

Four classes of physical fitness in German children and adolescents: only differences in performance or at-risk groups?

Lena Lämmle · Matthias Ziegler · Ilka Seidel · Annette Worth · Klaus Bös

Received: 6 April 2011 / Revised: 30 October 2012 / Accepted: 12 November 2012 / Published online: 30 November 2012
© Swiss School of Public Health 2012

Abstract

Objectives Regarding children's and adolescents' development of physical fitness (PF), information about whether there are different performance classes in PF would potentially be helpful. If such classes also differ in related health parameters, this might indicate 'at-risk' groups requiring different support measures.

Methods Factor mixture modeling was used to investigate whether participants of the German "Motorik"-Module (2003–2006) aged 6–17 years ($N = 4,529$) differ in their PF as described by endurance, strength, coordination, and flexibility.

Results Four different classes of PF emerged after controlling for age and gender. Participants in classes 1 and 2 either performed well in all coordination or in all strength tests. Participants in class 3 and class 4 performed poorly in

all coordination and in the push up (strength) tasks. Only participants in class 4 revealed significantly higher BMI values and less physical activity suggesting they are, in fact, an 'at-risk' group.

Conclusions Different support measures are needed to enable adequate development. There might be a risk group demonstrating comparably low coordination that cannot be spotted by looking at BMI or activity level.

Keywords Physical fitness · Physical activity · Body mass index · Skinfold thickness · Factor mixture modeling

Introduction

Physical fitness (PF), with the two types of health-related and skill/performance-related fitness, represents an essential mainstay of health in children and adolescents (Cantell et al. 2008; Hoekstra et al. 2008; Malina 2001). Nevertheless, in accordance with international results, recent longitudinal findings have shown a decreasing level of both PF types in German children and adolescents (Bös 2003; Kretschmer and Wirsing 2007). Thus, appropriate intervention programs are needed to reinforce levels of PF. To establish well-designed and well-implemented PF-promoting interventions, at least two sources of information are needed. First of all, certain performance classes for PF should be identified as target groups for individual interventions and, as a second step, associated relations to other health parameters such as physical activity (PA) or body mass index (BMI) should be examined (Cantell et al. 2008; Hands and Larkin 2001). If such relations could be found—for example that children and adolescents of lower performance classes for PF also have lower PA levels and

L. Lämmle (✉)
Department of Sports Psychology, Technische Universität München, Georg-Brauchle-Ring 60/62, 80992 Munich, Germany
e-mail: lena.laemmle@tum.de

M. Ziegler
Psychological Assessment, Humboldt Universität zu Berlin, Berlin, Germany

I. Seidel
Department for Strength-oriented and Technical-acrobatical Sports, Institute for Applied Training Science Leipzig, Leipzig, Germany

A. Worth
Institute for Movement and Sport, Pädagogische Hochschule Karlsruhe, Karlsruhe, Germany

K. Bös
Institute of Sport and Sports Sciences, Karlsruhe Institute of Technology, Karlsruhe, Germany

higher BMI scores—‘at-risk’ groups could be identified. Interestingly, even though there is consensus about the need to identify PF performance classes, there is no consensus about what the most efficient methodological approach is. Furthermore, research using other health indices based on performance classes to identify ‘at-risk’ groups is rare and mainly focused on such more highly specified performance classes such as children with motor learning difficulties (Hands and Larkin 2001; Cantell et al. 2008). Therefore, the first aim of the present paper is use factor mixture modeling (FMM) to identify performance classes of PF in German children and adolescents. If different performance classes can be found, the second aim is to examine whether other health parameters differ for such performance classes. Research with children up to the age of young adults has shown that PF is correlated with PA, body weight, and body fat (Deforche et al. 2003; Graf et al. 2004; Minck et al. 2000; Okely et al. 2004; Ruiz et al. 2007). In summary, if performance classes of PF emerge in the present data based on FMM analyses, we will also investigate whether there are differences of PA and body composition (BC; operationalized as body mass, body fatness or skinfold thickness; Brunet et al. 2007) for the identified classes of PF in order to establish whether some of the classes can be declared as ‘at-risk’ groups, i.e. a combination of poor PF and poor health indicators. If such relationships exist, they could provide meaningful insights for intervention programs. This is especially true if PF class differences can be related to PA and BC, where fruitful interventions can be made.

Methods

Sample and study

The following analyses were based on the German Health Interview and Examination Survey for Children and Adolescents (KiGGS), which was conducted between May 2003 and May 2006. The KiGGS survey¹ is a nationwide, cross-sectional study on the health status of children and adolescents between the ages of 0 and 17 years, conducted by the Robert Koch Institute of Berlin (RKI; Kurth 2007). The KiGGS survey was complemented by the “Motorik”-Module (MoMo, Karlsruhe 2003–2006)² for a more differentiated recording of physical activity and physical fitness in children and adolescents between the ages of 4

and 17 years (Bös et al. 2009). The present study focuses on this latter data set. Data were collected from 4,529 children and adolescents (2,244 girls and 2,285 boys) between the ages of 4 and 17 with an average age of 9.45 years (SD = 4.01) with the use of a stratified two-stage probability sampling procedure. The primary sample units were 150 research sites of all political communities in Germany. In the second step, participants were randomly selected from the official registers of local residents’ registration offices of these communities. Participation was voluntary. The response rate was 66.6 % with 5.3 % quality-neutral dropouts (Kamtsiuris et al. 2007). Because of the high demand on time participants were tested twice. They first filled out the KiGGS questionnaires, participated in physical examinations and tests as well as in a personal interview conducted by a physician. The MoMo data (Karlsruhe 2003–2006) about PA (questionnaire) and PF (measurements) were collected at a later date (Hölling et al. 2007).

