AUC Kinanthropologica, Vol. 60, No. 2, pp. 93–111

Test-retest reliability of self-report instruments for measuring leisure activities in adults

Melisa Schneiderová[1,2,*], Josef Mana[1], Jiří Lukavský[1], Radek Trnka[1], Klára Daďová[1,2], Hana Georgi[1]

[1] Prague College of Psychosocial Studies, Prague, Czech Republic
[2] Charles University, Prague, Czech Republic
* Corresponding author: melissaschneiderova@seznam.cz

ABSTRACT

Leisure activities contribute to the preservation of physical and mental capabilities in aging. Measuring leisure pursuits with psychometrically sound instruments is essential to understand the specific relationships between different activities and health outcomes. This study aimed to examine the test-retest reliability of self-report instruments for measuring leisure activities in adults. The sample consisted of 86 healthy adult participants. Participants completed two assessments, three weeks apart, using the following instruments: 1) a single-question Attitude to Physical Activity scale (A2PA); 2) a modification of a COBRA questionnaire on recent engagement in physical and mental activities (SA-COBRA Cognitive and Physical); 3) scales measuring engagement in social/spiritual activities over the past two years from the standard Victoria Longitudinal Study – Activity Lifestyle Questionnaire (VLS-ALQ-S+). All instruments exhibited acceptable test-retest reliability, demonstrating their ability to reflect consistent patterns of lifestyle over time. This finding supports their suitability for assessing recent engagement in leisure activities.

KEYWORDS

physical activities; cognitive activities; social activities; questionnaire; public health

DOI

10.14712/23366052.2025.3

INTRODUCTION

Leisure activities (LA) are defined as enjoyable pursuits undertaken during free time (Kleiber & Nimrod, 2009). A common categorization of LA is physical, cognitive, and social. Participation in LA has been consistently associated with a reduced risk of developing dementia or cognitive decline (Najar et al., 2019) and a higher perception of subjective happiness and life satisfaction (An et al., 2020). In particular, LA, along with other factors, such as educational attainment, and occupational demands, are believed to contribute to cognitive reserve (Alvares Pereira et al., 2022; Stern et al., 2020). It is a concept popularized by Yaakov Stern (2002) that describes the capability of the brain to cope with the consequences of changes or damage to the brain through pre-existing cognitive processes, which is crucial for healthy aging and may be a reason why some people keep excellent cognitive functions into old age. Cognitive reserve is cumulatively built over the course of life, with research continuously refining the factors that influence its development. In particular, engagement in physical activities has attracted a lot of attention as there is a growing evidence of their association with health including cognitive health in aging (Livingston et al., 2020). To reflect scientific evidence, the World Health Organization (WHO) issued physical activity recommendation for adults, which is to engage in moderate-intensity physical activity for at least 150 minutes (i.e. 2,5 hours) per week. Other guidelines add a recommendation of frequency of physical activity: at least three times a week (Izquierdo et al., 2021; U.S. Department of Health and Human Services, 2018).

Effectively investigating the impact of LA requires their measurement. To identify associations between LA and developmental changes, including aging, it is necessary to measure both retrospective and current engagement. Methods for measuring participation vary according to study objectives and time frames. The type and number of activities, frequency of participation, intensity, and duration are the most frequent measures used to assess engagement (Fallahpour et al., 2016). A brief overview of the most commonly used questionnaires focused on the main domains of LA is presented in Table 1 and elaborated further.

Few questionnaires document lifetime activities. Data on an individual's past are mostly acquired through self-reporting, as longitudinal studies are rare. The Historical Adulthood Physical Activity Questionnaire (HAPAQ) evaluates lifelong physical activity (Besson et al., 2010). The questionnaire is divided into two parts; the first examines activity over the last 15 years, and the second focuses on 10-year intervals from age 20 to the last 15 years. Each part includes closed questions on physical activity in the household, work, transportation, sports, and exercise. The HAPAQ shows acceptable validity (Besson et al., 2010), however, no study to date has investigated test-retest reliability. Another example is the Lifetime Physical Activity Questionnaire (LTPAQ) (Friedenreich et al., 1998), which estimates physical activity from childhood to the present. It focuses on frequency, duration, and intensity across occupational and volunteer activities, household tasks, and exercise/sports. Participants recall their first job (at least 8 hours/week for four months) and subsequent jobs, household tasks (at least 7 hours/week for four months), and exercise/sports activities (at least 2 hours/week for four months). The LTPAQ shows a high test-retest correlation after six to eight weeks (Friedenreich et al., 1998). However, most questionnaires evaluate

Table 1 Examples of leisure activity questionnaires

Source	Questionnaire	Activity type	Time frame	Test-retest reliability
Flora et al. (2023)	IPAQ	Physical	Over the past seven days	r = 0.71
Milton et al. (2010)	A single-item PA measure	Physical	Over the past week or past month	r = 0.72–0.82
Nevalainen et al. (2015)	COBRA	Physical, cognitive, social	Over the summer week	NA
Galvin et al. (2021)	CLAS	Physical, cognitive, social	Over the past year	NA
Jopp & Hertzog (2010)	VLS-ALQ	Physical, social, developmental, experiential activities, crafts, game playing, TV watching, travel, and technology use	Over the past two years	r = 0.65–0.70
Besson et al. (2010)	HAPAQ	Physical	Over the lifetime since age 20	NA
Friedenreich et al. (1998)	LTPAQ	Physical	Over the lifetime	r = 0.72–0.87

Note: HAPAQ = Historical Adulthood Physical Activity Questionnaire; LTPAQ = Lifetime Physical Activity Questionnaire; IPAQ = International Physical Activity Questionnaire; VLS-ALQ = Victoria Longitudinal Study – Lifestyle Activities Questionnaire; CLAS = Cognitive & Leisure Activity Scale; NA = Not available

current or recent participation in leisure activities over periods ranging from a few days to a year. The International Physical Activity Questionnaire (IPAQ) short form, for instance, focuses on the intensity and duration of physical activity over the past seven days (Craig et al., 2003). Single-item physical activity scales are also relevant; for example, Milton et al. (2011) developed one for screening physical activity over the past seven days or month: "In the past week/past month, on how many days have you done a total of 30 minutes or more of physical activity, which was enough to raise your breathing rate. This may include sport, exercise, and brisk walking or cycling for recreation or to get to and from places, but should not include housework or physical activity that may be part of your job."

