



# The Open Public Health Journal

Content list available at: <https://openpublichealthjournal.com>



## RESEARCH ARTICLE

### Test-retest Reliability of Physical Activity Behavior Questionnaire for Adults

Min-Haeng Cho<sup>1,\*</sup>

<sup>1</sup>Department of Sports & Leisure, College of Humanity, Daegu University, Gyeongsan, South Korea

#### Abstract:

#### Background:

There is a greater need than ever to assess daily physical activity more accurately and to monitor comprehensive trends in habitual physical activity to meet the physical activity recommendations of health organizations, such as the American Heart Association and the World Health Organization.

#### Objective:

The purpose of this study was to design a physical activity behavior questionnaire that reflected health organizations' physical activity recommendations and to assess the questionnaire's test-retest reliability.

#### Methods:

By using a quota sampling technique, 200 participants were selected to participate in the first survey, and 117 of them completed the second survey. All respondents were participating in recreational programs provided by the community centers (senior center, residential culture center, lifetime academy, and sports center). Participants ranged in age from 20 to 88 years old ( $M = 51.83$ ,  $SD = 21.70$ ). The test-retest repeatability was assessed using Cohen's kappa statistics.

#### Results:

The kappa score showed good agreement for all items (frequency, intensity, duration, and overall length) related to aerobic, muscular strength, and flexibility activity. The reliability coefficient showed good to high internal consistency throughout. The alpha statistics was 0.658 for questions on flexibility activity participation, 0.666 for questions on aerobic activity participation, and 0.935 for questions on muscular activity participation.

#### Conclusion:

This paper made an important methodological contribution to the assessment of physical activity by showing good test-retest reliability and internal consistency properties for the measurements of physical activity that reflect health organizations' recommendations for different dimensions, types, and domains of physical activity.

**Keywords:** Health, Physical activity, Questionnaire, Reliability, Types & dimensions of physical activity, Kappa score.

#### Article History

Received: August 13, 2020

Revised: December 24, 2020

Accepted: January 15, 2021

## 1. INTRODUCTION

Many countries, including South Korea, have updated their physical activity recommendations on the dose-response relationship between the frequency, duration, intensity, and type of physical activity conducive to health. Although the recommendations for physical activity vary somewhat, and therefore, so do the estimates of physical activity, more accurate

measurements of physical activity have become an interesting and important topic of study around the world. One study has noted the existence of 85 self-reported physical activity questionnaires for adults and 13 questionnaires for elders [1].

While many physical activity questionnaires have been developed in the past few decades for different populations in various countries, they are limited in terms of their reliability and validity [2 - 4]. Many self-reported physical activity questionnaires have been created to evaluate the frequency, duration, and intensity of physical activity in adults [5, 6]. However, the content of these measures differ widely and have

\* Address correspondence to this author at Department of Sports & Leisure, College of Humanity, Daegu University, South Korea 53, 201 Daegudae-ro, Gyeongsangbuk-do, 384 Republic of Korea; Tel: 010-4818-6447; Fax: 053-850-6094; E-mail: cho6447@hanmail.net

limited reliability and validity that is related to the objective measurement of the physical activity. Some questionnaires focus on a single type of activity, such as occupational, leisure, sport, or aerobic exercise [7]. Others either provide too narrow of a description of an individual's physical activity, or they offer a continuous, total, or subscale score in which the items used are often not comparable across the measures [8, 9]. The physical activity questionnaires are not sufficient to distinguish occupational activity from leisure activity [10]. Another deficiency of the questionnaires is that they ignore the relative importance of the performed aerobic, resistance, or flexibility activity.

