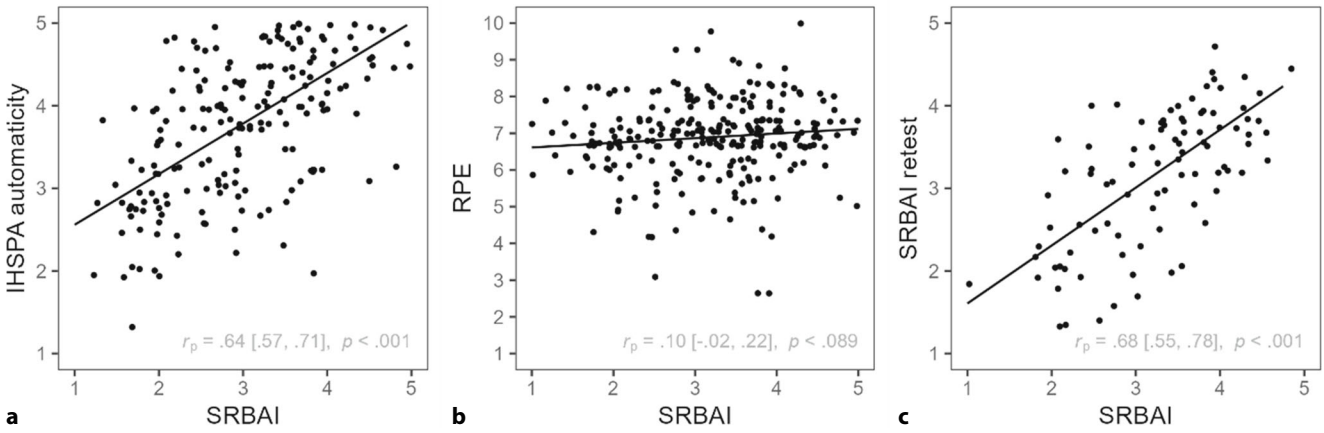

Fig. 2 ▲ Plots for **a** the convergent validity, **b** the discriminant validity, and **c** the test-retest reliability of SRBAI-G scores in Study 1, *Note.* Jittering was used to provide a better view of the relationships

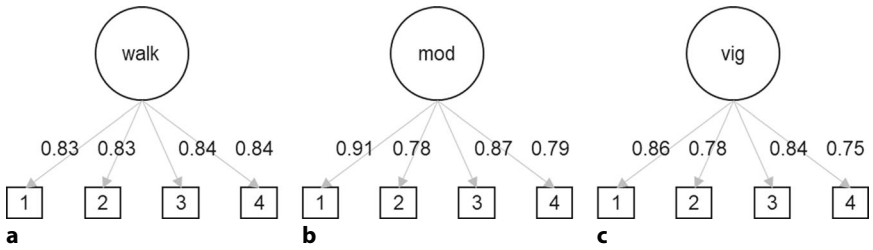

Fig. 3 ▲ Factor loadings for the CFAs with the SRBAI-G applied to **a** walking as well as **b** moderate and **c** vigorous PA of participants in Study 2

texts and the items within one context in a randomised order. At the end of the survey, participants had the opportunity to leave a comment and we asked them to indicate whether they had answered the survey seriously. Further items for exploratory purposes are reported as part of the survey documentation (German) in the repository (see supplemental material 5).

Sociodemographic data. As in Study 1, we assessed age, gender, and educational level to characterise the sample.

PA. Participants were requested to report the weekly frequency and duration of their time spent with moderate and vigorous PA as well as with walking. Moderate PAs were described as activities that involve physical exertion, where you breathe a little harder than normal (with examples such as carrying groceries, cycling as active transport, gardening, etc.). Descriptions and assessments were again based on the IPAQ (Craig et al., 2003). Additionally, we asked them to name the specific moderate and vigorous PA they typically spent the most time with and asked them to indicate on how many days in a typical week they engage with these activities on average, and for how long.