Beyond physical activity, the Victoria Longitudinal Study (VLS) questionnaire by Hultsch et al. (1993) assesses various leisure activities, including physical, social, and self-maintenance. The original questionnaire included 70 activities, with participants rating their frequency on a Likert scale from (0) never to (9) daily. It has been modified for research purposes, such as in Jopp and Hertzog's (2010) version, which restructured 57 items into categories like physical, craft, gaming, television, social and private, social and public, religious, technology use, developmental, experiential, and travel, forming the Activity Lifestyle Questionnaire (VLS-ALQ). The VLS-ALQ demonstrated good test-retest reliability over 16 months, except for the 'Travel' scale, which had a lower correlation (r = 0.41) (Jopp & Hertzog, 2010). More recently, Galvin et al. (2021) developed the Cognitive & Leisure Activity Scale (CLAS), which focuses on participation in cognitive activities over the previous year. The scale consists of 16 items, for example, 'Playing cards or Board Games', 'Socializing with friends',

and so on. Participants are asked to indicate how often they engage in each activity on a scale ranging from 0 (never) to 5 (daily). So far, no studies have looked at test-retest reliability for the CLAS.

Measuring engagement in leisure activities in the Czech Republic

Several instruments are available to measure leisure activities among the adult population in the Czech Republic. Studies focusing on physical activity often utilize the IPAQ (e.g., Lojdová et al., 2021; Mitáš et al., 2014; Vašíčková et al., 2012). Broader studies on leisure activities typically employ custom questionnaires tailored to specific needs. For example, Frantál et al. (2020) developed a questionnaire for older adults that included a section assessing the frequency and duration of recent leisure activities during a typical day, both inside and outside the home, including part-time work. The Cognitive SuperAging Study (Heissler et al., 2021) used a questionnaire based on a Swedish study Cognition, Brain and Aging (COBRA), which included three sections – section A) consisted of 18 items related to cognitive activities (e.g., reading books), section B) included 19 items assessing physical activities (e.g., walking), and section C) included 10 items related to social activities (e.g., time spent with family members). In each section, respondents selected activities they engaged in during a typical summer week, indicating the number of days and hours spent on each (Nevalainen et al., 2015). Further, respondents assessed whether they performed the activity less, the same, or more than six years ago and rated the difficulty of each activity on a scale from (0) not at all difficult to (5) extremely difficult. The questionnaire was designed for the Swedish population thus some activities may not well apply to the Czech population (e.g., sailing).

To our knowledge, no standardized Czech leisure activity assessment tool has been developed to comprehensively evaluate adult's leisure participation across physical, cognitive, and social domains over an extended time frame. To address this gap, we adapted a set of instruments designed to measure long-term engagement in leisure activities encompassing these domains.

The instruments included: (1) the single-item Attitude to Physical Activity (A2PA) to evaluate attitudes towards physical activity throughout life, (2) the SA-COBRA-Cognitive and-Physical to assess long-term engagement in the respective leisure activities, and (3) VLS-ALQ-S+ to assess participation in social-public, social-private, and religious activities in last two years. The aim of this study is to examine the test-retest reliability of those self-report instruments.

METHODS

Participants and procedure

The participants were healthy adults with Czech as a native language, they were recruited and assessed by the psychology students who underwent a training in the methods administration.

The sample consisted of 86 adult participants (40 males, 41 females, and 5 not stated) with the mean age of 55.6 ± 10.6 ranging from 40 years to 83 years. Higher education (college or secondary school) was represented by 79 (91.8%) participants, while lower education (primary or vocational school) was represented by 7 (8.2%) participants.

The assessments were carried out in two waves during the fall of 2022, three weeks apart, in the form of structured interview. The assessment procedure included inquiry on basic demographic data and the questionnaires regarding leisure activities. All participants were fully informed prior to participation and provided an informed consent in accordance with the Declaration of Helsinki. The study and the methods were approved by the Prague College of Psychosocial Studies (PVSPS) Institutional Review Board (reg. No. 3/2022 and 3/2021).

Instruments

Attitude to Physical Activity scale (*A2PA*) is a single-question screening scale estimating attitude to and participation in physical activity over the life course: "What has been your relationship to sport and physical activity during adulthood, from the age of 20 until now?" The scale was designed for the COSACTIW study of life style and cognition in older age. Six categorical options were developed based on a discussion among the research team to assess positive implicit attitudes to physical activity throughout life course and whether the WHO (2020) criteria had been met. The A2PA was administered as a self-report questionnaire (see Appendix A).

SA-COBRA is a modification of a comprehensive questionnaire originally developed for the COBRA study (Nevalainen et al., 2015). The questionnaire was translated into Czech in collaboration with its author Nina Nevalainen through translation and back-translation process. It was designed to assess regular physical and cognitive activity. It was administered as a structured interview, but it can be administered also as a self-rating questionnaire.

- SA-COBRA-Cognitive scale was narrowed to eight cognitive activities most typical for Czech SA (Heissler et al., 2021). It includes questions on pre-set types of mental activities most common among SuperAgers in healthy Czech population, for example, "Using a computer for purposes other than gaming (i.e., including emailing, photo editing, Skype communication, etc.)" (Appendix B).
- SA-COBRA-Physical scale was modified not to include pre-set activities. Instead, respondents were asked to recall and name physical activities of vigorous intensity, moderate intensity and/or light intensity that they have typically performed within a typical summer week (Appendix C). The modification reflects the criteria of the World Health Organization for physical activity in adults (WHO, 2020).

