Psychometric properties of the Physical Activity Questionnaire for older children (PAQ-C) in the UK

Abstract

Objectives: The Physical Activity Questionnaire for Older Children (PAQ-C) is a validated self-report questionnaire designed to assess moderate to vigorous physical activity in children. Currently however, there are no data supporting the use of the PAQ-C in British samples. *Design:* Two studies using independent samples assessed the psychometric properties of the PAQ-C in children aged 9-11 from the UK. *Method:* Study one (N = 336) examined general test score characteristics, internal reliability, factor structure and construct validity of the PAQ-C with the Self-Report Habit Index (SRHI). Study two (N = 131) re-examined the factor structure and assessed convergent validity with BMI and cardiovascular fitness (CVF). *Results:* The PAQ-C had acceptable item distribution, item total correlations (> .30) and internal reliability (α = .82 & .84). Exploratory factor analyses (EFA) identified two factors which appear to be sensitive to the context in which the activity is performed 'in school' and 'out of school'. The PAQ-C was related to the SRHI (r = .30) and inversely related to CVF (r = -.38) but not with BMI. *Conclusion:* With the exception of one problematic item; physical activity during PE, several analyses suggested that the PAQ-C had acceptable measurement properties in this group. Pragmatically, the ease of use and efficient format of the PAQ-C makes it a feasible option for large studies and/or when time, money and manpower are limited. That said, further development of the PAQ-C may be required for younger samples and its usefulness for intervention research has yet to be established.

Keywords: Factor analysis, Measurement, Psychometrics, Reliability, Validity.

Psychometric properties of the Physical Activity Questionnaire for older children (PAQ-C) in the UK

Physical activity is a behavioural category that involves a variety of actions including transport related behaviours, work related activities, leisure time activities and sport participation. For this reason self-report instruments are typically complex. The advantage of self-report questionnaires is that they are practical, economical and allow the researcher to test large numbers of participants in a relatively short space of time. Contextual prompts and items that query for location and/or purpose also improve the quality of data and provide important dimensions of physical activity not easily captured using objective measures such as heart rate monitors and accelerometers (Matthews, 2002). There are however concerns regarding the use of self-report instruments particularly in children because of the difficulty they have in correctly interpreting questions and accurately recalling activity (Janz, Lutuchy, Wenthe & Levy, 2008). For instance, children's activity is generally sporadic (Baquet, Stratton, Van Praagh & Berthoin, 2007) and thus may not be memorable in terms of frequency duration and intensity, which is the type of information that self-report questionnaires commonly ask for (Hussey, Bell & Gormley, 2007). For this reason physical activity questionnaires for use with children need to be designed in such a way that the impact of cognitive, memory and estimation skills is reduced to an acceptable minimum (Kremers, Visscher, Seidell, van Mechelen & Brug, 2005).

Because of the diversity in available questionnaires, it is not easy for researchers to decide which instrument is most suitable for his or her specific demands (Chinapaw, Mokkink, van Poppel, van Mechelen & Terwee, 2010). To this end a number of reviews (e.g. Chinapaw et al., 2010; Biddle et al., 2011; Tessier, Vuillemin & Briançion, 2008; Corder, Ekelund, Steele, Wareham & Brage, 2008; Sirard & Pate, 2001; Welk, Corbin & Dale, 2000; Welk & Wood, 2000) have been conducted attempting to select, synthesize and appraise available evidence concerning the general characteristics and psychometric properties of physical activity questionnaires. These reviews have found that few (if any) self report measures demonstrate validity, reliability and responsiveness (Chinapaw et al., 2010) and have concluded that further development and testing of self report measures in young people is required (Biddle et al., 2011).

One potentially valuable instrument identified by Tessier et al. (2008), Chinapaw et al. (2010) and Biddle et al. (2011) was the PAQ-C (Crocker, Bailey, Faulkner, Kowalski & McGrath, 1997). The PAQ-C is a self-administered seven-day recall questionnaire designed to assess MVPA in children aged 8-14 years. The purpose of the PAQ-C is to provide a general indication of children’s physical activity levels. It consists of ten items, nine of which are used to calculate a summary of activity scores. The other question assesses whether sickness or other events prevented the child from doing his/her regular activity in the last week. The first question in the PAQ-C is an activity checklist consisting of 22 common activities plus two blank spaces for ‘other’ physical activities. This question is scored as the mean of all activities using a five-point scale, with higher scores indicating higher levels of activity. The primary purpose of this question is to aid memory recall through the use of memory cues. The remaining eight items are organized using a segmented time-of-day or day-of-the-week strategy. These items are also scored using a five-point scale with higher scores indicating higher levels of activity. The summary score for the PAQ-C is the mean of the nine items.

So far the psychometric properties of the PAQ-C have been tested in largely white Canadian samples. There is no data showing its use in Europe (Biddle et al., 2011). In the former group, the PAQ-C has demonstrated good internal consistency, test re-test reliability and sensitivity to detect gender differences (Crocker et al., 1997; Kowalski, Crocker & Faulkner, 1997). It has also been shown to converge with teachers ratings of children’s physical activity (*r* = .45), the Godin and Shephard Leisure Time Exercise Questionnaire (*r* = .41), the Seven Day Physical Activity Recall Interview (*r* = .46, .43), physical activity measured via accelerometry (*r* = .39), the Canadian Home Fitness Test (step test) (*r* = .28), Athletic competence (*r* = .48) and more (Kowalski et al., 1997).