Habit. We used the SRBAI-G adapted to the individual main activities ("[Main moderate/vigorous activity] is something …") and one fixed item stem related to walking ("Walking in everyday life is something …"). In contrast to the individualised item stems which refer to activities within the same intensity level (moderate or vigorous) but not necessarily to the same behaviour (e.g. running or playing soccer are both vigorous PA), the fixed item stem refers to the same behaviour for each participant (i.e., walking). Each item was answered on a 5-point Likert-style scale ranging from 1 (*strongly disagree*) to 5 (*strongly agree*). Internal consistencies of the SRBAI-G scales were acceptable (McDonald's $\omega > 0.80$).

Data analyses

We used the same procedures as described in the data analyses section of Study 1 and applied the identical evaluation criteria (see **Table 1**). The assumption of multivariate normality was violated, and multivariate outliers were identified for all three PA contexts. Consequently, robust statistical procedures and sensitivity analyses for models without outliers were conducted. Missing data analysis revealed that after applying the exclusion criteria $n = 271$ cases were left for moderate PA, $n = 283$ for vigorous PA and $n = 302$ for walking. Missing data in the raw data did not vary across age groups, gender, or educational level.

Results

Item characteristics

The item characteristics for all three applications of the SRBAI-G are shown in **Table 4**. The internal consistency of each scale was acceptable (walking: $\omega = 0.86$ [0.83, 0.88], moderate PA: $\omega = 0.87$ [0.84, 0.89], vigorous PA: $\omega = 0.85$ [0.82, 0.88]).

Factorial validity

Table 5 provides the model fit indices for all three PA contexts of the SRBAI-G. Chi-square-tests of the CFAs were not significant, meaning that the model implied variance-covariance matrices did not vary significantly from the empirically assessed matrix. Misfit analyses revealed that the models tend to slightly overestimate the relationships between

Table 4 Item characteristics for the SRBAI-G in the context of walking as well as moderate and vigorous PA of participants in Study 2

Item	*M*	*SD*	Skew	Diff	Discr
Walking is something…					
1. I do automatically.	4.29	0.87	−1.27	0.86	0.71
2. I do without having to consciously remember.	4.23	0.96	−1.25	0.85	0.69
3. I do without thinking.	4.21	0.94	−1.33	0.84	0.71
4. I start doing before I realize I'm doing it.	3.96	1.11	−0.98	0.79	0.72
[My main moderate activity] is something…					
1. I do automatically.	3.44	1.31	−0.43	0.69	0.78
2. I do without having to consciously remember.	3.5	1.26	−0.43	0.70	0.68
3. I do without thinking.	3.51	1.28	−0.44	0.70	0.75
4. I start doing before I realize I'm doing it.	2.81	1.29	0.24	0.56	0.67
[My main vigorous activity] is something…					
1. I do automatically.	3.14	1.23	−0.13	0.63	0.74
2. I do without having to consciously remember.	3.18	1.3	−0.04	0.64	0.66
3. I do without thinking.	3.2	1.27	−0.03	0.64	0.72
4. I start doing before I realize I'm doing it.	2.43	1.21	0.59	0.49	0.64

McDonald's ω of each scale was acceptable (walking: ω = 0.86 [0.83, 0.88], moderate PA: ω = 0.87 [0.84, 0.89], vigorous PA: ω = 0.85 [0.82, 0.88])
Skew Skewness of the respective distribution of item responses, ***Diff*** Item difficulty, ***Discr*** Item discriminability

Table 5 Results of the (robust) CFAs for the SRBAI-G in the context of walking as well as moderate and vigorous activities in Study 2

Model	*Df*	χ^2	RMSEA [95% CI]	SRMR	CFI	TLI
Walking	2	0.75	0.00 [0.00, 0.07]	0.00	1.00	1.00
Moderate PA	2	7.14	0.10 [0.02, 0.19]	0.02	0.99	0.98
Vigorous PA	2	3.27	0.05 [0.00, 0.15]	0.01	1.00	0.99

None of the Chi-Square-Tests were significant

item 1 and 3 as well as between item 2 and 3. Nevertheless, each model met the cut-off criteria regarding the SRMR, the CFI and the TLI. While the RMSEAs for the models related to walking and vigorous PA met the cut-off criteria of RMSEA < 0.06, the model for moderate PA did not, RMSEA = 0.10 [0.02, 0.19]. As the RMSEAs CIs of each model included values which exceeded the cut-off value, random sampling variations could lead to RMSEA values above the cut-off in some cases. Factor loadings are shown in **Fig. 3**. Applying the pre-set criteria (see **Table 1**), factor loadings ranged from high (λ > 0.70–0.80) to very high (λ > 0.80).