Furthermore, in both SA-COBRA-Cognitive and – Physical, respondents are asked to specify the number of days and hours they dedicated to each activity per week. Additionally, they are asked to compare whether they performed the activity less, about the same, or more than six years ago and rate the subjective difficulty of performing the activity on a scale ranging from 1 (not at all difficult) to 5 (extremely difficult).

VLS-ALQ-S+ is an abridged version of an instrument assessing leisure activities VLS-ALQ developed by Jopp and Hertzog (2010). To assess social participation specifically, we used two scales: the 6-item social-public scale (e.g., "I invite friends to my house for dinner/lunch") and the 5-item social-private scale (e.g., "I volunteer"). Additionally, we included the item "I attend church services" from the religious scale as a measure of participation in religious activities. The questionnaire VLS-ALQ-S+ comprises 12 items, and the participants were asked to rate the frequency of their

engagement in each activity on a scale of 1 (never) to 9 (daily). The research team employed a collaborative translation process, followed by a meeting to discuss and reach consensus on a final translation.

The entire battery was pilot-tested with several functionally independent, cognitively healthy older adults to confirm clarity and comprehensibility before being used in this study.

Data analysis

A2PA

We presented responses to the A2PA scale using a contingency table, with wave 1 responses in rows and wave 2 in columns. To test for a presence of association between response frequencies across waves, we computed Pearson's χ^2 test of the null hypothesis that the joint distribution of the cell counts is the product of the row and column marginals. We estimated the test-retest reliability of A2PA via quadratically weighted Cohen's κ coefficient with values above.61 indicating substantial strength of agreement between waves (Landis & Koch, 1977).

SA-COBRA-Cognitive

First, we report the activities and means/medians of the corresponding responses in wave 1 to illustrate the nature of the responses that participants provided. To assess the test-retest reliability of the responses, we compared the data on total time using Pearson correlation coefficients. For frequency, subjective difficulty and subjective historical comparison, we report how often participants repeated their response in wave 2 exactly, and how often with minor deviations (+/–1).

SA-COBRA-Physical

Because SA-COBRA-Physical features open responses, we focused on the question of whether the total amount of self-reported activities could be considered sufficient according to the WHO guidelines ("at least 150–300 minutes of moderate-intensity aerobic physical activity; or at least 75–150 minutes of vigorous-intensity aerobic physical activity; or an equivalent combination of moderate- and vigorous-intensity activity throughout the week, for substantial health benefits" (WHO, 2020, p. 32). Additionally, we analysed to what extent this self-reported adherence was reliable across Wave 1 and 2. We calculated the numbers of minutes spent in moderate-intensity and vigorous-intensity activities. Then, we compared these values with recommended thresholds (separately for lower and upper range values).

VLS-ALQ-S+

We described each VLS-ALQ-S+ scale and VLS-ALQ-S+ item scores by their means, medians and standard deviations separately for wave 1 and wave 2. To assess the test-retest reliability of scales and item responses we calculated Pearson and Spearman correlation coefficients of the wave 1 and wave 2 data. Values above 0.7 were considered acceptable while values above 0.8 were considered to represent good test-retest reliability. Furthermore, to test for systematic changes in data distributions between waves that can be missed via correlation analysis only, we compared means and signed ranks

of wave 1 and wave 2 data via paired t-test and Wilcoxon signed rank test respectively. Scales/item scores with p values less than 0.05 were considered statistically clearly different between waves. Finally, we assessed factor structure and internal consistency of VLS-ALQ-S+ total score as well as internal consistency of VLS-ALQ-S+ private and public scales separately in wave 1 and wave 2 via Confirmatory Factor Analysis (CFA) and Cronbach's α respectively. Two CFAs were fitted, one to data from each wave, consisting of three separate correlated factors for private, public and religious items. The single-item religious scale was modelled as a single-indicator latent variable with zero observed variable variance. Models' fit was evaluated via the χ^2 test, the Tucker Lewis Index (TLI), the Comparative Fit Index (CFI) and the root-mean-square-error-approximation (RMSEA) with values of TLI > 0.9, CFI > 0.9, and RMSEA < 0.08 considered to indicate adequate fit. Values of Cronbach's α above 0.7 were considered acceptable while values above 0.8 were considered to represent good internal consistency. All statistical analyses were performed using R (Version 4.3.3, R Core Team, Vienna, Austria).

RESULTS

A2PA

Response counts relating wave 1 responses (rows) to wave 2 responses (columns), are presented in Table 2. We observed a strong statistically clear association of response type counts ($\chi^2(25) = 93.171$, $p < 0.001$) with substantial agreement between waves ($\kappa = 0.706$, 95% CI [0.567, 0.846]).

Table 2 Response counts to the A2PA screening scale

	I'm an athletic person	I enjoy movement/ exercise	I exercised at least 3 times a week	I don't avoid movement/ exercise	I'm not an athletic person	I had to stop doing sports (Injury)
I'm an athletic person	6	1	0	0	1	0
I enjoy movement/ exercise	4	11	4	1	0	0
I exercised at least 3 times a week	1	4	12	2	1	0
I don't avoid movement/exercise	0	1	4	12	2	2
I'm not an athletic person	0	0	1	4	4	0
I had to stop doing sports (Injury)	0	0	0	1	0	0

SA-COBRA-Cognitive

The most commonly reported activity was 'Using a computer' (see Table 3). 'Using a computer', 'Reading books', 'Playing cards/games', and 'Art activities' were reported by at least half of the participants – the remaining activities were pursued only by a small part of the sample. On average, participants reported a similar level of engagement in activities as they did six years ago (48.8%). There was no apparent trend toward a decrease over time with people reporting "more" and "less" similarly often (24.7% vs 26.1%). Participants engaged in subjectively easy activities (rank 1 in 63.4%, 2 in 19.0%, 3 in 13.2%, 4 in 2.4%).