Despite strong preliminary evidence for the PAQ-C in Canadian children, research suggests that it may lack external validity when 'exported' to other racial or ethnic contexts. For example, Moore, Hanes, Barbeau, Gutin, Trevino and Yin (2007) found no relationship between the PAQ-C, cardiovascular fitness (measured using a modified version of the Harvard step test) or BMI in African American or Hispanic children. There was however a small significant relationship between the PAQ-C and Athletic competence in European American (*r* = .14) and African American (*r* = .14) children and small to moderate relationships between the PAQ-C, BMI (*r* = .16), systolic blood pressure (*r* = .17) and cardiovascular fitness (*r* =.30) in European American children. The factor structure of the PAQ-C has also been shown to vary by race. For instance Janz et al. (2008) identified a one factor structure in a sample of children from the Midwest, whereas Moore et al. (2008) identified a two factor structure in Hispanic children.

Aims and objectives

Evidently, there is need for a psychometrically sound self-report instrument that can be used in large scale physical activity research with children. One potentially valuable instrument that has been identified for this purpose is the PAQ-C. What makes the PAQ-C so attractive is that it utilizes memory cues to enhance children's ability to recall their activity. The PAQ-C's measurement of general physical activity is also attractive because of the difficulty measuring intensity, frequency and duration, especially with self-report.

Currently there is no data supporting the use of the PAQ-C in British samples. This article will report on two studies which describe the basic psychometric properties of the

PAQ-C in a cohort of children aged 9-11 years from the UK. The data in study one comes from a larger study (Thomas & Upton, in press) and reports the general test score characteristics (i.e. distribution of scores, item descriptive statistics, corrected item total correlations [CITCs] and internal reliability), factor structure and construct validity of the PAQ-C with the Self-Report Habit Index (SRHI) (Verplanken & Orbell, 2003). The data in study two was collected from a second independent sample to report the internal reliability, results from confirmatory factor analysis (CFA) and convergent validity of the PAQ-C with BMI and cardiovascular fitness.

Construct validity evidence can be established by testing how well a physical activity measure is correlated with theoretically related constructs (Thomas & Nelson, 1990). In order to explain why the PAQ-C should be related to the SRHI, some features of habit should be highlighted. Verplanken and Aarts (1999) define habits as “learned sequences of acts that have become automatic responses to specific cues and are functional in obtaining certain end goals or states” (p.104). According to this definition then habits are learned. This refers to the fact that habits are established through a history of repetition. Although repetition is not the sole requirement for habit development (functionality and automaticity are also important features of the definition) it certainly plays an important role in the habituation process and this has been demonstrated empirically (e.g. Lally, Van Jaarsveld, Potts & Wardle, 2010). Given the proposed theoretical and empirical relationship between repetition and habit strength it is hypothesised that the PAQ-C and the SRHI will be correlated demonstrating construct validity.

Method

Study one

Participants

Participants in the study were part of a larger project examining environmental and psychosocial correlates of physical activity in youth. For the purpose of statistical analysis, this study only included participants for whom data were available on both the PAQ-C and SRHI. Schools selected from registers of schools held by the University of Worcester were contacted via e-mail, inviting them to take part in the project. Four state funded primary schools replied to the e-mail and were recruited for the study. From these schools four hundred and fifty eight pupils completed the PAQ-C and the SRHI. The sample consisted of both males (N = 239) and females (N = 235) aged 9-11 years with a mean age of 9.83 years (SD = .89). School statistics indicated that the majority of pupils were of White British ethnicity with only 2% from minority ethnic backgrounds. This is lower than average nationally and for the local authority area which is 28.5% and 11% respectively (Department for Education 2010). The ACORN population profile was used as a proxy for socio-economic status. Results indicate that the school was a category three, i.e. populated by children from working families with household incomes around the national average.

Measures

PAQ-C: Physical activity was measured using the PAQ-C. In order to assess clarity of wording and comprehensibility a small focus group was conducted (N = 10). Based on feedback from participants several minor modifications were made involving item wording (i.e. changing the word recess to break time) and cultural adaptations to the activities listed in the physical activity checklist. Changes to the checklist were as follows; Inline skating was changed to Roller skating, Aerobics to Group exercise, Baseball to Cricket, Soccer to Football, Street hockey to Hockey and Cross country skiing to Snow/dry slope skiing. Two activities were also removed from the checklist; Street hockey and Cross country skiing. The Flesch-Kincaid readability score for the final questionnaire was 5.5, indicating that it could be understood by an average pupil in the 5th grade (U.S grade level), or in the United Kingdom, a child aged 10 years.

SRHI: The SRHI is a 12-item self-report questionnaire designed to assess habit strength. Responses are given on a five-point Likert scale anchored by strongly agree (5) and strongly disagree (1). To date the SRHI has been used in a variety of areas including physical activity. Reported reliability for the instrument as a measure of habitual physical activity among children and adolescents is .84 (Kremers & Brug, 2008). Recently Thurn, Finne, Brandes and Bucksch (2014) also confirmed the validity of the SRHI against total activity measured via accelerometry (r = .53).

