Multi-level validation of the German physical activity barriers self-efficacy scale in a sample of female sixth-graders

Joachim Bachner (✉ joachim.bachner@tum.de)

Technische Universität München <https://orcid.org/0000-0002-8942-3055>

David Joseph Sturm

Technische Universität München

Stephan Haug

Technische Universität München

Yolanda Demetriou

Technische Universität München

Research

Keywords: multi-level, lavaan, confirmatory factor analysis, internal consistency, criterion validity, adolescents

Posted Date: December 5th, 2019

DOI: <https://doi.org/10.21203/rs.2.18261/v1>

License: This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Version of Record: A version of this preprint was published at BMC Public Health on June 22nd, 2020.
See the published version at <https://doi.org/10.1186/s12889-020-09096-4>.

Abstract

Background The physical activity behaviour of the majority of children and adolescents is insufficient. Self-efficacy is regarded as one of the most important determinants able to enhance physical activity. The purpose of this study is to validate the German version of the physical activity barriers self-efficacy scale by means of a multi-level approach. Factorial validity, internal consistency and criterion validity were examined for the individual and the class level.

Methods The final sample comprised 454 female sixth-graders of 33 classes. The original 8-item physical activity barriers self-efficacy scale was translated by a committee approach and pilot tested within the pretest procedure. To examine the factorial validity of the translated scale, a multi-level confirmatory factor analysis was conducted with the lavaan package in R. Internal consistency was estimated with the alpha function of the psych package. Criterion validity was examined by correlating self-efficacy with moderate-to-vigorous physical activity (MVPA) assessed with accelerometers.

Results In contrast to previous validation studies, two-dimensional models fit the data better than uni-dimensional models. The best overall fit exhibited a 2x2-model, with two factors both on the individual and the class level ($\chi^2 = 65.13$, CFI = .985, TLI = .977, RMSEA = .046, SRMR = .033). The latent factors correlated highly on both levels ($r = .87$; $r = .72$). Every item loaded significantly on its respective factor on both levels. Internal consistency for the full scale and the first subscale was good on the individual level and excellent on the class level. For the second subscale, Cronbach's alpha was low on level 1 and excellent on level 2. Weak relations between self-efficacy values and MVPA were found for level 1, strong associations were found for level 2.

Conclusions The validation speaks in favour of a two-dimensional scale measuring not only actual self-efficacy but also support from family and friends. Furthermore, the results argue for the relevance of the multi-level approach which is able to differentiate between self-efficacy on the individual level and collective efficacy on the class level.

Background

Regular physical activity (PA) contributes to the prevention of chronic diseases, such as diabetes mellitus, cancer or cardiovascular diseases, and lowers the risk of premature death (1–4). The World Health Organization (WHO) recommends children and youths aged five to seventeen years to accumulate at least 60 minutes of moderate-to-vigorous physical activity (MVPA) per day, with MVPA comprising every kind of activity which needs at least the amount of energy spent during ordinary walking (5). There are two reasons why it is important that children and adolescents fulfil the PA recommendation. One would be the positive short- and middle-term influences on their health and well-being (1, 6–8). Secondly, a tracking effect describes the role of adolescents' PA as a significant predictor of PA in adulthood: The more active a person is in adolescence, the higher the probability of an active lifestyle in adulthood (7, 9).

According to a questionnaire-based study, only 26% of children and adolescents in Germany aged between three and seventeen years reach daily 60 minutes of MVPA. Furthermore, less girls than boys (22.4% vs. 29%) fulfil the recommendation (10). In addition, PA levels in this population decline with increasing age (10, 11). A systematic review of Van Hecke et al. (12) supports the effects of gender and age. Although not even the most popular device-based approaches like accelerometry offer perfectly reliable PA data (12–14), a vast majority of studies indicates that PA behaviour in adolescence does not comply with the respective recommendation (15). Moreover, the WHO recommendation is merely regarded as a minimum value. Higher MVPA levels are associated with additional health benefits (16). Therefore, in any case it is worthwhile to promote PA from an early age.

At this point, the question arises which determinants should be focused on to enhance youth's PA behaviour. Ecological models suggest that PA behaviour is affected by several interacting levels of influence ranging from policy variables, like investments in public recreation facilities, to intrapersonal variables, like psychological constructs (17). Among these psychological constructs, self-efficacy concerning PA is of great importance. In a review of reviews by Bauman and colleagues (18) self-efficacy was the only psychological factor consistently identified as a positive correlate and determinant of PA in children and adolescents. This finding was confirmed by an umbrella systematic review specifically focussing on psychological constructs (19). Yet another systematic review (20) focused on the PA-related age effect and indicated that self-efficacy was one of very few constructs able to diminish the decline in PA between the age of ten and 18. Furthermore, two systematic reviews (21, 22) analysing intervention studies identified PA self-efficacy as the most promising mediator to increase PA.

The manifestation of self-efficacy, however, does not only occur on the individual level but also on the group level. Bandura (23) defined the phenomenon of collective efficacy as “a group's shared belief in its conjoint capabilities to organize and execute the courses of action required to produce given levels of attainment”. Although a group's collective efficacy cannot simply be seen as the sum of the members' individual self-efficacy perceptions (24), self-efficacy has been found to be a significant predictor of individually perceived collective efficacy in sports contexts (25, 26). Like individual self-efficacy (e.g., 19), collective efficacy is also related to performance. A meta-analysis found a significant positive relation between collective efficacy and performance in organizational settings including sports (27). Furthermore, collective efficacy also affects individual performances and attitudes. In two experiments manipulating the individual perception of collective efficacy, the participants of the low collective efficacy condition set lower performance goals, showed lower commitment and put less effort in a cycling task than the participants perceiving a high collective efficacy (28, 29). Thus, individual self-efficacy and collective efficacy are intertwined as individual self-efficacy acts as a predictor of collective efficacy (25, 26) and collective efficacy again influences the individual (28). This point implies specific consequences for analysing self-efficacy and collective efficacy perceptions of individuals nested in groups, like athletes in teams or students in classes. Under these circumstances, the assumption that individual perceptions are independent of one another cannot longer be maintained. Instead, the use of multi-level modeling is strongly recommended (24, 30).