The reported hours per week was highly correlated across the Waves 1 and 2 ($r = 0.785$). The results are similar if we drop all zero responses ($r = 0.749$).

When reporting frequency per week, participants often used extreme responses (0×: 57.1%, 7×: 9.0% in Wave 1; 0×: 60.9%, 7×: 8.4% in Wave 2). Participants were considerably consistent in their use of extreme values – 90.8% of people who reported a zero in Wave 1 for a given activity, reported zero also in Wave 2. Similarly, 64.5% of people reporting the maximum frequency (7×) repeated their response in Wave 2. Altogether, 70.1% of responses were repeated exactly in Wave 2, 84.4% of responses differed only in +/−1.

When comparing their activities with the time six years ago, 68.3% used the same response with no apparent trend in under- or over-estimating (10.5% increased their response in Wave 2, 10.5% decreased the response, 10.8% missing response in Wave 2). People reported identical difficulty judgments in 54.9% of responses, in 72.5% the responses differed only in +/−1 (in 21.7% of cases, people did not report difficulty in Wave 2).

Table 3 Reported activities in Wave 1 sorted by the mean times per week frequency. The means and medians are calculated only after excluding the zero responses.

Activity	Reports		Times per week		Hours per week	
	Never	At least once	Median	Mean	Median	Mean
Using the computer not for games	20	66	6	5.2	9.0	9.0
Reading books	23	63	3	3.3	3.0	5.1
Playing a musical instrument or singing	70	16	3	3.1	2.5	3.5
Crosswords	55	31	2	2.8	2.0	3.1
Playing cards or board games	43	43	2	2.8	4.0	4.7
Art activity and handcraft	42	44	2	2.7	3.0	4.8
Riddles sudoku	65	21	2	2.4	1.0	2.4
Puzzle	75	11	1	1.9	2.0	3.2

SA-COBRA-Physical

The vast majority of the sample passed the WHO guidelines for the amount of physical activities (in Wave 1: 91.9% met the lower criteria threshold and 87.2% met the higher threshold). The results were similar in Wave 2 (84.9% and 76.7%, respectively). In 82.5%, the comparison with lower criteria threshold led to the same evaluation in both waves (i.e. passed in both, failed in both). Comparison with the higher threshold yielded the same evaluation in 76.7% cases.

VLS-ALQ-S+

Sample description, as well as difference statistics, are presented in Table 4. We did not observe any statistically significant differences between waves' means or signed ranks. The CFA model of the first wave data showed good fit to the data (χ^2 (52) = 63.138, p = 0.139, TLI = 0.919, CFI = 0.936, RMSEA = 0.050, 90% CI [0.000, 0.089]). On the other hand, the CFA model of the second wave did not reach level of adequate model fit (χ^2 (52) = 82.286, p = 0.005, TLI = 0.754, CFI = 0.806, RMSEA = 0.082, 90% CI [0.046, 0.115]). Across waves, the Social – Total scale had acceptable internal consistency (= 0.729, 95% CI [0.636, 0.806]; = 0.706, 95% CI [0.606, 0.790]) while the Social – Private (= 0.507, 95% CI [0.326, 0.652]; = 0.510, 95% CI [0.331, 0.655]) and Social – Public (= 0.689, 95% CI [0.571, 0.782]; = 0.683, 95% CI [0.563, 0.778]) scales had poor internal consistency.

The test-retest reliability estimates are presented in Figure 1. The overall Social – Total score scale had good test-retest reliability while test-retest reliability of Social – Private and Social – Public scales was acceptable. Regarding items scores, four items ("Volunteer", "Visit friends or relatives", "Talk to friend on phone", and "Attend organized social events") had poor test-retest reliability, five items ("Give dinner for friends", "Attend parties", "Eat out at restaurant", "Engage in political activities", and "Attend club meetings") had acceptable test-retest reliability and three items ("Attend church service", "Give public talk", and "Go out with friends") had good test-retest reliability.

Table 4 Descriptive and difference statistics of VLS-ALQ-S+ scales and item scores

Scale/Item	N	M ± SD	Md (IQR)	t-test	Wilcoxon test
Social – Total	79 / 82	45.53 ± 9.79 / 45.43 ± 9.61	42 (13) / 44 (13)	t(75) = –0.632, p = 0.529	V = 921.0, p = 0.863
Social – Private	82 / 83	33.20 ± 5.10 / 33.07 ± 5.40	34 (6) / 33 (7)	t(78) = –0.802, p = 0.425	V = 894.5, p = 0.716
Social – Public	84 / 85	10.80 ± 5.64 / 10.67 ± 5.76	9 (8) / 9 (8)	t(82) = –0.693, p = 0.490	V = 794.0, p = 0.798
Go out with friends	86 / 86	5.76 ± 1.56 / 5.62 ± 1.60	6 (2) / 5 (1)	t(85) = –1.536, p = 0.128	V = 219.0, p = 0.152
Visit friends or relatives	86 / 86	5.86 ± 1.64 / 5.84 ± 1.71	6 (2) / 6 (2)	t(85) = –0.167, p = 0.868	V = 470.0, p = 0.766

