Short Assessment of the Big Five:
Robust Across Survey Methods Except Telephone Interviewing

Authors: Lang, Frieder R., John, Dennis, Lüdtke, Oliver, Schupp, Jürgen, & Wagner, Gert G.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

This file contains supplementary information material including additional analyses conducted in order to support findings reported in the manuscript entitled "*Short Assessment of the Big Five: Robust Across Survey Methods Except Telephone Interviewing"* (Behavior Research Methods, 2011. doi: 10.3758/s13428-011-0066-z).

Table SM1 shows that reported results remained stable after controlling for different sample sizes. Table SM2 displays invariance of factor loadings across assessment methods for the total sample. Small differences between factor loadings are due to standardization. Table SM3 shows factor correlations between NEOAC over assessment methods for the total sample that are considerably lower than factor correlations between NEOAC obtained with confirmatory factor analysis (e.g. Marsh et al., 2010). Tables SM4-6 show that pooling PAPI with CAPI and SELF-Alone with SELF-Not-Alone is justified. Table SM7 reports standardized factor loadings and item intercepts for older adults in the CATI assessment condition. As shown in the table, the Five-Factor-Model is not prevalent in the older adults sample in CATI. Table SM8 displays high latent test-retest stability coefficients for five assessment conditions. Table SM9 shows that all measurement invariance results for the 2005-wave remained essentially stable in the 2009-wave. Manifest test-retest stability coefficients over five years across method groups in young, middle-aged, and older adults are displayed in Table SM10. Table SM11 shows Cronbach`s Alpha for NEOAC across method groups in the total sample. Table SM12 shows mean level differences and standard errors for the manifest scale values in wave 2005 between young, middle-aged, and older adults that are in accordance with the literature (e.g. Donellan & Lucas, 2008). Explained variances for FACE-CATI-SELF in an exploratory factor analysis with Varimax rotation are displayed in Table SM13. Table SM14 shows correlated uniquenesses between the same item measured in wave-2005 and wave-2009. Tables SM15-16 show results for measurement invariance testing across three method conditions within three age-groups in one nine-group comparison (age-by-method).

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

First, we controlled for possible effects of different sample sizes across the three conditions of assessment (FACE, CATI, SELF). We repeated all analyses with equal sample sizes for the total sample (young, middle-aged, and older adults). For this purpose, we randomly draw subsamples of the larger FACE and SELF samples to obtain comparable sample sizes. As shown in Table SM1, the results remained stable.

Table SM1

Summary of Goodness-of-Fit Statistics for Measurement Invariance (FACE-CATI-SELF) With Equal Sample Sizes.

Model	MLR/df	Nfp	CFI	TLI	RMSEA
Configural invariance	517/120	285	.952	.874	.053
Weak measurement invariance	649/220	185	.948	.926	.041
Strong Measurement Invariance	766/240	165	.936	.916	.043
Strict Measurement Invariance	851/270	135	.930	.918	.043

Note. For FACE, *N* = 1,178; for CATI, *N* = 1,178; for SELF, *N* = 1,178. MLR/df = Maximum Likelihood Robust chi-square/degrees of freedom ratio; Nfp = Number of free parameters; CFI = Comparative fit index; TLI = Tucker-Lewis-Index; RMSEA = Root Mean Square Error of Approximation.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM2 shows high convergent and low divergent factor loadings for the strict measurement model over methods in the total sample (comparable to Table 5 in the original manuscript).

Table SM2

Standardized Factor Loadings and Unstandardized Intercepts For the Strict Measurement Invariance Model[a] for FACE-CATI-SELF.