Due to its high relevance, self-efficacy has been extensively examined in the field of PA. Over time, however, the definitions and the respective measures of youth PA self-efficacy have become more and more heterogeneous. Therefore, Voskuil and Robbins (31) conducted a concept analysis regarding the defining attributes, antecedents and consequences of the different conceptualisations. Eventually, they defined youth PA self-efficacy “as a youth’s belief in his/her capability to participate in PA and to choose PA despite existing barriers” (31). The conceptualization of barriers self-efficacy regarding PA by Dishman and colleagues (32, 33) considers the two main points of this definition by addressing both the self-perceived confidence in the capability to be physically active as well as the recognition of barriers to PA (31).

To date, no instruments exist which are specifically constructed and appropriately validated to examine PA self-efficacy of children and adolescents in secondary school in Germany. Questionnaires specifically designed for children are needed, especially regarding the wording of items. Twelve-year-olds produce a response quality worse than that of youths aged fourteen (34). Scott (35) even argues that adolescents cannot answer properly to adult items before the age of sixteen.

Therefore, the purpose of this study is to validate a German version of the barriers self-efficacy scale for PA (33) in a sample of sixth-graders. The validation is conducted in accordance with the multi-level approach described by Huang (36). Factor structure and scale dimensionality are analysed by means of a multi-level confirmatory factor analysis (MCFA). Internal consistency is also estimated for both the individual and group level, respectively. Furthermore, criterion validity is tested by examining the relation of PA barriers self-efficacy and actual PA behaviour on both levels.

Methods

Participants

The sample included 507 female sixth-graders recruited from 33 classes of fifteen secondary schools in Munich. The participants were part of the CReActivity project, a randomized controlled trial aiming to promote PA of female adolescents (37). Mean age was 11.61 years ($SD = .55$, $N = 430$). The girls were on average normal weight (mean BMI = 19.49, $SD = 3.68$, $N = 386$). The number of BMI values was diminished as parts of the sample refused to be weighed. Refusal was shown by both apparently overweight and normal weight girls. The sample comprised participants from households of low, medium and high socioeconomic status (SES; mean = 49.80, $SD = 15.96$, $N = 412$). SES was assessed by asking the children to name and describe their parents’ current jobs. The answers were classified referring to the International Socioeconomic Index of occupational status (ISEI) which is based on the International Standard Classification of Occupation 2008 (ISCO-08) (38). When the jobs of both parents could be classified, the job with the higher ISEI was considered (HISEI). Vague answers making definite classifications impossible reduced the number of HISEI values.

The study was approved by the ethics commission of the Technical University of Munich (155/16 S) and the Ministry of culture and education of the state of Bavaria in Germany.

Measures

Barriers self-efficacy

The barriers self-efficacy scale was used to assess the girls' perceived self-efficacy to be physically active (33). The scale contains eight items. The original items were validated in samples of sixth- and eighth-grade girls. Confirmatory factor analyses supported a unidimensional model (32, 33, 39). Participants responded on a five point Likert-type scale ranging from 1 ("Disagree a lot") to 5 ("Agree a lot"). The scale validated here was translated into German by means of the committee approach and was pilot tested with the pretest procedure (40, 41).

Physical activity

To assess leisure time MVPA, participants wore accelerometers (ActiGraph GT3X - wGT3X-BT) for seven consecutive days except during water-based activities. The device was placed on the right hip. Sampling rate was set to 30 Hz. Participants had to wear it on weekdays starting at the latest on their way to school until 9 pm or until they went to bed. On weekend days, the students had to put it on as soon as they woke up until 9 pm or until they went to bed.

Procedures

Several weeks before the beginning of the data assessments, students and their parents were informed in writing about the purpose and the procedure of the assessment. Students did not participate if they had not provided a written consent form before.

Data assessments took place at the beginning of a physical education (PE) lesson. Codes were used to ensure the anonymity of the participants. Before handing out the accelerometers, the assessment team explained how to put them on. At least 25% of the students of each class received a sheet with information regarding the accelerometers enabling them to serve as contact persons for their classmates. After the students had put on the accelerometers correctly, they filled out the questionnaire. The actual PE lesson did not start until the last student had completed the questionnaire.

Data analysis

Multi-level validation of the barriers self-efficacy for PA scale

As the sample examined in this study provides clustered data, the validation is based on the multi-level approach by Huang (36). Ignoring the clustered nature of the data can lead to wrong parameter estimates, standard errors and model fits. It is recommended to account for multilevel data even if intracluster correlations (ICC) of the single manifest variables are small (42). In nested data, factor

structures might not be the same for each level (36). MCFA provides the opportunity to examine individual- and group-level data simultaneously. To this end, the total population covariance matrix is split into a pooled within-group covariance matrix and a between-group covariance matrix. Thereby, both within- and between-group effects may be estimated at the same time. Huang (36) offers an R syntax to be used with the lavaan package (43) and a function for generating the required matrices based on the five MCFA steps outlined by Hox (44, Chap. 14).

