Adolescents and Self-Reported Physical Activity: An Evaluation of the Modified Godin Leisure-Time Exercise Questionnaire

JACQUELINE ZELENER*1, and MARGARET SCHNEIDER†2

1Program in Public Health, University of California, Irvine, Irvine, CA, USA; 2School of Social Ecology, University of California, Irvine, Irvine, CA, USA

*Denotes undergraduate student author, †Denotes professional author

ABSTRACT

International Journal of Exercise Science 9(5): 587-598, 2016. The Godin Leisure-Time Exercise Questionnaire assesses self-reported physical activity (PA) among adults, and was later modified for children. However, the modified version (Godin-Child Questionnaire) has not been validated among adolescents. This study evaluates the construct validity and test-retest reliability of the Godin-Child Questionnaire among adolescents. The study participants, sixth graders (age M = 11.06 ± 0.436; 48% males, 48% Latino), were assessed at 2 time points (fall and spring: N=139). First, the study determined whether adolescents accurately reported exercise intensity based on adolescents' ratings of perceived exertion (RPE) during moderate- and hard-intensity exercise tasks. Second, objective assessments of PA obtained using the Actigraph® accelerometer were correlated with PA self-reported on the Godin-Child Questionnaire for its reliability. Finally, interviews explored participants' interpretations of the Godin-Child Questionnaire. RPEs suggested that adolescents generally perceived exercise intensity accurately, with no significant differences between genders. There was a weak correlation between the Godin-Child Questionnaire and the Actigraph® during fall (r=.22, P<.05) and spring (r=.24, P<.05), with a significant correlation for males (P<.05) at both time periods. Test-retest correlations were acceptable (fall-spring r=.68; P<.001) and were equally strong among females and males. Interviews revealed that some respondents overlooked written directions. The Godin-Child Questionnaire may be a useful evaluative measure of self-reported physical activity for comparing activity levels across groups of adolescents, but may be less accurate for assessing physical activity on an individual level.

KEY WORDS: Physical activity, adolescents, fitness, self-report, health

INTRODUCTION

Considerable evidence exists for the importance and benefits of increasing physical activity (PA) as a means of maintaining health and preventing disease among youth (16, 19). The American College of Sports Medicine and the United States Department of Health and Human Services recommend that children and adolescents 6-17 years old participate in a minimum of one hour of moderate-to-vigorous PA (MVPA) daily, which includes “any movement
produced by skeletal muscles that results in caloric expenditure” (1, 4, 6, 19, 27). Regular PA may prevent chronic diseases, such as obesity and cardiovascular disease (2). To deter these conditions, it is imperative to start encouraging shorter durations of PA from infancy so that PA becomes an integral part of an individual’s lifestyle (12, 17).

Accurate measurements of PA are important for overall health, social, policy and fiscal reasons. Proper measurements of PA help to identify associations between PA and health, which are essential in preventing disease and improving quality of life (15). When the activity levels of specific subpopulations (e.g., different ethnic or socioeconomic groups) are not correctly estimated, the result may be inefficient allocation of resources or even misdirection of policies and public health priorities.

To properly assess PA, it is important to have a variety of tools to capture what is a complex behavior. The method appropriate to a specific study depends on the research question, the size of the sample and the target population. As such, it is beneficial to have a library of methods to allow researchers to select a tool specific to the study’s purpose (30). PA assessment methods can be categorized; criterion methods (e.g., doubly labelled water), objective methods (e.g., accelerometry) and subjective methods (e.g., surveys) (29, 30). The ongoing study from which the data for this report are drawn utilizes an Actigraph® accelerometer as an objective indicator and the modified Godin Leisure-Time Exercise Questionnaire because it is short, easy to use, designed for use among youth and yields an estimate of time spent in usual weekly activity at a moderate-to-vigorous intensity. The adapted Godin Leisure-Time Exercise Questionnaire (hereafter referred to as the Godin-Child Questionnaire) differs from the original questionnaire in that the child version contains illustrations and simplified response options (10).