	FACE						CATI						SELF					
I see Myself as Someone Who ...	N	E	O	A	C	iic	N	E	O	A	C	iic	N	E	O	A	C	iic
Worries a lot	.506	-.026	.049	.073	.115	4.73	.513	-.026	.042	.074	.097	4.73	.478	-.027	.044	.074	.113	4.73
Gets nervous easily	.786	.001	.015	-.019	-.018	3.63	.798	.001	.009	-.019	-.015	3.63	.764	.001	.014	-.019	-.018	3.63
Remains calm in tense situations[b]	.513	.001	-.271	-.101	-.076	3.39	.552	.001	-.246	-.109	-.068	3.39	.496	.001	-.251	-.104	-.076	3.39
Is talkative	.054	.656	.086	.087	.165	5.55	.057	.687	.077	.092	.146	5.55	.050	.673	.078	.087	.161	5.55
Is outgoing, sociable	.033	.643	.189	.129	.019	5.12	.035	.668	.167	.136	.016	5.12	.031	.661	.170	.130	.018	5.12
Is reserved[b]	-.099	.563	-.114	-.184	-.086	3.88	-.099	.560	-.096	-.185	-.072	3.88	-.091	.566	-.100	-.181	-.082	3.88
Is original, comes up with new ideas	-.096	.094	.600	-.200	.150	4.54	-.108	.106	.572	-.225	.141	4.54	-.096	.104	.578	-.214	.156	4.54
Values artistic, aesthetic experiences	.031	.030	.483	.077	-.009	4.12	.033	.032	.432	.081	-.008	4.12	.029	.031	.441	.078	-.009	4.12
Has an active imagination	.012	.095	.657	.019	-.048	4.76	.013	.072	.574	.021	-.044	4.76	.012	.071	.614	.020	-.048	4.76
Is sometimes rude to others[b]	-.118	-.018	-.135	.526	.002	5.17	-.122	-.018	-.118	.544	-.002	5.17	-.111	-.018	-.122	.527	-.002	5.17
Has a forgiving nature	.022	.101	.131	.361	.070	5.57	.022	.104	.115	.374	.061	5.57	.020	.103	.118	.359	.068	5.57
Is considerate and kind to almost everyone	.061	.055	.195	.642	.147	5.88	.063	.057	.169	.663	.126	5.88	.057	.057	.175	.641	.142	5.88
Does a thorough job	.004	-.008	-.022	-.081	.815	6.26	.005	-.009	-.021	-.091	.759	6.26	.004	-.008	-.020	-.082	.801	6.26
Tends to be lazy[b]	-.040	.075	-.218	.131	.444	5.87	-.043	.079	-.195	.139	.392	5.87	-.037	.077	-.195	.131	.429	5.87
Does things efficiently	-.030	.009	.168	.010	.606	5.85	-.034	.010	.133	.011	.565	5.85	-.029	.009	.155	.010	.604	5.85

Note. N = Neuroticism; E = Extraversion; O = Openness to Experience; A = Agreeableness; C = Conscientiousness; iic = Item intercepts.

[a] all age groups. [b] item was recoded (inversed).

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM3 shows low factor correlations for the strict measurement invariance model over method groups in the total sample (comparable to results presented in Table 6 of the original manuscript)

Table SM3

Standardized Factor Correlations For the Strict Measurement Invariance Model[a]

	FACE					CATI					SELF				
	N	E	O	A	C	N	E	O	A	C	N	E	O	A	C
N	1.00					1.00					1.00				
E	-.174	1.00				-.193	1.00				-.170	1.00			
O	-.029	.484	1.00			.153	.366	1.00			.032	.378	1.00		
A	-.080	-.049	.061	1.00		.043	-.061	.110	1.00		-.094	-.015	.131	1.00	
C	-.117	.244	.326	.325	1.00	-.003	.118	.189	.238	1.00	-.069	.187	.228	.326	1.00

Note. N = Neuroticism; E = Extraversion; O = Openness to Experience; A = Agreeableness; C = Conscientiousness.
[a] all age groups.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Tables SM4, SM5, and SM6 show results for the SELF-Alone/SELF-Not-Alone, PAPI/CAPI, and PAPI-without self-completion/CAPI comparisons. Based on the Chen-criteria (CFI change is less than .01 and/or RMSEA change is less than .15) all three comparisons show strict measurement invariance in young, middle-aged, and old adults. Therefore, pooling PAPI with CAPI on the one hand, and pooling SELF-Alone with SELF-Not-Alone is justified. PAPI (paper-assisted personal interviewing) and CAPI (computer-assisted personal interviewing) reflect different face-to-face interviewing situations (see Method description in the manuscript).

Table SM4

Measurement Invariance Between Two Self-Administered Assessment Conditions (SELF-Alone vs. Self-Not-Alone) Across Age Groups[a]: Summary of Goodness-of-Fit Statistics

	Young adults					Middle-aged adults					Older adults				
Model	MLR/df	Nfp	CFI	TLI	RMSEA	MLR/df	Nfp	CFI	TLI	RMSEA	MLR/df	Nfp	CFI	TLI	RMSEA
Configural Invariance	281/80	190	.974	.932	.040	361/80	190	.966	.911	.045	261/80	190	.953	.877	.054
Weak Measurement Invariance	310/130	140	.977	.962	.030	378/130	140	.970	.952	.033	318/130	140	.951	.921	.043
Strong Measurement Invariance	318/140	130	.977	.965	.029	382/140	130	.970	.955	.032	342/140	130	.948	.921	.043
Strict Measurement Invariance	336/155	115	.977	.968	.027	397/155	115	.971	.960	.030	394/155	115	.938	.916	.045

Note. Young adults: for SELF-Alone, *N* = 1,075; for SELF-Not-Alone, *N* = 2,053. Middle-aged adults: for SELF-Alone, *N* =1,043; for SELF-Not-Alone, *N* = 2,382. Old adults: for SELF-Alone, *N* = 533; for SELF-Not-Alone, *N* = 999. MLR/df = Maximum Likelihood Robust chi-square/degrees of freedom ratio; Nfp = Number of free parameters; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.