Considering that many countries estimate physical activity levels using questionnaires and analyze this information to identify health promotion strategies, there is a greater need than ever to assess daily physical activity more accurately and to monitor comprehensive trends in habitual physical activity to meet the recommendations of health organizations such as the World Health Organization. Since physical activity recommendations emphasize three main kinds of physical activity (aerobics, muscle strengthening, and flexibility) with regard to the type, intensity, and duration [11, 12], physical activity questionnaires should be consistent with the recommendations. Thus, physical activity questionnaires should not focus on 30-minute combination of any physical activity but rather contain items that survey each of the three types of recommendations: aerobic, strength, and flexibility activities. Furthermore, as long-term physical activity continues to accumulate [13], an item monitoring long-term physical activity is necessary. Therefore, to reflect health organizations' physical activity recommendations, the new physical activity questionnaire should be a tool to assess whether an individual is consistent with the activity recommendations by reporting the frequency, intensity, duration, and long-term aerobic, strength, and flexibility activities.

The purpose of this study was to design a physical activity behavior questionnaire that incorporates health organizations' physical activity recommendations, and to assess that questionnaire's test-retest reliability.

## 2. METHODS

### 2.1. Samples

This study utilized four samples collected at different times and locations to assess the reliability of the physical activity questionnaire. The author used a quota sampling technique to select the samples. Firstly, the author randomly selected one senior center, one residential culture center, one sports center, and one-lifetime academy in Samcheok, South Korea. Secondly, the author randomly selected 50 samples from each of the centers. The samples included Samcheok residents aged 20 and over who were participating in recreational programs provided by the community centers (senior center, residential culture center, lifetime academy, and sports center). At the centers, four research staff members conducted the initial face-to-face interview surveys of 200 samples during the survey period, which was conducted from April 4 to May 2 (6:30 a.m. to 11:30 a.m. and 2:00 p.m. to 9:00 p.m.), 2018. The retest was scheduled to be administered four weeks after the first face-to-

face interview. The retest period was June 1 to June 30, 2018. Of the 200 initial respondents, 117 completed the second survey. 44 (37.6%) were males, while 73 (62.4%) were females. Participants ranged in age from 20 to 88 years old ( $M = 51.83$ ,  $SD = 21.70$ ).

### 2.2. Physical Activity Pattern Questionnaire

The physical activity recommendation documents emphasize a combination of aerobic, resistance, and flexibility exercise and also offer evidence of connecting these physical activity recommendations to common diseases and conditions. The purpose of the developed questionnaire was to capture those recommendations. Of the study's items, 17 (5 demographic items and 12 physical activity items) of 23 were selected based on Cho's study [14] and the International Physical Activity Questionnaire (IPAQ) research committee [15].

An expert panel in the field of sports, leisure, and recreation was invited to classify the activities into three different types of physical activities. The categories were aerobic exercise and sports (walking, biking, jogging, swimming, aerobics, basketball, softball, soccer, golf, table tennis, badminton, football, *etc.*), flexibility exercises (stretching, yoga, Pilates, calisthenics, *etc.*), and muscular exercises (weight training, free weight training, *etc.*). An expert panel in this study determined items to assess aerobic, muscular, and flexibility activity behaviors. The panel also suggested that respondents choose from the given set of responses. For the assessment of aerobic activity patterns, four items (frequency, intensity, duration, and the long-term/overall length of physical activity during free time) were included on a five-point Likert-type scale. For example, questions such as, "How often do you participate in the activity in your free time?" were used to measure the frequency of physical activity, which was then categorized as "Almost every day" (5), "4-5 days/week," "3 days/week," "1-2 days/week," and "Very infrequently" (1).

Regarding the assessment of muscular activity patterns, again, four items (frequency, intensity, duration, and overall length of physical activity during leisure time) were included on a five-point Likert-type scale. For example, respondents were asked, "How intensely do you participate in the activity?" The intensity of activity participation was then categorized as "Very hard (5)," "Hard," "Moderate," "Light," and "Very light (1)."

For the assessment of flexibility activity patterns, four items (frequency, intensity, duration, and long-term/overall length of physical activity during free time) were included on a five-point Likert-type scale. For example, "How long do you do the activity in your free time?" The duration of physical activity participation was categorized as "Less than  $\geq 29$  minutes (1)," "30 - 59 minutes," "60 - 89 minutes," "90 - 119 minutes," and "More than  $\leq 120$  minutes (5)."