In step 1, a single-level factor analysis is performed using only the pooled within-group covariance matrix. In step 2, the null model, which assumes the factor structure of step 1 for both levels, is fitted. In this step, both the pooled within- and the between-group covariance matrices are used as input. Equality constraints for the two levels are applied, meaning that factor loadings, variances and covariances for every manifest variable and latent factor are assumed to be the same for the two levels. In step 3, new group-level latent variables are introduced to estimate the variance attributed to the groups. This step is referred to as the independence model since the newly introduced group-level variables are not allowed to covary. This constraint is eliminated in step 4, the so-called saturated model. All degrees of freedom at the between-group level are now used, making it a fully saturated model. Finally, in step 5, the model that is actually hypothesized is specified. At least one overall general factor is added for the between-group level which is defined responsible for the correlation of the latent group-level factors (36). For every model, small negative residual variances on the class level are fixed to zero to allow the model to fully converge. This common practice is particularly required when the number of groups on level 2 is small and ICCs are close to zero (44).

To evaluate model fit, several fit indices were considered (45): the χ^2 -likelihood ratio statistic, the comparative fit index (CFI), the Tucker-Lewis index (TLI), the root mean square error of approximation (RMSEA) and the standardized root mean square residual (SRMR). As the χ^2 -goodness of fit test tends to reject reasonably fitting models when applied to data of large samples, a variety of fit indices was used to estimate model fit (45). Whereas CFI and TLI values greater than .95 indicate a good model fit, values less than or equal to .08 suggest a good model fit when RMSEA and SRMR are considered (45).

Furthermore, as fit tends to improve by including more variables in the model, parsimony is another criterion taken into account when deciding for a preferred model. Akaike's information criterion (AIC) was considered in model comparison as it not only evaluates model fit but also penalizes an increasing amount of estimated parameters (46). Lower AIC values indicate better model fit. Eventually, it is aimed for a model explaining as much variance as possible with as few variables as necessary. Therefore, the optimal combination of model fit and parsimony is sought (47).

Scale reliability is indicated by Cronbach's alpha. Values were calculated for both levels separately by using the alpha function of the psych package in R (48). In case of non-positive definite matrices, alpha was calculated for the nearest positive definite matrix (49).

Criterion validity was examined by correlating self-efficacy values with the participants' MVPA values. Pearson r is indicated for both the pooled within-group correlation and the between-group correlation.

Physical activity

During PA data download the vector magnitude counts were summed over 1-second epochs (10-second epochs for GT3X because of lower memory and battery capacities). The low frequency extension filter was not used. Wear-time validation was conducted with the algorithm by Choi, Liu, Matthews and Buchowski (50). A participant's PA data was considered as valid when data of at least three weekdays and one weekend day were available with at least eight hours of wear time being required for a valid day. The wear-time validated PA data was analysed utilizing the cut points by Hänggi, Phillips and Rowlands (51) to eventually calculate the average duration of MVPA per day for each participant. The cut points by Hänggi et al. (51) were chosen because they provide a precise assessment and were validated by applying the same data sampling and processing criteria as the ones chosen for this study (52).

Results

Of the 507 participants originally included in the sample, 53 had missing values in at least one item of the self-efficacy scale. To avoid bias it was refrained from replacing missing values. The participants excluded from the analysis did not differ significantly from the valid sample regarding BMI, SES, self-efficacy and MVPA. Finally, 454 sixth-graders built the final sample.

The descriptive statistics of the eight items of the barriers self-efficacy scale are presented in Table 1. Means of the items ranged from 3.17 (SD = 1.30) to 3.96 (SD = 1.14). Skewness and kurtosis values were low to moderate. ICCs were small (< .05).

Table 1
Descriptive statistics for the items of the barriers self-efficacy scale (N = 454)

Item #	M	SD	Skewness	Kurtosis	ICC
1	3.83	1.06	-0.49	-0.65	.02
2	3.36	1.35	-0.30	-1.10	.04
3	3.96	1.14	-0.92	0.03	.04
4	3.86	1.10	-0.66	-0.30	.03
5	3.33	1.30	-0.19	-1.06	.03
6	3.64	1.21	-0.53	-0.65	– .01
7	3.71	1.10	-0.42	-0.69	– .01
8	3.17	1.30	0.00	-1.11	.01
Note. M = mean; SD = standard deviation; ICC = intraclass correlation.					

For the single-level one-factor model an acceptable fit was found (model A in Table 2). Depending on the chosen fit index, the fit of the null model improved or got worse compared to model A ($\chi^2 = 85.87$, CFI = .975, TLI = .975, RMSEA = .048, SRMR = .047, AIC = 10,425.47). Whereas model fit did not change substantially for the independence model ($\chi^2 = 75.50$, CFI = .977, TLI = .973, RMSEA = .050, SRMR = .044, AIC = 10,431.10), the fit of the saturated model again differed according to the respective fit index ($\chi^2 = 44.36$, CFI = .979, TLI = .942, RMSEA = .073, SRMR = .030, AIC = 10,455.96). In the last step of the algorithm outlined by Hox (44, Chap. 14) and Huang (36) model B was obtained, see Table 2. For this model, one overall general factor is added for the class level (1 × 1-model). Model B contains twice as many degrees of freedom as the single-level model A, which led to an increase of the χ^2 and AIC value. However, according to the CFI, TLI and RMSEA indices model fit improved. On level 1 all factor loadings were significant whereas on level 2 three out of eight items exhibited significant loadings (Table 3).