Measurements

Physical fitness. Participants were tested with eight tests of the MoMo test battery (Karlsruhe 2003–2006; for more detailed information see Lämmle et al. 2010). The test battery proportionally consists of tests for health-related fitness abilities including endurance, strength, and flexibility, as well as tests for skill-related fitness abilities including coordination under precision demands and under time pressure (see Table 1, Fig. 1). The objectivity, reliability and content-related validity of all test items were tested in several different age groups within the entire age range (Bös et al. 2009). The content-related validity of all tests was evaluated as good ($M_{\text{Significance}} = 1.9$, $M_{\text{Practicability}} = 1.7$) throughout with regard to significance and feasibility as based upon expert ratings. More precisely, the test development was based on an international expert questionnaire involving 40 selected fitness experts in 25 European countries who were asked about the relevance of the test contents and requirements in sport-motoric tests regarding the documentation of PF (Bös 1992). Subsequently, 13 experts evaluated the significance and practicability of the study exercises on a scale of 1 (very good) to 5 (very bad). Further, to determine test–retest reliability, the motoric tests were performed twice within 4 days on the same children, applying the same test situation and the same study investigator. Good test–retest reliability coefficients ($r_{\text{min}} = 0.74$, $r_{\text{max}} = 0.96$) were found (Bös et al. 2009).

Body composition. The body height (stature) was measured standing up with a calibrated portable Holtain-Stadiometer and an accuracy of up to 0.1 cm. The body weight (mass) was measured in underwear on a calibrated

¹ The KiGGS survey was sponsored by the Federal Ministry of Health (BMG) as well as the Federal Ministry of Education and Research (BMBF) in the years 2002–2006.

² The MoMo study is sponsored by the Federal Ministry for Families, Senior Citizens, Women and Adolescents (BMFSFJ) from the years 2002 to 2008.

Table 1 Measures of physical fitness in the MoMo test battery (Karlsruhe 2003–2006)

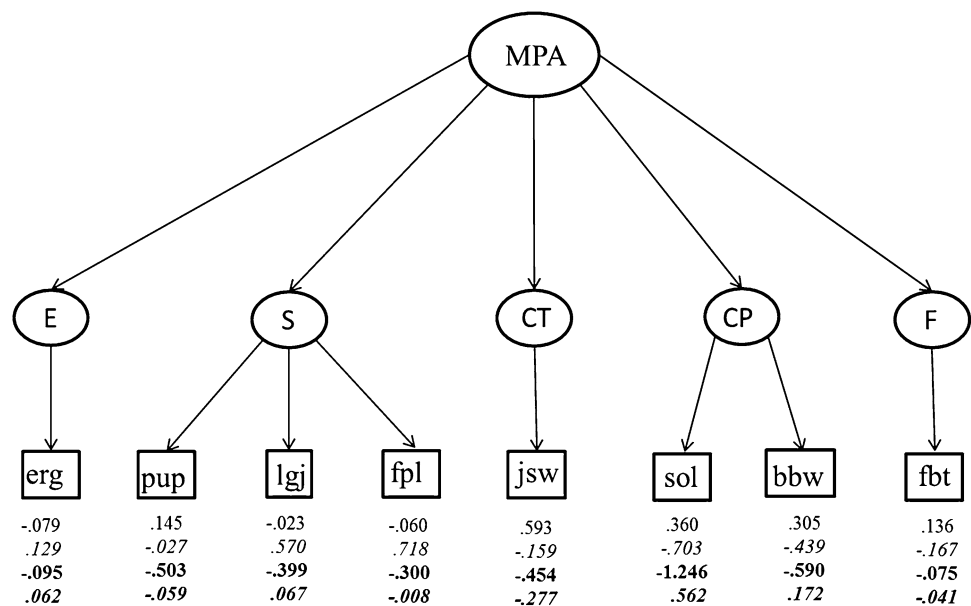
Types of PF	PF abilities	Tests
HRF	E	Bicycle endurance test [aerobic fitness performance, cardiovascular system; The test starts with a calculated initial load of 0.5 W kg ⁻¹ body weight. This is then followed by an increase in load of a further 0.5 W kg ⁻¹ body weight every 2 min. The test is then discontinued for any one of the following three reasons: (1) when a load heart rate of 190 (6–10 years) or 180 (11–17 years) beats min ⁻¹ is observed; (2) when the frequency of rotation decreases below 50 revolutions min ⁻¹ for a period of at least 20 s; (3) when the subject stops due to exhaustion. The variable used for analysis is the wattage associated with a heart rate of 170 divided by the body weight (relative PWC 170). Since the health-related MPA should be assessed, PWC 170 and not the maximum heart rate load has been chosen as an endurance criterion]
	S	Standing long jump [jumping power, springiness of the leg muscles; The task to be performed by the test subject is to jump as far as possible using both legs together. The subject stands with their legs parallel and with bent knees on the starting line. One is allowed to increase the propulsion only by swinging one's arms. The jump is performed using both legs and by landing on both feet, while one is not permitted to grasp backwards with one or both hands. The distance from the starting line to the heel of the foot furthest back after landing is measured (in cm). The best of two jumps was used for analysis]. Force plate jump (capability of the leg extenders for demonstrating springiness; The subject stands in place on the measurement plate while keeping his/her hands positioned on the hips. No additional propulsion is to be achieved by using the arms. The subject must obtain his/her propulsion from the standing position only by bending their legs. The test participants are supposed to jump upwards in a vertical direction as high as possible and land again on the platform. The power-time history of the reaction strength on the force plate is measured and evaluated with the help of a computer employing an analog/digital converter. The height jumped in meters is then computed as an important parameter. The variable used for analysis was the best out of three attempts. Between each of the three high jumps, there was a rest of 30 s each). Push-ups (dynamic muscular endurance in the upper extremities; The test participants had to perform as many push-ups as possible within 40 s. The number of correctly performed push-ups is counted. The test subject lies in a prone position on his/her stomach and the hands grasp one another on the buttocks. The hands are released from behind the back, placed beside the shoulders, which are then pressed toward the floor until the arms are extended and the body leaves the ground. Subsequently, one hand is released from the ground and claps onto the other hand. During this process, only the hands and the feet have contact with the ground. The trunk and the legs are extended. A lordosis should be avoided. Afterwards, the arms are kept flexed toward the body while still in a prone position and the initial position is assumed once again. Before a new push-up is performed, the test subject must grasp his/her hands behind their back. From this position, the push-ups are counted. The variable used for analysis was the number of correctly performed push-ups in 40 s)
	F	Forward bending of the trunk (flexibility of the trunk, elasticity of the back and leg muscles; The subject stands on a wooden box and slowly bends forward at the waist. The arms and hands must reach as far as possible downward. Hereby, the legs must remain extended. The maximum position of bending is held for 2 s. The better of two attempts is recorded in centimeters)
SRF	CP	Standing on one leg (sensomotoric regulation while performing an exercise involving precision, coordination for the entire body; The task for the test subject is to stand on one leg, chosen by him/herself, on a T-shaped balancing bar for 1 min. If the free foot touches the ground, the test subject must try to immediately recapture the balancing position on the T-bar. The timer continues to run during this short contact with the ground. However, if one completely leaves the bar, the timer is stopped until the subject is able to resume the same initial position. The person is not allowed to change the standing leg during the test. The objective is to avoid touching the ground for as long as possible. The arms can be used to aid balance. The number of contacts made with the ground is counted. The variable used for analysis was the number of ground contacts of the leg being held up during the 1 min). Balancing backwards (sensomotoric regulation while performing an exercise involving precision, coordination for the entire body; The test participants must walk backwards and keep their balance on three different-sized beams. The test always begins from a start platform. The widths of the beams are 3, 4.5 and 6 cm and they are each 300 cm long. The goal is to stay on each of the beams during the course of two valid attempts. A total of six successful attempts are evaluated. The number of steps made while walking backwards is counted. The variable used for analysis was the sum of steps made over all six attempts while walking backwards)
	CT	Jumping sideways (total body coordination under time pressure, speed, muscular endurance capabilities of the lower extremities; Over the course of 15 s, the test participants must jump with both legs at the same time, as quickly as possible, sideways over the middle line of a carpet mat. Two 15 s sets are performed with a 1 min break between the two phases. Evaluated is the number of jumps made over the course of the two sets. The variable used for analysis was the mean value of the two attempts at jumping from side to side)