Because it is important to assess physical activity behavior to identify whether individuals meet physical activity recommendations, the author decided to use a summated rating scale, designed to produce scores that indicated the intensity, duration, frequency, and overall length of a person's judgments about their participation in those activities. In this scale, individuals indicate their agreement with each item.

### 2.3. Procedure and Analysis

After obtaining ethical clearance from the Daegu University Ethics Committee, a copy of the survey questionnaire was distributed to each of the 200 participants in the four different community centers in Samcheok, South Korea. The written consent forms and questionnaires were distributed to participants by the research staff. The participants were informed that participation in the study was voluntary, and they were free to withdraw from the study at any time. To prevent the possibility of introducing interviewer bias [16], the author provided basic instructions for interviewers to follow when they conducted the survey: read each question exactly, don't interpret the question, and offer to repeat the question.

During retesting, four trained survey interviewers conducted one-on-one interviews with 117 respondents. Of the 200 initial respondents, 83 did not participate in the retest because of a change in the shift pattern or scheduling a summer vacation during the survey period or being dropped out of the program before its completion. Therefore, the study sample was comprised of 117 residents who agreed to participate in the study.

The test-retest repeatability was assessed using Cohen's kappa statistics with 95% confidence intervals. According to Fleiss and Cohen, an agreement of less than 0.40 is considered as poor, between 0.41 and 0.59 as fair, between 0.60 and 0.74 as good, and of 0.75 or more as excellent [15]. Internal consistency was assessed by Cronbach's alpha coefficient. According to Nunally and Bernstein, an alpha over 0.7 indicates high reliability, over 0.5 indicates moderate reliability, and below 0.2 indicates low reliability [17, 18]. Data processing was performed using SPSS version 24. The significance level was set at  $p < 0.05$ .

### 3. RESULTS

200 subjects participated in the first survey, and 117 subjects participated in the follow-up questionnaire survey. Table 1 shows the characteristics of the study population. The respondents ranged in age from 20 to 88 years old. Forty-four respondents (37.6%) were male, while 73 (62.4%) were female. Forty-eight respondents (41.0%) lived in urban areas, and 69 (59.0%) lived in rural areas.

For the sake of clarity, the repeatability of the responses was shown according to the questions with five-point Likert-type scale answers. For the determination of the repeatability of the responses to the aerobic activity questions, Cohen's kappa statistic was used. The kappa values of the four items ranged from 0.878 to 0.852, indicating almost perfect agreement (Table 2).

For the sake of clarity, the repeatability of the responses on the frequency, duration, intensity, and overall length of muscular activity participation is presented in Table 2. The repeatability of questions pertaining to frequency, duration, and overall length of the muscular activity participation was excellent, with kappa values of 0.795 to 0.751. The repeatability of responses on the intensity of muscular activity participation was good, with a kappa value of 0.679.

The repeatability of questions pertaining to flexibility activity was excellent for three questions (frequency, duration, and intensity), with kappa values greater than 0.75. One question (overall length) showed a kappa value of 0.725, indicating substantial agreement (Table 2).

The reliability coefficient showed good to high internal consistency throughout. The alpha statistic was 0.658 for questions on flexibility activity participation, 0.666 for questions on aerobic activity participation, and 0.935 for questions on muscular activity participation.

**Table 1. Demographic characteristics of study participants.**

Variable		Frequency (n=117)	Percent (%)
Age	20 – 29	27	23.1
	30 – 39	12	10.2
	40 – 49	13	11.2
	50 – 59	19	16.1
	60 – 69	17	14.0
	70 – 79	15	13.0
	Over 80	13	11.2
Gender	Male	44	37.6
	Female	73	62.4
Community center	Senior center	41	35.0
	Residential culture	23	19.7
	Sport center	42	35.9
	Lifetime Academy	11	9.4
Living environment	Urban	48	41.0
	Rural	69	59.0
Marital status	Single	49	41.8
	Married/widowed	68	58.2