Discussion

In Study 2, the factorial validity of the SRBAI-G across the contexts walking in everyday life as well as moderate and vigorous PAs was investigated. Analyses showed that item 4 ("I start doing before I realize I'm doing it.") was unexpectedly positively skewed with a relatively low item difficulty. Although it seems possible that the German translation of realize (*bemerken*) rather targets consciousness than awareness, CFA supported the factorial validity and internal consistencies were acceptable. Our hypotheses can be accepted for each context except that the model for moderate PA exceeded the recommended cut-off criteria. We consider the following reasons for this result: First, simulations showed that even a correctly specified model shows exceeded RMSEAs when degrees of freedom are relatively low and that the CI should be interpreted (Kenny, Kaniskan, & McCoach, 2015). Values in- and outside the cut-off value therefore indicate further validation with a larger sample size.

Second, due to the adaptable nature of the SRBAI-G not only interindividual differences in the perception of the same activity (walking) were assessed but also individualised PAs (main moderate/vigorous PA). Participants reported confusion regarding how to answer the items when the scale is applied to activities that encompass several sub-actions which the participants are aware of. One participant mentioned that cycling is something which can be done automatically in the sense that they do not have to think about pedalling, but at the same time must be aware of the traffic, and therefore cannot cycle automatically or without thinking. Thus, items such as "Cycling to the gym in the morning is something I do without thinking" might be not answered or with ambiguity. As the order of the SRBAI-G for the three contexts was randomised and cycling was the most frequently mentioned main moderate PA, this could be one reason for the larger proportion of missing values for moderate PA. Formulating more precise items referring to instigation or execution of a behaviour (Gardner, 2022) and adding potential cues to the item stem might offer a solution to attenuate this problem (Sniehotta & Presseau, 2012). However, in most cases the problem cannot be entirely eliminated as it is not possible to represent each individually relevant cue in a single questionnaire. Ecological momentary assessment could be used to identify the most salient cues which could then be considered within the SRBAI or for the development of single items (for a discussion, see Gardner et al., 2024a).

Third, broad categories such as moderate PA might encompass actions which are characterised by distinctive automaticity features. For example, in comparison to cycling, lawn mowing is less complex as it is a closed skill where a behaviour is performed in a stable or largely predictable setting and therefore can follow a rigid script. While for lawn

mowing, a sense of efficiency might be dominant, in the case of pedalling, the sense of automaticity might stem from the feeling that no conscious decision is needed.

The model fit indices for walking were better than for the main moderate and main vigorous PA reported by the participants. This was to be expected from the line of argumentation above: Relatively simple actions seem to be more easily associated with concepts of automaticity compared to complex action patterns (Gardner, 2022). Additionally, as over 25% if the respondents reported the maximum habit strength in the context of walking, ceiling effects influenced the model fit.

General discussion

As the use of German translations of the SRBAI has increased in recent years (e.g. Jekauc et al., 2024; Labudek et al., 2023), the goal of this investigation was to validate the SRBAI-G. Moreover, since the implications of recent studies regarding the measurement practice of the SRHI were inconclusive (Opwis et al., 2023; Thurn et al., 2014), an additional goal was to re-evaluate the SRHI-G. Therefore, we conducted two preregistered studies. Overall, we found the SRBAI-G to be a reliable and valid instrument as the test-retest reliability and internal consistencies as well as factorial, convergent (i.e. IHSPA automaticity subscale), and discriminant validity (i.e. RPE) were as expected. Unexpectedly, the RMSEA of the model in the context of moderate PAs exceeded the cut-off criteria, demonstrating substantial deviations of the empirical data from the model. We explain this finding with the evidence that simulation studies demonstrated a high variability for the RMSEA in CFAs with lower degrees of freedom (Kenny et al., 2015), potential confusion caused by actions which encompass several sub-actions (i.e. cycling), and the broad variety of moderate PAs which can be associated with different perceptions of automaticity.