PF types of physical fitness, PF abilities physical fitness abilities, HRF health-related fitness, SRF health-related fitness, E endurance, S strength, F flexibility, CP coordination with precision requirement, CT coordination under time pressure

scale with an accuracy of up to 0.1 kg. The body mass index (BMI) was calculated as an index for the estimation of the BC [body weight (kg)/height (m²)]. Because the

BMI only insufficiently reflects the excessive increase of fatty tissue and its distribution, the skinfold thickness (SFT) was assessed by Holtain-Caliper using two

Fig. 1 A two-level model of motor performance ability (Karlsruhe 2003–2006). *MPA* motor performance ability, *E* endurance, *S* strength, *CT* coordination under time pressure, *CP* coordination with precision requirement, *F* flexibility, *erg* bicycle endurance test, *pup* push-ups, *lgj* long jump, *fpl* force plate, *jsw* jumping sideways, *sol* standing on one leg, *bbw* balancing backwards, *fbt* forward bending of the trunk. Words in standard class 1, words in *italic* class 2, words in **bold and italic** class 3, words in **bold** class 4



measurement points on the right side: triceps (the back of the upper arm) and subscapula (beneath the edge of the shoulder blade), in order to estimate the body fat.

Physical activity. PA was assessed using a questionnaire containing previously validated items involving the duration, frequency and intensity of PA in leisure time, school, and sports clubs in the past year. It consists of 35 items and has been developed in orientation to existing questionnaires (Bös et al. 2009). The test–retest reliability in the previous 7 day longitudinal study ranged from $r_{tt} = 0.72$ (PA in leisure time) to $r_{tt} = 0.93$ (PA in sports clubs). Construct validity-related evidence was gathered by looking at the correlation with a multisensory electronic monitor including a biaxial accelerometer measuring PA (SenseWear Pro 2). Results have shown higher correlations ($r_{min} = 0.56$, $r_{max} = 0.66$) in comparison with other PA questionnaires (e.g., Treuth et al. 2005).

In accordance with international recommendations (President’s Council on Physical Fitness and Sports 2008; World Health Organization (WHO) 2006) for activity guidelines with at least 60 min of moderate to vigorous activities daily, a PA parcel was built. For school, leisure time, and sports club activity, only time spent from at least moderate PA in minutes (min) was regarded: number of school sports hours \times 30 min + frequency of first sport \times duration first sport \times number of months first sport is undertaken/12 (months a year) + frequency of second sport \times duration second sport \times number of months second sport is undertaken/12 (months a year) + frequency of third sport \times duration third sport \times number of months third sport is undertaken/12 (months a year) + frequency of fourth sport \times duration fourth sport \times number of months fourth sport is undertaken/12 (month a year).

Sustained moderate to vigorous PA was documented as associated with positive health outcomes (Sallis et al. 2000).

