A Simple Method to Assess Exercise Behavior in the Community

by
G. Godin
Ecole des Sciences Infirmières
Université Laval
Ste-Foy (Québec)
and
R. J. Shephard
School of Physical and Health Education, and
Department of Preventive Medicine and Biostatistics
Univeristy of Toronto
Toronto, Ontario

GODIN, G. and SHEPHARD, R. J., A Simple Method to Assess Exercise Behavior in the Community. **Can. J. Appl. Spt. Sci.** *10:3 141-146, 1985. The reliability and concurrent validity of a simple questionnaire to assess leisure time physical activity has been investigated on 306 self-selected healthy adults of both sexes (163 M; 143 F). Values of body fat (BF) and maximum oxygen intake ($\dot{V}O_2$ max) expressed as percentiles of appropriate age and sex categories were used as criteria of validity for the questionnaire. BF and $\dot{V}O_2$ max were predicted from the Durnin and Womersley skinfold equations, and the laboratory version of the Canadian Home Fitness Test respectively. The strongest correlation was between $\dot{V}O_2$ max (percentile) and reported strenuous exercise ($r = 0.35$). The optimum discriminant function for $\dot{V}O_2$ max was based on a combination of reported strenuous and light activity. This yielded a correct 2-way classification of 69% of the subjects. A combination of sweat-inducing and moderate exercise yielded a correct 2-way classification of BF for 66% of subjects. The reliability coefficients for the optimum discriminant functions classifying $\dot{V}O_2$ max and BF were 0.83 and 0.85 respectively. We conclude that this simple instrument has potential value for the assessment of leisure time exercise behavior, offering the possibility of examining changes in behavior following the implementation of health and physical fitness promotion programmes in the community.*

Reliability, validity, questionnaire, behavior, physical activity

La fidélité et la validité concurrente d'un questionnaire pour mesurer le comportement de pratique d'activités physiques pendant les périodes de loisirs ont été investiquées auprès de 306 adultes volontaires des deux sexes (163 H; 143 F). Les valeurs de pourcentage de graisse (PG) et de la consommation maximale d'oxygène ($\dot{V}O_2$ max), exprimées en percentiles selon les catégories d'âges et le sexe, ont été utilisées comme critère de validité pour le questionnaire. Le PG et la $\dot{V}O_2$ max ont été prédit respectivement à partir des équations pour plis adipeux de Durnin et Womersley, et de la version laboratoire du Physitest Canadien. La corrélation la plus élevée a été mesurée entre la $\dot{V}O_2$ max (percentile) et la quantité d'activités intenses rapportée ($r = 0.35$). La fonction discriminante optimum pour la $\dot{V}O_2$ max était basée sur une combinaison des activités intenses et légères. Cette fonction a permis de classer correctement 69% des sujets. La fonction discriminante permettant de classer les sujets selon leur PG a été précise pour 66% des sujets. Les coefficients de fidélité pour les fonctions discriminantes permettant de classer les sujets selon leur $\dot{V}O_2$ max et leur PG étaient respectivement de 0.83 et 0.85. Nous concluons que ce simple instrument est approprié afin d'établir le comportement de pratique d'activités physiques pendant les périodes de loisir. Ce questionnaire offre également la possibilité d'observer les changements de comportement à la suite de la promotion d'un programme communautaire de santé et de condition physique.

Validité, fidélité, comportement, activité physique, questionnaire

INTRODUCTION

For almost a century, research workers have been interested in studying human habitual physical activity and the corresponding energy requirements (Durnin and Passmore, 1967; Shephard, 1977, 1978; Andersen et al., 1978). Recent epidemiological studies of the relationship between habitual physical activity and health have stressed the

importance of using appropriate instruments for such measurements (LaPorte et al., 1979). Attention has been directed particularly to the reliability and validity of devices and questionnaires used when investigating the exercise habits of large populations.

Electrical and mechanical devices usually modify personal behavior (Wessel et al., 1965), temporarily changing normal activity patterns. Moreover, the costs of instrumentation, operation and maintenance generally preclude making such measurements on any large scale.