The Godin Leisure-Time Exercise Questionnaire has been widely used to gather information about self-reported PA participation among adults (2, 8, 9). However, in a review of self-report instruments, this questionnaire was found to have weak test-retest reliability with younger populations (2). The Godin-Child Questionnaire was subsequently modified for use with children and adolescents, but evidence for the validity and reliability of the modified instrument is lacking (11). The Godin-Child Questionnaire has not yet been verified for its construct validity or test-retest reliability among adolescents. Construct validity refers to the extent to which the tool measures what it is supposed to measure, and test-retest reliability assesses the consistency of test scores over time (21). Therefore, it is important to rigorously evaluate the Godin-Child Questionnaire to determine whether researchers may confidently use this tool with adolescents.

To examine how much error may be expected using the Godin-Child Questionnaire, this study aims to both quantitatively and qualitatively evaluate the construct validity and test-retest reliability of the Godin-Child Questionnaire among 6th grade students. Four sources of data were used. The current study used data collected in the course of a larger study. Firstly, data from adolescents’ ratings of perceived exertion during exercise tasks of known intensity were analyzed to assess the accuracy of their perceptions. Secondly, the Actigraph's® objective data
on quantity of MVPA were compared to the amount of MVPA reported on the Godin-Child Questionnaire. Thirdly, the Godin-Child Questionnaire was evaluated for its test-retest reliability. Finally, qualitative data obtained from interviews were used to explore how well the respondents understood the instructions on the questionnaire.

METHODS

Participants
The larger study from which the data are drawn is an intervention study that focuses on encouraging adolescents to exercise at a level that generates positive affect (26). The larger study is ongoing and samples healthy adolescent 6th graders from a public middle school (N=139). The present analysis examines data from participants to date (mean age= 11.04±0.436); 48% male; 48% Latino, 19% Non-Latino White, 13% African American, 6% Multiracial, 9% Asian, 5% Other). Exclusion criteria included the following: medical exemption from Physical Education (PE), asthma, students on sports teams or involved in individual competitive sports when initially recruited, left-handedness, and a history of depression or head trauma. The handedness, depression and head trauma exclusion criteria were in place because of a component of the study that involved collecting electroencephalograms. The present study does not utilize the electroencephalogram data, so these methods are not described below.

Protocol
A University-based Institutional Review Board and the School District Research Review Committee reviewed and approved all protocols. Students provided signed assent and parents/guardians provided signed consent prior to participation. All assessments were administered in an empty classroom that had been converted into a clinical laboratory at the school. The following assessments were included in the present study, and were assessed in both fall (September - October) and spring (April - May): Godin-Child Questionnaire, Cardiovascular Fitness Test, Moderate-Intensity task (MOD) and Actigraph®.

Participants performed a cardiovascular fitness test on a stationary bicycle at a gradually increasing intensity (31). A mouthpiece that measures exhaled gases was used to measure the participant’s ventilation (breathing rate), oxygen uptake and carbon dioxide output, which were analyzed using a Sensor Medics® metabolic system to quantify the participant’s peak amount of oxygen consumed (VO₂-peak).

Participants completed the Moderate-intensity exercise task (MOD) approximately 7 days after the fitness test. The intensity of the MOD task was calibrated to the individual’s peak VO₂, normalizing the physiological load of the MOD to an intensity that is generally considered to be moderate. The MOD intensity was determined as 50% of the participant’s oxygen uptake reserve, which was determined as the difference between the individual peak VO₂ and the estimated resting VO₂. The average participant’s resting VO₂ was used in this equation.

During the fitness test and the MOD, perceived levels of exertion were solicited using the Borg Rating of Perceived Exertion (RPE) Scale every 3 minutes (3). When prompted, the participants
were instructed to point to the number, on a large-size poster of the scale, which corresponded to his or her RPE.

Participants were instructed to wear the Actigraph® accelerometer on the right hip for 7 consecutive days, exclusive of sleeping, swimming or bathing.

The Godin-Child Questionnaire was administered by a research assistant in an empty classroom that was converted into an exercise lab.

