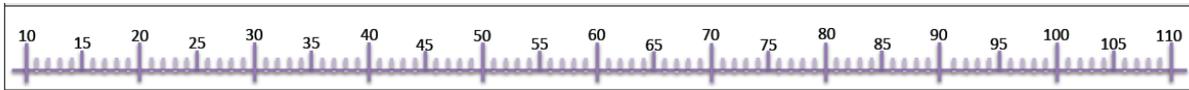


## Supplementary Materials

### 1. SHH-Q items (translation of the German final version)

Each SHH-Q item was presented above a timescale ranging from 10 to 110 years.



#### *Control items*

1. Please estimate how old you will approximately become: ...
2. Up to which age will you maintain your independence?
3. Please assess up to which age you will feel healthy.
4. From which age do you think you will be chronically sick?

#### *SHH-Q items on bodily fitness (Body)*

5. Up to which age do you think you will be able to maintain your current physical performance?
6. Up to which age do you believe you will be able to maintain or even increase your fitness by training?
7. Up to which age will you be able to incorporate additional physical activity into your daily life (e.g., bicycle tours, nordic walking, jogging, swimming, working out, hiking)?
8. From which age onward do you think your motivation to do sports will decline?
9. Please judge at what age your physical capabilities will have decreased by more than half (e.g., from which age onward will you only be able to jog, swim, or walk half of the current distance?).
10. Up to which age will you be able to motivate yourself to train more intensively than at present in order to maintain or improve your fitness?
11. Up to which age will you be able to practice sports such as jogging, cycling, nordic walking, or tennis? Please assess yourself even if you will not actually carry it out.
12. Up to which age will you be able to practice sports such as yoga, pilates, relaxation, gymnastics? Please assess yourself even if you will not actually carry it out.
13. Up to which age will you be able to practice sports such as strengthening muscles or weight training? Please assess yourself.

#### *SHH-Q items on occupational perspectives (Work)*

14. Considering your present health, up to which maximum age could you imagine working?
15. Considering your present health, up to which maximum age could you imagine starting a completely new job?

16. From which age would cardiovascular complaints hinder you from working in a new occupational field?
17. From which age would your mental fitness hinder you from working in a new occupational field?
18. From which age would loss of motivation or a lack of energy hinder you from working in a new occupational field?

*SHH-Q items on overall goals and plans (Life Goals)*

19. How far does the overall planning of your goals reach into the future?
20. How far ahead do you generally plan activities? (e.g., visiting your family next month; going on a round-the-world trip in 5 years; learning a new language in the next 10 years)
21. In view of your physical fitness, up to which age do you generally plan your future ahead?

*SHH-Q items on novelty-oriented exploration (Novelty)*

22. Up to which age do you feel physically and mentally able to learn new sports skills independently (e.g., dancing, sailing, mountaineering, diving)?
23. Up to which age do you feel physically and mentally able to be creative in a new way (e.g., theater group, painting, digital photography, etc.)?
24. Up to which age do you feel physically and mentally able to emigrate to an foreign country or to move to a new place independently?
25. Up to which age do you feel physically and mentally able to learn something mentally new (e.g., learn a new language, extend your computer knowledge, learn new historical facts, participate in a book club, etc.)?
26. Up to which age do you feel physically and mentally able to explore new countries and cultures actively and independently (e.g., city tours, expeditions, round trips)?
27. Up to which age do you feel physically and mentally able to learn new skills independently (e.g., a new craft, juggling)?
28. Up to which age do you feel physically and mentally able to come to grips with new technology independently?
29. Up to which age do you feel physically and mentally able to try new things in your daily life independently (e.g., taking a new route, visiting a new mall)?
30. Up to what age do you think you will be physically and mentally able to meet new people (e.g., by joining a sports group, club, or other interest groups)?