**Table 2. Repeatability of answers on aerobic, muscular strength, and flexibility activity questionnaire.**

Physical Activity Types	Items	Questions	Cohen's kappa	p-value
Aerobic	Frequency	How often do you participate in the activity?	0.856	.001
	Duration	How long do you do the activity?	0.871	.001
	Intensity	How intensely do you participate in the activity?	0.852	.001
	Long-term/Overall length	How long have you been performing the activity?	0.878	.001
Muscular strength	Frequency	How often do you participate in the activity?	0.795	.001
	Duration	How long do you do the activity?	0.785	.001
	Intensity	How intensely do you participate in the activity?	0.679	.001
	Long-term/Overall length	How long have you been performing the activity?	0.751	.001
Flexibility	Frequency	How often do you participate in the activity?	0.762	.001
	Duration	How long do you do the activity?	0.807	.001
	Intensity	How intensely do you participate in the activity?	0.783	.001
	Long-term/Overall length	How long have you been performing the activity?	0.725	.001

#### 4. DISCUSSION

The test-retest method is used to estimate the components of measurement error by repeating the measurement process on the same subjects under conditions that are as similar as possible [19]. The kappa scores were observed in items assessing the frequency, intensity, duration, and overall length of aerobic, muscular, and flexibility exercise. The author obtained good agreement for all items related to aerobic activity patterns: “How often do you participate in the activity?” “How intensely do you participate in the activity?” “How long do you do the activity?” and “How long have you been performing the activity?” The questions on the muscular strength activity pattern had an excellent agreement for three questions (frequency, duration, overall length) and a good agreement for one (intensity). The questions on the flexibility activity behavior were also repeatable. The kappa values were in the good range for one question (overall length) and in the excellent range for three questions (intensity; duration; frequency).

The alpha coefficient was used to test internal consistency. A Cronbach's alpha coefficient of  $< 0.40$  indicates that the measurement instrument is not reliable, whereas values of  $0.60$ – $0.80$  indicate substantial reliability [18]. The alpha statistic was  $0.658$  for questions on flexibility activity participation,  $0.666$  for questions on aerobic activity participation, and  $0.935$  for questions on muscular activity participation. None of the alpha statistics were classified in the poor or fair categories. Items relating to aerobic exercise, muscular activity, and flexibility exercise all proved to be reliable and satisfactory. The author obtained good or excellent agreement for 12 questions and good internal consistency of items relating to three types of physical activity questionnaire. It may be that such good results are, in part, due to the more suitable context for respondents to answer the questionnaire. Firstly, this questionnaire has a specific definition of physical activity. Although many nations are using a modified version of the IPAQ, which includes physical activity in other domains (work, transportation, and household) in which physical activity occurs [19, 20], this questionnaire focused on specific forms of free time behavior that consciously aim at improving physical activity, sport, and exercise. Some researchers revealed that the IPAQ and other subjective surveys tend to overestimate the amount of physical activity [21]. But, this specific-domain questionnaire allows the respondents to record more concisely the total amount of physical activity they

perform, with the primary purpose of maintaining or increasing health. Secondly, this questionnaire classified physical activity into three types (aerobic, muscular strength, and flexibility) and provided example activities in each type of physical activity to make it easier for respondents to answer. Also, for each type of physical activity, the questions included all three components (frequency, duration, and intensity), and items relating to sustainable long-term active behavior that can be generalized and converted to public health recommendations were also added. Considering the fact that health benefits are gained by the increased intensity, duration, and frequency of activity/exercise [22], continuing long-term physical activity appeared to play an important role in maintaining and enhancing health. Based on the findings of this study, this questionnaire was proven to be an acceptable and concise way to understand different physical activity intensity levels, to indicate all three components of physical activity, to classify physical activity into types (aerobic, muscular, flexibility), and to provide the total amount of physical activity. In conclusion, in this sample of subjects aged 20 to 88 who reside in the Samcheok community in South Korea, the author found excellent or good repeatability at a four-week interval for all questions that surveyed physical activity behaviors.