Most existing activity questionnaires are also lengthy and correspondingly difficult to use. The usual approach in a large-scale survey is to record physical activities, subsequently converting the reported information into corresponding energy expenditures (Taylor et al;. 1978). Some authors such as Montoye et al. (1976) have insisted that individual subjective responses need careful verification by experienced observers. However, data on the energy costs of many reported activities are either lacking or out of date (Reiff et al., 1967), so that only rough approximations of daily energy expenditures can be made. A further difficulty is that the energy cost of any given task varies with the environment and with the performer's age, sex, body mass, skill and level of fatigue (Durnin and Passmore, 1967). Only the average metabolic cost is known for most leisure activities, so that it becomes unrealistic to do more than classify people into one of several simple physical activities categories (Shephard and McClure, 1965; Shephard and Callaway, 1966).

Often, too little attention has been paid to the quality of epidemiological data, particularly the responses obtained through questionnaires and/or interviews (Gordis, 1979; LaPorte et al., 1979). Many systems of classifying the vigour of daily activity have been proposed (Wessel et al., 1975; Yasim et al., 1967; Morris et al., 1973; Montoye et al., 1976; Paffenbarger et al., 1978; Taylor et al., 1978; Magnus et al., 1979; Furrie and Stephens, 1983), but in general there is little information on the reliability and validity of the suggested methodology (Taylor et al., 1978; Gordis, 1979). Our objective was thus to develop a physical activity questionnaire that was reliable, valid, and easy to complete quickly without detailed checking by professional staff. The purpose of the questionnaire was to classify people into several activity categories, with a view to examining this aspect of behavior in relation to psychosocial variables before and after implementation of community health and physical fitness promotion programmes.

METHODS AND PROCEDURES

i) Concurrent validity

A group of 306 self-selected healthy adult volunteers (163 M; 143 F, aged 18 to 65 years) were recruited from the physical fitness evaluation service offered by the Department of Athletics and Recreation. Participants from the University of Toronto were evaluated on campus, and other white collar workers were tested at their normal place of employment.

All subjects first filled out an approved consent form and then responded to a simple activity questionnaire (Appendix). The questions posed were chosen after a preliminary study of items used by previous authors (Morris et al., 1973; Montoye et al., 1976; Paffenbarger et al., 1978; Taylor et al., 1978; Magnus et al., 1979; Williams et al., 1980). Each of the selected items discriminated successfully between very active and sedentary people in a univariate analysis.

Skinfold thicknesses were measured at four standard sites (biceps, triceps, subscapular, suprailiac) and the percentage of body fat (BF) was predicted, using the equations of Durnin and Womersley (1974). The Canadian Home Fitness step test was also carried out, with electrocardiographic determination of heart rate (Shephard et al., 1976; Shephard, 1979). Maximum oxygen intake ($\dot{V}O_2$ max) was estimated using the equation of Jetté et al. (1974). BF and $\dot{V}O_2$ max were each expressed as percentiles of appropriate age and sex norms for the Canadian population (Fitness and Amateur Sport, 1979; Jetté et al., 1980) to provide two relatively independent criteria of concurrent validity for the questionnaire.

Subjects were divided into two categories for each of the two criteria.[1] Individuals with a predicted $\dot{V}O_2$ max over the 50th percentile for their age and gender were assigned to a first (FIT) category, while the remaining subjects were assigned to a second (UNFIT) category. In the same manner, subjects with an estimated BF above the 50th percentile for their age and gender were classified as THIN, while the remaining subjects were categorized as FAT. Stepwise discriminant analyses tested how far the questionnaire was able to discriminate between FIT and UNFIT, and between THIN and FAT subjects. The results not only indicated the combined discriminatory power of all questionnaire items used, but also showed which questions were the most potent discriminators. Finally, the agreement between the classification as established by discriminant analysis and the a priori categorizations of fitness and fatness was tested by the coefficient of agreement for nominal scales (κ, Cohen, 1960).

ii) Reliability

A two week test-retest reliability study was conducted on a separate but similar population. The questionnaire was completed by 53 subjects on each of two occasions. The coefficient of reliability (Safrit et al., 1978), defined as the ratio between the variance of the true scores and the variance of the obtained scores, was calculated using an analysis of variance for repeated measures.