Table 2
Fit indices for five models tested via Multilevel Confirmatory Factor Analysis

Model					
Fit index	A: One factor at Level 1	B: One factor on both levels, nested	C: One factor on both levels, not nested	D: Two factors on both levels, nested	E: Two factors on both levels, not nested
χ^2	44.36	61.36	64.81	56.94	56.13
df	20	40	40	38	38
CFI	.977	.982	.979	.984	.985
TLI	.968	.975	.971	.976	.977
RMSEA	.054	.049	.052	.047	.046
SRMR	.032	.038	.035	.035	.033
AIC	9,624.68	10,432.96	10,436.41	10,432.54	10,431.73
Note. df = degrees of freedom; CFI = comparative fit index; TLI = Tucker-Lewis index; RMSEA = root mean square error of approximation; SRMR = standardized root mean square residual; AIC = Akaike's information criterion.					

Unlike the approach proclaimed by Huang (36), model C (1 × 1-model) was specified without a nested structure, meaning that the measurement model for level 1 is not included in level 2. Except for SRMR, the values in every fit index got slightly worse. Now, however, each factor loading on level 2 was significant (see Table 3). For the models D and E (2 × 2-models), a second latent factor was introduced on both levels which is modelled by items 2 (in the original version by Dishman and colleagues (33): “I can ask my parent or other adult to do physically active things with me.”) and 5 (original item: “I can ask my best friend to be physically active with me during my free time on most days.”). Responses to these two items rather depend on the social environment of the adolescents and not solely on themselves. The idea of creating a separate factor comprising these two items was further supported as they exhibited the lowest

factor loadings in the single-level model A. In line with this, their correlations with the other items were below average. Specifying two factors on each level for models D and E decreased the number of degrees of freedom since two more parameters had to be estimated than in models B and C. However, models D and E had a better model fit with respect to each index, including the AIC. In model E the two factors were specified without a nested structure which ended in the best overall fit of all models. In model E every item exhibited a statistically significant loading on its respective latent factor on both levels. In model D only one out of six items loaded significantly on factor 1 on level 2. The correlation of the latent factors in models D and E was 0.87 on level 1 and 0.69 and 0.72 on level 2, respectively.

Table 3
Standardized factor loadings and correlations of latent
factors

Model							
	A	B	C	D	E		
Items	F1	F1	F1	F1	F2	F1	F2
Individual level							
1	.72	.73	.72	.72		.72	
2	.45	.45	.45		.51		.51
3	.64	.65	.64	.64		.64	
4	.66	.68	.66	.67		.66	
5	.49	.49	.49		.55		.56
6	.61	.60	.61	.60		.61	
7	.76	.76	.76	.76		.77	
8	.76	.76	.76	.76		.76	
Class level							
1		.65	.87	.91		.86	
2		.82	.92		.84		.83
3		.12	.75	.57		.77	
4		.33	.81	.49		.82	
5		1.04	.83		.96		.72
6		.30	.57	.27		.57	
7		.32	.75	.36		.75	
8		.17	.81	.43		.83	
$R_{\text{individual}}$ (F1, F2)				.87		.87	
R_{class} (F1, F2)				.69		.72	
Note. Values in bold are statistically significant ($p < .05$).							

Aiming at a well-fitting and at the same time parsimonious solution, models with two factors on level 1 and only one factor on level 2 were specified as well (2×1 -models). The fit of both the nested version ($\chi^2 = 58.34$, CFI = .984, TLI = .976, RMSEA = .047, SRMR = .036, AIC = 10,431.94) and the version without a nested structure ($\chi^2 = 61.95$, CFI = .981, TLI = .972, RMSEA = .051, SRMR = .034, AIC = 10,435.55) were worse than that of model E.

Cronbach's alpha for the eight-item scale was 0.84 on the individual level and 0.91 on the class level. In the two-factor solution, the first subscale exhibited an alpha value of 0.85 on the individual level and 0.90 on level 2. Cronbach's alpha values for the remaining two items were 0.44 on level 1 and 0.96 on level 2, using the nearest positive definite matrix for level 2.

Average MVPA per day was 80.44 minutes (SD = 21.01, N = 374). The pooled within-group correlation between average MVPA per day and self-efficacy measured by the eight-item scale was 0.19 ($p < .001$, N = 345). Correlations with the items constituting subscales 1 ($r = 0.19$, $p < .001$, N = 350) and 2 ($r = 0.14$, $p < .01$, N = 359) were similar. Considering the 33 classes on level 2, the between-group correlations of MVPA per day and self-efficacy measured by the eight-item scale was $r = 0.65$ ($p < .001$, N = 33). Correlations of MVPA with the first ($r = 0.57$, $p < .001$) and the second subscale ($r = 0.59$, $p < .001$) were comparable.

Discussion

The guidelines for PA (5) are only fulfilled by a minority of children, adolescents and adults (e.g., 12, 15). As individual PA behaviour is often sustained from adolescence to adulthood (e.g., 9), interventions trying to enhance PA of children and adolescents are of great importance. To improve young people's PA behaviour, individual self-efficacy is one of the most important determinants to focus on (e.g., 19, 20). The barriers self-efficacy scale (33) assesses the individual self-efficacy regarding PA of adolescents and implies the findings of the concept analysis of Voskuil and Robbins (31).