Statistical analyses

One possible way of identifying performance classes of PF is to categorize people using cut-off points on the score distribution (Dwyer 1996). Another way to verify performance classes of PF is to use cluster analysis (Wishart 2006). However, both techniques do not provide information indicating the relevant (PF) measures that might be causing the main differences between the groups. Both sorts of information are given in Mixed-Rasch analysis (Rost 2004), which is a combination of a Rasch model and Latent Class analysis. Unfortunately, the Mixed-Rasch analysis cannot easily be applied to continuous data such as that in the present study. An alternative approach is the Factor Mixture Model (FMM; Lubke and Muthén 2005). This method can handle continuous data but also includes a classification method. The basic idea is that latent variables can be modeled using structural equation modeling. Once such models have been established and tested, a Latent Class analysis checks whether the population is heterogeneous with regard to the latent variable. In other words, we test whether there are qualitatively different classes with regard to the latent variables captured by the measurement instruments. A further advantage of FMM is that the approach then allows us to test whether the heterogeneity really originates from differences in the latent variable or rather from differences in the indicators. To this end tests for measurement invariance can be used (Chen et al. 2005; Sass 2011). In this study we tested whether the model specified as representing the relations between the

measures satisfies the conditions for configural, factorial, and strong factorial invariance across the classes found.

In summary, FMMs consist of latent class and common factor models and are used to explore unobserved population heterogeneity. Based on this combination it is possible to test whether population heterogeneity can be found on latent variables (endurance (E), strength (S), coordination under precision demands (CP), under time pressure (CT), flexibility (F)) or on observed variables (E: bicycle endurance test; S: force plate jump, standing long jump, push-ups; CT: jumping sideways; CP: standing on one leg, balancing backwards; F: forward bending of the trunk; Lubke and Muthén 2005).

Descriptive statistics, regression analysis, analysis of variance and factor mixture models are tested in MPlus 5.2 (Muthén and Muthén 1998–2007) and SPSS 19. As gender and age have considerable (Pissanos et al. 1983) but no health-related impacts on PF, both were treated as covariates in this study. Regression analyses using enter method with age and gender as predictors for each PF measure as criterion were conducted and the standardized residuals were saved and used for FFM tests. A visual inspection of the residuals revealed only minor violations of the normal distribution. Therefore, for all FMM model tests a robust maximum likelihood estimator (MLR) is used to correct standard errors for possible violations of the multivariate normal distribution. To assess the model fit cutoff suggestions by Beauducel and Wittmann (2005) and Hu and Bentler (1999) are applied. In addition to the χ^2 value and its significance we therefore also consider the SRMR (which should be below 0.11), the RMSEA (which should be equal to or less than 0.06), and the CFI (which should be equal to or higher than 0.95). In order to decide which class solution to chose for the FMM, we used the adjusted likelihood ratio test (aLRT) statistic Lo-Mendell-Rubin as well as the Consistent Akaike's Information Criteria (CAIC; Lubke and Neale 2006). After determining the

number of classes using FMM, classes are compared with regard to measurement invariance. If the assumptions of configural invariance hold, comparable but not identical latent variables are measured in the classes. Factorial invariance (equal loadings in the classes) would mean that the same latent variable is measured using the same metric but with a different origin. Strong factorial invariance (equal loadings and intercepts) would mean the same latent variable is measured using the same metric and the same origin. In that case all observed differences between the classes would be due to differences on the latent variable. If no strong factorial invariance can be established, class differences can also lie within the intercepts of the indicators. That would mean that the classes found differ with respect to the performance that is considered average within each class. To test for invariance we applied the cutoffs suggested by Chen (2007): $\Delta CFI > 0.010$ and $\Delta RMSEA > 0.015$.

Results

Descriptive statistics and correlations

For sample description and further interpretations, Table 2 provides raw-score means, and standard deviations of the PF items, BMI, SFT and PA. Further results of the regression analyses (adjusted R^2 and standardized β -coefficients) are also presented in Table 2.

Three performance classes of PF and one 'at-risk' group

The baseline model fit the data well $\chi^2(18) = 149.77$, $p < 0.001$, RMSEA = 0.040, 90 % confidence interval: 0.034–0.046, CFI = 0.99. The Lo-Mendell-Rubin adjusted likelihood ratio test was significant for all solutions

Table 2 Means (M) and standard deviations (SD) of the physical fitness items, body mass index (BMI), skinfold thickness (SFT) and physical activity (PA); adjusted R^2 and standardized β -coefficients for endurance (bet), strength (fpl, lgj, pup), coordination under pressure

	bet	fpl	lgj	pup	jsw	sol	bbw	fbt	BMI	SFT	PA
M	1.99	0.27	133.27	11.74	23.38	-10.70	26.79	-0.09	18.11	22.20	345.23
SD	0.49	0.09	38.64	3.84	11.07	9.42	12.30	7.48	3.73	10.90	199.62
R^2	0.17	0.66	0.65	0.19	0.73	0.42	0.38	0.06	0.37	0.19	0.05
β_{age}	0.17	0.80	0.78	0.43	0.85	0.64	0.61	-0.03	0.61	0.39	0.13
β_{gender}	-0.38	-0.13	-0.19	-0.07	0.02	0.08	0.09	0.25	0.01	0.19	-0.19

of time (jsw), coordination under precision demands (sol, bbw) and flexibility (fbt) as criterion and gender and age as predictors (Karlsruhe 2003–2006)

bet bicycle endurance test (watt), *fpl* force plate jump (cm), *lgj* standing long jump (cm), *pup* push-ups (number over the course of 40 s), *jsw* jumping sideways (number over the course of 15 s), *sol* standing on one leg (number of not allowed contacts made with the ground on a T-bar over the course of 60 s), *bbw* balancing backwards (steps on a beam; maximum of possible steps: 8 steps), *fbt* forward bending of the trunk (cm), *BMI* body mass index (kg/m^2), *SFT* skinfold thickness (mm), *PA* physical activity (min)

except the five-class solution ($p = 0.08$), which was therefore discarded. According to the CAIC information criterion, the solution with four classes fitted the data best (class 1: 85,734.94, class 2: 71,336.92, class 3: 70,644.25, class 4: 70,034.51). The first class contained $N = 153$, the second class $N = 502$, the third class $N = 904$ and the fourth class contained $N = 1,585$ of children and adolescents. The assumed PF structure for the four-class solution revealed an acceptable degree of overall model fit, $\chi^2(62) = 411.63$, $p < 0.001$, RMSEA = 0.061, 90 % confidence interval: 0.055–0.067, CFI = 0.96.