The Principal Investigator of the study and a research assistant conducted interviews with 18 of the participants between the fall and spring assessment periods. During these interviews, participants completed the Godin-Child Questionnaire, and then a research assistant asked a series of closed and open-ended questions to elicit information about how the participant arrived at the responses provided on the questionnaire. The interviews were conducted until the responses reached convergence, and the research assistants no longer received new information from the participants.

For this study, we used the RPE obtained at the end of the fitness test and at minute 9 of the MOD. Since the fitness test is terminated at the point when the participant feels s/he cannot continue, the RPE at the end of the fitness test is presumed to correspond to a time when the participant is working at a very hard intensity level. The RPE at 9 minutes into the moderate-intensity task has been shown to be representative of the RPE during the MOD, and was selected to reflect RPE during moderate-intensity exercise (25). Nine minutes is a long enough duration into the task that the participant was fully warmed up, yet not so far into the task that s/he was beginning to experience fatigue. When providing an RPE, participants responded between the minimum value of “6,” representing no exertion at all, and “20,” representing maximal exertion.

Data from the Actigraph® (Manufacturing Technology Inc., Florida) were downloaded to a computer and then analyzed using Actilife data analysis software developed by Manufacturing Technology, Inc. We used the validated Freedson algorithm to transform the raw data into estimates of time spent in MVPA (7). For a day to be included in the computation, a minimum of 8 valid hours must have been recorded. A valid hour is an hour in which there is not a string of 30 minutes or more with zero activity recorded. A minimum of 4 valid days (including at least 1 weekend day) of data was also required. The variable used in this study was the average minutes per day of MVPA recorded by each participant.

The Godin-Child Questionnaire is an overall assessment of participants’ self-reported usual level of PA. Participants were asked to respond with the number of days in a week (0-7 days) and how many minutes each day, in 10-minute increments (0-60+ minutes), that described how much mild, moderate and strenuous PA s/he did in a typical week (outside of school). The total number of minutes of activity in a typical week was calculated by multiplying the number of days per week by the number of minutes per day (for each intensity category) and then summing these values. Self-reported data are collected separately for strenuous activity
(e.g. running), moderate activity (e.g. baseball) and mild activity (e.g. golf) (11). Strenuous activity is defined as "it makes my heart beat quickly, and makes me sweat." Some examples and images of strenuous activity provided on the Godin-Child Questionnaire include running, fast bicycling, aerobic dance and football. Moderate activity is explained as "it doesn't make me tired and makes me sweat just a little." Examples such as fast walking, easy swimming, weight lifting and tennis are included. Finally, mild activity is given the description: "it makes me use little effort, and doesn't make me sweat." Some examples of mild activity include easy walking, bowling, fishing and yoga. Since the activities listed under "mild" and "moderate" on the Godin-Child Questionnaire overlap with the activities that fit within the Actigraph’s® moderate-intensity category, both mild and moderate activities were included when calculating the estimate of weekly MVPA from the Godin-Child Questionnaire.

The interviews were designed to identify adolescents’ thought processes as they responded to the Godin-Child Questionnaire. In particular, interviews assessed whether respondents included physical education in their report of activity levels, as the instructions on the instrument stated that PE should not be included. After completing the Godin-Child Questionnaire, participants were asked, “When filling out this questionnaire, did you include information about your activity during PE?” The research assistant then reviewed participant’s responses to each section of the questionnaire and asked the participant to describe what specific activities s/he was thinking about when providing the information.

**Statistical Analysis**

The current study has 4 main objectives: the first objective focuses on determining how accurate adolescents are in reporting their intensity of activity; the second focuses on how precise adolescents are in estimating their own usual amount of time spent in MVPA; the third part examines the Godin-Child Questionnaire for its test-retest reliability; and the fourth uses information from interviews to qualitatively assess the quality of the information gathered with the Godin-Child Questionnaire.

To validate adolescents’ perceptions of high- and moderate-intensity activity, means and standard deviations of the RPE at the end of the fitness test and at minute 9 on the MOD task were examined in relation to the known task intensity for the combined group and by gender.

Correlations were run between the average minutes of activity per day as provided by the Godin-Child Questionnaire and the average minutes of MVPA per day obtained from the Actigraph® to validate adolescent’s self-reported activity duration. This analysis was repeated for fall and spring. Correlations were run separately by gender to check for gender differences.