In this study a German version of the barriers self-efficacy scale was validated in terms of its factorial validity, internal consistency and criterion validity. Self-efficacy does not only act on the individual level but also on a group level, in this case named collective efficacy (23). Additionally, individual self-efficacy and collective efficacy interact (e.g., 25, 28). Therefore, and since the scale was validated with clustered data, analysis was conducted based on a multi-level framework (36). This way, a mismatch between the constitution of self-efficacy and its assessment and analysis was circumvented. The barriers self-efficacy scale can be applied to measure the construct both on the individual and the group level at the same time by applying the summary index model (53). It suggests that the aggregated variable on the group level can be the sum or the average of a variable assessed at the individual level.

Self-efficacy of our sample was comparable to the self-efficacy of the sample of sixth-graders used to validate the original scale by Dishman and colleagues (33) in terms of the means (3.61 vs. 3.74, see Table 1). Standard deviation was almost identical (0.83 vs. 0.79), kurtosis of the items was similar (-1.10 to 0.03 vs. -1.05 to 0.65).

The fit of the single-level one-factor model A (Table 2) was acceptable which built the legitimization for conducting the subsequent steps of the MCFA. The small ICCs of the items suggested no significant variance between the classes. The fits of the null model, independence model and saturated model did not allow for a clear-cut inference about a statistically significant group-level variance (36, 44).

Concerning the fit indices which are less sensitive to the number of parameters to be estimated, the fits of the multi-level models B and C were better than the one of the single-level model A (see Table 2), which justifies conducting a MCFA. The introduction of a second factor on both levels finally led to the best fit of model E in which it was refrained from specifying a nested structure. The correlation of the two latent factors was large on both levels. Aiming at a preferably parsimonious solution, this would speak in favour of a model with two factors on level 1 and only one factor on level 2, as they fit only slightly worse than model E. However, the decision for model E is justifiable for several reasons.

The wording of the items 2 and 5 addresses the family and peers of the participant. The answers to these items mainly depend on circumstances which cannot be fully controlled by an adolescent. If the parents both work full time and, on top of that, are not interested in being physically active, the child lacks the means to change these circumstances. Similarly, if the best friend does not like to be active and is not reachable within a manageable distance for a child, chances of regularly engaging in PA together are low. Thus, an actually self-efficacious adolescent can disagree with these items while agreeing with the remaining items which refer to more personally controllable aspects and attitudes. The fact that items 2 and 5 show the lowest loadings in the single-level model A and exhibit comparatively low correlations with the remaining items indicate that this scenario occurred in a considerable number of cases which finally argues for a second factor on the individual level.

The sample of this study contained only adolescents attending school in the city of Munich. Living in an urban area with good infrastructure, they are provided with relatively good possibilities to visit their friends on their own by foot, bike or public transport. In a sample including students both from urban and rural areas, the possibilities of visiting the best friend on one's own might differ to a greater extent between classes. In this case, between-group variance specifically concerning item 5 would increase. This speaks in favour of a second factor on level 2 as well, which makes a multi-level approach and finally a 2 × 2-solution even more appropriate.

These content-related considerations need to be corroborated by empirical evidence. Hence, it is of great importance that the parameter estimates of the MCFA support a solution with two factors on both levels. Model E offers the combination of the best fit and the highest loadings with all of them being statistically significant and is thus deemed the best solution. The striving for a more parsimonious solution should be seen critically as it involves a loss of precision and well-interpretable information. On both levels two latent factors can be identified and – although being highly correlated – differentiated in terms of their meaning.

Six items refer to the actual self-efficacy of the adolescent. These items assess the core of the construct. The items 2 and 5 relate to the personal environment regarding support for being physically active. This

interpretation holds for both levels. Individuals as well as school classes can differ in their (average) self-efficacy and in their (average) social resources regarding support of PA.

The high correlations of the two latent factors on both levels ($r \geq 0.72$) reflect their interdependence. An adolescent's self-efficacy for being physically active is associated with the attitudes and behaviour of parents and peers. The concept of vicarious experience (23) suggests that if a person observes another person performing successfully, it can enhance the confidence in the own ability to succeed in the same task. This can lead to the effect that an adolescent's PA behaviour influences his/her best friend and vice versa. Furthermore, the attraction paradigm (54) proclaims that perceived similarity to a peer is a major factor deciding whether a relationship turns into a close friendship or not. Taken together, it can be assumed that close friends often think the same way about being physically active as their similarity led to their friendship in the first place (54) and vicarious experiences help in further assimilating to each other (23). This could explain the correlation of item 5 with the six items assessing the actual PA self-efficacy.

Speaking of vicarious experience it is to mention that, "the greater the assumed similarity, the more persuasive are the models' successes and failures" (23). Since adolescents normally perceive their parents as being less similar to them as their friends, the parents' modeling of PA is not sufficient to enhance their children's PA self-efficacy. Instead, parental PA support finally has a significant positive effect (55). For sixth-graders, particularly parents' emotional and instrumental social support have an effect on the adolescents' PA self-efficacy (56). These findings could explain why responses to item 2 correlate highly with actual self-efficacy.

Finally, combining the possible scenarios described above with the empirical results regarding factor loadings and the correlations of items and latent factors, a strong positive association between actual PA self-efficacy and support of parents and peers can be assumed for many, but not for all adolescents. Thus, although previous validations of the barriers self-efficacy scale supported a unidimensional model (32, 33, 39), the recognition of a conceptually and statistically distinguishable factor assessing PA-related support by parents and peers is regarded advisable.