## 2. SHH-Q factorizability

In order to assess factorizability, modification indices of the confirmatory factor analysis (CFA) of the four-factor model were used to detect items with high cross-loadings as well as items with low loadings on latent factors. While most items showed only minor cross-loadings, and thus demonstrate a good factor structure, we excluded items 10 and 16 (not shown in section 1) from the initial questionnaire because they exhibited cross-loadings above 0.50. All confirmatory factor analyses (CFAs) conducted to validate the questionnaire were carried out with the 26 remaining items. The full factor structure of model A is given in table S1.

**Table S1.** Factor structure of the SHH-Q based on confirmatory factor analysis.

SHH-Q Factor	Novelty	Body	Work	Life Goals
	Item (Loading)	Item (Loading)	Item (Loading)	Item (Loading)
	i24: 0.76	i5: 0.63	i15: 0.72	i21: 0.68
	i25: 0.83	i6: 0.72	<i>i16*</i>	i22: 0.93
	i26: 0.76	i7: 0.87	i17: 0.82	i23: 0.95
	i27: 0.84	i8: 0.80	i18: 0.84	
	i28: 0.85	i9: 0.75	i19: 0.87	
	i29: 0.82	<i>i10*</i>	i20: 0.83	
	i30: 0.83	i11: 0.77		
	i31: 0.89	i12: 0.85		
	i32: 0.83	i13: 0.84		
		i14: 0.73		
Cronbach's $\alpha$	0.95	0.93	0.92	0.89

*Note.* The corresponding model is shown in figure 1a.  $\chi^2_{293} = 1,050.828$ ; CFI = 0.943; RMSEA = 0.045; 90% CI = 0.042–0.048; SRMR = 0.041 ( $\chi^2$  = Chi-square statistics; CFI = Comparative Fit Index; RMSEA = root mean square error of approximation; CI = confidence interval).

\* Items (i) 10 and 16 were excluded from the initial questionnaire because they exhibited cross-loadings of above 0.50.

### 3. Cognitive tasks and latent cognitive model

**Cognitive tasks.** The participants were invited to two cognitive sessions with an exact interval of seven days and at the same time of day to avoid circadian confounding effects on performance. Each session lasted about 3.5 h. Participants were tested in groups of 4–6 individuals. The cognitive battery included measures of learning and memory performance, attention/processing speed, working memory, executive functioning, and perceptual speed. The group was instructed via a standardized session protocol. Each task started with a practice trial to make sure that all participants understood the task. Responses were given via button boxes, the computer mouse, or keyboard. For the purpose of the present study, only cognitive tasks included in CFA as manifest variables are described in detail below.

#### **Episodic memory (EM)**

*Scene Encoding* (abbreviated as Scene in figure S1). Participants performed an incidental encoding task with 88 complex, gray-scaled images (44 indoor and 44 outdoor scenes; mean grey value 127, SD = 75) of neutral emotional valence sequentially presented each for 2.5 s. Participants performed an indoor/outdoor judgment on each image by button presses (left index finger for inside scenes and right index finger for outdoor scenes). In the retrieval tests, 44 old images (22 indoor – 22 outdoor) presented at encoding task were randomly presented together with 44 novel distractor images (22 indoor – 22 outdoor) for 2.5 s. After each presentation participants rated their confidence of recognition memory on a scale ranging from 1 to 5: 1 = sure that new; 2 = may be new; 3 = unsure whether old or new; 4 = may be old; 5 = sure that old). There were no time restrictions for confidence ratings. Recognition memory performance for scenes (hits minus false alarms) was assessed after a delay of approximately 2.5 h (short-delay memory), and used as the manifest variable in the model.

*Verbal Learning and Memory Test* (List). This task assesses auditory verbal learning of 15 words including five learning trials, early recall, an interference list after five learning trials, free recall tests directly after an interference list as well as 30 min later (late recall), and a final recognition test [3]. The sum of items recalled across trials 1 to 5 provides a measure of overall learning performance that was used here as the manifest variable.