To examine the test-retest reliability of the Godin-Child Questionnaire, correlations (Pearson’s r) were run for the total minutes of exercise reported on the Godin-Child Questionnaire between data collection in the fall and spring. The correlation was examined for the total sample and for males and females separately.
Responses to the yes/no question in the interviews regarding the inclusion of activity during PE were tallied to determine how carefully participants read the directions. The interviews were also qualitatively reviewed for common themes and patterns across participants to gain a broader understanding of how adolescents self-report how long they were physically active at different levels of intensity.

RESULTS

The study sample included 139 sixth-grade students. The average age was 11.06 years ±.43, and 48% of the sample was male. The sample was ethnically diverse (48% Latino). Table 1 shows means for cardiorespiratory fitness and body composition at baseline. T-tests comparing means across gender indicated that males had a higher VO$_2$ (L/min) and VO$_2$ peak (mL/kg/min) than females (p< .05), but there was no difference between genders for BMI Percentile.

Table 1. Participant characteristics (Mean±Standard Deviation).

<table>
<thead>
<tr>
<th></th>
<th>Total (N= 139)</th>
<th>Female (N= 72)</th>
<th>Male (N=67)</th>
</tr>
</thead>
<tbody>
<tr>
<td>VO$_2$peak (L/min)*</td>
<td>1.64±.31</td>
<td>1.58±.29</td>
<td>1.71±.32</td>
</tr>
<tr>
<td>VO$_2$peak (mL/kg/min)*</td>
<td>37.07±7.45</td>
<td>35.35±6.92</td>
<td>38.92±7.59</td>
</tr>
<tr>
<td>BMI Percentile (%)</td>
<td>63.11±31.45</td>
<td>61.09±32.36</td>
<td>65.28±330.54</td>
</tr>
</tbody>
</table>

* significant gender difference at p<.05

Mean RPEs at the end of the fitness test were 16.78±2.70 in fall and 16.35±2.68 in spring, with no significant differences between genders. Figure 1 illustrates the spread of the fall RPE data at the end of the fitness test, and shows that although the majority of the participants rated the task in the realm of “hard” activity, there were individual differences. The range of responses was from 9 to 20, meaning that some individuals perceived the high-intensity activity to be at a relatively low intensity.

![Figure 1](http://www.intjexersci.com)  
**Figure 1.** Distribution of Borg Rating of Perceived Exertion (RPE) scale at End of Fitness Test for fall.

![Figure 2](http://www.intjexersci.com)  
**Figure 2.** Distribution of Borg Rating of Perceived Exertion (RPE) scale at 9 minutes during the Moderate Intensity Exercise Task for fall.
Mean RPEs during the MOD task were 13.28±3.19 in fall and 12.55±3.29 in spring. There were no significant differences between males and females. These mean values indicate that the majority of the adolescents rated the MOD task as “somewhat hard.” As shown in the data collected in the fall (Figure 2), however, there was again considerable individual variability. Some adolescents reported an RPE of 6 (corresponding to a perceived exertion level of “no exertion at all”) and others reported an RPE of 20 (corresponding to a perceived exertion level of “maximal exertion”).

Correlations were run between the responses on the Godin-Child Questionnaire and average daily MVPA on the Actigraph® at fall and spring. There was a weak yet significant correlation between the Godin-Child Questionnaire and the Actigraph® during the fall (r=.22, P< .01) and spring (r=.24, P<.01). Correlations between the Child-Godin Questionnaire MVPA and Actigraph® MVPA by gender indicated a significant correlation for males (p<.05) in fall and spring, and no significant correlation among females at either time period. There were no ethnic differences in any of the analyses when non-Latinos were compared to Latinos.

Self-reports of total minutes of exercise for fall and spring were significantly positively correlated (r=.68; P<.001). Test-retest correlations indicated that the Child-Godin Questionnaire was equally reliable among females and males. Overall, the Godin-Child Questionnaire was found to be reliable.