Scale/Item	N	M ± SD	Md (IQR)	t-test	Wilcoxon test
Attend parties (e.g., birthday)	86 / 86	4.49 ± 1.27 / 4.50 ± 1.43	4 (1) / 4 (1)	t(85) = 0.116, p = 0.908	V = 348.0, p = 0.810
Talk to friend on phone	84 / 85	7.99 ± 1.47 / 8.00 ± 1.48	8 (1) / 8 (1)	t(82) = 0.000, p = 1.000	V = 385.0, p = 0.607
Give dinner for friends	85 / 86	4.07 ± 2.00 / 3.99 ± 1.73	4 (3) / 4 (2)	t(84) = –0.323, p = 0.748	V = 330.0, p = 0.549
Eat out at restaurant	85 / 84	5.34 ± 1.69 / 5.21 ± 1.94	5 (3) / 5 (3)	t(82) = –0.560, p = 0.577	V = 324.0, p = 0.891
Engage in political activities	86 / 86	1.97 ± 1.54 / 1.83 ± 1.48	1 (2) / 1 (1)	t(85) = –1.228, p = 0.223	V = 105.5, p = 0.199
Give public talk	86 / 85	2.15 ± 2.12 / 2.21 ± 2.23	1 (2) / 1 (2)	t(84) = 0.472, p = 0.638	V = 68.5, p = 0.645
Attend club meetings	84 / 86	2.25 ± 1.88 / 2.26 ± 1.90	1 (3) / 1 (3)	t(83) = –0.453, p = 0.652	V = 111.0, p = 0.887
Attend organized social events	86 / 86	2.81 ± 1.52 / 2.63 ± 1.54	3 (3) / 2 (3)	t(85) = –1.085, p = 0.281	V = 456.5, p = 0.249
Volunteer	86 / 86	1.87 ± 1.62 / 1.90 ± 1.61	1 (1) / 1 (1)	t(85) = 0.168, p = 0.867	V = 125.0, p = 0.751
Attend church service	85 / 86	1.89 ± 1.82 / 1.94 ± 1.90	1 (1) / 1 (1)	t(84) = 0.000, p = 1.000	V = 36.5, p = 0.784

Note: M = Mean; SD = Standard Deviation; Md = Median; IQR = Interquartile range

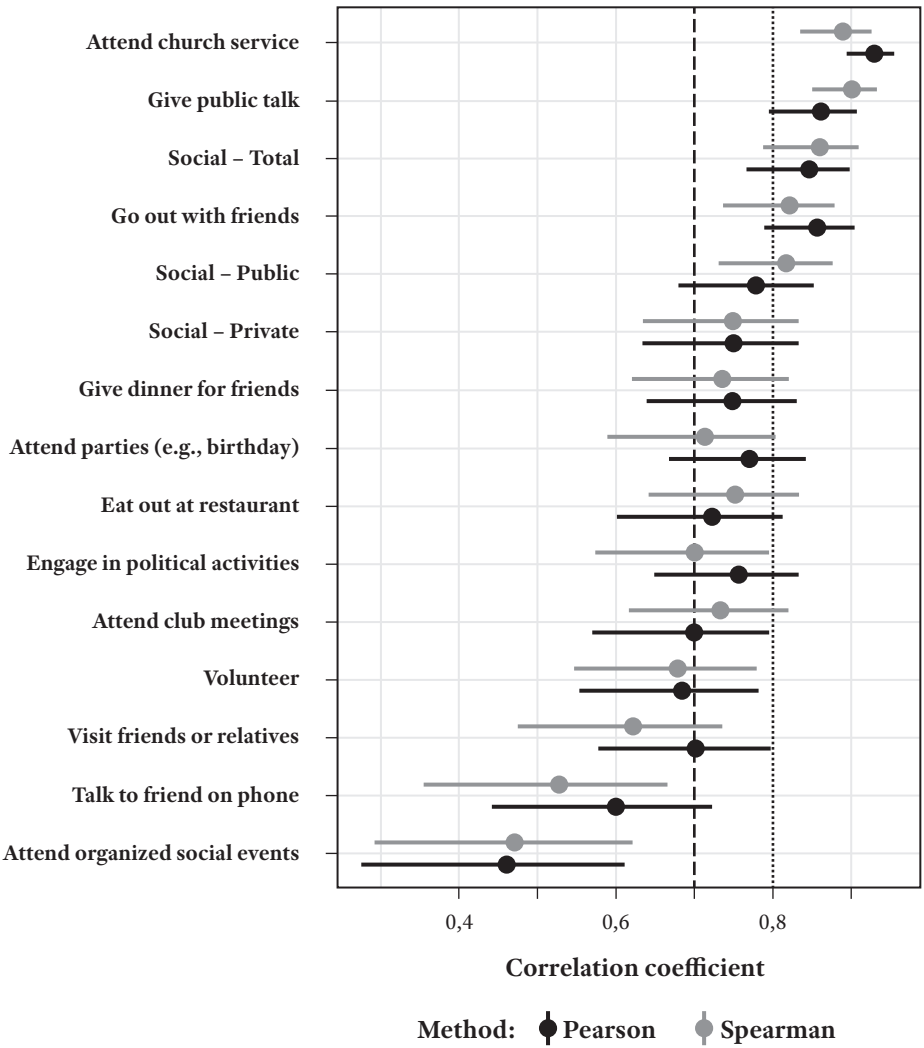

Figure 1 Test-retest reliability of VLS-ALQ-S+ scales and item scores

Note: Correlations above 0.7 were considered acceptable (dashed line).

DISCUSSION

This study aimed to evaluate the temporal stability, or test-retest reliability, of the leisure activity self-report instruments. Test-retest reliability is a critical psychometric aspect of measurement instruments, as it assesses their ability to provide consistent and accurate results over time. The findings revealed that the instruments showed acceptable test-retest reliability, suggesting their suitability for assessing individuals' leisure activities.

A2PA

A single-item screening scale exhibited a robust statistically significant association between response type counts across a 3-week interval. These findings indicate that the instrument consistently produces similar responses over time. Thus, we conclude that A2PA is reliable and can be used to assess one's attitude towards physical activity. Several studies have examined the reliability of single-item instruments for assessing physical activity, reporting high repeatability ICC = 0.75 (Scott et al., 2015) and strong correlations ranging from 0.82 to 0.88 (Li et al., 2000; Milton et al., 2011). Our results align with the established notion of test-retest reliability and compare favourably to previous findings.