*Face–Profession task* (Face). This task assesses associative binding on the basis of recognition of incidentally encoded face–profession pairs. During the study phase 45 face–profession pairs were each presented for 3.5 s on the computer screen and the participants had to indicate via button presses whether the faces matched the profession or not. After a 3-min delay between study and test phase 54 face–profession pairs consisting of 27 old pairs, 9 new pairs, and 18 newly arranged pairs

were presented (in newly arranged pairs the shown face is the same, but is associated with a new profession). The participants were asked to decide whether they had seen a given face–profession combination before or not and to rate the confidence of their decision on a three-point scale ranging from 1 = not sure to 3 = very sure. Recognition memory for the rearranged face–profession pairs (hit minus false alarms) served as the manifest variable in the model.

*Object Location task (Object)*. In this task, sequences of 12 colored photographs of real-world objects were displayed at different locations in a 6-by-6 grid. After presentation, objects appeared at the side of the screen and had to be moved to the correct locations by clicking on the objects and the locations with the computer mouse. One practice trial and two test trials were included. The sum of correct placements across the two test trials is used as the manifest variable (see also [4]).

### **Working memory (WM)**

*Spatial Updating (Spatial)*. In each block of this task, a display of two or three 3-by-3 grids was shown for 4 s. In each of these, one blue dot was presented in one of the nine locations. Those two or three locations had to be memorized and updated according to shifting operations that were indicated by arrows appearing below the corresponding field. The presentation time of the arrows was 2.5 s with an interstimulus interval of 0.5 s. After six updating operations, the two or three grids reappeared and the resulting end positions had to be clicked on. After ten practice blocks with memory loads of two and three grids, ten test blocks with load two and three were conducted and used for scoring. The average percentage of correct placements was used as one of the manifest variables for the working memory (WM) factor (see also [4]).

*Letter Updating (LetterU)*. In this task subjects were presented with 7, 9, 11, or 13 letters in a sequence. Once a sequence stopped, subjects had to report the last three letters in correct order by pressing buttons on the button box corresponding to A, B, C, and D.

*Number-N-Back (Number)*. Three one-digit numbers (ranging from 0 to 9) were presented sequentially in three cells situated horizontally followed by the next sequence of three digits. This cycle was repeated 30 times. In each cycle, two-choice decisions on whether the current stimulus matched the stimulus shown three steps earlier had to be made. Four practice trials including 30 runs were followed by 6 test trials with 30 runs. Subjects made their decision via button-box presses with their left and right index fingers (see also [4]).

**Fluid intelligence (Gf)**

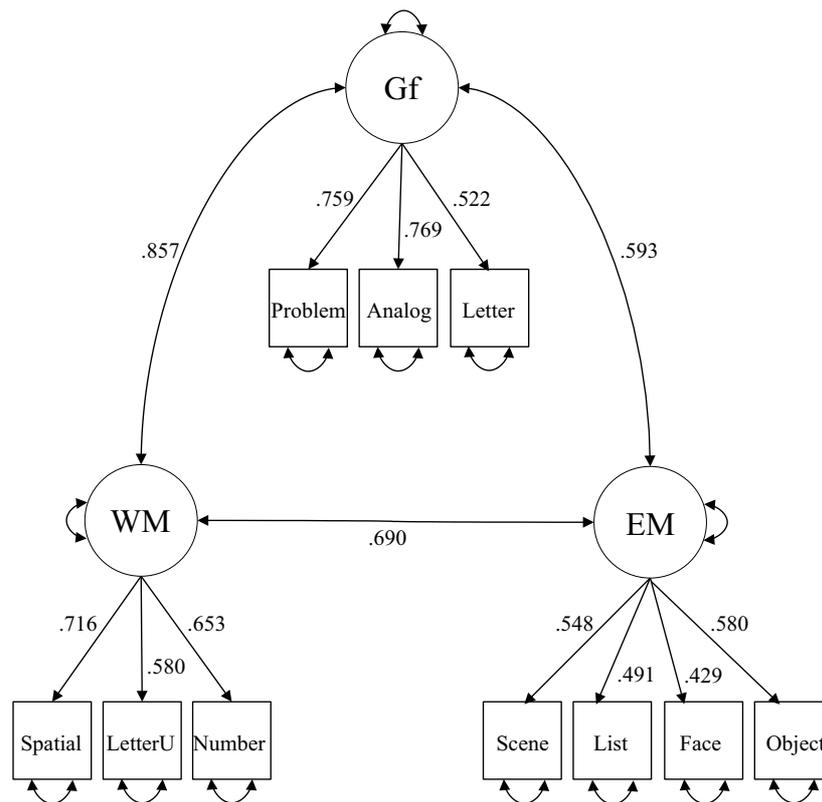
*Practical Problems* (Problem). This task consisted of 12 items depicting everyday problems such as the times in a bus schedule, instructions for medication, a warranty for a technical appliance, a rail map, as well as other forms and tables. For each item, the problems were presented in the upper part of the screen, and five alternative responses were shown in the lower part. Subjects responded by clicking on one of the five alternatives with the computer mouse. A single practice item was provided. The test phase was terminated when subjects made three consecutive errors, or when they reached the maximum time limit of 10 minutes, or after they had answered the last test item. Items were ordered by difficulty (see also [5]).