A qualitative analysis of adolescents’ responses showed that out of the 18 interviews conducted, only 4 participants clarified that s/he did not include PE when responding to the Godin-Child Questionnaire, despite written instruction to exclude PE. When the participants were asked to further elaborate on the types of physical activities that they reported usually engaging in, many of the participants responded using similar language as that provided on the Godin-Child Questionnaire (e.g. “easy walking,” “fast walking,” “running”). Moreover, the adolescents tended to use examples of PA that were either listed or pictured on the Godin-Child Questionnaire, which suggests that their responses had been influenced by the wording and pictures provided on the instrument.

DISCUSSION

This study examined the reliability and validity of the Godin-Child Questionnaire, which purports to assess usual activity levels among youth by self-report. By analyzing the findings collected from the RPE and Actigraph® data, along with the subjective responses from the Godin-Child Questionnaire and interviews, this study examined the extent to which the Godin-Child Questionnaire may be confidently used when examining self-reported activity levels in adolescents. The findings suggest that the Godin-Child Questionnaire is both valid and reliable for comparing groups, but less reliable and less valid for comparing individuals.

Based on the data, adolescents on average are able to perceive and report accurate levels of intensity of PA. The Godin-Child Questionnaire also appears to be reliable in that the amount of activity reported remains stable over time. Correlations between the amount of activity
reported on the Godin-Child Questionnaire and the amount recorded by the Actigraph®, however, suggest that youth are not accurate reporters of activity duration. Moreover, interviews with students revealed that they did not read the directions carefully and appeared to be influenced by the prompts on the survey in terms of which activities they reported. Overall, our results suggest that the Godin-Child Questionnaire may be useful for comparing activity levels between groups and examining change in activity levels of groups over time, but should not be used to evaluate individual activity levels or quantify activity in relation to metabolic expenditure.

The aim of the first analysis was to assess how accurate adolescents are as reporters of their exercise intensity. RPE ratings at the end of the fitness test, when participants were working at a very high intensity, provided evidence that most students did report the intensity to be very hard. Similarly, RPE ratings during the moderate-intensity exercise task demonstrated that, on average, adolescents did accurately report their exertion levels as “somewhat hard.” These data support the presumption that at the group level, youth will be accurate reporters of prior participation in activities that are “strenuous” and/or “moderate.” Examination of the distribution of responses, however, revealed that some students under-reported the intensity of the clinic exercise tasks and some over-reported, which suggests that personal experiences may influence how a person perceives PA intensity (27). We therefore conclude that it would be inadvisable to place too much confidence in individual scores on the Godin-Child Questionnaire.

While self-report questionnaires are valuable tools in measuring PA because they are cost-effective and place a low burden on participants, there is difficulty in administering these tools to adolescents because of potential recall inaccuracies (14, 22, 23). The results of this study may suggest ways in which the Child-Godin Questionnaire may be improved to improve the accuracy of self-reported PA in adolescents. Overall, while questionnaires such as the Child-Godin Questionnaire may be useful in categorizing average levels of intensities of PA, individual responses regarding intensity may be less accurate (27, 29).

Based on the data collected from the Godin-Child Questionnaire and the Actigraph®, the findings suggest that adolescents are poor at estimating the amount of time spent engaged in PA. Consistent with the literature, younger populations may be unable to estimate time accurately, and thus may be inefficient when recalling the amount of time spent in PA (24).

The weak association found between the Godin-Child Questionnaire and the Actigraph® overall might be attributed to several factors. One possible explanation is that the weak correlation is driven by the different time frames specified by the multiple assessments. That is, the assessments were not assessing activity across the same days (the Godin-Child Questionnaire asks about a “typical week,” whereas the Actigraph® recorded data from a specific week). Thus, if the days on which the adolescent wore the Actigraph® were non-representative, then the correlation between the two assessments would be weakened. Another possible source of error in the self-reporting of the students was the inconsistency in whether they did or did not include activity that was part of the PE program at school. This
inter-individual variability introduces another source of error that may have contributed to the weak correlation between the Godin-Child Questionnaire and the Actigraph®. Finally, the interviews conducted with students suggested that some respondents double-counted an activity as both “strenuous” and “moderate,” thus doubling their estimate of time spent doing physical activity. For example, one student explained that when she plays soccer, she is at times exercising strenuously and at other times just walking, so she counted her soccer time as both strenuous and moderate. The above suggestions indicate that the accuracy of responses may be improved by providing more explicit directions and tailoring the assessment tool for adolescents.

