Electronic Supplement to:

Self-Reported Mid- to Late-Life Physical and Recreational Activities: Associations with Late-Life Cognition

Brandon E. Gavett[1,3], Keith F. Widaman[2], Cathryn McKenzie[3], Fransia S. De Leon[4], Evan Fletcher[3], Sarah Tomaszewski Farias[3], Dan Mungas[3]

1. Department of Neurology, University of California, Davis, Sacramento, California, USA

2. Graduate School of Education, University of California, Riverside, Riverside, California USA

3. School of Psychological Science, University of Western Australia, Perth, Western Australia, Australia

4. School of Medicine, University of California, Davis, Sacramento, California, USA

Address correspondence to: Brandon Gavett, 4860 Y St. Ste. 3700, Sacramento, California, 95817 United States. Email: bgavett@ucdavis.edu.

Measurement Invariance Testing for Physical and Recreational Activities at Age 40 and Current

Physical and recreational activity participation was measured using the Life Experiences Assessment Form (LEAF; see Brewster et al., 2014). Participants used the LEAF to answer questions about their current physical and recreational activity engagement as well as their retrospectively-rated physical and recreational activity engagement at age 40. Physical activity items on the LEAF ask participants to use a 5-point Likert scale to rate their frequency of activity engagement, ranging from 1 (“Never”) to 5 (“Every day or almost every day”), on tasks described as light work-related tasks, heavy work-related tasks, light house/yard work, vigorous house/yard work, light exercise, and vigorous exercise. Recreational activity items on the LEAF ask about reading, complex cooking, writing, taking classes, performance arts, games or puzzles, cultural events, arts and crafts, socializing, and attendance at clubs or meetings, using the same 5-point Likert scale. Item response categories were collapsed when necessary to avoid sparse cells. Before testing the time invariance of physical and recreational activities factors, we first established unidimensional models for each activity type, as shown in Figure 1. Because the indicator variables were ordinal, we used the WLSMV estimator in Mplus and delta parameterization. The fit statistics of the unidimensional models, when applied separately to each time epoch, are shown in Table 1.

Subsequently, we combined the data from both time epochs into a single two-factor latent variable model, for physical and recreational activities separately. We included residual covariance terms to account for item autocorrelations across time epochs (e.g., age 40 reading with current reading). The fits of the two-factor models are shown in Table 2.

Figure 1. Path diagrams showing hypothesized factor structures for physical activities (A) and recreational activities (B).