#### CONCLUSION

In summary, this paper made an important methodological contribution to the assessment of physical activity by showing good repeatability at a four-week interval for all questions that surveyed physical activity behaviors using a questionnaire based on health organizations' recommendations.

#### LIMITATIONS AND FUTURE STUDY

There are some considerations and limitations that lead us to interpret the results of this study carefully. As this questionnaire is a subjective method to assess physical activity, the results are highly dependent on the respondents' cognitive and recall biases [16]. In addition, the sample used in this study is comparatively small compared to those used in survey research. The fact that the sample was drawn from community centers is another limitation of the current study. The results are from a small sample, which cannot be generalized to the general population. The reliability of this questionnaire should be tested in large samples with different demographics. Also, this study is limited due to the lack of a parallel form reliability analysis. Therefore, the parallel form test should be used to test

the reliability of the questionnaire. Also, a future study should focus on the development of a physical activity index using this questionnaire. Future studies should also consider evaluating the reliability and validity of this study questionnaire when administered *via* different survey methods and in different populations.

## ETHICS APPROVAL AND CONSENT TO PARTICIPATE

Ethical clearance was obtained from the ethics committee of Daegu University, South Korea (Approval no. 20190549).

## HUMAN AND ANIMAL RIGHTS

No Animals were used in this research. All human research procedures followed were in accordance with the ethical standards of the committee responsible for human experimentation (institutional and national), and with the Helsinki Declaration of 1975, as revised in 2013.

## CONSENT FOR PUBLICATION

A written informed consent was obtained from all the participants.

## AVAILABILITY OF DATA AND MATERIALS

The data that support the findings of this study are available from the corresponding author [M.H.C] upon reasonable request.

## FUNDING

This research was supported by Daegu University Research Grant, 2019.

## CONFLICT OF INTEREST

The author declares no conflict of interest, financial or otherwise.

## ACKNOWLEDGEMENTS

Declared none.