In a next step, we tested for measurement invariance. Results can be found in Table 3.

As can be seen both criteria for factorial invariance are satisfied but not for strong factorial invariance. Thus, differences between the classes could be due to differences in mean task performance. We, therefore, compared the intercepts for the four classes in a next step. Unstandardized intercepts and their bias corrected confidence intervals for the four classes are depicted in Fig. 2 and Table 4.

All measures for the four classes vary within $\pm 1SD$ except the standing on one leg measure for coordination

Table 3 Measurement invariance (Karlsruhe 2003–2006)

	χ^2	df	P^a	RMSEA	CFI ^b	SRMR	Δ RMSEA	Δ CFI
Configural	354.09	73	<0.001	0.058	0.966	0.04		
Factorial	395.51	79	<0.001	0.06	0.962	0.044	0.002	-0.004
Strong factorial	1,343.2	88	<0.001	0.112	0.849	0.082	0.052	-0.113

^a Bollen–Stine p value
^b 90 % confidence interval

Fig. 2 Unstandardized intercepts for the four classes (Karlsruhe 2003–2006). *erg* bicycle endurance test, *pup* push-ups, *lgj* standing long jump, *fpl* force plate jump, *jsw* jumping sideways, *bbw* balancing backwards, *sol* standing on one leg, *fbt* forward bending of the trunk

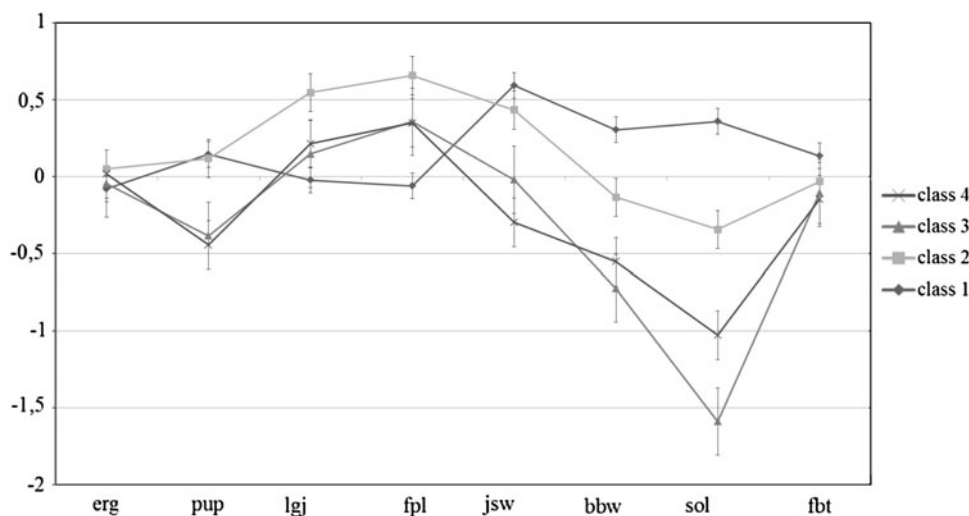


Table 4 Unstandardized intercepts for the four classes (Karlsruhe 2003–2006)

	Class 1			Class 2			Class 3			Class 4		
	I	95 % CI		I	95 % CI		I	95 % CI		I	95 % CI	
		LL	UL		LL	UL		LL	UL		LL	UL
erg	-0.079	-0.13	-0.03	0.129	0.03	0.23	-0.095	-0.24	0.06	0.062	0.00	0.12
pup	0.145	0.10	0.20	-0.027	-0.12	0.06	-0.503	-0.61	-0.40	-0.059	-0.11	-0.01
lgj	-0.023	-0.08	0.03	0.570	0.47	0.67	-0.399	-0.46	-0.34	0.067	0.02	0.11
fpl	-0.060	-0.11	-0.01	0.718	0.61	0.83	-0.300	-0.36	-0.24	-0.008	-0.05	0.04
jsw	0.593	0.54	0.65	-0.159	-0.25	-0.07	-0.454	-0.51	-0.41	-0.277	-0.32	-0.24
bbw	0.305	0.26	0.35	-0.439	-0.53	-0.35	-0.590	-0.64	-0.53	0.172	0.12	0.22
sol	0.360	0.33	0.39	-0.703	-0.79	-0.62	-1.246	-1.28	-1.21	0.562	0.53	0.60
fbt	0.136	0.08	0.19	-0.167	-0.27	-0.07	-0.075	-0.14	-0.02	-0.041	-0.09	0.00

I unstandardized intercept, *CI* confidence interval, *LL* lower limit, *UL* upper limit, *erg* bicycle endurance test, *pup* push-ups, *lgj* standing long jump, *fpl* force plate jump, *jsw* jumping sideways, *bbw* balancing backwards, *sol* standing on one leg, *fbt* forward bending of the trunk

Table 5 Means (M) and standard errors (SE) of body mass index (BMI), skinfold thickness (SFT) and physical activity (PA) for the four classes (Karlsruhe 2003–2006)