[a]Factor variances and factor covariances are invariant for all three age groups.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM5.

Measurement Invariance Between PAPI and CAPI Assessment Conditions Among Age Groups[a]*: Summary of Goodness-of-Fit Statistics*

	Young adults					Middle-aged adults					Older adults				
Model	MLR/df	Nfp	CFI	TLI	RMSEA	MLR/df	Nfp	CFI	TLI	RMSEA	MLR/df	Nfp	CFI	TLI	RMSEA
Configural Invariance	355/80	190	.965	.909	.047	321/80	190	.976	.937	.038	353/80	190	.973	.929	.041
Weak Measurement Invariance	401/130	140	.966	.945	.036	396/130	140	.973	.957	.031	400/130	140	.973	.957	.032
Strong Measurement Invariance	435/140	130	.963	.944	.036	444/140	130	.969	.954	.032	445/140	130	.970	.955	.033
Strict Measurement Invariance	468/155	115	.960	.946	.036	471/155	115	.968	.957	.031	469/155	115	.969	.958	.032

Note. Young adults: for PAPI, *N* = 1,883; for CAPI, *N* = 1,281. Middle-aged adults: for PAPI, *N* = 2,359; for CAPI, *N* = 1,772. Old adults: PAPI, *N* = 2,290; for CAPI, *N* = 1,681. MLR/df = Maximum Likelihood Robust chi-square/degrees of freedom ratio; Nfp = Number of free parameters; CFI = Comparative fit index; TLI = Tucker-Lewis-Index; RMSEA = Root Mean Square Error of Approximation.

[a]Factor variances and factor covariances are invariant for all three age groups.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM6.

Measurement Invariance Between PAPI-Without Self-Completion and CAPI Assessment Conditions Among Age Groups [a]: Summary of Goodness-of-Fit Statistics

	Young adults					Middle-aged adults					Older adults				
Model	MLR/df	NFP	CFI	TLI	RMSEA	MLR/df	NFP	CFI	TLI	RMSEA	MLR/df	NFP	CFI	TLI	RMSEA
Configural Invariance	326/80	190	.963	.904	.048	292/80	190	.976	.936	.039	314/80	190	.975	.934	.040
Weak Measurement Invariance	380/130	140	.963	.940	.038	381/130	140	.971	.953	.033	373/130	140	.974	.958	.032
Strong Measurement Invariance	443/140	130	.955	.933	.040	451/140	130	.964	.946	.036	419/140	130	.970	.955	.033
Strict Measurement Invariance	503/155	115	.948	.930	.041	480/155	115	.963	.949	.034	447/155	115	.969	.958	.032

Note. Young adults: for PAPI-without self-completion, *N* = 1,397; for CAPI, *N* = 1,281. Middle-aged adults: for PAPI-without self-completion, *N* = 1,787; for CAPI, *N* = 1,772. Old adults: PAPI-without self-completion, *N* = 1,953; for CAPI, *N* = 1,681. MLR/df = Maximum Likelihood Robust chi-square/degrees of freedom ratio; NFP = Number of free Parameters; CFI = Comparative fit index; TLI = Tucker-Lewis-Index; RMSEA = Root Mean Square Error of Approximation.
[a]Factor variances and factor covariances are invariant for all three age groups.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM7 reports standardized factor loadings and item intercepts (IIC) for older adults in the CATI assessment condition. As shown in the table, factor loadings are not consistent with the five-factor model. There were high cross-loadings for older adults for items of N, E, A, and C.

Table SM7

Standardized Factor Loadings and Unstandardized Intercepts (IIC) for Older Adults (CATI)

