

Differences in self-reported perceived and objective measures of duration and intensity of physical activity for adults in skiing

George Tzetzis¹, Andreas Avgerinos², Nickolas Vernadakis¹ & Efthimis Kioumourtzoglou¹

¹*Department of Physical Education and Sport Science, Aristotelian University of Thessaloniki, Greece;* ²*Department of Physical Education, Sports Sciences and Recreation Management, Loughborough University, UK*

Accepted in revised form 17 May 2001

Abstract. The effect of physical activity (PA) on health is well documented. The assessment of PA is a valuable and important issue, however, there are several methodological issues among the available methods of measurement that may have implications for the prevention of specific diseases. The purpose of this study was to examine the differences between an objective method of measurement and the subjective estimation of the PA for novice skiers. Seventy-five students aged 19–21 years old with no previous experience in ski participated in this study. Participants wore a heart rate monitor during practice in order to record the exercise intensity. Simultaneously, a trained observer recorded their time on task. A day after the objective measurement, the participants filled a questionnaire in order to estimate their per-

ceived exertion during practice as well as their perceived time on task. The results showed (1) differences between the observed time on task and the perceived recalled time, (2) no differences overall between the recorded and perceived recalled intensity of exercise but when groups were split according to their objective intensity a difference was found for each group respectively. Participants overall had overestimated the time on task, but they had underestimated the intensity of their effort recalled 1 day after their practice. These results raise the question whether a questionnaire as method of measuring PA is reliable, especially when it is used to estimate energy expenditure. However, further studies must be made in order to examine the implications of such a question.

Key words: Duration, Intensity, Perceived exertion objective measures, Perceived time, Physical activity

Introduction

Valid measurement of physical activity (PA) is a challenging task because there are several health-related dimensions and attributes associated with health status [1–5]. The measurement of PA can be defined in several ways such as the amount of work performed, the time period of the activity, the units of movement or even as numerical scores derived from responses to a questionnaire [6]. In the more sophisticated questionnaires, PA is expressed in terms of energy expenditure. However, according to the existing scientific literature, the effects of two markers of PA are the most important for health outcomes: (1) the amount of PA performed as indicated by estimation of the weekly caloric expenditure based on the equation described by Kriska and Caspersen [7], and (2) the effect of exercise intensity, independent to the overall amount of activity.

According to the formula described from Kriska and Caspersen [7], the total time of PA of each individual comprises the base for his/her classification according to ACSM (1995) guidelines for health. However, the exact nature, duration and intensity of activity required to produce health benefits is yet a controversial issue.

In order to assess PA there is a number of different methods available, each one tailored for a specific use. They can be divided in laboratory and field methods. The advantages and disadvantages of these methods depend upon the population being studied and the research objectives [6]. Field methods are further divided in objective measurements of PA (such as behavioral observation, motion analyses, movement assessment devices and physiologic response to activity), and self-report measurements (such as diaries, retrospective reports on self-completed forms and retrospective interviewer-conducted forms). However, techniques based on physiological assessment used in laboratory settings or complicated electronic devices are inappropriate for measuring PA in epidemiological studies because of their intrusiveness and cost [8]. Self-report measures, and especially questionnaires, have been commonly used to measure PA because they are inexpensive, unobtrusive and easy to administer.

Self-report instruments vary in the specificity with which mode, duration, intensity and frequency of PA are assessed, in the period of time covered by the report, and the nature of the collected data. Moreover, the data is reported as ratings, activity scores

(indexes), times, calories expenditure, or other summary scores [9]. Nevertheless, instruments that classify PA behavior in indexes, activity scores, ratings or other arbitrary summary scores, have limited value because of the lack of specificity in terms of the characteristics of the imposed stimulus, and often they have not been fully examined for evidence of reliability and validity [10]. The most promising self-reports estimate PA by deriving data for summing the time spent in PA and the time weighted by an estimation of the intensity of that activity covering a period of time of 24 hours/7 days in different seasons of the year.