Based on the results of Study 1, researchers are advised to carefully define the automaticity features relevant to their particular research question and to select or develop a measure which covers these features. If the goal is to focus on lack of awareness as represented by the SRBAI-G items, Studies 1 and 2 demonstrated the SRBAI-G to be a valid and reliable instrument. However, recent research suggest that the SRBAI oversimplifies the multidimensional concept of automaticity (Gardner et al., 2024a). Moreover, little is known about the meta-cognitive approaches of study participants to answer items related to automaticity features and especially to the same features in different contexts. For instance, think aloud studies found that the intended meaning by researchers deviated from the interpretation by the participants for the SRHI (Gardner & Tang, 2014) and for habit in general (Brown, Hagger, & Hamilton, 2024). Likewise, we may ask how, doing something *automatically*, is mentally conceptualised in specific contexts, as compared to doing something frequently. Think-aloud studies regarding the items covering the automaticity features of relevant behaviours could help to further improve the validity of the habit and automaticity measures.

Nonetheless, the main problem is how to assess habit as a cognitive process where a response is automatically triggered by a cue. First, objective measurement approaches such as assessing behavioural repetition cannot serve as a valid marker for habit strength as the activation of an association is key, not the (non-)usage. Other methods that may include behavioural or neuroscience measures are therefore needed to complement self-report data. Second, multiple studies question the validity of self-reports aimed at measuring mental processes with some of them focusing specifically on the measurement of habit (e.g. Hagger, Rebar, Mullan, Lipp, & Chatzisarantis, 2015, Mazar & Wood, 2022). Additional research is needed to further enhance the measurement practices and subsequently to develop effective interventions to support individuals' intentions to be physically active on a regular basis.

Dual process models in the PA domain focus on explicit cognitive processes such as intentions on one side, and on implicit cognitive processes such as habits on the other side to explain PA behaviour (e.g. Brand & Ekkekakis, 2018; Strobach et al., 2020). Recent evidence supports the idea that intentions and habit conjointly predict PA behaviour (Hagger et al., 2023; Jekauc et al., 2024). It is argued that the reflective process of regulating oneself to follow through an PA intention requires self-regulatory competences and the willingness to exert the effort to apply them, whereby in the presence of habits, preparatory actions and PA itself will be carried out more automatically (Pfeffer, Englert, Strobach, Ljubic, & Jekauc, 2024). Given that many individuals do not follow through their PA intentions, the formation of PA habits is seen as a powerful tool to sustainably bridge the IBG (Galla & Duckworth, 2015).

Despite the contributions of this investigation, the following limitations should be acknowledged. First, as the convenience samples mainly consisted of young adults which were highly educated, much more physically active than average, and drawn from a single geographic region, it is not representative for the general population. Studies with more diverse samples are required to generalise the findings across populations. Second, there is the potential for common methods bias as the findings are solely built on self-report data. Measures based on self-reports are built on the premise that people are aware of their habits. When habitual behaviours are executed without awareness, a self-report may reflect other automatic processes than those related to habit and may include retrospective speculations about the antecedents of potential consequences (Hagger et al., 2015; Mazar & Wood, 2022; Sniehotta & Presseau, 2012). Additionally, we only used activities which are carried out frequently while in the context of PA less frequent behaviours (e.g. vacuuming, carrying groceries) are also of interest. Future studies could investigate whether the SRBAI-G can discriminate between individuals with a strong vs. weak habit for a certain action when frequency is balanced.

The present investigation demonstrated the SRBAI-G to be a valid and reliable measure with the flexibility to adapt it to different contexts and the opportunity to combine this feature

with individualisation. To ensure the effective utilisation of the SRBAI-G, it is essential to analyse the behaviour of interest, for example in terms of relevant sub-behaviours or cues to create an item stem which reduces ambiguity. As the longer SRHI-G has been shown to exceed standard criteria, we recommend determining which habit definition and automaticity features are suited for a particular research question and subsequently apply or develop a measure which covers this definition most accurately. Future research (e.g. regarding criterion validity) will gain a deeper understanding of how habits can be assessed in order to, for instance, use habit theory to support individuals in their efforts to adopt a physical active lifestyle.