Additionally, the findings indicate that the Child-Godin Questionnaire was correlated with the Actigraph® for males, but not among females. This difference may be an indication that males are generally active at a higher intensity, and perhaps more consistently active over time. Individuals are more likely to accurately recall high-intensity exercise bouts than lower-intensity (20), so if males are engaging in greater amounts of high-intensity activity, they may be more accurate in their recall of physical activity participation. Moreover, the self-report asked about “usual activity,” but the accelerometer assessed activity during a specific 7-day period. Adolescents who maintain a more consistent level of PA from week to week would be expected to report usual activity that would more closely mirror the activity monitored by the accelerometer.

Similar to the original Godin Leisure-Time Exercise Self-Reported Questionnaire, the Godin-Child Questionnaire was found reliable based on test-retest reliability (5, 13, 27), in that it consistently yielded similar results over time. Both the original and modified questionnaires were also found reliable among females and males based on test-retest correlations for each gender (5). Despite the individual variability in responses on the Godin-Child Questionnaire outlined above, the responses are consistent when considering the study population as a whole. Thus, consistent with similar PA Questionnaires, the Godin-Child Questionnaire may be a valuable assessment of PA in large groups (10). Since participants’ responses at different assessment periods yielded stable responses over time, there is reason to believe that the Godin-Child Questionnaire may be used to estimate changes in physical activity over time between groups.

To better understand individual responses, interviews were conducted with the participants to determine why participants responded in the way that they did. During the interviews, participants were asked about the specific activities s/he had in mind when responding to the Godin-Child Questionnaire. Participants often responded with activities that were already listed as examples on the Godin-Child Questionnaire, rather than alternative activities. The interviews indicated that the consistency of the participants’ responses is partially dependent on the participants’ ability to follow instructions (18, 23). The interviews indicate that the participants, overall, did not follow written instructions, suggesting that the responses on the Godin-Child Questionnaire may not have been genuine for some participants. Similar to the literature, we can see that participants may be inconsistent when responding to self-reported measurements (10).
Future studies should consider the accuracy in reporting lifestyle activity versus planned exercise. The study from which these data were obtained does not consider activities such as walking to the bus stop or biking to the grocery store, both lifestyle activities that engage participants, yet are not organized sports.

Additionally, since the present study was conducted in a school rather than a lab, the research assistants did not have control over the ambient temperature during assessments; however, it was the study’s policy to reschedule testing on days when the temperature was above a comfortable level. Future studies may benefit from collecting data in a lab setting, which would permit temperature regulation and perhaps yield a narrower range of RPE responses during the standardized exercise tasks. Studies may also conduct a thorough analysis into the causes for recall bias and potential methods to reduce the bias when administering self-report questionnaires. Administering a self-report questionnaire as close in time as possible to the PA may help reduce recall bias.

Furthermore, the interview responses from the present study suggest considerable error in comprehending the instructions for the self-reported questionnaire. Thus, future work should consider administering the Godin-Child Questionnaire in the form of an interview rather than a written survey. Finally, since the intensity and type of PA may vary between assessment periods, it is important that future studies examine the same time interval when using multiple assessment tools to reduce variability. In the present study, the Actigraph® measured a specific time frame, yet the Godin-Child Questionnaire assessed a “typical” week. It is possible that the days monitored using the Actigraph® may not be representative of a participant’s typical physical activity regimen.

ACKNOWLEDGEMENTS

The project described was supported by the National Institutes of Diabetes and Digestive and kidney Diseases, National Institutes of Health, through Grant RO1 DK088800; the National Center for Research Resources; and the National Center for Advancing Translational Sciences, National Institutes of Health, through Grant UL1 TR000153. The project was also facilitated by the contributions of Wendy Starks and the staff and faculty at Marshall Academy of the Arts.

REFERENCES