Procedure

All study procedures and related documents were approved by a University ethics committee. Letters explaining the study were sent to parents and guardians and informed consent was obtained. The letter also ensured confidentiality and anonymity of individual results. Both the PAQ-C and the SRHI were administered to pupils during school time in the autumn term. Each class was assessed separately in the presence of class teachers. Class sizes ranged from 20 to 30 pupils. Pupil assent was obtained verbally, immediately prior to the study. The target behaviour MVPA was defined for participants as “Sports or dance that make you sweat or make your legs feel tired, or games that make you breathe hard like tag, skipping, running, climbing and others” (Crocker et al., 1997). Participants were then asked to provide their own examples of MVPA to the class to ensure that they understood. Both questionnaires were read aloud. A whiteboard and overhead projector was also used; this ensured that each item was properly explained, read and completed before students moved on to the next question. A research assistant was on hand to support pupils and facilitate questionnaire administration. The PAQ-C took approximately 20 minutes to complete while the SRHI took approximately 7 minutes to complete.

Data analysis

Item analysis was conducted using descriptive statistics. To ensure that items can detect all values within the range and are able to discriminate between individuals at different levels of the construct, items should have means close to the centre of the range and relatively high variance. The item/scale relationship was evaluated using CITCs. According to Field (2005) CITCs should exceed .30 if items are measuring aspects of the same thing, i.e. total activity. Cronbach's alpha (Cronbach, 1951) was used for the reliability analysis. As a general rule of thumb alpha levels between .70 and.90 indicate good reliability. Exploratory factor analysis (EFA) was performed to identify the factor structure of the PAQ-C in British youth and to provide evidence for construct validity. Construct validity in EFA can be established if items cluster into meaningful groups (Field, 2003) which reflect the behavioural domain or theoretical constructs the questionnaire was designed to measure. Construct validity was also examined using multiple regression analysis to explore the relationship between the PAQ-C and the SRHI. In this way construct validity is established if the two measures are correlated after controlling for school level clustering effects. Significance levels were set at $P < .05$. Data were analysed using IBM SPSS 20.

Results

Seventy three participants indicated that sickness or other events prevented them from performing their usual activities during the previous week and a further 49 questionnaires were incomplete, leaving 336 cases available for data analysis. The final sample consisted of both males ($n = 168$) and females ($n = 168$) aged 9–11 years. The mean age of the sample was 9.93 years ($SD = 0.80$). Thirty five percent of the sample were nine years old, 37% were 10 years old and 28% were 11 years old. There were no significant age or sex differences for those retained vs. excluded from the analysis.

Descriptive statistics

Descriptive statistics for the PAQ-C items, summary score and SRHI can be seen in table 1. Each item was scored on a five-point scale with higher values indicating higher activity/stronger habits. Concerning the PAQ-C, most of the items (and the summary score) had means close to the centre of the range and demonstrated adequate variability. The checklist had a relatively low mean and variance mostly because of the large number of activities in the checklist that individuals had not participated in. However this item plays the important function of enhancing memory recall by cueing specific physical activities.

Insert Table 1.

Corrected item total correlations and internal reliability

The item/scale relationship was evaluated using CITCs. All CITCs were >.30 (see table 1). Reliability was examined using Cronbach's alpha. Scale reliability was good (α = .82). These findings suggest that the PAQ-C has acceptable item/scale properties in this population.

Exploratory factor analysis

In order to establish the construct validity of the PAQ-C an EFA employing Principle Axis Factoring and Direct Oblimin rotation was performed on the data. Results of the EFA suggested the existence of two distinct factors with Eigenvalues greater than one. Cumulatively they accounted for 55.77% of the variance. The items loading on factor one described physical activity conducted away from the school setting (e.g. right after school, on the last weekend). The other items on factor one were more general asking about physical activity during the respondent's free time or for each day last week. The items on factor two described physical activity during break time or at lunch. The item referring to physical activity during P.E did not load, as one might expect, on factor two with the other 'in school items', and instead loaded on factor one. The correlation between the two factors was .46.

Construct validity

Construct validity was also assessed by examining the relationship between the PAQ-C and the SRHI (mean score = 3.95, SD = .66, α = .88) controlling for school level clustering effects (see table 2). This was done by regressing schools onto the outcome variable (total activity), adding SRHI to the regression model and computing the change in R_2 (which represents the association between total PA and SRHI after accounting for school-level differences in total PA). Because the separate factors identified in the EFA appear to be tapping the variance resulting from the setting in which physical activity is conducted, an 'in school' activity summary score and an 'outside of school' activity summary score were also regressed on the SRHI. Results showed that the SRHI was positively related to total activity (R^2 change = .30) and activity performed outside of school (R^2 change = .30). The relationship between the PAQ-C and activity performed in school was negligible (R^2 change = .08) after controlling for school level clustering effects.

Insert Table 2.

Study two

Study two utilised a primary data source to cross validate the two factor structure of the PAQ-C in a second independent sample, as well as to examine its relationship with body composition and cardiovascular fitness. BMI and the ½ mile walk run test were used to provide convergent validity evidence for the PAQ-C.

Participants

One hundred and fifty six pupils were recruited from one primary school in the UK. The school was recruited following similar procedures to those set out in study one. Participants consisted of both males (n = 84) and females (n = 85) aged 9 – 11 years. The majority of pupils (88%) were of White British ethnicity. The ACORN population profile was used as a proxy for socio-economic status. Results indicate that the school was a category three, i.e. populated by children from working families with household incomes around the national average.