Reliability was estimated for the individual and the class level separately (36). Cronbach's alpha for the eight-item scale was good on level 1 and excellent on level 2 (57). Cronbach's alpha is positively associated with the number of items (58). Alpha values for the shorter six-item subscale representing actual PA self-efficacy yet were not diminished, which speaks for an even higher internal consistency of this sub-group of items compared to the complete scale. Cronbach's alpha for the two-item support factor was low on level 1 and excellent on level 2. Thus, the association of support from family and peers becomes less ambiguous when the nesting of students in classes is considered. Higher reliability values on the group level were expected since reliability tends to increase and measurement error tends to decrease when measures are aggregated across students within the same classes (59).

Likewise, the use of aggregated measures on the class level normally affects factor loadings and correlations to be higher on this level (59). Whereas in the final model E, loadings for most items are

indeed higher on the class level than on the individual level, the correlations of the two latent factors do not meet this assumption. Although the number of classes included in this study fulfils the minimum amount for conducting a MCFA (60), it might still have diminished the group-level correlation.

Finally, the scores on the complete scale and its subscales were correlated with the actual PA behaviour to evaluate the criterion validity. The participants' scores on both the complete eight-item scale and the two subscales exhibited a significant positive relation to their actual PA behaviour. This is in line with previous research emphasising the role of self-efficacy as an important determinant of healthy PA behaviour of children and adolescents (e.g., 18, 19). The correlations were clearly higher for the class level which again justifies the multi-level approach. Furthermore, it underlines the differentiation between self-efficacy on the individual and the group level, pointing towards the relevance of collective efficacy (23).

Conclusions And Future Directions

Thoroughly validated scales with good psychometric criteria are essential for sound evaluations of cross-sectional studies and intervention programmes. This multi-level validation of the German version of the barriers self-efficacy scale (33) led to a model with two factors on the individual and on the class level, respectively. The factors of actual self-efficacy and support from family and friends are correlated but statistically and conceptually distinguishable. Therefore, it should be discussed if the scale should be kept considered unidimensional. Furthermore, it could be argued if the construct of self-efficacy should be conceptualized as unidimensional or if the support of the social environment like parents and friends is an indispensable, constituent part of self-efficacy.

Whereas the eight-item scale as well as the items of the first factor exhibited good to excellent Cronbach's alpha values, the internal consistency of the second factor is poor on the individual level. While this could have affected the correlation between the two factors as well as the criterion validity (61), it should also be discussed, if the second factor items should be arranged on one factor, since the answers to these items depend on two different groups of people (family and friends).

Additionally, the findings should be verified in a more diverse sample comprising girls and boys of different age and from both rural and urban background. Further research about the construct or, more specifically, about the barriers self-efficacy scale (33) should include more classes on the group level and also more students per class.

Declarations

Additional files

Additional file 1: file name: sample; format: .docx; description: info regarding how the sample was recruited, how representative the sample was of the target group, how the analysed sample differed from the recruited sample and how any missing data were handled

Additional file 2: file name: data set; format: .sav; description: minimal dataset necessary to replicate the analysis

Ethics approval and consent to participate

The study was approved by the ethics commission of the Technical University of Munich (155/16 S) and the Ministry of culture and education of the state of Bavaria in Germany.

Students did not participate if they had not provided a consent form signed by them and their parents before.

Consent for publication

Not applicable.

Availability of data and materials

All data analysed during this study are included in the supplementary information files of this article.

Competing interests

The authors declare that they have no competing interests.

Funding

The study is funded by the German Research Foundation (DE 2680/3–1). The researchers are independent of the funders who have no influence on study design, conduct, analyses, or interpretation of the data, the decision to submit the results or the preparation of the manuscript.

Authors' contributions

JB collected, analysed and interpreted the data and drafted the manuscript. DJS also collected data and contributed to the data analysis. SH lent substantial support in the analysis and interpretation of the data. YD supervised the cooperation of the authors and revised the manuscript critically for important intellectual content. All authors read and approved the final manuscript.

Acknowledgements

The authors would like to thank the student assistants for their support during the data assessments.

Abbreviations

AIC: Akaike's information criterion; BMI: body mass index; CFI: comparative fit index; df: degrees of freedom; HISEI: Highest International Socioeconomic Index of occupational status; ICC: intracluster correlation; ISCO: International Standard Classification of Occupation; ISEI: International Socioeconomic Index of occupational status; M: mean; MCFA: multi-level confirmatory factor analysis; MVPA: moderate-to-vigorous physical activity; PA: physical activity; PE: physical education; RMSEA: root mean square error of approximation; SD: standard deviation; SES: socioeconomic status; SRMR: standardized root mean square residual; TLI: Tucker-Lewis index; WHO: World Health Organization

References

1. Granger E, Di Nardo F, Harrison A, Patterson L, Holmes R, Verma A. A systematic review of the relationship of physical activity and health status in adolescents. *European journal of public health*. 2017;27(suppl_2):100-6.
2. Lee IM, Shiroma EJ, Lobelo F, Puska P, Blair SN, Katzmarzyk PT. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. *The Lancet*. 2012;380(9838):219-29.
3. McKinney J, Lithwick DJ, Morrison B, Nazzari H, Isserow SH, Heilbron B, et al. The health benefits of physical activity and cardiorespiratory fitness 2016. 131-7 p.
4. Warburton DER, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. *CMAJ : Canadian Medical Association Journal*. 2006;174(6):801-9.
5. World Health Organization. Global action plan on physical activity 2018–2030: more active people for a healthier world. Geneva, Switzerland: World Health Organization; 2018.
6. Archer T. Health Benefits of Physical Exercise for Children and Adolescents. *Journal of Novel Physiotherapies*. 2014;04(02).
7. Hallal PC, Wells JC, Reichert FF, Anselmi L, Victora CG. Early determinants of physical activity in adolescence: prospective birth cohort study. *BMJ*. 2006;332(7548):1002-7.
8. Loprinzi PD, Cardinal BJ, Loprinzi KL, Lee H. Benefits and Environmental Determinants of Physical Activity in Children and Adolescents. *Obesity Facts*. 2012;5(4):597-610.