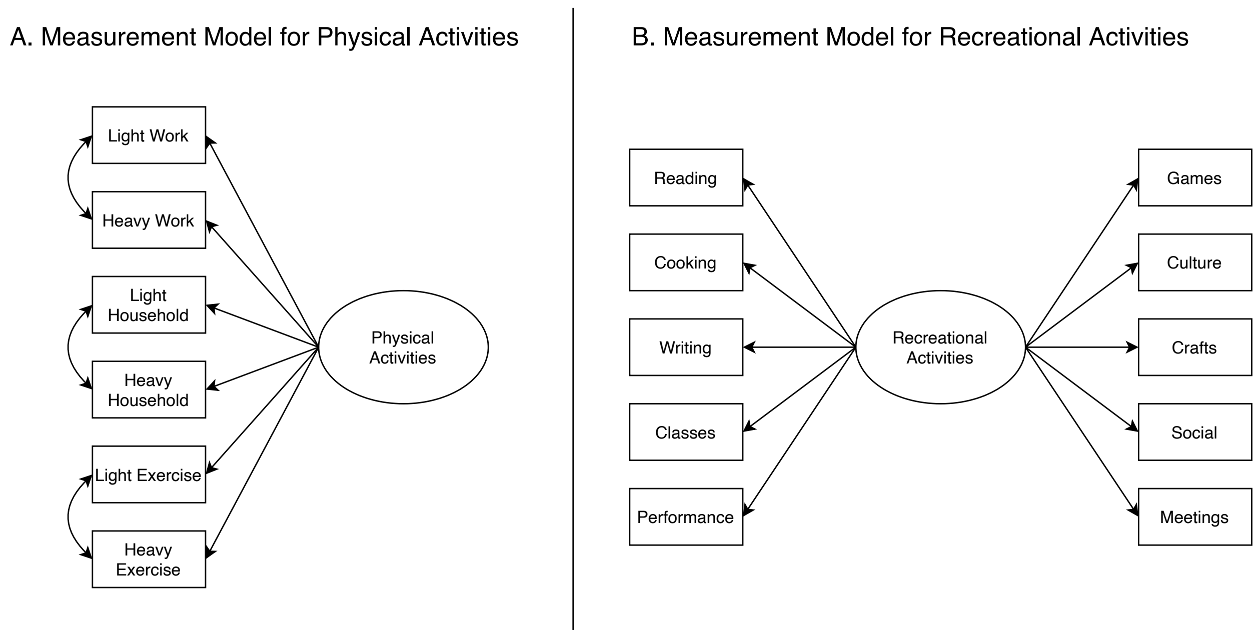

Table 1. Fits of Single Factor Models for Physical and Recreational Activities at Each Time Epoch

Factor	Time	χ^2	*df*	*p*	RMSEA [90% CI]	CFI	TLI
Physical Activities	Age 40	20.33*	6	< .01	0.062 [0.034, 0.093]	.986	.964
Physical Activities	Current	16.29*	6	.01	0.053 [0.023, 0.085]	.985	.963
Recreational Activities	Age 40	162.87*	35	< .01	0.079 [0.067, 0.092]	.849	.806
Recreational Activities	Current	112.32*	35	< .01	0.061 [0.049, 0.074]	.900	.871

Note. RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index; TLI = Tucker-Lewis Index.

Table 2. Fits of Two-Factor Models (Age 40 and Current together) for Physical and Recreational Activities

Factor	χ^2	*df*	*p*	RMSEA [90% CI]	CFI	TLI
Physical Activities	105.13*	41	< .01	0.051 [0.039, 0.063]	.965	.943
Recreational Activities	406.67*	159	< .01	0.051 [0.045, 0.057]	.914	.897

Note. RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index; TLI = Tucker-Lewis Index.

The single factor models for physical activities fit the data well at both time points. The fit of the models to the recreational activities data was suboptimal, especially at age 40, when judged based on CFI and TLI; in contrast, the RMSEA showed reasonably good fit (Table 1). However, when combined into a single two-factor model, the fit of the recreational activities model improved to acceptable standards, and the physical activities model continue to fit well (Table 2). Therefore, we judged the models to be well-fitting enough to warrant continued investigation of their measurement invariance properties and proceeded to test the time-invariance of the two models.

Although not a substantive focus of analysis, we also sought to determine whether self-reported physical and recreational activities – at age 40 and current – could be modeled equally well in participants with and without cognitive impairment. To perform these analyses, we used the two-factor models described above (factors: current and age 40), separately for both physical and recreational activity engagement, and constrained the factor loadings and thresholds to equality (scalar invariance) across two groups: those with normal cognition ($n = 407$) and those with cognitive impairment (clinical diagnosis of MCI or Dementia; $n = 188$). The fits of the scalar invariance models are shown in Table 3. For recreational activities, scalar invariance was largely supported (i.e., no pronounced decrements in the CFI, TLI, and RMSEA fit statistics), aside from a significant chi-square difference test. For physical activities, the scalar invariance model was statistically significantly different from the configural invariance model based on the chi-square difference test and the decrease in CFI was more pronounced than conventional guidelines recommend (i.e., -.01). In contrast, the change in TLI and the change in RMSEA favored the scalar invariance model. Because of the mixed findings when comparing the scalar invariance model to the configural model, and because the absolute fits of the scalar invariance models were reasonably good (Table 3), we continued with our planned analyses.

Table 3. Fits of Scalar Invariance Two-Factor Models (factors for Age 40 and Current) for Physical and Recreational Activities across Cognitively Normal and Cognitively Impaired Groups

Activity Type	χ^2	*df*	*p*	RMSEA [90% CI]	CFI	TLI	$\Delta\chi^2$	Δdf	*p*	ΔCFI	ΔTLI	ΔRMSEA
Physical	227.21*	122	< .01	0.054 [0.043, 0.065]	.937	.932	77.58*	40	< .01	-.019	.002	.001
Recreational	613.96*	388	< .01	0.045 [0.038, 0.051]	.914	.916	90.95*	70	< .05	.003	.022	.005

Note. RMSEA = root mean square error of approximation; CI = confidence interval; CFI = comparative fit index; TLI = Tucker-Lewis Index. Delta statistics are reported in comparison to the configural model.