Measures

PAQ-C: This was the same as in study one.

BMI: Body mass index was calculated as weight in kilograms divided by height in metres squared. Height and weight were measured without shoes in normal school clothes. Height was measured to the nearest 0.1 cm using a stadiometer. Weight was measured to the nearest 0.1kg using a Seca scale.

Cardiovascular fitness: The ½ Mile walk run test was used to measure cardiovascular fitness. The tests were performed on a 400 m track in the school playing field. The time of completion was recorded to the nearest second using a stop watch. Several fitness test batteries in the U.S have included the ½ mile walk run test as a suitable test to assess cardiovascular fitness in young people, for example, the Amateur Athletic Union Test Battery, the President's Challenge: Health Fitness and the President's Challenge: Physical Fitness.

Procedure

The PAQ-C was administered following the procedures set out in study one. After completing the PAQ-C pupils were filtered through one at a time to a separate room where measurements of their height and weight were taken. Pupils then completed the ½ mile walk run test in groups of 10-15 (depending on class size) and their times were recorded. Three researchers were on hand to assist with data collection. Participants were encouraged to complete the distance as fast as possible, although walking was permitted if they could not keep running.

Data analysis

Following study one, the item/scale relationship was evaluated using CITCs and internal reliability was examined using Cronbach's alpha. A CFA using AMOS 20 was used to confirm the two factor structure of the PAQ-C. Because the χ^2 test statistic is sensitive to sample size, the comparative fit index (CFI) and the root mean square error of approximation (RMSEA) were used to assess model fit. A CFI > .95 is indicative of a well fitting model (Hu & Bentler, 1999), while an RMSEA < .05 is indicative of good (Blunch, 2008) fit and <.06 - .08 is considered acceptable (Schreiber, Stage, King, Nora & Barlow, 2006). Pearson's *r* was used to assess the convergent validity of the PAQ-C against the ½ mile walk run test and BMI. Convergent validity is established if the measures are inversely correlated. Significance levels were set at P < .05.

Results

Thirteen children were absent for the walk run test, a further eight children indicated that sickness or other events prevented them from performing their usual activities during the previous week and a further three questionnaires were incomplete or spoiled. This resulted in a final sample size of 132. The final sample consisted of both males (*N* = 69) and females (*N* = 63) aged 9-11 years (Mean = 10.30, SD – 0.62). Eight percent of the sample were nine years old, 51% were 10 years old and 42% were 11 years old. There were no significant differences with regard to age, sex or BMI for those retained vs. excluded from the analysis.

Descriptive statistics, corrected item total correlations and reliability analysis

Descriptive statistics, CITCs and Cronbach's alpha for the PAQ-C in study two can be seen in table 1. The range, mean and standard deviation for each item was similar to those reported in study one. All CITCs were > .30. Scale reliability was good (α = .84). The findings provide further evidence that the PAQ-C has acceptable item/scale properties in this population.

Confirmatory factor analysis

Given that the EFA in study one identified a two factor structure (broadly relating to physical activity performed ‘in school’ and ‘outside of school/general activity’) this model was tested. Since the PAQ-C was designed to provide a measure of total activity, the PE item was retained in the analysis and regressed on the ‘out of school’ factor in line with results of the EFA. This approach is justified since PE shares similar characteristics with organised sports and clubs that children attend outside of school. For example, both are often structured, rule governed and consist of activities that are typically more complex than those performed during break or at lunch.

Results of the CFA indicated that the two factor model identified in study one did not provide a good fit to the data. However inspection of modification indices suggested that model fit could be improved by regressing the PE item on the ‘in school’ factor. Re-running the model in this way provided a good fit to the data, χ^2 (26) = 41.78, CFI = .96, RMSEA = .068. All parameter estimates were significant and ranged from .48 - .82 (see Table 1). The correlation between the two factors was .52.

Convergent validity

Convergent validity was assessed by correlating the PAQ-C summary score with cardiovascular fitness and BMI. The separate factors identified in the CFA were also correlated with cardiovascular fitness and BMI. The PAQ-C summary score was inversely correlated with the ½ mile walk run test[1] (r = -.38, P < .01), but not with BMI. The ‘in school’ (r = -.27, P < .001) and ‘outside of school’ (r = -.37, P < .01) summary scores were also inversely correlated with the ½ mile walk run test but not with BMI. The results provide some support for the convergent validity of the PAQ-C.

[1] Because cardiovascular fitness is represented by a lower time to completion

Discussion

This study sought to determine the psychometric properties of the PAQ-C in children from the UK. With the exception of the problematic PE item, several analyses suggested that the PAQ-C had acceptable measurement properties in this group. Pragmatically, the ease of use and efficient format of the PAQ-C makes it a feasible option for large studies and/or when time, money and manpower are limited.

First, the Flesch-Kincaid readability score for the questionnaire was examined in order to assess comprehension difficulty. Statistics indicated that the PAQ-C could be understood by an average pupil in the 5th grade (U.S grade level), or in the United Kingdom, a child aged 10 years. This may be problematic considering that the PAQ-C was designed to assess MVPA in children aged as young as 8. Since word and sentence length are the core measures used to determine the Flesch-Kincaid score, future research may wish to address this issue by simplifying the readability of the instrument even further.