RESULTS

The physical characteristics of the subjects are presented in Table 1. Average percentile scores for BF and $\dot{V}O_2$ max were, respectively, 59.2 and 62.2. Findings were similar for both sexes, the sample being slightly thinner and fitter than the reference population.

Leisure-time exercise habits are presented in Table 2. A total weekly score of 45.8 arbitrary units has been derived by summing the reported weekly frequency of participation at each of three intensity levels multiplied by the corresponding anticipated MET value (9, 5, or 3 METS).

TABLE 1

Physical Characteristics of Sample Used In Validity Test

Sample Characteristics	Male	Female	Total
Number of Subjects	163	143	306
Age ($\bar{X} \pm S$), Years	31.1 ± 9.6	30.3 ± 10.0	30.7 ± 9.8
% BF ($\bar{X} \pm S$)	17.3 ± 5.2	25.9 ± 5.0	
BF Percentile ($\bar{X} \pm S$)	59.1 ± 28.1	59.2 ± 26.0	59.2 ± 27.1
$\dot{V}O_2$ max ($\bar{X} \pm S$) ($ml \cdot kg^{-1} \cdot min^{-1}$)	43.8 ± 7.1	34.7 ± 4.7	
$\dot{V}O_2$ max Percentile ($\bar{X} \pm S$)	63.3 ± 31.6	60.9 ± 27.2	62.2 ± 29.6

TABLE 2

Leisure-Time Exercise Habits of Validity Sample

Leisure Time Exercise	Times Per Week
Strenuous	2.2
Moderate	2.9
Light	3.9
Total (T)*	45.8*

Frequency of Sweat-Inducing Exercise	Proportion of Subjects
Often	97/306
Sometimes	143/306
Never/Rarely	66/306

*T = (9 x Strenuous) + (5 x Moderate) + (3 x Light); Arbitrary Units

The strongest individual correlations between subjective and objective data related $\dot{V}O_2$ max percentile to reported strenuous exercise ($r = 0.38$; $p < 0.001$) and BF percentile to the self-rating of sweat-inducing exercise ($r = 0.21$, $p < 0.001$) (Table 3). Using the two-way categorizations of fitness and fatness, it was possible to generate a discriminant function for each of the two dependent variables (Table 4). The $\dot{V}O_2$ max function was determined mainly by strenuous exercise, with a small contribution from light exercise. The BF function was generated by the self-rating of strenuous and a sweat-inducing exercise.

Table 5 presents weighted means for each group and each discriminant function. The signs and weights of the discriminating variables indicate that the probability of being FIT was increased when the subject reported a combination of strenuous exercise and light exercise; a Wilkes λ of 0.88 shows that the variables used in this function accounted for about 11.9% of the between-group

TABLE 3

Pearson Correlation Coefficients Between Reported Leisure Activity and Objective Measures of Physical Condition

MEASURE OF PHYSICAL CONDITION	REPORTED LEISURE ACTIVITY				
	STRENUOUS	MODERATE	LIGHT	TOTAL	SWEAT-INDUCING
BF Percentile	0.21***	0.08	0.06	0.13**	0.21***
$\dot{V}O_2$ max Percentile	0.38***	0.03	0.04	0.24***	0.26***

** $p < 0.01$

*** $p < 0.001$

TABLE 4

Standardized Weightings of Variables Contributing to Discriminant Functions For $\dot{V}O_2$ max and BF

Activity Description	$\dot{V}O_2$ max Function	BF Function
Strenuous	0.98	0.73
Moderate	—	—
Light	0.18	—
Sweat-Inducing	—	0.37

variance in $\dot{V}O_2$ ($p < 0.001$). Likewise, the weights and signs of the variables used in the BF function indicate that the probability of being THIN was increased when the subject reported a combination of strenuous exercise and frequent sweat-inducing exercise. Nevertheless a Wilkes λ of 0.95 indicates that these two variables explained only about 5.3% of the between-group variance in BF ($p < 0.001$).