However, a number of researchers have reported various problems and limitations of self-report instrument [10–14]. Problems usually arise when respondents are asked to estimate the time and the intensity of PA that they performed. Although many researchers have examined the validity and reliability of various self-report instruments against objectives measures of PA [11, 15–18], very few studies have examined the magnitude of the error correlated with the perceived time and estimation of intensity, when respondents recall their involvement in PA and/or exercise. Also, there is a limited number of studies that explored how the perceived intensity influences the perceived time on task or PA. This is a critical issue because the accuracy of the measure may have implications on the prevention of specific diseases [19] and furthermore, measures of PA are used in various settings to describe PA habits in populations, classify PA levels for intervention efforts, evaluate intervention effectiveness, assess changes in PA over time, and identify behavioral correlates of PA. The purpose of this study was to compare the accuracy of the recalled information with the objective measurement of time and intensity for activities performed by the students.

Methods

Sample

Seventy-five physical education students, 42 girls and 33 boys, 19–21 years of age ($M = 20.2$) enrolled in the ski course as part of their undergraduate program. All subjects had no previous experience in skiing and they voluntarily participated in the study.

Procedure

The study was conducted in a skiing resort and a sample of 10 subjects each day for 8 days was under inspection. Subjects participated in two 45 min practice sessions each day and their practice included specific fundamental skiing exercises. Prior to the starting of the practice session, each subject wore a

polar heart rate monitor (polar HRM) that recorded the intensity of the practice. After giving the prescribed instructions, subjects were engaged in the ski practice session. The data from the HRM were collected at the end of the two practice sessions each day. All sessions were videotaped and three trained observers recorded and evaluated the practice time (time on task) of each individual by completing an observation protocol.

The following day subjects completed a short PA questionnaire [20]. They were asked to report their perceived intensity and their time on task of each of the two practice sessions of the previous day. Subjects were then divided in three groups according to their mean heart rate (HR) scores during the practice. The cut-points for the three groups suggested by Grosser and Starischaka [21] and Zintl [22] and corresponded to adults. The first group ($n_1 = 41$) was consisted of students whose mean HR was fewer than 130 pulses per minute (p/min) (low intensity group). Students who worked in a range of mean intensity between 130–170 p/min ($n_2 = 28$) comprised the medium intensity group. Finally, students with a HR mean more than 170 p/min comprised the high intensity group ($n_3 = 6$).

Measures

Measure 1: Objective assessment of exercise intensity

In order to objectively assess the intensity of exercise, the polar HRM recorded the HR during practice. This device, compared to laboratory methods, has been proved to be a valid and reliable method of estimating work performance as well as functional and stable in a variety of field conditions [23, 24].

Measure 2: Objective time assessment

The most basic level of duration recording was the observation of the time on task which was described as the time students spent in a state of motion devoted to the fitness activity [25] a single behavior, event, or episode [26]. For this reason subjects were recorded on videotape [25] and observers evaluated the duration of the PA behavior.

Measure 3: The questionnaire

Subjects completed a questionnaire 24 hours after the practice and therefore, they reported recalled information. They were asked for: (1) their perceived exertion and (2) their perceived time on task during their practice the previous day. The questionnaire was consisted of the following items: (1) demographic information, (2) previous experience in ski, (3) Borgs' scale (four very light–20 very hard) in order to estimate the intensity of perceived exertion [27], (4) a question concerning their estimation of the time on task.

Tasks

Three different ski skills were used in this experiment according to the Rogers' (unpublished) skiing test.

Turn

The students assumed a starting position with ski tips behind the starting line. On the 'go' signal the students skied towards a slalom, lope 30 feet away, then had to process through six gates and cross the finish line. Two trials were given. The students' time needed to complete the run.

Stop

The students assumed a starting position with ski tips behind the starting line and had to stop after skiing straight down the fall line for 30 feet, executing a snow plow stop and stop completely as quick as possible. Three trials were allowed. Trials were repeated if the students fell or began to stop before reaching the stop line. Each trial was scored on a point basis, depending on the distance it took the skiers to stop.

Climb

To begin the student stood below the starting gate with ski tips behind the first gate. On the 'go' signal, the student walked the first 25 feet (level base of the slope) to gate 1. At gate 1, the student performed a step turn to the right and half steps the 50 feet to gate 2. At gate 2 a kick turn to the left was performed followed by half steps covering 50 feet to gate 3. From gate 3 to gate 4 the student performed a herringbone step. At gate 4 a step turn was performed followed by a side step to the finish line with leading. Proper climbing technique required proper use of the ski poles. Students were striving towards time.