Most of the questionnaire items had means close to the centre of the range and demonstrated adequate variability, one exception was the physical activity checklist; this was attributed to the large number of activities that individuals had not participated in. Janz et al. (2008) suggests rescaling this item to reflect a range consistent with the other items in the questionnaire.

Ceiling effects were observed on both the PE, break and lunch time items. This may reflect social desirability in responding given that the questionnaire was completed in the school setting. However, studies utilising direct observation have shown that although relatively brief, it is during these times that children carry out their highest daily levels of activity with regard to intensity (Sleap & Warburton, 1996). Likewise, in a series of accelerometer studies involving over 200 primary school children aged 10–12 years, Cooper and Page (2006) found that the lunch break provides the highest amount of MVPA achieved at any time during the week. Whilst ceiling effects may not be too problematic in discriminatory studies this has greater implications for intervention research. For example, children with high scores may have a substantial improvement in their MVPA which cannot be detected. Consequently, the questionnaire could be improved by modifying/rescaling these items.

Construct validity was established by testing the extent to which the PAQ-C related to the SRHI. First of all, concerning the SRHI, results indicated that the instrument had a high mean and low variance similar to results reported by Garder, De Bruijn and Lally (2011) who showed that habit scores are often above the mid-point for studies on PA. Concerning habit – physical activity relations, the relationship between the PAQ-C and SRHI in Gardener's study was .42. Nonetheless, only two of the five studies in Gardner's review involved primary school children and only one reported habit behaviour correlations. Looking at this study, which used a similar questionnaire to the PAQ-C in that it employed a segmented time-of-day/day-of-the-week strategy, the correlation between SRHI and physical activity was .31 (Kremers & Brug, 2008). After controlling for school level clustering effects in the present study the relationship between the PAQ-C and the SRHI was .30, similar to that reported by Kremers and Brug (2008).

Concerning the correlations between the SRHI and the separate factors which emerged from the EFA, stronger correlations were observed between the SRHI and 'general physical activity' i.e. physical activity conducted away from school' than for physical activity conducted in school, i.e. at break or during lunch. First, the SRHI and the 'in school' factor measure behaviour at different levels of specificity and generality. For example, the 'in school' factor measures physical activity performed in the school environment, i.e. during break or at lunch. By default it also measures behaviour at a specific time of day and for a limited period of time. The SRHI on the other hand measures behaviour irrespective of the environment, time or reference period in which it occurs. This contextual and temporal mismatch violates the principle of compatibility and any low correlations could be due to this. Also, the extent to which children are able to meta-cognitively reflect on their habits and aspects of automaticity at this age is questionable and may compromise the validity of the SRHI. Nonetheless, Schraw and Moshman (1995) argue that children as young as six can reflect with accuracy on their cognitions.

Construct validity was further examined using an EFA. Results of the EFA suggested the existence of two distinct factors with Eigen values greater than one. With the exception of PE based activity, items loading on factor one described general physical activity conducted away from the school setting; those loading on factor two described physical activity during break time or at lunch. It is worth mentioning that the number of factors retained for analysis was based on achieving an Eigenvalue ≥ 1. However the possibility of a third factor cannot be ruled out since an Eigenvalue of .93 accounting for an extra 10.29% of the variance was also identified. Since the PE item did not load as anticipated with the other 'in school' items and because it had made a relatively weak contribution to the 'out of school' factor it may be that activity performed during PE is a single item factor which represents a more structured physical activity environment.

Clearly the PE item presents a problem for the analysis, however there are several other reasons why this maybe an anomaly. First, children can't choose whether or not to take PE, but they can usually chose whether they are active (or not) during their break or at lunch. This may explain why the PE item failed to load with the other 'in school' items. Another explanation may lie in the complexity of activities performed during PE compared to those performed during lunch or at break. For example, children usually participate in relatively simple activities like active play or tag during their lunch or at break. During PE however children may participate in more complex, organised activities such as swimming or gymnastics. Again this may account for the inconsistency between the PE item and the 'in school' factor. Finally, given that the data was collected in four separate schools analysis should take into account school level clustering effects since variability within schools, particularly when it comes to the provision of PE, is likely to be smaller than variability between schools. For example, different schools are likely to have different policies surrounding PE particularly regarding frequency and duration of provision. This could explain why the PE item did not load onto the anticipated factor. Although multi-level modelling is usually employed to account for within cluster associations, with only four schools this type of analysis is not possible.

Overall, the CFA in study two supported the initial two factor solution, however this time the item describing physical activity during PE loaded, as one might expect with the other 'in school items'. However, inspection of squared multiple correlations showed that the 'in school' factor only explained 23.1% of the variance in PE based activity compared to 67.4% and 46.6% of the variance for physical activity at lunch time or during break. From this, and the results of the EFA, it appears that the contribution of the PE item to general, or even school based activity, is relatively small. This is not surprising given that curricular PE amounts to less than 1% of a child's waking time (Fox, Cooper & McKenna, 2004). Taken together the results suggest that the PAQ-C is sensitive to the context in which the activity is performed. These results are consistent with Moore et al. (2007) who identified a two factor structure in Hispanic children broadly relating to 'in school' and 'out of school' activity. They are also consistent with the ecological framework which suggests that physical activity may be tied to the setting in which it takes place.