Corresponding address

Phil Ljubic
Department of Sports Sciences, Goethe University Frankfurt
Ginnheimer Landstr. 39,
60487 Frankfurt/Main,
Germany
ljubic@sport.uni-frankfurt.de

Author Contribution. Conceptualization: PL, SW, and CE; Investigation: PL, BB, and CE; Resources: CE; Supervision: SW and CE; Data curation, formal analysis, project administration, software, visualization, writing—original draft: PL; Writing—review & editing: PL, SW, and CE.

Funding. Open Access funding enabled and organized by Projekt DEAL.

Data Availability. Data, code, and additional online materials are openly available at the project's Open Science Framework page (https://osf.io/4u9hs/).

Declarations

Conflict of interest. P. Ljubic, B. Broda, S. Windmann and C. Englert declare that they have no competing interests.

All procedures performed in studies involving human participants or on human tissue were in accordance with the ethical standards of the institutional and/or national research committee and with the 1975 Helsinki declaration and its later amendments or comparable ethical standards. Informed consent was obtained from all individual participants included in the study.

References

Bagozzi, R. P. (1981). Attitudes, intentions, and behavior: a test of some key hypotheses. *Journal of Personality and Social Psychology*, *41*(4), 607–627. https://doi.org/10.1037/0022-3514.41.4.607.

Bargh, J. A. (1992). The ecology of automaticity: toward establishing the conditions needed to produce automatic processing effects. *The American Journal of Psychology*, *105*(2), 181–199. https://doi.org/10.2307/1423027.

Bargh, J. A. (1994). The four horsemen of automaticity: awareness, intention, efficiency, and control in social cognition. In R. Wyer & T. Srull (Eds.), *Handbook of social cognition* (pp. 1–40). Lawrence Erlbaum.

Boateng, G. O., Neilands, T. B., Frongillo, E. A., Melgar-Quiñonez, H. R., & Young, S. L. (2018). Best practices for developing and validating scales for health, social, and behavioral research: a primer. *Frontiers in Public Health*, *6*, 149. https://doi.org/10.3389/fpubh.2018.00149.

Borg, G. A. V. (1982). Psychophysical bases of perceived exertion. *Medicine & Science in Sports & Exercise*, *14*(5), 377–381. https://doi.org/10.1249/00005768-198205000-00012.

Brand, R., & Ekkekakis, P. (2018). Affective–reflective theory of physical inactivity and exercise: foundations and preliminary evidence. *German Journal of Exercise and Sport Research*, *48*(1), 48–58. https://doi.org/10.1007/s12662-017-0477-9.

Brauer, K., Ranger, J., & Ziegler, M. (2023). Confirmatory factor analyses in psychological test adaptation and development: a nontechnical discussion of the WLSMV estimator. *Psychological Test Adaptation and Development*, *4*(1), 4–12. https://doi.org/10.1027/2698-1866/a000034.

Brown, D. J., Hagger, M. S., & Hamilton, K. (2024). A qualitative investigation of individuals' lay representations of habit. *Psychology & Health*. https://doi.org/10.1080/08870446.2024.2412572.

Bull, F. C., Al-Ansari, S. S., Biddle, S., Borodulin, K., Buman, M. P., Cardon, G., Carty, C., Chaput, J.-P., Chastin, S., Chou, R., Dempsey, P. C., DiPietro, L., Ekelund, U., Firth, J., Friedenreich, C. M., Garcia, L., Gichu, M., Jago, R., Katzmarzyk, P. T., & Willumsen, J. F. (2020). World Health Organization 2020 guidelines on physical activity and sedentary behaviour. *British Journal of Sports Medicine*, *54*(24), 1451–1462. https://doi.org/10.1136/bjsports-2020-102955.

Cheval, B., & Boisgontier, M. P. (2021). The theory of effort minimization in physical activity. *Exercise and Sport Sciences Reviews*, *49*(3), 168–178. https://doi.org/10.1249/JES.0000000000000252.

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International physical activity questionnaire: 12-country reliability and validity. *Medicine and Science in Sports and Exercise*, *35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB.

De Houwer, J. (2019). On how definitions of habits can complicate habit research. *Frontiers in Psychology*, *10*, 2642. https://doi.org/10.3389/fpsyg.2019.02642.