*Figural Analogies* (Analog). Items in this test followed the format "A is to B as C is to ?". One figure pair was presented in the upper left part of the screen and a single figure was shown beside it. Participants had to use the same rule as the one applying to the complete figure pair to choose one of the five alternative responses presented below. Subjects entered their response by clicking on one of the five alternatives with the mouse. Before the test phase, instructions and three practice items were given. The test phase was terminated when subjects made three consecutive errors, when they reached the maximum time limit (10 min), or after they had worked on the last test item. Items were ordered by difficulty (see also [5]).

*Letter Series* (Letter). The task consisted of 22 items. Each item contained five letters followed by a question mark (e.g., c e g i k ?). Items were displayed in the upper half of the screen, and five response alternatives were presented in the lower half. Items followed simple rules such as +1, -1, +2, or +2 +1. Subjects entered their response by touching one of the five answer alternatives. The score was based on the total number of correct responses. Instructions and three practice items were given before the test phase. The test phase was terminated when subjects made three consecutive false responses, when they reached the maximum time limit (6 min), or after they had answered the last item of the test. Items were ordered by difficulty. Sample items were used with respect to tests related to speed, reasoning, and knowledge (cf. (see also [5]).

**Confirmatory factor analyses of three latent cognitive factors**

CFA was applied to define a three-factor model of cognitive abilities in latent space using measures of specific cognitive tasks selected from the cognitive battery of BASE-II as manifest variables. The model shown in figure S1 fits the data well ( $\chi^2_{32} = 112.1$ ; CFI = 0.97; RMSEA = 0.043; 90% CI = 0.034–0.050; SRMR = 0.031).



**Figure S1.** Confirmatory factor analysis (CFA) of three cognitive abilities, working memory (WM), episodic memory (EM), and fluid intelligence (Gf).

*Note.* Manifest variables loading on each of the latent factors are represented by squares. Numbers next to single-headed arrows represent significant standardized path coefficients, and numbers next to double-headed arrows refer to standardized covariances between the latent factors.

#### 4. Transformation of SHH-Q items and robustness checks

To test for multivariate normality of the age-adjusted SHH-Q items, we calculated Mardia's coefficient (<http://www.biosoft.hacettepe.edu.tr/MVN>). Mardia's coefficient indicates a significant deviation from normality ( $g_2p = 1,409.12$ ;  $z_{kurtosis} = 52.65$ ;  $p_{kurtosis} < 0.001$ ). Therefore we decided to employ a Box-Cox transformation in an attempt to restore multivariate normality. Even after a Box-Cox optimal transformation, however, the data exhibited a significant deviation from normality ( $g_2p = 838.32$ ;  $z_{kurtosis} = 25.12$ ;  $p_{kurtosis} < 0.001$ ). To test the robustness of our main analyses, however, we nevertheless repeated all analyses with the transformed SHH-Q items. Table S2 shows that none of our results changed as a result of the non-normality multivariate distribution issue.