	Class 1		Class 2		Class 3		Class 4	
	M	SD	M	SD	M	SD	M	SD
SFT	21.81	0.29	23.31	0.29	18.40	0.55	23.29	0.35
BMI	17.94	0.09	17.94	0.09	17.87	0.16	18.69	0.10
PA	354.35	5.76	364.35	5.76	360.61	10.83	300.89	7.01

SFT skinfold thickness (mm), BMI body mass index (kg/m²), PA physical activity (min)

under precision demands in classes 3 and 4. Here, children and adolescents performed significantly worse. Children and adolescents in class 1 are good in all coordination tests, but also revealed lowest strength in their legs. Children and adolescents in class 2 performed well in all strength tests as well as the jumping sideways test (coordination under time pressure), but performed worse in the coordination under precision demands test. Children and adolescents in class 3 and class 4 had the lowest coordination skills and also had problems to solve the push up task. Nevertheless, these children and adolescents possess comparatively good strength in their legs. All four classes performed nearly equivalent in the bicycle endurance and forward bending of the trunk test. Thus, the contribution of both tasks to distinguishing between classes seems to be low.

In a next step, it is of interest to see whether these four classes also differ in BMI, SFT and PA (see Table 5). Analyses of variance with gender and age as covariates revealed small but significant differences ($p < 0.001$) for BMI ($\eta^2_{\text{partial}} = 0.01$, $1-\beta > 0.90$), SFT ($\eta^2_{\text{partial}} = 0.02$, $1-\beta > 0.90$) and PA ($\eta^2_{\text{partial}} = 0.01$, $1-\beta > 0.90$) over the four classes. Bonferroni corrected pair-wise comparisons showed that children and adolescents of the classes 1–3 had significantly lower BMI values and were also more active than class 4 ($p < 0.001$). However, with regard to SFT all comparisons were significant at the 0.01 level except the comparison for classes 2 and 4 with nearly equivalent high SFT scores.

Discussion

Against the background of decreasing levels of PF in German children and adolescents (Bös 2003; Kretschmer and Wirsching 2007), the aim of the present paper was to provide two essential sources of information for future reasonable PF-promoting interventions. Firstly, if certain performance classes of PF in German children and adolescents aged 6 to 17 years existed, they should be identified to develop targeted interventions. Using factor mixture models four performance classes of PF emerged: class 1 is characterized by coordination skills with

simultaneously minimal leg power; class 2 is characterized mainly by leg power (and also strength in the upper extremities) and coordination under time pressure skills, but simultaneously not by coordination under precision demand skills; classes 3 and 4 are comparable except for the standing on one leg task. Here, class 4 outperformed class 3. However, both classes achieved lowest scores in all coordination tasks and for the push up task. The second source of information needed for future reasonable PF-promoting interventions pertains to the possible identification of ‘at-risk’ groups within these performance classes of PF. In other words, whether other health parameters can be found that differ for these classes too. Thus, we investigated whether the four classes found also differ with regard to BMI, SFT and PA. Earlier findings documented relationships between varying measurements of PF and PA (Ara et al. 2004; Deforche et al. 2003; Gray and Smith 2003; Moliner-Urdiales et al. 2010). Participation in PA can improve the development of PF (Katic et al. 2002; Korsten-Reck et al. 2007; Koutedakis and Bouziotas 2008). Besides, larger weight—maybe as a consequence of less PA—can lead to poorer PF (Okely et al. 2001). Indeed, analyses revealed significant but small differences in BMI, SFT and PA for the four classes. In detail, the bivariate comparisons showed that class 4 is not only characterized by comparably low PF performance skills, but also by high BMI and SFT scores as well as low PA and might thus be regarded as an ‘at-risk’ group. Interestingly, even though class 3 partially performed even worse than class 4 in some PF tests, children and adolescents in this class had the lowest BMI and SFT scores and are nearly equally as active as children and adolescents from classes 1 and 2. Thus, future research should investigate why this group has relatively larger coordination problems despite being slim and active. Neither PA nor BC can thus explain these comparatively low performances in PF. Equally interesting is the finding that class 2 is comparably physically active but has as nearly as high SFT scores as class 4 (the BMI is not inconspicuous). This could potentially indicate faulty exercises. Finally, neither PA nor BC can explain the performance differences of PF found for classes 1–3. Thus, one possible explanation is that these classes imply

different PF performance classes. Another explanation could be, though, that variables such as sedentary behavior, eating patterns, forms of sport, or when children and adolescents started being active etc. which were not considered here cause the class differences. Consequently, further research with more (health related) parameters is needed to explain performance classes of PF. Nevertheless, it should be noted that nearly all measures for the four classes vary within ± 1 SD except the standing on one leg measure for coordination under precision demands in the classes 3 and 4. Thus, in most cases the differences in PF for the four classes are within the range of average performance. However, due to factor mixture modeling, information on discrimination function of PF tests can also be extracted from the results: while the standing on one leg task discriminates mostly between the four classes, neither the endurance task (bicycle ergometer test) nor the flexibility task (forward bending of the trunk) distinguished between one of the classes. The finding that strength tasks contributed to the differentiation of the classes is in accordance with previous findings, which showed strength as a major PF ability (e.g., Deforche et al. 2003). The finding that the bicycle ergometer test did not contribute to the differentiation of the four classes might be explained by the fact that within the bicycle ergometer test, larger body mass does not inhibit performance to the same degree as, for instance, when running. Riddiford-Harland et al. (2000), for example, showed that obese children find it more difficult to move their larger body mass against gravity. In the Eurofit tasks, obese persons demonstrated inferior performances on all tests requiring propulsion or lifting of the body mass, which can be assumed to be weight-bearing tasks (Deforche et al. 2003). Nevertheless, further research using endurance tests, which have been shown to be negatively related to body weight (Hoekstra et al. 2008), is necessary to back up the claim that endurance tasks may also be a possible cause for different classes of PF. The finding that coordination tasks contributed to the differentiation of the classes to that extent is of high interest as a review of fitness reliability studies from Ortega et al. (2008) revealed that studies on coordination (but also flexibility) are lacking (Ortega et al. 2008).