Feil, K., Fritsch, J., & Rhodes, R. E. (2023). The intention-behaviour gap in physical activity: a systematic review and meta-analysis of the action control framework. *British Journal of Sports Medicine*, *57*(19), 1265–1271. https://doi.org/10.1136/bjsports-2022-106640.

Fernández, R. B., Monge-Rojas, R., Solano López, A. L., & Cardemil, E. (2019). Re-evaluating the self-report habit index: the cases of physical activity and snacking habits. *Psychology & Health*, *34*(10), 1161–1178. https://doi.org/10.1080/08870446.2019.1585852.

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*(3), 508–525. https://doi.org/10.1037/pspp0000026.

Garcia, L., Pearce, M., Abbas, A., Mok, A., Strain, T., Ali, S., Crippa, A., Dempsey, P. C., Golubic, R., Kelly, P., Laird, Y., McNamara, E., Moore, S., de Sa, T. H., Smith, A. D., Wijndaele, K., Woodcock, J., & Brage, S. (2023). Non-occupational physical activity and risk of cardiovascular disease, cancer, and mortality outcomes: a dose–response meta-analysis of large prospective studies. *British Journal of Sports Medicine*, *57*(15), 979–989. https://doi.org/10.1136/bjsports-2022-105669.

Gardner, B. (2015). A review and analysis of the use of 'habit' in understanding, predicting and influencing health-related behaviour. *Health Psychology Review*, *9*(3), 277–295. https://doi.org/10.1080/17437199.2013.876238.

Gardner, B. (2022). Habit and behavioural complexity: habitual instigation and execution as predictors of simple and complex behaviours. *Current Research in Behavioral Sciences*, *3*, 100081. https://doi.org/10.1016/j.crbeha.2022.100081.

Gardner, B., & Lally, P. (2023). Habit and habitual behaviour. *Health Psychology Review*, *17*(3), 490–496. https://doi.org/10.1080/17437199.2022.2105249.

Gardner, B., & Tang, V. (2014). Reflecting on non-reflective action: an exploratory think-aloud study of self-report habit measures. *British Journal of Health Psychology*, *19*(2), 258–273. https://doi.org/10.1111/bjhp.12060.

Gardner, B., de Bruijn, G.-J., & Lally, P. (2011). A systematic review and meta-analysis of applications of the Self-Report Habit Index to nutrition and physical activity behaviours. *Annals of Behavioral Medicine*, *42*(2), 174–187. https://doi.org/10.1007/s12160-011-9282-0.

Gardner, B., Abraham, C., Lally, P., & de Bruijn, G.-J. (2012). Towards parsimony in habit measurement: testing the convergent and predictive validity of an automaticity subscale of the self-report habit index. *International Journal*

of Behavioral Nutrition and Physical Activity, *9*(1), 102. https://doi.org/10.1186/1479-5868-9-102.

Gardner, B., Lally, P., & Rebar, A. L. (2024a). Can—and should—automaticity be self-reported using a single item? A secondary analysis of 16 datasets. *Applied Psychology: Health and Well-Being*, *16*(4), 2438–2457. https://doi.org/10.1111/aphw.12600.

Gardner, B., Rebar, A. L., de Wit, S., & Lally, P. (2024b). What is habit and how can it be used to change real-world behaviour? Narrowing the theory-reality gap. *Social and Personality Psychology Compass*. https://doi.org/10.1111/spc3.12975.

Grove, J. R., Zillich, I., & Medic, N. (2014). A process-oriented measure of habit strength for moderate-to-vigorous physical activity. *Health Psychology and Behavioral Medicine*, *2*(1), 379–389. https://doi.org/10.1080/21642850.2014.896743.

Hagger, M. S., Rebar, A. L., Mullan, B., Lipp, O. V., & Chatzisarantis, N. L. D. (2015). The subjective experience of habit captured by self-report indexes may lead to inaccuracies in the measurement of habitual action. *Health Psychology Review*, *9*(3), 296–302. https://doi.org/10.1080/17437199.2014.959728.

Hagger, M. S., Hamilton, K., Phipps, D. J., Protogerou, C., Zhang, C.-Q., Girelli, L., Mallia, L., & Lucidi, F. (2023). Effects of habit and intention on behavior: Meta-analysis and test of key moderators. *Motivation Science*, *9*(2), 73–94. https://doi.org/10.1037/mot0000294.