TABLE 5

Mean Scores (Centroids) on the Two Discriminant Functions When Subjects Were Divided Into Two Sub-Groups on the Basis of Physical Measurement to Be Predicted

Sub-Group	$\dot{V}O_2$ max Function	BF Function
Over 50th Percentile	0.28	0.18
50th Percentile or Under	−0.49	−0.31

When the actual number of individuals in each $\dot{V}O_2$ max category was compared with the number predicted from the discriminant function, 69% of individuals were correctly classified (Table 6). The value $_k$, which describes the amount of agreement beyond that expected by chance, was 0.30 (95% confidence limits 0.18 to 0.42)[2]. Table 7 presents parallel information for the BF function. Sixty six percent (66%) of all subjects were correctly assigned, with a $_k$ value of 0.17 (95% confident limits 0.04 to 0.30). In this case $_k$ was only slightly above the level of agreement anticipated by chance.

Two week test-retest reliability coefficients were respectively 0.94, 0.46, 0.48, and 0.80 for self-reports of strenuous, moderate, light, and sweat-inducing exercise (Table 8). Reliability coefficients for the optimum discriminant functions classifying $\dot{V}O_2$ max percentile and BF percentile were 0.83 and 0.85 respectively.

TABLE 6

Comparison of Actual Number of Cases in Fit and Unfit Categories Relative to the Number of Subjects Assigned to These Categories Using the $\dot{V}O_2$ max Discriminant Function

Group	Actual Number of Cases	Predicted Group Membership Fit	Unfit
Fit	194	156	38
Unfit	110	57	53

NOTE: $k = 0.30$; $d_k = 0.06$;
95% Confidence Limits = 0.18 ←→ 0.42;
Overall Accuracy of Classification = 69%

TABLE 7

Comparison of Actual Number of Cases in Thin and Fat Categories Relative to the Number of Subjects Assigned to These Categories Using the BF Discriminant Function

Group	Actual Number of Cases	Predicted Group Membership Thin	Fat
Thin	192	170	22
Fat	112	82	30

NOTE: $k = 0.17$; $d_k = 0.066$;
95% Confidence Limits = 0.04 ←→ 0.30;
Overall Accuracy of Classification = 66%

TABLE 8

Two Week Test-Retest Reliability Coefficients for Self-Reports of Leisure Activity and Corresponding Discriminant Functions

Self-Reports of Activity	Reliability Coefficients*
Strenuous	0.94
Moderate	0.46
Light	0.48
Total	0.74
Sweat-Inducing	0.80
$\dot{V}O_2$ max Percentile Classification	0.83
BF Percentile Classification	0.85

* $R_{TT} = \dfrac{MS_{IND} - MS_{RES}}{MS_{IND} + MS_{RES}}$

DISCUSSION

The purpose of this study was to explore the ability of a simple questionnaire to assess patterns of exercise behavior during a subject's leisure time. Discriminant function analyses indicate that the selected questionnaire items were able to discriminate the majority of FIT from UNFIT people (Table 6), and the majority of THIN from FAT subjects (Table 7). Scores yielded by the two discriminant functions were also very stable over a two-week period (Table 8).