Statistical analysis

Since homogeneity test for all variables were not significant *t*-test analysis was used to determine the differences between (1) the perceived time on task and the objective time, (2) the perceived intensity and the objective intensity for the whole sample. The group was split in three homogeneous groups (1) a low, (2) medium and (3) high according to their objective intensity. An analysis of variance (ANOVA) was used to estimate the difference between (1) the perceived and objective time on task and (2) intensity for the three groups. Each variable was tested using an alpha level of 0.05.

Results

There were significant differences $t(73) = 22.706$, $p < 0.05$ between the objective ($M = 39.28$,

$SD = 14.98$) and the perceived time on task ($M = 65.55$, $SD = 37.12$) (Figure 1). The *t*-test indicated that overall subjects overestimated their time on task.

The scores from the objective intensity from HRM and the questionnaire were transformed into *z*-scores in order to be compared. The results showed that there were no significant differences $t(73) = 22.706$, $p < 0.05$ between the objective intensity ($M = 3.2$, $SD = 0.09$) and the perceived intensity ($M = 5.7$, $SD = 1$) (Figure 2).

After the forming of the three groups according to subjects' intensity levels, a comparison among the groups was made for the objective time on task and the perceived time on task. There were no significant differences $F(2.73) = 1.2$, $p > 0.05$ on the objective time on task when groups split according to their intensity levels however, there were significant differences $F(2.73) = 5.253$, $p < 0.05$ among the three groups for the perceived time on task (Figure 3). The Scheffé post hoc analysis indicated, that the high intensity group ($M = 75$, $SD = 21.21$) perceived higher practice time than the other two groups of medium ($M = 62.88$, $SD = 31.05$) and low ($M = 66.85$, $SD = 38.37$) intensity. There were no differences for the objective time on task.

A comparison was made among the three groups for the objective and the subjective intensity. It was found that even if the groups were significantly different $F(2.73) = 76.74$, $p < 0.05$ in the levels of intensity according to the objective measurement they

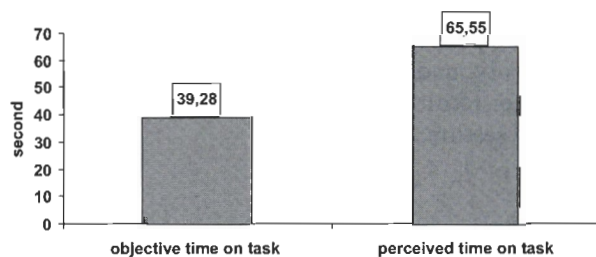


Figure 1. Subjects' scores of objective and subjective time on task.

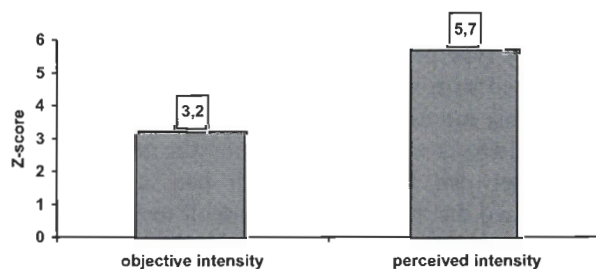


Figure 2. Subjects' scores of perceived and objective intensity.

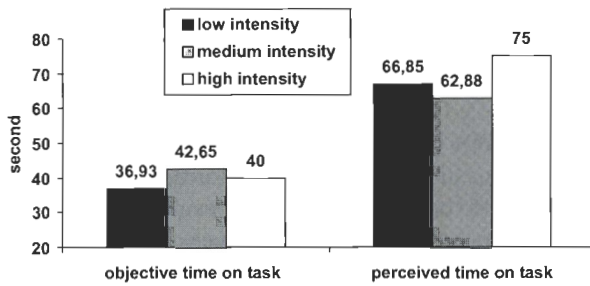


Figure 3. Difference among the three intensity groups for the objective and the perceived time on task.