Hewig, J. S. (2024). Intentionality and the power of habit. In L. R. Ingerslev & K. Mertens (Eds.), *Phenomenology of Broken Habits* (1st edn., pp. 71–92). Routledge. https://doi.org/10.4324/9781003332466-6.

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

Hull, C. L. (1943). *Principles of behavior: an introduction to behavior theory*. Appleton-Century.

Imboden, C., Claussen, M., Seifritz, E., & Gerber, M. (2021). Physical activity for the treatment and prevention of depression: a rapid review of meta-analyses. *Deutsche Zeitschrift Für Sportmedizin/German Journal of Sports Medicine*, *72*(6), 280–287. https://doi.org/10.5960/dzsm.2021.499.

Jekauc, D., Gürdere, C., Englert, C., Strobach, T., Bottesi, G., Bray, S., Brown, D., Fleig, L., Ghisi, M., Graham, J., Martinasek, M., Tamulevicius, N., & Pfeffer, I. (2024). The contribution and interplay of implicit and explicit processes on physical activity behavior: empirical testing of the physical activity adoption and maintenance (PAAM) model. *BMC Public Health*, *24*, 1239. https://doi.org/10.1186/s12889-024-18589-5.

Jorgensen, T. D., Pornprasertmanit, S., Schoemann, A. M., & Rosseel, Y. (2012). semTools: Useful Tools for Structural Equation Modeling (p. 0.5-6) (Computer software). https://CRAN.R-project.org/package=semTools

Kenny, D. A., Kaniskan, B., & McCoach, D. B. (2015). The performance of RMSEA in models with small degrees of freedom. *Sociological Methods & Research*, *44*(3), 486–507. https://doi.org/10.1177/0049124114543236.

Labudek, S., Fleig, L., Jansen, C.-P., Kramer-Gmeiner, F., Nerz, C., Clemson, L., Klenk, J., Becker, C., & Schwenk, M. (2023). Changes in psychological determinants of behavior change after Individual versus group-based lifestyle-integrated fall prevention: results from the LiFE-is-LiFE trial. *Gerontology*, *69*(2), 212–226. https://doi.org/10.1159/000524701.

Lakens, D. (2017). Equivalence tests. *Social Psychological and Personality Science*, *8*(4), 355–362. https://doi.org/10.1177/1948550617697177.

Mazar, A., & Wood, W. (2022). Illusory feelings, elusive habits: people overlook habits in explanations of behavior. *Psychological Science*, *33*(4), 563–578. https://doi.org/10.1177/09567976211045345.

Mittal, B. (1988). Achieving higher seat belt usage: the role of habit in bridging the attitude-behavior gap. *Journal of Applied Social Psychology*, *18*(12), 993–1016. https://doi.org/10.1111/j.1559-1816.1988.tb01189.x.

Moors, A., & De Houwer, J. (2006). Automaticity: a theoretical and conceptual analysis. *Psychological Bulletin*, *132*(2), 297–326. https://doi.org/10.1037/0033-2909.132.2.297.

Morean, M. E., DeMartini, K. S., Foster, D., Patock-Peckham, J., Garrison, K. A., Corlett, P. R., Krystal, J. H., Krishan-Sarin, S., & O'Malley, S. S. (2018). The Self-Report Habit Index: assessing habitual marijuana, alcohol, e-cigarette, and cigarette use. *Drug and Alcohol Dependence*, *186*, 207–214. https://doi.org/10.1016/j.drugalcdep.2018.01.014.

Opwis, M., Bartel, E. C., Salewski, C., & Schmidt, J. (2023). Sorry—Bad Habit! Validation of the German Self-Report Habit Index with a test for its relation to potentially addictive forms of health-risk behaviors. *International Journal of Mental Health and Addiction*, *22*, 3414–3432. https://doi.org/10.1007/s11469-023-01057-3.

Orbell, S., & Verplanken, B. (2015). The strength of habit. *Health Psychology Review*, *9*(3), 311–317. https://doi.org/10.1080/17437199.2014.992031.

Ouellette, J. A., & Wood, W. (1998). Habit and intention in everyday life: the multiple processes by which past behavior predicts future behavior. *Psychological Bulletin*, *124*(1), 54–74. https://doi.org/10.1037/0033-2909.124.1.54.