9. Telama R. Tracking of physical activity from childhood to adulthood: a review. *Obes Facts*. 2009;2(3):187-95.
10. Finger JD, Varnaccia G, Borrmann A, Lange C, Mensink G. Körperliche Aktivität von Kindern und Jugendlichen in Deutschland – Querschnittergebnisse aus KiGGS Welle 2 und Trends2018.
11. Dumith SC, Gigante DP, Domingues MR, Kohl HW, 3rd. Physical activity change during adolescence: a systematic review and a pooled analysis. *Int J Epidemiol*. 2011;40(3):685-98.
12. Van Hecke L, Loyen A, Verloigne M, van der Ploeg HP, Lakerveld J, Brug J, et al. Variation in population levels of physical activity in European children and adolescents according to cross-European studies: a systematic literature review within DEDIPAC. *International Journal of Behavioral Nutrition and Physical Activity*. 2016;13(1):70.
13. Hägggi JM, Phillips LR, Rowlands AV. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *J Sci Med Sport*. 2013;16(1):40-4.
14. Migueles JH, Cadenas-Sanchez C, Tudor-Locke C, Löff M, Esteban-Cornejo I, Molina-Garcia P, et al. Comparability of published cut-points for the assessment of physical activity: Implications for data harmonization. *Scandinavian Journal of Medicine & Science in Sports*. 2019;29(4):566-74.
15. Hallal PC, Andersen LB, Bull FC, Guthold R, Haskell W, Ekelund U. Global physical activity levels: surveillance progress, pitfalls, and prospects. *The Lancet*. 2012;380(9838):247-57.
16. World Health Organization. Global recommendations on physical activity for health. Geneva, Switzerland: World Health Organization; 2010.
17. Sallis JF, Owen N, Fisher E. Ecological Models of Health Behavior. In *Health Behavior and Health Education: Theory, Research and Practice* 2008. 465-82 p.
18. Bauman AE, Reis RS, Sallis JF, Wells JC, Loos RJF, Martin BW. Correlates of physical activity: why are some people physically active and others not? *The Lancet*. 2012;380(9838):258-71.
19. Cortis C, Puggina A, Pesce C, Aleksovskaja K, Buck C, Burns C, et al. Psychological determinants of physical activity across the life course: A "DEterminants of Diet and Physical ACTivity" (DEDIPAC) umbrella systematic literature review. *PLoS One*. 2017;12(8):e0182709.
20. Craggs C, Corder K, van Sluijs EM, Griffin SJ. Determinants of change in physical activity in children and adolescents: a systematic review. *Am J Prev Med*. 2011;40(6):645-58.
21. van Stralen MM, Yildirim M, te Velde SJ, Brug J, van Mechelen W, Chinapaw MJ. What works in school-based energy balance behaviour interventions and what does not? A systematic review of mediating mechanisms. *Int J Obes (Lond)*. 2011;35(10):1251-65.
22. Lubans DR, Foster C, Biddle SJ. A review of mediators of behavior in interventions to promote physical activity among children and adolescents. *Prev Med*. 2008;47(5):463-70.
23. Bandura A. Self-efficacy: The exercise of control. New York, NY, US: W H Freeman/Times Books/ Henry Holt & Co; 1997. ix, 604-ix, p.
24. Feltz DL, Short SE, Sullivan PJ. Self-efficacy in sport. Champaign, IL, US: Human Kinetics; 2008. x, 405-x, p.

25. Magyar TM, Feltz DL, Simpson IP. Individual and Crew Level Determinants of Collective Efficacy in Rowing. *Journal of Sport & Exercise Psychology*. 2004;26(1):136-53.
26. Wang S-L, Lin SSJ. The effects of group composition of self-efficacy and collective efficacy on computer-supported collaborative learning. *Computers in Human Behavior*. 2007;23(5):2256-68.
27. Gully SM, Incalcaterra KA, Joshi A, Beaubien JM. A meta-analysis of team-efficacy, potency, and performance: Interdependence and level of analysis as moderators of observed relationships. *Journal of Applied Psychology*. 2002;87(5):819-32.
28. Greenlees I, Graydon JK, Maynard IW. The impact of collective efficacy beliefs on effort and persistence in a group task. *J Sports Sci*. 1999;17(2):151-8.
29. Greenlees I, Graydon J, Maynard I. The impact of individual efficacy beliefs on group goal selection and group goal commitment 2000. 451-9 p.
30. Ede A, Hwang S, Feltz DL. Current directions in self-efficacy research in sport. *Revista Iberoamericana de Psicología del Ejercicio y el Deporte*. 2011;6(2):181-201.
31. Voskuil VR, Robbins LB. Youth physical activity self-efficacy: a concept analysis. *Journal of advanced nursing*. 2015;71(9):2002-19.
32. Motl RW, Dishman RK, Trost SG, Saunders RP, Dowda M, Felton G, et al. Factorial validity and invariance of questionnaires measuring social-cognitive determinants of physical activity among adolescent girls. *Prev Med*. 2000;31(5):584-94.
33. Dishman RK, Hales DP, Sallis JF, Saunders R, Dunn AL, Bedimo-Rung AL, et al. Validity of social-cognitive measures for physical activity in middle-school girls. *J Pediatr Psychol*. 2009;35(1):72-88.
34. Borgers N, de Leeuw E, Hox J. Children as Respondents in Survey Research: Cognitive Development and Response Quality 1. *Bulletin of Sociological Methodology/Bulletin de Méthodologie Sociologique*. 2000;66(1):60-75.
35. Scott J. *Children as Respondents: Methods for Improving Data Quality* 1997.
36. Huang F. *Conducting Multilevel Confirmatory Factor Analysis Using R* 2017.
37. Demetriou Y, Bachner J. A school-based intervention based on self-determination theory to promote girls' physical activity: study protocol of the CReActivity cluster randomised controlled trial. *BMC public health*. 2019;19(1):519.
38. Ganzeboom H. A new International Socio-Economic Index (ISEI) of occupational status for the International Standard Classification of Occupation 2008 (ISCO-08) constructed with data from the ISSP 2002–2007. 2010.
39. Dishman RK, Motl RW, Saunders RP, Dowda M, Felton G, Ward DS, et al. Factorial invariance and latent mean structure of questionnaires measuring social-cognitive determinants of physical activity among black and white adolescent girls. *Prev Med*. 2002;34(1):100-8.
40. Brislin RW. Back-Translation for Cross-Cultural Research. *Journal of Cross-Cultural Psychology*. 1970;1(3):185-216.