The results of the factor analyses highlight many of the difficulties of assessing physical activity in children via self report. With open ended questionnaires, accurate recall in primary school children becomes questionable. With structured questionnaires such as the

PAQ-C the choice of questions relative to contextual factors, (e.g. at school at home, in P.E) and psychometric concerns can become problematic (Moore et al., 2007). However, physical activity is a multi-dimensional construct (Miles, 2007). With this in mind the different dimensions emerging from the factor analyses are likely indicators of a higher order factor, i.e. total physical activity. Indeed the correlation between the two factors in both studies showed that they share a moderate amount of common variance, e.g. 21 and 27%. Future research may wish to explore the hierarchical structure of the PAQ-C further by subsuming the differentiated components by a global higher order factor – total activity.

Although a global higher order factor ‘total activity’ would provide a more parsimonious model of physical activity, the separate factors identified in the present study offer additional information on physical activity patterns and/or levels within specific behaviour settings. This is important as it permits the estimation of the relative contribution of MVPA in a particular domain to total MVPA, not to mention the possibility of examining inter-domain MVPA relationships. Few, if any existing questionnaires currently explore the type of environment in which the individual normally undertakes physical activity.

Regarding convergent validity, the PAQ-C summary score was significantly correlated with cardiovascular fitness. What’s more, the magnitude of the correlation was far higher than those reported in previous studies (e.g. Moore et al., 2008 & Kowalski et al., 1997); however the difference most likely reflects the use of different indices of cardiovascular fitness, i.e. ½ mile walk run test vs. step test. Examination of the separate factors emerging from the CFA in study two showed that physical activity conducted away from school showed a stronger relationship with cardiovascular fitness than activity performed within the school setting. Although children carry out their highest daily levels of activity with regard to intensity at break and during lunch (Sleap & Warburton, 1996), the limited time for physical activity during these periods (compared to evenings and weekends) may explain these results.

Studies examining the validity of the ½ mile test are lacking. To our knowledge only one study has been conducted in healthy children which tested the criterion related validity of the ½ mile test for estimating VO2 peak. In this study Castro-Pinero, Ortega, Mora, Sjostrom and Ruiz (2009) reported correlation coefficients of .55 and .53 in children aged 6-17 years. Despite the lack of studies examining the validity of the ½ mile walk run test, the test was chosen since it is thought to reduce the influence that important variables have in running performance, especially in early years (e.g. willingness to accept strenuous effort, motivation and monotony). No relationship was found between the PAQ-C and BMI. Similar results were reported by Moore et al. (2008) who found that the PAQ-C was unrelated to BMI in both Hispanic and African American children. There was however a weak relationship between the PAQ-C and BMI (r =-.16) in European American children.

Aside from the issues mentioned already the PAQ-C has some definite drawbacks. First, the sporadic short-burst nature of children's physical activity makes it difficult to capture via self-report methods; the PAQ-C is not immune to this and may therefore provide an underestimate of children's true activity levels. Second, the PAQ-C was developed to assess general levels of physical activity; it does not provide specific frequency and time information. As a result recommended physical activity levels are not represented in the PAQ-C scores. Nevertheless, children generally have difficulty when recalling the frequency and/or duration of activities, and this has been well documented (Hussey et al., 2007). According to Welk, et al. (2000), the PAQ-C's general measurement is beneficial for studies that do not need estimates of specific time or frequency. Finally, the PAQ-C cannot discriminate between moderate and vigorous activity and cannot assess physical activity during school holidays.

Study limitations and directions for future research

Several limitations of the research warrant consideration. First, although the PAQ-C is designed for children aged 8-14 years, the age range in the present study was 9-11 years. This limits the generalisability of the results for older children and early adolescents. Given that the readability score for the PAQ-C was 10 years and above, future research may wish to examine the suitability of the PAQ-C for younger age groups in particular. Second, the PAQ-C is not validated against an objective measure of physical activity; instead a measure of fitness is used. The issue of whether fitness can be used as a proxy for physical activity in children is questionable as a significant amount of fitness test performance can be explained by heredity and maturation (Pangrazi, 2000). For this reason future research should seek to validate the PAQ-C in British youth against an objective measure of physical activity. Third, the sample size in study two was relatively small for a validation study. Limited time and man power to measure and record height, weight and fitness data resulted in a sample size-practicality trade off. Collecting data from children also presents a more challenging predicament than it does with adults. Related to the issue of sample size is the attrition rate. In study one, 15% of participants were excluded from the study because they indicated that sickness or other events prevented them from performing their usual activities during the previous week. In study two, 5% of participants were excluded for the same reason. The issue of whether participants need to be removed from the sample on this basis is an interesting question related to the validity of the scale, i.e. sensitivity to change. Future research may wish to address this aspect of the PAQ-C, examining whether or not those who would normally be excluded from the analysis due to self-reported sickness actually report reductions in their level of activity. Fourth, the EFA in study one was conducted on the raw item scores rather than pooled within-school between-item correlations. The former technique fails to account for within cluster associations whereas latter accounts for clustering by partialing out school level effects. Finally, the data in study one comes from a larger data set that was not designed specifically to validate the PAQ-C. Although the use of secondary data is a cost effective way to explore the utility of the PAQ-C it has several disadvantages. For example, the researcher has less control over the study population and the exact measures employed.