SA-COBRA-Cognitive + Physical

The SA-COBRA is a modified questionnaire from the COBRA study (Nevalainen et al., 2015), and this study is the first to examine its reliability.

Regarding part Cognitive of the SA-COBRA questionnaire, the findings revealed no significant discrepancies between Wave 1 and Wave 2 responses. Participants consistently reported similar levels of cognitive activity in both waves. Moreover, their responses regarding their involvement in activities six years ago also remained stable, without any notable increases or decreases. Participants consistently reported similar frequencies of weekly activity participation throughout both waves. Furthermore, the reported hours of weekly activity involvement exhibited a high correlation across waves ($r = 0.785$). These findings indicate that the SA-COBRA questionnaire demonstrates consistency and reliability over time.

Regarding part Physical, since the responses were open, we focused on determining whether the total amount of self-reported activities met the criteria outlined in the WHO guidelines. The majority of our sample adhered to the WHO guidelines regarding the level of physical activity. In the first wave, 91.9% met the criteria for the lower threshold ("at least 150–300 minutes of moderate-intensity aerobic physical activity", and 87.2% met the criteria for the higher threshold ("at least 75–150 minutes of vigorous-intensity aerobic physical activity", (WHO, 2020). The findings were similar in the second wave, with 84.9% for lower and 76.7% for higher threshold. Consistency between waves was found in 82.5% of cases when compared to the lower threshold, and in 76.7% of cases when compared to the higher threshold. These results suggest that SA-COBRA-Physical is capable of accurately assessing adherence to the WHO physical activity guidelines.

VLS-ALQ-S+

The test-retest reliability analyses after a three-week interval revealed good overall reliability, with a correlation coefficient of 0.85 for the total scores. Individual scales also exhibited acceptable reliability, with the social-public scale demonstrating a correlation of 0.78 and the social-private scale of 0.75. These findings align with previous research by Jopp and Hertzog (2010), who reported similar values of 0.78 and 0.70 for the social-public and social-private scales, respectively.

Interestingly, the CFA revealed a good model fit for Wave 1 but a weaker fit for Wave 2. This disparity may be attributed to several factors, including potential measurement error. Despite this, the model still provides an adequate representation of

the data. Furthermore, our study found notably lower internal consistency values for the social-private scale, with a Cronbach's alpha of 0.51 in both the first and second waves. This contrasts with the higher consistency observed by Jopp and Hertzog, who reported alphas of 0.75 and 0.78 during the respective waves. Similarly, our social-public scale values were lower in the first wave ($\alpha = 0.69$) and higher in the second wave ($\alpha = 0.68$), deviating from Jopp and Hertzog's results ($\alpha = 0.75$ in the first wave; $\alpha = 0.61$ in the second wave).

Jopp and Hertzog employed a sample of 267 participants with a mean age of 49.58 ± 17.32 years, while our study involved 86 participants with a mean age of 55.6 ± 10.6 years. This notable discrepancy in sample size may have contributed to the weaker model fit in wave 2 and lower internal consistency in our study. Nevertheless, the VLS-ALQ-S+ instrument demonstrates acceptable reliability overall.

The instruments appear to be appropriate for assessing engagement in leisure activities over extended time frame. However, it is important to note that our testing primarily focused on reliability aspects, particularly through test-retest analysis. While test-retest reliability is considered a robust indicator of instrument stability, it is challenging to eliminate all sources of measurement error (Polit, 2014). For instance, despite employing a three-week interval between assessments, participants may exhibit response shifts due to reassessment of their internal beliefs (Sprangers & Schwartz, 1999) or experience memory inaccuracies (Frank et al., 2023). This could be particularly relevant for the A2PA scale, which assesses a person's general attitude towards physical activity. However, one-item scales, despite their simplicity, can provide valuable overall assessments of various aspects of one's life (Bowling, 2005).

Limitations and future research directions

This study acknowledges several limitations, the most significant of which is relatively small sample size. This limitation may have altered the accuracy of reliability assessments, particularly for the VLS-ALQ-S+ social-private scale, where weaker model fit in wave 2 and low internal consistency was observed. Furthermore, individuals with lower education levels were under-represented in our sample. Together with higher age, we could not claim the sample reflects the population structure, but it well represents the samples often used in aging studies. Notably, the higher homogeneity of the sample and consequent smaller range of responses could limit the observed reliability of the methods. Future research should address these limitations by using a larger sample size with a more diverse population. This would allow for more generalizable conclusions to be drawn and more accurate assessments of reliability to be made. In addition, future research would benefit from including more measures of reliability and validity. These measures would ensure that the instruments are assessed comprehensively.

CONCLUSION

This study assessed the test-retest reliability of the leisure activity instruments for adults (A2PA, SA-COBRA Cognitive + Physical, VLS-ALQ-S+). The results indicated adequate test-retest reliability for the instruments, suggesting their suitability for assessing leisure activity engagement. However, to comprehensively investigate

the impact of leisure activities on public health, it is crucial to include retrospective measures of engaging in leisure activities over life-course. Objective long-term data on lifestyle parameters are currently unavailable. Therefore, retrospective subjective methods remain essential for advancing knowledge in this area. Drawing generalizable conclusions from these methods requires understanding their parameters, such as retest reliability.

ACKNOWLEDGEMENTS

The study is a part of the Cognitive SuperAging in Physically Active Women (COSACTIW) project GA22-24846S funded by Czech Science Agency.

REFERENCES

Alvares Pereira, G., Silva Nunes, M. V., Alzola, P., & Contador, I. (2022). Cognitive reserve and brain maintenance in aging and dementia: An integrative review. *Applied Neuropsychology: Adult*, *29*(6), 1615–1625. https://doi.org/10.1080/23279095.2021.1872079.