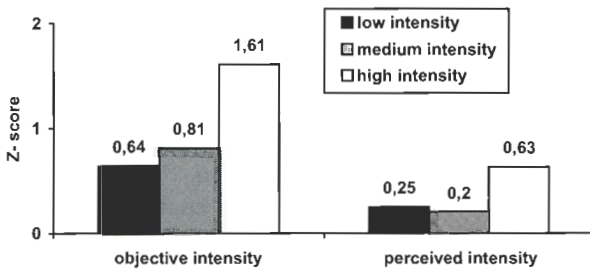


Figure 4. Differences between the objective and the perceived intensity for the different intensity levels.

did not significantly differ $F(2.73) = 2.75, p > 0.05$ in their perception of the intensity (Figure 4). For each corresponding level of intensity a comparison was made between objective and subjective intensity scores. Significant differences were found for all three pairs of comparisons, namely, low intensity ($t = 4.50, p < 0.05$), medium intensity ($t = 5.30, p < 0.05$) and high intensity ($t = 8.45, p < 0.05$). This finding implies that participants, for each corresponding level of intensity, underestimated their exertion when they asked to recall this information a day after the practice session.

Discussion

The measurement of PA has raised a considerable debate. This is because health is a multi-factorial construct and PA has been shown that can affect numerous health parameters in different ways [19]. Thus, different researchers have measured different aspects of PA such as caloric expenditure associated with any type of activity [28], while others have been concerned with measuring aerobic activity of specific durations, intensities, and frequencies [29]. Furthermore, there is an agreement that the true relationships between PA and health can not be fully established due to the current state of techniques used in the measurement of PA [30].

Many studies have shown the diversity between the various methods, caused probably by the limited use of the 'gold standard' (doubly labeled water) and the

practical field limitations of other methods [15, 31, 32]. The need for explanation of this variation marks the importance of the validity and the accuracy both of the objective and the subjective PA measurements. Although the questionnaire is a widespread method of measuring PA in epidemiological studies, it is characterized by a number of limitations that may affect its potential use as a valid and accurate research instrument. In this study an effort was made to estimate the accuracy of recalled information for perceived time and intensity of exercise.

The results indicated that subjects overall, independently of the intensity of their effort, overestimated their time on task. They perceived that the time they exercised was much longer than the real time. In addition, it was found that the high intensity group recalled the time on task with less accuracy. This finding indicates that the accuracy of the recalled time on task was negatively associated with the level of activities' intensity performed by the participants. Therefore, self-report PA instruments that ask from responders' precise time judgements in order to classify them in activity categories, it is likely to overestimate the actual activity. This is an important finding because the duration of exercise is a key indicator in the measurement of PA.

In this research the estimation of the perceived intensity, retrospectively recalled, for the whole sample did not differ from the objective intensity of the exercise measured by the HRM. However, when participants split into three groups according to their intensity levels the discrepancies of intensity existed as measured by a HRM did not exist when measured by the questionnaire. This finding indicates that participants for each corresponding level of intensity underestimated their exertion when they asked to recall this information a day after the practice session. Borgs' scale is a valid instrument to measure perceived exertion simultaneously with the exercise, however, it seems that it was difficult for the participants to recall the intensity of their effort or to express it in a more quantitative way one day after practice. The Paffenbargers' physical activity questionnaire [33] is an example of a self-report instrument in which applied the idea of a 10-point scale in order to quantify the level of exertion 'When you are exercising in your usual fashion, how would you rate your level of exertion (degree of effort)? Please, circle one number'. According to the findings of this study, this description did not provide more accurate information than more general descriptions used in other questionnaires (such as 'huff and puff activity' [10], as 'make you breathe heavily and make your heart beat fast' ('modifiable activity questionnaire for adolescents') [34], as 'regularly engagement in strenuous activity or hard physical labor' (the lipid research clinics' physical activity questionnaire), [15]). It is expected that measures of PA that require from responders to recall intensity of activities retrospec-

tively for prolonged periods of time (e.g. 12 months) to provide only general estimations [30].

The discussion so far indicates that self-reports may be ineffective instruments to provide accurate information concerning the time and intensity of PA. An important question that emerges comes with the magnitude of the associated error in the measurement. This is a particular difficult answer taking into account the methodology used in this study, namely, (1) the error of the time and intensity estimation may be more obvious when the recall time is longer than used in this study (for example 7 days recall), (2) the type of the activity (daily living activities or occupational activities vs. organized or sport activities) may influence the process of the recalled information.