Despite our use of a relatively simple index of cardiorespiratory condition the $\dot{V}O_2$ max function was quite successful in distinguishing FIT individuals (more than 80% correct), indicating that self-reports of a high frequency of both strenuous and light exercise are significantly associated with the Canadian Home Fitness Test index of physical condition. The limited success in categorizing UNFIT subjects (48.1% correct) suggests that self-reports of a low frequency of exercise are not closely associated with a poor fitness level. Possibly, unfit subjects have a less accurate perception of their leisure behavior. Other unmeasured variables such as constitutional differences of aerobic power and physical activity undertaken at work may also exert a counter-influence, particularly in subjects who lie at the lower end of the fitness spectrum. In the same way, the BF function is quite successful in predicting THIN subjects (88.5% correct), indicating that self-reports of high frequency of both strenuous exercise and frequent sweat-inducing exercise are significantly related with this expression of physical condition. The lower accuracy of categorization for FAT subjects (26.7% correct) shows that failure to report strenuous or sweat-inducing exercise does not necessarily lead to fatness. Presumably other unmeasured factors again come into play-accuracy of reporting, along with differences of constitution, metabolic rate and eating habits.

The validity of the present questionnaire compares favorably with other simple techniques of predicting fitness, such as the 5-level categorization of habitual activity proposed by Shephard and McClure (1965). By a combination of various objective physiological measurement, these authors explained 58% of the variance in the subjective grading of the level of normal exercise activity. In terms of predicting fitness, our questionnaire could possibly be improved further by adding items relating to occupational activity and constitution, but this would not improve its validity with respect to our primary objective, that of evaluating current leisure behavior.

The limitations of our sample must finally be stressed. The questionnaire has to date been validated only on a self-selected group of educated individuals, white collar employees with some inclination towards physical activity. Nevertheless, the present results suggest that our simple questionnaire can provide information on activity patterns that will be useful when evaluating the baseline behavior of a population.

Received 15 September, 1983.
Accepted in final form 22 March, 1985.

1 Since our validity criteria were «percentiles» values based on Canadian norms, we do not have to consider either «sex» or «age» since by definition «percentile norms» are standardized for both «sex» and «age».

2 A value of 0.40 is recommended for predictive validity (Fleiss, 1981). However, considering that this is a concurrent validity study, our value of 0.30 is high enough for eliminating the possibility that the observed classification would have occured by chance.

ACKNOWLEDGEMENT

The cooperation of Dr. M. Cox, of the Fitness Testing Unit, Department of Athletics and Recreation, University of Toronto, is acknowledged with thanks.

REFERENCES

Andersen, K. L., Masironi, R., Rutenfranz, J., and Seliger, V. **Habitual physical activity and health.** World Health Organization, Copenhagen, 1978.

Cohen, J. A coefficient of agreement for nominal scales. **Educ. and Psych. Measurement**, 20(1): 37-46, 1960.

Durnin, J. V. G. A. and Passmore, R. **Energy, work and leisure,** Heinemann Educational Books, London, 1967.

Durnin, J. V. G. A. and Womersley, J. Body fat assessed from total body density and its estimation from skinfold thichness: measurements on 481 men and women aged from 16 to 72 years. **Br. J. Nutr.**, 32: 77-97, 1974.

Furrie, A. D. and Stephens, T. Energy expenditure patterns in the Canadian population. In: F. Landry (Ed.) **Health risk estimation, risk reduction and health promotion.** Proceedings of the 18th annual meeting, The Society of Prospective Medicine, pp. 104-119, 1983.

Fitness and Amateur Sport. **Physitest normalisé; normes et notes en percentile.** Minister of State, Federal Government, 1979.

Fleiss, J. **Statistical methods for rates and proportions.** Toronto: John Wiley & Sons, 1981.

Gordis, L. Assuring the quality of questionnaire data in epidemiological research. **Am. J. Epidemiology**, 109(1)): 21-24, 1979.

Jetté, M., Campbell, J., Mongeon, J., and Routhier, R. The Canadian home fitness test as a predictor of aerobic capacity. **Can. Med. Ass. J.**, 114: 680-682, 1974.

Jetté, M., Chartier, P., Lauzon, R. et al. Reformulation of norms for percentage of body fat. Research report submitted to Fitness and Amateur Sport, December 1980.

LaPorte, R. E., Kuller, L. H., Kupper, D. J. et al. An objective measure of physical activity for epidemiologic research. **Am. J. Epidemiology**, 109(2): 158-168, 1979.