An, H.-Y., Chen, W., Wang, C.-W., Yang, H.-F., Huang, W.-T., & Fan, S.-Y. (2020). The Relationships between Physical Activity and Life Satisfaction and Happiness among Young, Middle-Aged, and Older Adults. *International Journal of Environmental Research and Public Health*, *17*(13), 1–10. https://doi.org/10.3390/ijerph17134817.

Bowling, A. (2005). Just one question: If one question works, why ask several? *Journal of Epidemiology & Community Health*, *59*(5), 342–345. https://doi.org/10.1136/jech.2004.021204.

Craig, C. L., Marshall, A. L., Sjöström, M., Bauman, A. E., Booth, M. L., Ainsworth, B. E., Pratt, M., Ekelund, U., Yngve, A., Sallis, J. F., & Oja, P. (2003). International Physical Activity Questionnaire: 12-Country Reliability and Validity. *Medicine & Science in Sports & Exercise*, *35*(8), 1381–1395. https://doi.org/10.1249/01.MSS.0000078924.61453.FB.

Fallahpour, M., Borell, L., Luborsky, M., & Nygård, L. (2016). Leisure-activity participation to prevent later-life cognitive decline: A systematic review. *Scandinavian Journal of Occupational Therapy*, *23*(3), 162–197. https://doi.org/10.3109/11038128.2015.1102320.

Frank, C. C., Mundy, L. M., & Smith, J. (2023). Life course engagement in enriching activities: When and how does it matter for cognitive aging? *Psychology and Aging*, *38*(4), 263–276. https://doi.org/10.1037/pag0000744.

Frantál, B., Klapka, P., & Nováková, E. (2020). When home becomes a cage: Daily activities, space-time constraints, isolation and the loneliness of older adults in urban environments. *Moravian Geographical Reports*, *28*(4), 322–337. https://doi.org/10.2478/mgr-2020-0024.

Galvin, J. E., Tolea, M. I., & Chrisphonte, S. (2021). The Cognitive & Leisure Activity Scale (CLAS): A new measure to quantify cognitive activities in older adults with and without cognitive impairment. *Alzheimer's & Dementia: Translational Research & Clinical Interventions*, *7*(1). https://doi.org/10.1002/trc2.12134.

Heissler, R., Kopeček, M., & Georgi, H. (2021). Leisure activities of SuperAgers. *Ageing 2021*, 77–86. http://www.konferencestarnuti.cz/files/Starnuti_2021_sbornik.pdf.

Izquierdo, M., Duque, G., & Morley, J. E. (2021). Physical activity guidelines for older people: Knowledge gaps and future directions. *The Lancet Healthy Longevity*, *2*(6), e380–e383. https://doi.org/10.1016/S2666-7568(21)00079-9.

Jopp, D. S., & Hertzog, C. (2010). Assessing adult leisure activities: An extension of a self-report activity questionnaire. *Psychological Assessment*, *22*(1), 108–120. https://doi.org/10.1037/a0017662.

Kleiber, D. A., & Nimrod, G. (2009). 'I can't be very sad': Constraint and adaptation in the leisure of a 'learning in retirement' group. *Leisure Studies*, *28*(1), 67–83. https://doi.org/10.1080/02614360802260820.

Landis, J. R., & Koch, G. G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, *33*(1), 159–174.

Li, S., Carlson, E., & Holm, K. (2000). Validation of a Single-Item Measure of Usual Physical Activity. *Perceptual and Motor Skills*, *91*(2), 593–602. https://doi.org/10.2466/pms.2000.91.2.593.

Livingston, G., Huntley, J., Sommerlad, A., Ames, D., Ballard, C., Banerjee, S., Brayne, C., Burns, A., Cohen-Mansfield, J., Cooper, C., Costafreda, S. G., Dias, A., Fox, N., Gitlin, L. N., Howard, R., Kales, H. C., Kivimäki, M., Larson, E. B., Ogunniyi, A., ... Mukadam, N. (2020). Dementia prevention, intervention, and care: 2020 report of the Lancet Commission. *The Lancet*, *396*(10248), 413–446. https://doi.org/10.1016/S0140-6736(20)30367-6.

Lojdová, M., Kvintová, J., Štěrbová, D., & Krol, P. (2021). Assessment of life satisfaction and its selected components in relation to the level of physical activity. *Journal of Physical Education and Sport*, *3*(21), 1346–1353. https://doi.org/10.7752/jpes.2021.03171.

Milton, K., Bull, F. C., & Bauman, A. (2011). Reliability and validity testing of a single-item physical activity measure. *British Journal of Sports Medicine*, *45*(3), 203–208. https://doi.org/10.1136/bjsm.2009.068395.

Mitáš, J., Ding, D., Frömel, K., & Kerr, J. (2014). Physical Activity, Sedentary Behavior, and Body Mass Index in the Czech Republic: A Nationally Representative Survey. *Journal of Physical Activity and Health*, *11*(5), 903–907. https://doi.org/10.1123/jpah.2012-0277.

Najar, J., Östling, S., Gudmundsson, P., Sundh, V., Johansson, L., Kern, S., Guo, X., Hällström, T., & Skoog, I. (2019). Cognitive and physical activity and dementia: A 44-year longitudinal population study of women. *Neurology*, *92*(12), e1322–e1330. https://doi.org/10.1212/WNL.0000000000007021.

Nevalainen, N., Riklund, K., Andersson, M., Axelsson, J., Ögren, M., Lövdén, M., Lindenberger, U., Bäckman, L., & Nyberg, L. (2015). COBRA: A prospective multimodal imaging study of dopamine, brain structure and function, and cognition. *Brain Research*, *1612*, 83–103. https://doi.org/10.1016/j.brainres.2014.09.010.