**Table S2a.** CFA testing robustness of factorial validity of four-factor solution of the SHH-Q using the Box-Cox transformed SHH-Q items.

<b>Model</b>	$\chi^2$	<b>d.f.</b>	<b>p</b>	<b>CFI</b>	<b>RMSEA/ 90% CI</b>	<b>SRMR</b>	<i>Chi-square comparison of Model A to... (<math>\chi^2</math>/d.f./p)</i>
<i>Model A</i>	1,213.227	293	<0.001	0.944	0.040/0.046	0.040	
<i>Model B</i>	2,463.324	295	<0.001	0.869	0.074/0.072	0.059	71.51/2/<0.001
<i>Model C</i>	1,803.317	295	<0.001	0.909	0.062/0.059	0.050	282.93/2/<0.001
<i>Model D</i>	4,532.386	299	<0.001	0.743	0.103/0.101	0.077	2361.7/6/<0.001

*Note.* CFA = confirmatory factor analysis;  $\chi^2$  = Chi-square statistics; d.f. = degrees of freedom; p = probability value; CFI = Comparative Fit Index; RMSEA = root mean square error of approximation; CI = confidence interval; Model A = four-factor model; Models B & C = three-factor models; Model D = single-factor model.

**Table S2b.** Stepwise regression models for cognition, physical fitness, and metabolic status. Each model includes three stepwise regressions, including: (1) covariates: age, education, and gender, (2) the Future Time Perspective (FTP) score, and (3) two SHH-Q composite scores, namely SHH-Q *Novelty* and SHH-Q *Body*. Results for each stepwise increment of the single predictor is shown.

<i>Stepwise Latent Regression Models for Cognition</i>							
Dependent variables	Predictors	Coefficients	SE	p	R <sup>2</sup> /R <sup>2</sup> change	Incremental validity <sup>a,b</sup> ( $\chi^2$ /d.f./p)	
<b>EM</b>	1. Age	-0.136*	0.042	<0.001	0.140		
	Gender	0.166*	0.037	<0.001			
	Education	0.283*	0.039	<0.001			
	2. FTP scale	-0.069	0.040	0.086	0.140/0.000		4.59/3/0.204
	3. SHH-Q <i>Novelty</i> <sup>1</sup>	0.334*	0.064	0.001	0.194/0.054		57.38/12/<0.001
	SHH-Q <i>Body</i> <sup>1</sup>	-0.063	0.053	0.236			
<b>WM</b>	1. Age	-0.101*	0.050	0.007	0.134		
	Gender	-0.083*	0.035	0.008			
	Education	0.290*	0.036	<0.001			
	2. FTP scale	-0.096*	0.038	0.012	0.137/0.003		4.59/3/0.204
	3. SHH-Q <i>Novelty</i> <sup>1</sup>	0.076	0.060	0.208	0.161/0.024		57.38/12/<0.001
	SHH-Q <i>Body</i> <sup>1</sup>	0.073	0.051	0.152			
<b>Gf</b>	1. Age	-0.100*	0.045	0.003	0.169		
	Gender	-0.073*	0.032	0.022			
	Education	0.324*	0.041	<0.001			
	2. FTP scale	-0.121*	0.037	0.001	0.173/0.004		4.59/3/0.204
	3. SHH-Q <i>Novelty</i> <sup>1</sup>	0.146*	0.054	0.007	0.218/0.045		57.38/12/<0.001
	SHH-Q <i>Body</i> <sup>1</sup>	-0.008	0.048	0.998			

Table S2b continued overleaf...

... table S2b continued.