Magnus, K., Matroos, A., and Strackee, J. Walking, cycling, or gardening, with or without seasonal interruption, in relation to acute coronary events. **Am. J. Epidemiology**, 110(6): 724-733, 1979.

Montoye, H. J., Block, W. D., Metzner, H. L., and Keller, J. B. Habitual physical activity and serum lipids: males, age 16-64 in a total community. **J. Chron. Dis.**, 29: 697-709, 1976.

Morris, J. N., Chave, S. P. W., Adam, C., and Sirey C. Vigorous exercise in leisure-time and the incidence of coronary heart-disease. **Lancet**, February 17, 331-339, 1973.

Paffenbarger, R. S., Wing, A. L., and Hyde, R. T. Physical activity as an index of heart attack risk in college alumni. **Am. J. Epidemiology**, 108: 161-175, 1978.

Reiff, G. G., Montoye, H. L., Remington, R. D. et al. Assessment of physical activity by questionnaire and interview in: **Physical activity and the heart**, Karvonen, M. J. and A. J. Barry, Thomas publisher, pp. 336-371, 1967.

Salfrit, M. J., Atwater, A. E., Baumgartner, T. A., and West, C. **Reliability Theory**. A.A.H.P.E.R., 1976.

Shephard, R. J., Bailey, D. A., and Mirwald, R. L. Development of the Canadian Home Fitness Test. **Can. Med. Ass. J.**, 114: 675-679, 1976.

Shephard, R. J. **Endurance Fitness**, 2nd edition, Toronto, University Press, 1977.

Shephard, R. J. **Human Physiology Work Capacity**, Cambridge, University Press, 1978.

Shephard, R. J. Current Assessment of the home fitness test. **S. Af. J. Spt. Sci.**, 2: 19-35, 1979.

Shephard, R. J. and McClure, R. L. The prediction of cardio-respiratory fitness. **Int. Z. Angew. Physiol.**, 21: 212-223, 1965.

Shephard, R. J. and Callaway, S. Principal component analysis of the responses to standard exercise training. **Ergonomics**, 9(2): 141-154, 1966.

Taylor, H. L., Jacobs, D. R., Schucker, B. et al. A questionnaire for the assessment of leisure time physical activities. **J. Chron. Dis.**, 31: 741-755, 1978.

Wessel, J. A., Montoye, H. L., and Mitchell, H. Physical activity assessment: By recall record. **Am. J. Pub. Health**, 55(9): 1430-1436, 1965.

Williams, C. L., Carter, B. J., and Eng, A. The «Know Your Body» program: A developmental approach to health education and disease prevention. **Prev. Med.**, 9: 371-383, 1980.

Yasin, S., Alderson, M. R., Marr, J. W. et al. Assessment of habitual physical activity apart from occupation. **Brit. J. Prev. Soc. Med.**, 21: 163-169, 1967.

APPENDIX

LEISURE TIME EXERCISE QUESTIONNAIRE

1. Considering a **7-day period** (a week), how many times on the average do you do the following kinds of exercise for **more than 15 minutes** during your **free time** (write in each circle the appropriate number).

	TIMES PER WEEK
a) **STRENUOUS EXERCISE (HEART BEATS RAPIDLY)** (i.e. running, jogging, hockey, football, soccer, squash, basketball, cross country skiing, judo, roller skating, vigorous swimming, vigorous long distance bicycling)	◯
b) **MODERATE EXERCISE (NOT EXHAUSTING)** (i.e. fast walking, baseball, tennis, easy bicycling, volleyball, badminton, easy swimming, alpine skiing, popular and folk dancing)	◯
c) **MILD EXERCISE (MINIMAL EFFORT)** (i.e. yoga, archery, fishing from river bank, bowling, horseshoes, golf, snow-mobiling, easy walking)	◯

2. Considering a **7-day period** (a week), during your **leisure-time**, how often do you engage in any regular activity long enough to **work up a sweat** (heart beats rapidly)?

OFTEN	SOMETIMES	NEVER/RARELY
1. ☐	2. ☐	3. ☐