It could be argued that the PAQ-C does not really ‘categorise’ children into high, medium and low levels of activity; it simply provides a general score of physical activity. For this reason future research may wish to examine the values on the PAQ-C and whether they accurately represent high active or low active children as it is this type of data that researchers are often looking for when choosing a physical activity measure. Finally, future research may also wish to examine competing factor structures, i.e. one factor vs. two factor model (where PE is regressed on the ‘in school’ factor) or test for invariance of the factor structure across age, gender or socioeconomic status.

Conclusion

The PAQ-C was designed by researchers to provide a global measure of physical activity in children, however it also appears to tap into the variance resulting from the setting in which the activity takes place. Currently the PAQ-C includes an item for activity during PE that would be assumed to load with the other ‘in school’ items; this was not the result of the EFA. This inconsistency should be considered especially in studies that sample from different schools. Aside from this, the PAQ-C had acceptable item and test score characteristics, CITCs and Cronbach’s alpha. Construct and convergent validity between the PAQ-C, the SRHI and cardiovascular fitness was also established. Further development of the PAQ-C may be required for younger samples to reduce comprehension difficulty given the Flesch-Kincaid readability score. Finally, the responsiveness of the PAQ-C is yet to be established in British youth and so its usefulness for intervention research is yet to be determined.

References

Baquet, G., Stratton, G. Van Praagh, E., & Berthoin, S. (2007). Improving physical activity assessment in children with high frequency accelerometry monitoring: A methodological issue. *Preventive Medicine, 44,* 143–147.

Biddle, S., Gorely, T., Pearson, N., & Bull, F. (2011). An assessment of self-reported physical activity instruments in young people for population surveillance: Project ALPHA. *International Journal of Behavioural Nutrition and Physical Activity, 8:1,* Retrieved 10th September 2010 from: http://www.ijbnpa.org/content/8/1/1

Blunch, N. (2008). *Introduction to Structural Equation Modelling using SPSS and AMOS.* London: Sage.

Castro-Pinero, J., Ortega, F.B., Mora, J., Sjostrom, M. & Ruiz, J.R. (2009). Criterion related validity of ½ mile walk run test for estimating VO2 peak in children aged 6-17 years. In*ternational Journal of Sports Medicine, 30(5),* 366-71.

Chinapaw, M., Mokkink, L., van Poppel, M., van Mechelen, W., & Terwee, C. (2010). Physical Activity Questionnaires for Youth: A Systematic Review of Measurement Properties. *Sports Medicine, 40(7),* 539-563.

Cooper, A., & Page, A. (2006). Childhood obesity, physical activity and the environment. In N. Cameron, N. Norgan and G. Ellison (eds). *Childhood Obesity: Contemporary Issues* (119–134). London: Taylor and Francis.

Crocker, P., Bailey, D., Faulkner, R., Kowalski, K., & McGrath, R. (1997). Measuring general levels of physical activity: Preliminary evidence for the Physical Activity Questionnaire for Older Children. *Medicine and Science in Sports and Exercise, 29(10),* 1344–1349.

Cronbach, L. (1951). Coefficient alpha and the internal structure of tests. *Psycometrika, 16,* 297–334.

Corder, K., Ekelund, U., Steele, R., Wareham, N., & Brage, S. (2008). Assessment of physical activity in youth. *Journal of Applied Physiology, 105,* 977–987.

Field, A. (2005). *Discovering statistics using SPSS: and sex and drugs and rock 'n' roll (2nd Edn.).* London: Sage.

Field, A. (2003). Designing a good questionnaire. Retrieved 20th December 2013 from: http://www.statisticshell.com/docs/designing_questionnaires.pdf

Fox, K.R., Cooper, A. & McKenna, J. (2004). The school and promotion of children's health-enhancing physical activity: Perspectives from the United Kingdom. *Journal of Teaching in Physical Education, 23,* 338-358.

Gardner, B., De Bruijn, G., & Lally, P. (2011). A systematic review and meta-analysis of applications of the self-report habit index to nutrition and physical activity behaviours. *Annals of Behavioural Medicine (42*), 174-187.

Hu, L., & Bentler, P. (1999). Cut off criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives, *Structural Equation Modelling, 6*(1), 1-55.

Hussey, J., Bell, C., & Gormley, J. (2007). The measurement of physical activity in children. *Physical Therapy Reviews, 12,* 52–58.

Janz, K., Lutuchy, E., Wenthe, P., & Levy, S. (2008). Measuring activity in children and adolescents using self-report: PAQ-C and PAQ-A. *Medicine and Science in Sports and Exercise, 40, 4,* 767–772.

Kowalski, K., Crocker, P., & Faulkner, R. (1997). Validation of the Physical Activity Questionnaire for Older Children. *Paediatric Exercise Science, 9,* 174–186.

Kremers, S., & Brug, J. (2008). Habit strength of physical activity and sedentary behaviour among children and adolescents. *Paediatric Exercise Science, 20,* 5–17.

Kremers, S., Visscher, T., Seidell, J., van Mechelen, W., & Brug, J. (2005). Cognitive determinants of energy balance related behaviours: Measurement issues. *Sports Medicine, 35(11),* 923–933.