## REFERENCES

- [1] Lee PH, Macfarlane DJ, Lam TH, Stewart SM. Validity of the international physical activity questionnaire short form (IPAQ-SF): A systematic review. *Int J Behav Nutr Phys Act* 2011; 8: 115. [http://dx.doi.org/10.1186/1479-5868-8-115] [PMID: 22018588]
- [2] Chinapaw MJ, Mokkink LB, van Poppel MN, van Mechelen W, Terwee CB. Physical activity questionnaires for youth: a systematic review of measurement properties. *Sports Med* 2010; 40(7): 539-63. [http://dx.doi.org/10.2165/11530770-000000000-00000]
- [3] Doma K, Speyer R, Leicht AS, Cordier R. Comparison of psychometric properties between usual-week and past-week self-reported physical activity questionnaires: a systematic review. *Int J Behav Nutr Phys Act* 2017; 14(1): 10. [http://dx.doi.org/10.1186/s12966-017-0470-6] [PMID: 28137268]
- [4] Helmerhorst HJ, Brage S, Warren J, Besson H, Ekelund U. A systematic review of reliability and objective criterion-related validity of physical activity questionnaires. *Int J Behav Nutr Phys Act* 2012; 9(1): 103. [http://dx.doi.org/10.1186/1479-5868-9-103] [PMID: 22938557]
- [5] Oliver M, Badland H, Mavoa S, Duncan MJ, Duncan S. Combining GPS, GIS, and accelerometry: methodological issues in the assessment of location and intensity of travel behaviors. *J Phys Act Health* 2010; 7(1): 102-8. [http://dx.doi.org/10.1123/jpah.7.1.102] [PMID: 20231761]
- [6] Pereira MA, FitzGerald SJ, Gregg EW, *et al.* A collection of Physical Activity Questionnaires for health-related research. *Med Sci Sports Exerc* 1997; 29(6): S1-S205. [PMID: 9243481]
- [7] Montoye HJ. Estimation of habitual physical activity by questionnaire and interview. *Am J Clin Nutr* 1971; 24(9): 1113-8. [http://dx.doi.org/10.1093/ajcn/24.9.1113] [PMID: 5094483]
- [8] Booth M. Assessment of physical activity: an international perspective. *Res Q Exerc Sport* 2000; 71(2)(Suppl. 2): 114-20. [http://dx.doi.org/10.1080/02701367.2000.11082794]
- [9] Pols MA, Peeters PH, Ocké MC, Slimani N, Bueno-de-Mesquita HB, Collette HJ. Estimation of reproducibility and relative validity of the questions included in the EPIC Physical Activity Questionnaire. *Int J Epidemiol* 1997; 26(1)(Suppl. 1): S181-9. [http://dx.doi.org/10.1093/ije/26.suppl\_1.S181] [PMID: 9126546]
- [10] Shephard RJ. Limits to the measurement of habitual physical activity by questionnaires. *Br J Sports Med* 2003; 37(3): 197-206. [http://dx.doi.org/10.1136/bjism.37.3.197] [PMID: 12782543]
- [11] Warburton DE, Nicol CW, Bredin SSD. Health benefits of physical activity: the evidence. *CMAJ* 2006; 174(6): 801-9. [http://dx.doi.org/10.1503/cmaj.051351] [PMID: 16534088]
- [12] U.S. Department of Health and Human Services. Physical Activity Guidelines for Americans. Washington, DC: US Department of Health and Human Services 2018. [https://health.gov/sites/default/files/2019-09/Physical\\_Activity\\_Guidelines\\_2nd\\_edition.pdf](https://health.gov/sites/default/files/2019-09/Physical_Activity_Guidelines_2nd_edition.pdf)
- [13] Hallal PC, Martins RC, Ramirez A. The Lancet Physical Activity Observatory: promoting physical activity worldwide. *Lancet* 2014; 384(9942): 471-2. [http://dx.doi.org/10.1016/S0140-6736(14)61321-0] [PMID: 25110267]
- [14] Cho MH. Physical activity levels of Korean elderly participants in the community recreational center programs. *CEPAL Rev* 2017; 121: 2884-90.
- [15] IPAQ Research Committee. Guidelines for data processing and analysis of the International Physical Activity Questionnaire (IPAQ)—short and long forms 2005 <https://sites.google.com/site/theipaq/scoring-protocol> [Accessed June 13 2018];
- [16] West BT, Kreuter F, Jaenichen U. Interviewer effects in face-to-face surveys: a function of sampling, measurement error or nonresponse? *J Off Stat* 2013; 29: 277-97. [http://dx.doi.org/10.2478/jos-2013-0023]
- [17] Fleiss L, Cohen J. The equivalence of weighted kappa and the intraclass correlation coefficient as measures of reliability. *Educ Psychol Meas* 1973; 33(3): 613-9. [http://dx.doi.org/10.1177/001316447303300309]
- [18] Nunnally JC, Bernstein IR. *Psychometric Theory*. New York: McGraw-Hill 1994.
- [19] Evenson KR, McGinn AP. Test-retest reliability of a questionnaire to assess physical environmental factors pertaining to physical activity. *Int J Behav Nutr Phys Act* 2005; 2: 7-10. [http://dx.doi.org/10.1186/1479-5868-2-7] [PMID: 15958168]
- [20] Finger JD, Tafforeau J, Gisle L, *et al.* Development of the European Health Interview Survey - Physical Activity Questionnaire (EHIS-PAQ) to monitor physical activity in the European Union. *Arch Public Health* 2015; 73: 59. [http://dx.doi.org/10.1186/s13690-015-0110-z] [PMID: 26634120]
- [21] Finger JD, Gisle L, Mimiilidis H, *et al.* How well do physical activity questions perform? A European cognitive testing study. *Arch Public Health* 2015; 73: 57. [http://dx.doi.org/10.1186/s13690-015-0109-5] [PMID: 26629340]
- [22] Katzmarzyk PT, Tremblay MS. Limitations of Canada's physical activity data: implications for monitoring trends. *Can J Public Health* 2007; 98(Suppl. 2): S185-94. [PMID: 18213948]