Limitations

It should be noted that because the data were drawn from a cross-sectional survey, there is no clear advice concerning causal relationships. It was not possible to clarify if lower PF is a cause or outcome for lower levels of PA and more body fat. However, longitudinal findings revealed that changes in PA as well as changes in body fat were associated with changes in PF (Aires et al. 2010). Furthermore,

we would be disposed to expect differences for education in the present analyses. However, due to the fact that the underlying sample was representative with regard to education, this aspect has not been taken into account. Future research should focus on this aspect.

Conclusions

We were able to show that there are different performance classes of PF: one more skill-related fitness class with simultaneously lower health-related fitness (class 1); one more health-related fitness class with simultaneously lower skill-related fitness (class 2); and two classes with mainly no skill-related fitness but reduced health-related fitness (classes 3 and 4). We were also able to show that one of the classes namely class 4 seems to be an 'at-risk' group since children and adolescents not only have lower performance levels of PF, but also higher SFT and BMI scores and are less physically active. However, neither BMI, nor SFT or PA can explain the differences between the other three classes in the present analyses. Thus, on the one hand this study provides the fruitful information confirming that different performance classes of PF in German children and adolescents exist. This finding implies that different prevention and intervention measures are needed to promote PF in its entirety. For example, children and adolescents in class 1 should be strengthened in their health-related fitness. Especially for children and adolescents in class 4 we can additionally conclude that not only their skill as well as their health-related fitness but also their PA level should be promoted. Furthermore, intervention programs for children and adolescents in class 4 should also have a focus on reducing their body fat and weight. On the other hand this study suggests that further research is needed to explain these detailed findings on PF performance. For example, based on the present data we cannot explain why class 2 has as nearly as high SFT scores as the 'at-risk' group. We further cannot explain why some children and adolescents perform well in health but not skill-related PF tasks. More (health) parameters than PF and BC should, therefore, be considered in future research. Finally this study also emphasizes the relevance of coordination tasks in future PF fitness (health) studies, which is still an unexplored field of research.

Acknowledgments This research was supported by the Federal Ministry of Health (BMG) as well as the Federal Ministry of Education and Research (BMBF) in the years 2002–2006 and by the Federal Ministry for Families, Senior Citizens, Women and Adolescents from the years 2002 to 2008.

Conflict of interest The authors declare that they have no competing interests.

Ethical standard The study has been carried out according to the ethical guidelines laid down by the medical ethics committee.