To investigate the time-invariance of the physical and recreational activities models, we applied a series of constraints on the loading and threshold parameters, one indicator at a time. For example, the physical activities factor has six indicators, so we generated six different models, with each model using one of the six indicators as the linking item. When an item is used as a linking item, its loadings are constrained to be equal to one another across age epochs, and its thresholds are also constrained to be equal to one another across age epochs. The loadings and thresholds of the non-linking items are freely estimated. After examining the fit of all six linking items (see Table 3), we identified the best linking item as the one whose model had the smallest χ^2 value. If the overall model fit was good, as judged by standard fit statistics, then the first linking item identified was considered to have loadings and thresholds that are invariant to the time epoch (age 40 and current).

After identifying the first linking item, we repeated that same process to determine whether a second linking item could be added to the model without causing a decrease in fit. If the change in χ^2 ($\Delta\chi^2$) was non-significant, then the second item was also considered to have loadings and thresholds that are invariant to time epoch.

This iterative process was repeated until a significant $\Delta\chi^2$ was observed. If the $\Delta\chi^2$ statistic was never significant at any step, then full invariance of loadings and thresholds would be achieved. In contrast, if the $\Delta\chi^2$ statistic reached significance, then any item not already identified as invariant would be considered non-invariant, resulting in a measurement model with partial invariance. A partially invariant model contains some items whose loadings and thresholds may be constrained to be equal over time, but requires other loadings and thresholds to be freely estimated to provide a valid estimate of the trait being measured and to estimate the amount of change in that trait over time.

The results for this measurement invariance testing procedure are shown in Tables 3-5 (Physical Activities). Repeating that same process for recreational activities yielded the results shown in Table 6.

Table 4. Step 1 of Physical Activities Measurement Invariance Testing

	Item						Model Fit		
Model	Light Work	Heavy Work	Light Housework	Heavy Housework	Light Exercise	Heavy Exercise	χ^2	CFI	RMSEA
P1a							**108.321***	**0.964**	**0.05**
P1b							141.977*	0.946	0.06
P1c							108.541*	0.964	0.05
P1d							118.755*	0.959	0.053
P1e							119.255*	0.958	0.053
P1f							112.426*	0.962	0.05

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation. Blue boxes indicate when an item is used as a linking item (its loadings are constrained to be equal across the two time epochs, and its thresholds are also constrained to be equal across the two age epochs). Gray boxes indicate that an item's parameters are freely estimated. Bolded fit statistics are used to highlight the best fitting model.

* $p < .05$

Table 5. Step 2 of Physical Activities Measurement Invariance Testing

	Item						Model Fit			
Model	Light Work	Heavy Work	Light Housework	Heavy Housework	Light Exercise	Heavy Exercise	χ^2	$\Delta\chi^2$	CFI	RMSEA
P2a							148.32	9.4	.945	0.058
P2b							152.86	37.4*	.941	0.061
P2c							**123.05**	**7.7**	**.959**	**0.051**
P2d							249.87	89.3*	.888	0.083
P2e							137.64	21.7*	.950	0.055

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation. Blue boxes indicate when an item is used as a linking item (its loadings are constrained to be equal across the two time epochs, and its thresholds are also constrained to be equal across the two age epochs). Gray boxes indicate that an item's parameters are freely estimated. Bolded fit statistics are used to highlight the best fitting model. $\Delta\chi^2$ refers to the change in the $\Delta\chi^2$ value obtained in the previous step. Non-significant $\Delta\chi^2$ values indicate that model fit did not significantly worsen when the additional constraints were added, and thus provides evidence to suggest that the new linking loading and threshold parameters are invariant to time.

* $p < .05$

Table 6. Step 3 of Physical Activities Measurement Invariance Testing

	Item						Model Fit			
Model	Light Work	Heavy Work	Light Housework	Heavy Housework	Light Exercise	Heavy Exercise	χ^2	$\Delta\chi^2$	CFI	RMSEA
P3a							155.16*	32.6*	.944	0.055
P3b							225.70*	109.4*	.905	0.073
P3c							325.43*	165.3	.849	0.092
P3d							**149.67***	**25.2***	**.947**	**0.055**

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation. Blue boxes indicate when an item is used as a linking item (its loadings are constrained to be equal across the two time epochs, and its thresholds are also constrained to be equal across the two age epochs). Gray boxes indicate that an item's parameters are freely estimated. Bolded fit statistics are used to highlight the best fitting model. $\Delta\chi^2$ refers to the change in the $\Delta\chi^2$ value obtained in the previous step. Non-significant $\Delta\chi^2$ values indicate that model fit did not significantly worsen when the additional constraints were added, and thus provides evidence to suggest that the new linking loading and threshold parameters are invariant to time.