Polit, D. F. (2014). Getting serious about test-retest reliability: A critique of retest research and some recommendations. *Quality of Life Research*, *23*(6), 1713–1720. https://doi.org/10.1007/s11136-014-0632-9.

Scott, J. J., Morgan, P. J., Plotnikoff, R. C., & Lubans, D. R. (2015). Reliability and validity of a single-item physical activity measure for adolescents. *Journal of Paediatrics and Child Health*, *51*(8), 787–793. https://doi.org/10.1111/jpc.12836.

Sprangers, M. A. G., & Schwartz, C. E. (1999). Integrating response shift into health-related quality of life research: A theoretical model. *Social Science & Medicine*, *48*(11), 1507–1515. https://doi.org/10.1016/S0277-9536(99)00045-3.

Stern, Y. (2002). What is cognitive reserve? Theory and research application of the reserve concept. *Journal of the International Neuropsychological Society*, *8*(3), 448–460. https://doi.org/10.1017/S1355617702813248.

Stern, Y., Arenaza-Urquijo, E. M., Bartrés-Faz, D., Belleville, S., Cantilon, M., Chetelat, G., Ewers, M., Franzmeier, N., Kempermann, G., Kremen, W. S., Okonkwo, O., Scarmeas, N., Soldan, A., Udeh-Momoh, C., Valenzuela, M., Vemuri, P., Vuoksimaa, E., & and the Reserve, Resilience and Protective Factors PIA Empirical Definitions and Conceptual Frameworks Workgroup (2020). Whitepaper: Defining and investigating cognitive reserve, brain reserve, and brain maintenance. *Alzheimer's & Dementia*, *16*(9), 1305–1311. https://doi.org/10.1016/j.jalz.2018.07.219.

U.S. Department of Health and Human Services (2018). *2018 Physical activity guidelines advisory committee scientific report*. https://health.gov/sites/default/files/2019-09/PAG_Advisory_Committee_Report.pdf.

Vašíčková, J., Roberson, D., & Frömel, K. (2012). The Education Level and Socio-Demographic Determinants of Physical Activity in Czech Adults. *Human Movement*, *13*(1). https://doi.org/10.2478/v10038-012-0005-6.

WHO (2020). *WHO guidelines on physical activity and sedentary behaviour*. World Health Organization. https://iris.who.int/bitstream/handle/10665/336656/9789240015128-eng.pdf?sequence=1.

APPENDIX A

A2PA

Attitude towards physical activity throughout life

What has been your attitude to sport and physical activity during your adult life, **from the age of 20** until now? **Select one answer.**

1. I am a very sporty person; I seek sports and physical activity. I am an active member of a sports club or organization (including a tourist organization), I used to play sports as an amateur. I certainly had at least three times a week of vigorous exercise (including brisk walking) and for at least 2.5 hours per week. Rather a lot more.
2. I like to move. I have had at least three times a week of vigorous exercise during my adult life, and I have often played sports for fun. I usually get more than 2.5 hours of exercise per week.
3. At least three times a week I had more intense movement, including brisk walking. I have had stretches of time when I did more sport.
4. I'm not a sporty person, but I don't shy away from exercise. I am not sure I can say that I have had at least three times a week of physical activity (including brisk walking) for most of my life and at least 2.5 hours of physical activity in total per week.
5. I am not a sporty person, I enjoy other things. I almost certainly did not have physical activities three times a week for a total of 2.5 hours per week.
6. I used to be an athlete, but due to an injury or illness I had to stop regular exercise. I miss sport. – If you choose this answer, please indicate in the following "Age" section by a number at what age you had to stop or significantly reduce your regular physical activity.
   Age .....

APPENDIX B

COBRA-Cognitive

Reflect on what your typical week looks like in the summer. How many hours per week do you usually spend on the listed activities? (circle the answer – the number representing the hours per week). **Over the past 6 years, has this amount of time remained approximately the same, or is it now more or less than before?** (mark the appropriate column with an X). **How challenging are these activities for you usually?** (circle: 1: not challenging at all – 5: extremely challenging/don't know). *For each intensity, ask for the specific activities the participant engages in and all the details.*

Activity	Times p/w	Hours p/w	Now less (0)	Same (1)	Now more (2)	Difficulty	Don't know (0)
Computer use for non-gaming purposes (including writing emails, editing photos, communicating via Skype, etc.)?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Reading fiction?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Crosswords?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Solving quizzes, Sudoku, etc.?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Assembling puzzles (e.g., jigsaw puzzles)?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Art activities or handwork?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Playing card or board games?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Playing a musical instrument or singing?		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
Other:		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0

APPENDIX C

COBRA-Physical

Reflect on what your typical week looks like in the summer. How many times per week do you engage in any activity of the described intensity for at least 10 minutes? (write down the number). **How many hours per week do you usually spend on the listed activities?** (circle the answer – the number representing the hours per week). **Over the past 6 years, has this amount of time remained approximately the same, or is it now more or less than before?** (mark the appropriate column with an X). **How physically challenging are these activities for you usually?** (circle: 1: not challenging at all – 5: extremely challenging/don't know). *For each intensity, ask for the specific activities the participant engages in and all the details.*

Activity	Times p/w	Hours p/w	Now less (0)	Same (1)	Now more (2)	Difficulty	Don't know (0)
VERY INTENSE physical activity – rapid heartbeat, sweating, unable to speak fluently during it. E.g. running or very fast walking, aerobics, intense swimming, cycling uphill or fast, etc.							
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
MODERATE INTENSITY physical activity – accelerated heartbeat, light sweating, able to speak fluently. E.g. brisk walking, tennis doubles, leisure cycling, yoga, volleyball, badminton, light swimming, or intensive gardening, etc.							
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0

LIGHT physical activity – no sweating, able to sing during it. E.g. slow walking, stretching, bowling, pétanque, light housework or gardening, etc.							
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0
		0 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15+				1 2 3 4 5	0