References

- Aires L, Andersen LB, Mendonca D, Martins C, Silva G, Mota J (2010) A 3-year longitudinal analysis of changes in fitness, physical activity, fatness and screen time. *Acta Paediatr* 99:140–144
- Ara I, Vicente-Rodriguez G, Jimenez-Ramirez J, Dorado C, Serrano-Sanchez JA, Calbet JAL (2004) Regular participation in sports is associated with enhanced physical fitness and lower fat mass in prepubertal boys. *Int J Obes* 28:1585–1593
- Beauducel A, Wittmann WW (2005) Simulation study on fit indexes in CFA based on data with slightly distorted simple structure. *Struct Equ Model* 12:41–75
- Bös K (1992) Sport international—the relevance of fitness tests and fitness programs in European countries—results from a questionnaire with fitness experts. *Int J Phys Educ* 29:37–39
- Bös K (2003) Motorische Leistungsfähigkeit von Kindern und Jugendlichen (Motor performance ability of children and adolescents). In: Schmidt W, Hartmann-Tews I, Bretschneider WD (eds) *Erster Deutscher Kinder- und Jugendsportbericht* (First German children's and adolescent's sport report). Schorndorf, Germany, pp 85–108
- Bös K, Worth A, Opper E, Oberger J, Romahn N, Woll A, Jekauc D (2009) Das Motorik-Modul: motorische Leistungsfähigkeit und körperlich-sportliche Aktivität von Kindern und Jugendlichen in Deutschland. (The Motorik-Module: motor performance ability and physical activity of children and adolescents in Germany). Nomos-Verlag, Baden-Baden
- Brunet M, Chaput JP, Tremblay A (2007) The association between low physical fitness and high body mass index or waist circumference is increasing with age in children: the 'Québec en Forme' Project. *Int J Obes* 31:637–643
- Cantell M, Crawford SG, Doyle-Baker PK (2008) Physical fitness and health indices in children, adolescents and adults with high or low motor competence. *Hum Mov Sci* 27:344–362
- Chen FF (2007) Sensitivity of goodness of fit indexes to lack of measurement invariance. *Struct Equ Model* 14(3):464–504
- Chen FF, Sousa KH, West SG (2005) Testing measurement invariance of second-order factor models. *Struct Equ Model* 12: 471–492
- Deforche B, Lefevre J, De Bourdeaudhuij I, Hills AP, Duquet W, Boukaert J (2003) Physical fitness and physical activity in obese and nonobese Flemish youth. *Obes Res* 11:434–441
- Dwyer C (1996) Cutoff scores and testing: statistics, judgment, truth, and error. *Psychol Assess* 8:360–362
- Graf C, Koch B, Kretschmann-Kandel E, Falkowski G, Christ H, Coburger S, Dordel S (2004) Correlation between BMI, leisure habits and motor abilities in childhood (CHILT-project). *Int J Obes* 28:22–26
- Gray A, Smith C (2003) Fitness, dietary intake, and body mass index in urban Native American youth. *J Am Diet Assoc* 103: 1187–1191
- Hands B, Larkin D (2001) Using the Rasch measurement model to investigate the construct of motor ability in young children. *J Appl Meas* 2:101–120
- Hoekstra T, Boreham CA, Murray LJ, Twisk JWR (2008) Associations between aerobic and muscular fitness and cardiovascular disease risk: the Northern Ireland Young Hearts Study. *J Phys Act Health* 5:815–829
- Hölling H, Kamtsuris P, Lange M, Thierfelder W, Thamm M, Schlack R (2007) Der Kinder- und Jugendgesundheitsurvey (KiGGS): studienmanagement und Durchführung der Feldarbeit (The German health interview and examination survey for children and adolescents (KiGGS): study management and conduct of field work). *Bundesgesundhbl Bundesgesundhforsch Gesundheitschutz* 50:557–566
- Hu LT, Bentler PM (1999) Cut off criteria for fit indexes in covariance structure analysis: conventional criteria versus new alternatives. *Struct Equ Model* 6:1–55
- Kamtsuris P, Lange M, Schaffrath Rosario A (2007) Der Kinder- und Jugend-gesundheitssurvey (KiGGS): Stichprobendesign, Response und Nonresponse-Analyse. (The German Health Interview and Examination Survey for Children and Adolescents (KiGGS): Sample design, response and nonresponse analysis). *Bundesgesundhbl Gesundheitsforsch Gesundheitschutz* 50:547–556
- Katic R, Males B, Miletic D (2002) Effect of 6-month athletic training on motor abilities in seven-year-old schoolgirls. *Coll Antropol* 26:533–538
- Korsten-Reck U, Kaspar T, Korsten K, Kromeyer-Hauschild K, Bös K, Berg A, Dickhuth HH (2007) Motor abilities and aerobic fitness of obese children. *Int J Sports Med* 28:762–767
- Koutedakis Y, Bouziotas C (2008) National physical education curriculum: motor and cardiovascular health related fitness in Greek adolescents. *Br J Sports Med* 37:311–314
- Kretschmer J, Wirsing D (2007) Mole. Motorische Leistungsfähigkeit von Grundschulkindern in Hamburg. Abschlussbericht zum Forschungsprojekt (Mole. Motor performance ability of primary school children in Hamburg. Final report with regard to research project). Hamburg, Germany, moeve: 71–146
- Kurth B (2007) Der Kinder- und Jugendgesundheitsurvey (KiGGS): ein Überblick über Planung, Durchführung und Ergebnisse unter Berücksichtigung von Aspekten eines Qualitätsmanagement. [The German Health Interview and Examination Survey for Children and Adolescents (KiGGS): an overview of its planning, implementation and results taking into account aspects of quality management]. *Bundesgesundhbl Gesundheitsforsch Gesundheitschutz* 50:533–546
- Lämmle L, Tittlbach S, Oberger J, Worth A, Bös K (2010) A two-level model of motor performance ability. *J Exerc Sci Fit* 8:41–49
- Lubke GH, Muthén BO (2005) Investigating population heterogeneity with factor mixture models. *Psychol Methods* 10(1):21–39
- Lubke G, Neale MC (2006) Distinguishing between latent classes and continuous factors: resolution by maximum likelihood? *Multivar Behav Res* 41(4):499–532
- Malina RM (2001) Physical activity and fitness: pathways from childhood to adulthood. *Am J Hum Biol* 13:162–172
- Minck MR, Ruitter LM, Van Mechelen W, Kemper HC, Twisk JW (2000) Physical fitness, body fatness, and physical activity: the Amsterdam Growth and Health Study. *Am J Hum Biol* 12: 593–599
- Moliner-Urdiales D, Ortega FB, Vicente-Rodriguez G, Rey-Lopez JP, Gracia-Marco L, Widhalm K, Ruiz JR (2010) Associations of physical activity with muscular strength and fat-free mass in adolescents: the HELENA study. *Eur J Appl Physiol* 109:1119–1127
- Okely AD, Booth ML, Patterson JW (2001) Relationship of physical activity to fundamental movement skills among adolescents. *Med Sci Sports Exerc* 33:1899–1904
- Okely AD, Booth ML, Chey T (2004) Relationships between body composition and fundamental movement skills among children and adolescents. *Res Q Exerc Sport* 75:238–247
- Ortega FB et al (2008) Reliability of health-related physical fitness tests in European adolescents. The HELENA Study. *Int J Obes* 32:49–57
- Pissanos BW, Moore JB, Reeve TG (1983) Age, sex, and body composition as predictors of children's performance on basic motor abilities and health-related fitness items. *Percept Mot Skills* 56:71–77
- Riddiford-Harland DL, Steele JR, Storlien LH (2000) Does obesity influence foot structure in prepubescent children? *Int J Obes* 24:541–544

- Rost J (2004) Lehrbuch Testtheorie—Testkonstruktion. (Textbook on test theory—test construction). Huber, Bern
- Ruiz JR, Ortega FB, Rizzo NS, Villa I, Hurtig-Wennlöf A, Oja L, Sjöström M (2007) High cardiovascular fitness is associated with low metabolic risk score in children: the European youth heart study. *Pediatr Res* 61:350–355
- Sallis J, Prochaska JJ, Taylor WC (2000) A review of correlates of physical activity of children and adolescents. *Med Sci Sports Exerc* 32(5):963–975
- Sass DA (2011) Testing measurement invariance and comparing latent factor means within a confirmatory factor analysis framework. *J Psychoeduc Assess* 29(4):347–363
- Treuth MS, Hou N, Young DR, Maynard LM (2005) Validity and Reliability of the Fels Physical Activity Questionnaire for Children. *Med Sci Sports Exerc* 37:448–495
- Wishart D (2006) *ClustanGraphics primer: a guide to cluster analysis*. Clustan Limited, Edinburgh