<i>Stepwise Latent Regression Models for Physical Fitness</i>							
<b>Dependent variables</b>	<b>Predictors</b>	<b>Coefficients</b>	<b>SE</b>	<b>p</b>	<b>R<sup>2</sup>/R<sup>2</sup> change</b>	<b>Incremental validity<sup>a</sup> (<math>\chi^2</math>/d.f./p)</b>	
<b>Fitness</b>	1. Age	-0.102*	0.025	<0.001	0.826		
	Gender	-0.900*	0.014	<0.001			
	Education	0.004	0.022	0.503			
	2. FTP scale	0.017	0.022	0.746	0.828/0.002		4.75/1/0.029
	3. SHH-Q <i>Novelty</i> <sup>1</sup>	0.075	0.037	0.073	0.831/0.003		13.03/4/0.011
	SHH-Q <i>Body</i> <sup>1</sup>	0.049	0.033	0.193			
<i>Stepwise Latent Regression Models for Metabolic Status</i>							
<b>Dependent variables</b>	<b>Predictors</b>	<b>Coefficients</b>	<b>SE</b>	<b>p</b>	<b>R<sup>2</sup>/ R<sup>2</sup> change</b>	<b>Incremental validity<sup>a</sup> (<math>\chi^2</math>/d.f./p)</b>	
<b>MetS</b>	1. Age	-0.109*	0.036	0.002	0.175		
	Gender	-0.396*	0.047	<0.001			
	Education	-0.091*	0.034	0.007			
	2. FTP scale	-0.075*	0.035	0.022	0.181/0.006		7.67/1/0.006
	3. SHH-Q <i>Novelty</i> <sup>1</sup>	0.187*	0.054	0.003	0.204/0.023		35.74/4/0.000
	SHH-Q <i>Body</i> <sup>1</sup>	-0.255*	0.054	<0.001			

*Note.* SE= standard error;  $\chi^2$  = Chi-square statistics; d.f. = degrees of freedom; p = probability value. <sup>1</sup> Scores based on Box-Cox-transformed SHH-Q composites for each SHH-Q dimension; <sup>a</sup> Incremental validity is used to determine whether entering FTP first and SHH-Q second increases the predictive ability of the dependent variables of each model by applying a  $\chi^2$  difference test. <sup>b</sup> Incremental validity values are equal for all three cognitive factors because they are based on the same model. \* p < 0.05.

## **5. Fit indices used in CFA**

The Chi-square statistic assesses the absolute model fit. For large sample sizes, however, the test is very powerful in the detection of even minor differences between the observed and model implied covariance matrix, thus rejecting the null hypothesis (good model fit) even in cases where model misspecification is practically negligible [1]. The root mean square error of approximation (RMSEA) describes closeness of fit. Values of the RMSEA  $< 0.06$  reflect close and acceptable fit of the model. The 90% confidence interval (CI) around the RMSEA point estimate should also contain zero to indicate a good fit. The Comparative Fit Index (CFI) tests the relative improvement in fit by comparing the proposed model with a baseline model. Values for the CFI around 0.90 are considered acceptable, whereas values around 0.95 indicate a good fit [2].

## **6. Descriptive information**

Tables S3 and S4 provide an overview of the main variables of interest and their correlations. The scaling of the SHH-Q scores refers to age-adjusted years. See also Methods section in the article for further information. For all variables of interest, data were missing occasionally. Thus, to conduct further structural equation modeling (SEM) analyses including all available variables of interest, we used Full Information Maximum Likelihood (FIML). FIML is a SEM-based missing data estimation approach that yields unbiased parameter estimates and standard errors. This method estimates a likelihood function for each individual based on the variables that are present, such that all available data are used.

**Table S3.** Means and standard deviations (SD) of characteristics, SHH-Q factors, objective health, metabolic measures, and examples of cognitive and psychosocial variables shown.