* $p < .05$

Table 7. Results of Recreational Activities Measurement Invariance Testing

Model	Linking Item(s)	χ^2	$\Delta\chi^2$	CFI	RMSEA
1	Cultural Activities	409.34*	--	.914	0.051
2	1 + Cooking	431.07*	1.5	.908	0.052
3	2 + Social Activities	425.73*	2.4	.911	0.050
4	3 + Classes	432.87*	9.5	.910	0.050
5	4 + Performance	464.08*	30.8*	.901	0.052

Note. CFI = comparative fit index; RMSEA = root mean square error of approximation.

Based on the results above, partial time invariance was found for the factor loadings and thresholds of both measurement models: physical and recreational activities. In the physical activities model, items with time-invariant loadings and thresholds include light work-related activities and heavy house/yard work activities (see Model P2c in Table 4). In contrast, heavy work-related activities, light house/yard work, light exercise, and vigorous exercise were not invariant to time.

In the recreational activities model, items with time-invariant loadings and thresholds include cultural events, complex cooking, socializing, and taking classes (see Table 6). In contrast, reading, writing, performance arts, games or puzzles, arts and crafts, and attendance at clubs or meetings were not invariant to time.

Additional Details About the Multilevel Modeling Analyses

Estimation of the full structural models regressing cognitive outcomes on activity engagement was performed using the multilevel modeling platform in Mplus. Specific technical details relevant to these analyses include the following. Participant identification number was treated as the CLUSTER variable. WITHIN variables included time, prior evaluation (0 = no, 1 = yes) to account for practice effects, and a Spanish language (0 = no, 1 = yes) x prior evaluation interaction term to account for differential practice effects that have been observed for Spanish-speaking individuals in this cohort. BETWEEN variables included the physical and recreational activity scores (treated as CATEGORICAL), male (0 = no, 1 = yes), age (centered at age 70 years), years of education (centered at 12), Spanish language (0 = no, 1 = yes), referral source (0 = community, 1 = clinic), and binary indicator variables to code for race/ethnicity: Black/African American (0 = no, 1 = yes), Hispanic/Latinx (0 = no, 1 = yes), and other (0 = no, 1 = yes).

The analysis used a TWOLEVEL RANDOM multilevel model with the BAYES estimator (using the GIBBS(PX1) sampler, 2 chains, and 100000 iterations). Point estimates were derived from the median values of the posterior distribution. A PROBIT link function was used by default. On the WITHIN level of the model, we identified random slopes for each of the four cognitive domains (verbal episodic memory, semantic memory, executive functioning, and spatial skills) by regressing the observed cognitive scores on a variable representing time (e.g., VMSL | VM_ST ON TIME;), where VMSL is the random slope for verbal memory, VM_ST is the observed verbal memory scores (standardized) and TIME is number of years after the baseline study visit. To continue with the verbal memory example, in the WITHIN part of the model, VM_ST was regressed on an indicator of prior evaluation – to account for practice effects – and the prior evaluation by Spanish language interaction term to allow practice effects to differ by Spanish-speaking status. VM_ST is also by default the label used for the verbal memory random intercept term. The study's primary hypotheses were all based on the parameters estimated in the BETWEEN part of the model. The 95% credible intervals from the posterior sampling distribution were used to make inferences about population parameters. For variables with residual variances estimated in both the WITHIN and BETWEEN parts of the model (i.e., random intercepts for verbal memory, spatial memory, executive functioning, and spatial skills), we divided each variable's residual variance estimated in the BETWEEN part of the model by the sum of the BETWEEN + WITHIN residual variances to obtain intra-class correlation coefficients, which are reported below.

Table 8. Intra-class correlation coefficients for random effects in two-level multilevel models.

Random Effect	Physical Activities Model	Recreational Activities Model
Verbal Memory Intercept	0.70	0.67
Semantic Memory Intercept	0.76	0.74
Executive Functioning Intercept	0.76	0.73
Spatial Intercept	0.54	0.52