Pfeffer, I., & Strobach, T. (2018). Behavioural automaticity moderates and mediates the relationship of trait self-control and physical activity behaviour. *Psychology & Health*, *33*(7), 925–940. https://doi.org/10.1080/08870446.2018.1436176.

Pfeffer, I., Hamilton, K., Hannan, T., & Wegner, M. (2023). Models to explain and change health behavior and physical activity. In J. Schüler, M. Wegner, H. Plessner & R. C. Eklund (Eds.), *Sport and exercise psychology* (pp. 617–647). Springer. https://doi.org/10.1007/978-3-031-03921-8_25.

Pfeffer, I., Englert, C., Strobach, T., Ljubic, P., & Jekauc, D. (2024). Das Physical Activity Adoption and Maintenance (PAAM) Modell: Aktuelle Überlegungen und Erweiterungen zur Erklärung der Intentions-Verhaltens-Lücke. In C. Englert, I. Pfeffer, K. Staufenbiel & C. Vater (Eds.), *Sportpsychologie: Grundlagen und Trends* (pp. 223–292). utb. https://doi.org/10.36198/9783838562902.

Pornprasertmanit, S., Miller, P., Schoemann, A., & Jorgensen, T. D. (2012). simsem: SIMulated Structural Equation Modeling (p. 0.5-16) (Computer software). https://CRAN.R-project.org/package=simsem

Rhodes, R. E., McEwan, D., & Rebar, A. L. (2019). Theories of physical activity behaviour change: a history and synthesis of approaches. *Psychology of Sport and Exercise*, *42*, 100–109. https://doi.org/10.1016/j.psychsport.2018.11.010.

Sheeran, P. (2002). Intention—Behavior relations: a conceptual and empirical review. *European Review of Social Psychology*, *12*(1), 1–36. https://doi.org/10.1080/14792772143000003.

Sniehotta, F. F., & Presseau, J. (2012). The habitual use of the self-report habit index. *Annals of Behavioral Medicine*, *43*(1), 139–140. https://doi.org/10.1007/s12160-011-9305-x.

Strain, T., Flaxman, S., Guthold, R., Semenova, E., Cowan, M., Riley, L. M., Bull, F. C., & Stevens, G. A. (2024). National, regional, and global trends in insufficient physical activity among adults from 2000 to 2022: a pooled analysis of 507 population-based surveys with 5·7 million participants. *The Lancet Global Health*, *12*(8), 1232–1243. https://doi.org/10.1016/S2214-109X(24)00150-5.

Strobach, T., Englert, C., Jekauc, D., & Pfeffer, I. (2020). Predicting adoption and maintenance of physical activity in the context of dual-process theories. *Performance Enhancement & Health*, *8*(1), 100162. https://doi.org/10.1016/j.peh.2020.100162.

Thurn, J., Finne, E., Brandes, M., & Bucksch, J. (2014). Validation of physical activity habit strength with subjective and objective criterion measures. *Psychology of Sport and Exercise*, *15*(1), 65–71. https://doi.org/10.1016/j.psychsport.2013.09.009.

Verplanken, B., & Orbell, S. (2003). Reflections on past behavior: a self-report index of habit strength. *Journal of Applied Social Psychology*, *33*(6), 1313–1330. https://doi.org/10.1111/j.1559-1816.2003.tb01951.x.

Verplanken, B., Aarts, H., van Knippenberg, A., & van Knippenberg, C. (1994). Attitude versus general habit: antecedents of travel mode choice. *Journal of Applied Social Psychology*, *24*(4), 285–300. https://doi.org/10.1111/j.1559-1816.1994.tb00583.x.

Watson, P., O'Callaghan, C., Perkes, I., Bradfield, L., & Turner, K. (2022). Making habits measurable beyond what they are not: a focus on associative dual-process models. *Neuroscience & Biobehavioral Reviews*, *142*, 104869.

Wood, W., Quinn, J. M., & Kashy, D. A. (2002). Habits in everyday life: thought, emotion, and action. *Journal of Personality and Social Psychology*, *83*(6), 1281–1297. https://doi.org/10.1037/0022-3514.83.6.1281.