Sample characteristics		
N <sub>whole sample</sub>	1,371	
Gender [% female] <sub>N = 1,371</sub>	50.1 %	
Variable name	Mean	(SD)
Age [years] <sub>N = 1,371</sub>	70.1	(3.8)
Years of education <sub>N = 1,153</sub>	14.2	(2.9)
Body Mass Index (BMI) <sub>N = 1,178</sub>	22.7	(3.6)
SHH-Q <i>Novelty</i> <sup>1</sup> <sub>N = 1,277</sub>	5.25	(2.6)
SHH-Q <i>Body</i> <sup>1</sup> <sub>N = 1,292</sub>	5.80	(2.5)
SHH-Q <i>Work</i> <sup>1</sup> <sub>N = 1,085</sub>	4.32	(2.8)
SHH-Q <i>Life Goals</i> <sup>1</sup> <sub>N = 1,226</sub>	4.67	(2.9)
FTP scale <sub>N = 1,258</sub>	2.6	(0.7)
Grip strength [kg] <sub>N = 1,301</sub>	30.5	(8.9)
Forced expiratory volume in one second (FEV <sub>1</sub> ) [ml] <sub>N = 840</sub>	3,395.9	(893.7)
Trunk fat [g] <sub>N = 1,350</sub>	13,861.6	(4698.5)
High density lipoprotein (HDL) [mg/dl] <sub>N = 1,253</sub>	62.2	(17.4)
Triglycerides [mg/dl] <sub>N = 1,251</sub>	113.7	(62.1)
Fasting glucose (plasma) [mg/dl] <sub>N = 1,257</sub>	96.00	(19.7)
Episodic memory (EM) <sup>1</sup> <sub>N = 1,371</sub>	5.6	(1.6)
Working memory (WM) <sup>1</sup> <sub>N = 1,371</sub>	17.7	(5.3)
Fluid intelligence <sup>1</sup> <sub>N = 1,371</sub>	10.2	(3.8)

*Note.* <sup>1</sup> = unit-weighted composite scores. The numbers of observations for each variable differ because of missing cases and are shown in the subscript numbers.

**Table S4.** Correlations (r) among SHH-Q subscales, the FTP scale, subjective health, current age, gender, and years of education.

	SHH-Q				FTP	Subj. health	Age	Gender	Education
	<i>Novelty</i>	<i>Body</i>	<i>Work</i>	<i>Life Goals</i>					
SHH-Q <i>Novelty</i>	1	0.74*	0.72*	0.52*	0.36*	0.19*	-0.29*	-0.01	0.06
SHH-Q <i>Body</i>		1	0.62*	0.48*	0.39*	0.23*	-0.29*	-0.00	0.01
SHH-Q <i>Work</i>			1	0.38*	0.28*	0.14*	-0.17*	-0.08	0.06
SHH-Q <i>Life Goals</i>				1	0.27*	0.11*	-0.09*	-0.05	-0.00
FTP					1	0.26*	-0.23*	-0.03	-0.14*
Subj. health						1	-0.09*	-0.02	0.12*
Age							1	0.02	-0.02
Gender								1	-0.15*
Education									1

*Note.* SHH-Q = Subjective Health Horizons Questionnaire, FTP = Future Time Perspective scale.  
\*  $p < 0.01$  (2-tailed).

### References

1. Bollen KA: Structural Equation Models with Latent Variables. New York, Wiley, 1989.
2. Bentler PM: Comparative fit indexes in structural models. *Psychol Bull* 1990;107:238–246.
3. Helmstaedter C, Durwen HF: The Verbal Learning and Retention Test: a useful and differentiated tool in evaluating verbal memory performance. *Schweiz Arch Neurol Psychiatr* 1990;141:21–30.
4. Schmiedek F, Lövdén M, Lindenberger U: Hundred days of cognitive training enhance broad cognitive abilities in adulthood: findings from the COGITO Study. *Front Aging Neurosci* 2010;2:27.
5. Lindenberger U, Mayr U, Kliegl R: Speed and intelligence in old age. *Psychol Aging* 1993;8:207–220.