There was no evaluation of the motivation that students had during their reporting of their intensity levels. Participants' motivation to provide accurate results may influence responses and future research is needed to assess the degree to which accuracy of PA information is affected by subjects' motivational level.

A second limitation related to the type of activity that may also influence participants' responses. In this study, a well-organized activity (ski) was chosen in a particular setting. The findings of the present study must be confirmed using other forms of PA such as occupational, daily or leisure activities performed in different settings. However, organized activities are easier to be recalled. It is assumed that with less organized activities it is more difficult to evaluate accurately the intensity or the time of the activity.

Participants of this study were physical education students and were more familiar with evaluation of time or intensity of their effort than the typical population. It is expected therefore, that the estimation of time and intensity may be different with groups of different characteristics in terms of age, educational level, activity level and obesity.

Finally, the accuracy of the recalled information may be different if: (1) subjects had to recall longer time (more than the 45 min of exercise) and (2) subjects completed the questionnaire more than the 24 hours. Some measures of PA require individuals to recall activities retrospectively for a year or more [30]. In the present study, PA was observed for a short period of time, and, such as, these results should not be over generalized. Longer observation and assessment periods may be needed for more accurate results.

The use of questionnaires is a favorable procedure in epidemiological studies, however, the discussion so far showed that they may not be the most accurate instrument to measure quantitatively PA levels. Therefore, when questionnaires are used to quantify PA (in terms of time or energy expenditure) in order to evaluate the effectiveness of intervention efforts, to assess changes in PA over time and to identify be-

havioral correlates of PA, should be considered with some degree of caution. The findings of the present study support that a combination of a subjective and an objective measures should be used, especially when use of objective measures only is not possible. Finally, it is clear that before the relation between health and PA can be established, research should focus on the development of simple, low cost, reliable, and valid PA assessments [35].

References

1. Blair SN. Physical activity, fitness and coronary heart diseases. In: Bouchard C, Shephard RJ, Stephens T (eds), *Physical Activity, Fitness and Health: International Proceeding and Consensus Statement*, Champaign, IL: Human Kinetics, 1994: 579–590.
2. Boreham C, Savage JM, Primor D, Cran G, Strain JJ. Coronary risk factors in schoolchildren. *Arch Dis Child* 1993; 68: 182–186.
3. Caspersen CJ. Physical activity epidemiology: Concepts, methods, and applications to exercise science. *Exer Sports Sci Rev* 1989; 17: 423–473.
4. Foret CM, Clemons MJ. The elderly's need for physical activity. *JOPERD* 1996; 67(4): 57–61.
5. Must A. Morbidity and mortality associated with elevated body weight in children and adolescents. *Am J Clin Nutr* 1996; 63(Suppl): 445S–447S.
6. Montoye HJ, Kemper HCG, Saris WHM, Washburn RA. *Measuring Physical Activity and Energy Expenditure*. Champaign, IL: Human Kinetics. 1996.
7. Kriska AM, Caspersen CJ. Introduction to a collection of physical activity questionnaires. *Med Sci Sports Exerc* 1997; 29(Suppl 6): S5–S9.
8. LaPorte RE, Montoye HJ, Caspersen CJ. Assessment of physical activity in epidemiologic research: Problems and prospects. *Pub Health Rep* 1985; 100: 131–146.
9. Sallis JF. Self-report measures of children's physical activity. *J Sch Health* 1991; 61: 215–219.
10. Cale L. Self-report measures of children's physical activity: Recommendations for future development and a new alternative measure. *Health Edu J* 1994; 53: 439–453.
11. Baranowski T, Dworkin RJ, Cieslic JC, et al. Reliability and validity of self-report of aerobic activity: Family health project. *Res Q Exerc Sport* 1984; 55(4): 309–317.
12. Saris WHM. The assessment and evaluation of daily physical activity in children. A review. *Acta Paediatr* 1985; 318(Suppl): 37–48.
13. Klesges R, Eck L, Mellon M, Fulliton W, Somes G, Hanson C. The accuracy of self-report of physical activity. *Med Sci Sports Exerc* 1990; 22(5): 690–697.
14. Sallis JF, Condon SA, Goggin JK, Lolody B, Alcaraz JE. The development of self-administered physical activity surveys for 4th grade students. *Res Q Exerc Sport* 1993; 64: 25–31.
15. Ainsworth BE, Jacobs DR, Leon LA. Validity and reliability of self reported physical activity status: The Lipid Research Clinics questionnaire. *Med Sci Sports Exerc* 1993; 25(1): 92–98.