Lally, P., Jaarsveld, C., Potts, H. & Wardle, J. (2010). How are habits formed: Modelling habit formation in the real world. *European Journal of Social Psychology, 40(6),* 998-1009.

Matthews, C. (2002). Use of self-report to assess physical activity. In G. Welk (Ed.). *Physical activity assessments for health related research.* Champaign (IL): Human Kinetics, 107–125.

Miles, L. (2007). Physical activity and Health. British Nutrition Foundation *Nutrition Bulletin, 32, 314–363.*

Moore, J., Hanes, J., Barbeau, P., Gutin, B., Trevino, P., & Yin, Z. (2007). Validation of the Physical Activity Questionnaire for Older Children in children of different races. *Paediatric Exercise Science, 19,* 6–19.

Pangrazi, R.P. (2000). Promoting physical activity for youth. *The ACHPER Healthy Life Styles Journal, 47(2),* 18-21.

Schrieber, J., Stage, F., King, J., Nora, A., & Barlow, E. (2006). Reporting Structural Equation Modelling and Confirmatory Factor Analysis Results: A Review. *Journal of Educational Research, Vol. 99(6),* 323-337.

Schraw, G., & Moshman, D. (1995). Metacognitive theories. *Educational Psychology Review*, *7*(4), 351-371.

Sirard, J., & Pate, R. (2001). Physical activity assessment in children and adolescents. *Sports Medicine, 31(6),* 439–454.

Sleap, M., & Warburton, P. (1996). Physical activity levels of 5 – 11 year old children in England: Cumulative evidence from three direct observation studies. *International Journal of Sports Medicine,* 17, 248–253.

Tessier, S., Vuillemin, A., & Briançion, S. (2008). Review of physical activity questionnaires validated for children and adolescents. *Science and Sports 23,* 118–125.

Thomas, J., & Nelson J. (1990). *Research Methods in Physical Activity,* 2nd ed. Champaign, IL: Human Kinetics.

Thomas, E.L., & Upton. D. (in press). Automatic and motivational predictors of children's physical activity: Integrating habit, the environment and the Theory of Planned Behavior. *Journal of Physical Activity and Health, 11(3).*

Thurne, J., Finne, E., Brandes, M. & Bucksch, J. (2014). Validation of physical activity habit strength with subjective and objective criterion measures. *Psychology of Sport and Exercise, 15(1),* 65-71.

Verplanken, B. (2006). Beyond frequency: Habit as a mental construct. *British Journal of Social Psychology, 45,* 639–656.

Verplanken, B. & Aarts, H. (1999). Habit, attitude and planned behaviour: Is habit and empty construct or an interesting case of goal-directed automaticity? *European Review of Social Psychology, 10,* 101–134.

Verplanken, B., & Orbell, S. (2003). Reflections on past behaviour: A self-report index of habit strength. *Journal of Applied Social Psychology, 33,* 1313–1330.

Welk, G., Corbin, C., & Dale, D. (2000). Measurement issues in the assessment of physical activity in children. *Research Quarterly for Exercise and Sport*, *71,* 59–73.

Welk, G., & Wood, K. (2000). Physical activity assessments in physical education: A practical review of instruments and their use in the curriculum. *Journal of Physical Education, Recreation and Dance*, *71*, 30-40.

Table 1. Descriptive statistics, corrected item total correlations and factor loadings for the PAQ-C in study one (N = 336) and study two (N = 132)

	Study one						Study two					
					EFA						***CFA***	
Item	*Range*	*Mean*	*SD*	*CITCs*	*Factor 1*	*Factor 2*	*Range*	*Mean*	*SD*	*CITCs*	*Factor 1*	*Factor 2*
Checklist	1–4.67	1.87	0.59	.39	.42		1.10-3.39	1.83	.42	.61	.63	
PE	1-5	4.14	0.80	.32	.37		2-5	4.18	.74	.38		.48
Break	1-5	4.30	0.96	.41		.82	1-5	4.25	.94	.41		.68
Lunch	1-5	4.10	1.15	.41		.67	1-5	4.13	.97	.47		.82
After school	1-5	3.31	1.32	.60	.72		1-5	3.15	1.28	.64	.78	
Evenings	1-5	3.27	1.24	.62	.73		1-5	3.12	.92	.61	.70	
Weekends	1-5	3.48	1.12	.63	.72		1-5	3.35	.92	.59	.62	
Describes	1-5	3.41	1.27	.59	.64		1-5	3.53	1.16	.70	.82	
Week	1.14–5.00	3.51	0.92	.72	.80		1.29-5	3.26	.88	.74	.81	
PAQ-C	1.79–4.77	3.49	0.68				1.76-4.74	3.36	.67			

*EFA – Exploratory factor analysis, CFA – Confirmatory factor analysis, SD – standard deviation, CITCs - corrected item total correlations

**Factor 1 and 2 – factor loadings from EFA and CFA. Factor loadings < .3 are suppressed

Table 2. Relationship between the PAQ-C and SRHI controlling for school level clustering effects

	Beta	R^2	R^2 change
Total activity			
School	-.02	.00	
SRHI	.55***	.30	.30***
Out of school activity			
School	.02	.00	
SRHI	.56***	.30***	.30***
In school activity			
School	-.09	.01	
SRHI	.28***	.08	.08***

*** $P < .001$