41. Cha ES, Kim KH, Erlen JA. Translation of scales in cross-cultural research: issues and techniques. *Journal of advanced nursing*. 2007;58(4):386-95.
42. Julian MW. The Consequences of Ignoring Multilevel Data Structures in Nonhierarchical Covariance Modeling. *Structural Equation Modeling: A Multidisciplinary Journal*. 2001;8(3):325-52.
43. Rosseel Y. lavaan: An R Package for Structural Equation Modeling. *Journal of Statistical Software*; Vol 1, Issue 2 (2012). 2012.
44. Hox JJ. Multilevel analysis: Techniques and applications, 2nd ed. New York, NY, US: Routledge/Taylor & Francis Group; 2010. x, 382-x, p.
45. Hu Lt, Bentler PM. Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*. 1999;6(1):1-55.
46. Korner-Nievergelt F, Roth T, von Felten S, Guélat J, Almasi B, Korner-Nievergelt P. Chapter 11 - Model Selection and Multimodel Inference. In: Korner-Nievergelt F, Roth T, von Felten S, Guélat J, Almasi B, Korner-Nievergelt P, editors. *Bayesian Data Analysis in Ecology Using Linear Models with R, BUGS, and STAN*. Boston: Academic Press; 2015. p. 175-96.
47. Preacher KJ, Zhang G, Kim C, Mels G. Choosing the Optimal Number of Factors in Exploratory Factor Analysis: A Model Selection Perspective. *Multivariate Behavioral Research*. 2013;48(1):28-56.
48. Revelle WR. *psych: Procedures for personality and psychological research*. 2017.
49. Higham N. Computing the Nearest Correlation Matrix - a Problem From Finance. *IMA Journal of Numerical Analysis*. 2002;22.
50. Choi L, Liu Z, Matthews CE, Buchowski MS. Validation of accelerometer wear and nonwear time classification algorithm. *Med Sci Sports Exerc*. 2011;43(2):357-64.
51. Hanggi JM, Phillips LR, Rowlands AV. Validation of the GT3X ActiGraph in children and comparison with the GT1M ActiGraph. *J Sci Med Sport*. 2013;16(1):40-4.
52. Migueles JH, Cadenas-Sanchez C, Ekelund U, Delisle Nystrom C, Mora-Gonzalez J, Lof M, et al. Accelerometer Data Collection and Processing Criteria to Assess Physical Activity and Other Outcomes: A Systematic Review and Practical Considerations. *Sports medicine (Auckland, NZ)*. 2017;47(9):1821-45.
53. Chen G. A FRAMEWORK FOR CONDUCTING MULTI-LEVEL CONSTRUCT VALIDATION. In: Mathieu John E, Francis JY, Fred D, editors. *Multi-level Issues in Organizational Behavior and Processes. Research in Multi-Level Issues*. 3: Emerald Group Publishing Limited; 2005. p. 273-303.
54. Byrne D. *The attraction paradigm*: Academic press New York; 1971.
55. Trost SG, Sallis JF, Pate RR, Freedson PS, Taylor WC, Dowda M. Evaluating a model of parental influence on youth physical activity. *Am J Prev Med*. 2003;25(4):277-82.
56. Peterson MS, Lawman HG, Wilson DK, Fairchild A, Van Horn ML. The association of self-efficacy and parent social support on physical activity in male and female adolescents. *Health Psychol*. 2013;32(6):666-74.

57. Kubinger K. Psychologische Diagnostik: Theorie und Praxis psychologischen Diagnostizierens (2., überarb. und erw. Aufl.)2009.
58. Tang W, Cui Y, Babenko O. Internal consistency: Do we really know what it is and how to assess it? 2014. 205-20 p.
59. Byrne BM. Structural equation modeling with Mplus: Basic concepts, applications, and programming. New York, NY, US: Routledge/Taylor & Francis Group; 2012. xvii, 412-xvii, p.
60. Kreft I, de Leeuw J. Introducing multilevel modeling. Thousand Oaks, CA, US: Sage Publications, Inc; 1998.
61. Dedrick R, Greenbaum P. Multilevel Confirmatory Factor Analysis of a Scale Measuring Interagency Collaboration of Children's Mental Health Agencies. Journal of emotional and behavioral disorders. 2011;19:27-40.