16. Baranowski T. Validity and reliability of self-report of physical activity: An information-processing perspective. *Res Q* 1988; 59(4): 314–327.
17. Coleman JK, Saelens BE, Wiedrich-Smith MD, Finn JD, Epstein H. Relationship between TriTrac-R3D vectors, heart rate, and self-report in obese children. *Med Sci Sports Exerc* 1997; 29: 1535–1542.
18. Kohl HW, Fulton JE, Caspersen CJ. Assessment of physical activity among children and adolescents: A review and synthesis. *Prev Med* 2000; 31: S54–S76.
19. Pate RR. Physical activity and health: Dose-response issues. *Res Q Exerc Sport* 1995; 66(4): 313–317.
20. Borg G. Psychophysical bases of perceived exertion. *Med Sci Sports Exerc* 1982; 14: 377–387.
21. Grosser M, Starischaka S. *Das neue konditions training*. Munchen: BLV Verlagsgesellschaft mbH. 1998.
22. Zintl F. *Ausdauertraining*. Munchen: BLV Verlagsgesellschaft mbH. 1990.
23. Leger L, Thivierge M. Heart rate monitors: Validity, stability and functionality. *Phys Sports Med* 1988; 16: 143–151.
24. Kalkwarf HJ, Haas JD, Belko AZ, et al. Accuracy of heart-rate monitoring and activity diaries for estimating energy expenditure. *Am J Clin Nutr* 1989; 49: 37–43.
25. Chung TB. The time on fitness instrument (TOFI) Time on task in Nonschool Settings. In: Darst PW, Zakrajsek DB, Mancini VH (eds), *Analyzing Physical Education and Sport Instruction*, Champaign, IL: Human Kinetics. 1989.
26. Cooper JO, Heron TE, Heward WL. *Applied Behavior Analysis*, Columbus, OH: Merrill, 1987.
27. Noble BJ, Robertson RJ. *Perceived Exertion*, United States: Human Kinetics, 1996.
28. Sallis JF, Buono MJ, Roby JJ, Micale FG, Nelson JA. Seven-day recall and other physical activity self-reports in children and adolescents. *Med Sci Sports Exerc* 1993; 25(1): 99–108.
29. Simons-Morton BG, O' Hara NM, Parcel GS, Wei Huang I, Baranowski T, Wilson B. Children's frequency of participation in moderate to vigorous physical activities. *Res Q Exerc Sport* 1990; 61(4): 307–314.
30. Williams E, Klesges RC, Hanson CL, Eck LH. A prospective study of the reliability and convergent validity of three physical activity measures in a field research trial. *J Clin Epidemiol* 1989; 42: 1161–1170.
31. Dishman RK, Steinhardt M. Reliability and construct validity for a 7-day re-call of physical activity in college students. *Med Sci Sports Exerc* 1988; 20(1): 14–25.
32. Verheul AM, Prins AN, Kemper HCG, Kardinaal AFM, Erp-Bart MAJ. Validation of a weight-bearing physical activity questionnaire in a study of bone density in girls and women. *Pediatr Exer Sci* 1998; 10: 38–47.
33. Lee IM, Paffenbarger J, Hsieh CC. Time trends in physical activity among college alumni, 1962–1988. *Am J Epidemiol* 1992; 108: 161–175.
34. Aaron DJ, Kriska AM, Dearwater JA, Cauley JA, Metz KF, LaPorte RE. Reproducibility and validity of an epidemiologic questionnaire to assess past-year physical activity in adolescents. *Am J Epidemiol* 1995; 142: 191–201.
35. LaPorte RE, Adams LL, Savage DD, Branes G, Dearwater S, Cook T. The spectrum of physical activity, cardiovascular disease and health: The epidemiologic perspective. *Am J Epidemiol* 1984; 120: 507–517.

Address for correspondence: George Tzetzis, Christopoulou 9, 54635 Thessaloniki, Greece
E-mail: tzetzis@phed.auth.gr