I see Myself as Someone Who ...	N	E	O	A	C	iic
Worries a lot (N)	0.240	-0.009	0.077	-0.102	0.030	4.98
Gets nervous easily (N)	**1.201**	0.007	-0.004	-0.003	0.007	3.90
Remains calm in tense situations (N-)	0.309	-0.137	-0.023	0.074	-0.176	3.15
Is talkative (E)	0.005	0.541	0.176	-0.022	0.144	5.86
Is outgoing, sociable (E)	0.007	0.739	0.050	0.010	0.151	5.53
Is reserved (E-)	-0.015	0.560	-0.060	-0.029	**-0.434**	3.39
Is original, comes up with new ideas (O)	-0.123	0.123	0.535	-0.089	0.054	4.96
Values artistic, aesthetic experiences (O)	0.020	0.026	0.369	0.099	-0.004	5.16
Has an active imagination (O)	0.022	-0.007	0.735	0.018	-0.021	5.23
Is sometimes rude to others (A-)	-0.001	0.001	0.000	**1.808**	-0.001	4.45
Has a forgiving nature (A)	0.048	**0.294**	0.087	**0.031**	0.059	5.96
Is considerate and kind to almost everyone (A)	0.047	0.064	0.141	**0.137**	**0.437**	6.00
Does a thorough job (C)	0.013	0.042	-0.043	-0.012	0.603	6.24
Tends to be lazy (C-)	-0.176	0.104	**-0.298**	0.080	0.290	5.36
Does things efficiently (C)	0.027	0.111	0.041	-0.024	0.684	5.96

Note. N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness; iic = Item Intercepts.
Parameters printed in bold indicate unexpected/inacceptable (too low/high) factor loadings.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM8 shows high latent test-retest stability coefficients over five years across five method groups in young, middle-aged, and old adults (cmparable to Table 7 of the original manuscript, in which some of the method contexts below were concatenated).

Table SM8

Five-Year Latent Factor Score Test-Retest Stability Coefficients Differentiating the Two Face-To-Face Interviewing Conditions (PAPI, CAPI) and the Two Self-Administration Conditions (SELF-Alone, SELF-Not-Alone)

	Young adults					Middle-aged adults					Older adults				
	N	E	O	A	C	N	E	O	A	C	N	E	O	A	C
Change[a]	.700	.753	.679	.763	.598	.790	.857	.725	.777	.614	.729	.707	.673	.606	.600
PAPI	.821	.685	.727	.730	.743	.818	.843	.684	.866	.748	.783	.737	.769	.704	.676
CAPI	.854	.770	.640	.684	.696	.749	.849	.706	.830	.685	.790	.820	.741	.793	.733
SELF-Alone	.888	.895	.777	.827	.776	.895	.933	.821	.903	.693	.878	.791	.690	.736	.521
SELF -Not-Alone	.812	.873	.761	.744	.727	.934	.891	.844	.886	.767	.881	.954	.785	.908	.646

Note. Young adults: Change, *N* = 1,151; PAPI, *N* = 784; CAPI, *N* = 617; SELF-Alone, *N* = 677; SELF with Interviewer, *N* = 1,003. Middle-aged adults: Change, *N* = 1,266; PAPI, *N* = 1,131; CAPI, *N* = 1,033; SELF-Alone, *N* = 688; SELF-Not-Alone, *N* = 1,385. Old adults: Change, *N* = 760; PAPI, *N* = 1,128; CAPI, *N* = 951; SELF-Alone, *N* = 341; SELF-Not-Alone, *N* = 544.
N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness.
[a]Change condition includes participants that changed between PAPI, CAPI, SELF-Alone, and SELF-Not-Alone.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM9 summarizes the results for the Measurement Invariance Testing at Time 2 (2009-wave). Five method groups were tested against each other in young, middle-aged, and old adults: (1) PAPI0509: Participants responded to the BFI via PAPI in 2005 and 2009; (2) CAPI0509: Participants responded to the BFI via CAPI in 2005 and 2009; (3) SELF-Alone0509: Participants responded to the BFI via SELF-Alone in 2005 and 2009; (4) SELF-Not-Alone 0509: Participants responded to the BFI via SELF-Not-Alone in 2005 and 2009; (5) Changed Method: Participants in the Change-Method Group responded to the BFI using one of the four aforementioned methods in 2005 and one of the remaining three methods in 2009. Based on the Chen-criteria (CFI change is less than .01 and/or RMSEA change is less than .15) all measurement invariance results for the 2005-wave remained essentially stable in the 2009-wave.

Table SM9

Summary of Goodness-of-Fit Statistics for Measurement Invariance in Young, Middle-aged, and Old Adults at Time 2 in 2009 (PAPI0509 - CAPI0509 – SELF-Alone 0509 – SELF-Not-Alone0509- Changed Method)[a]

	Young adults					Middle-aged adults					Older adults				
Model	MLR/df	NFP	CFI	TLI	RMSEA	MLR/df	NFP	CFI	TLI	RMSEA	MLR/df	NFP	CFI	TLI	RMSEA
Configural Invariance	563/200	475	0.966	0.91	0.046	625/200	475	0.967	0.913	0.044	603/200	475	0.956	0.885	0.052
Weak Measurement Invariance	775/400	275	0.965	0.953	0.033	885/400	275	0.962	0.950	0.033	841/400	275	0.952	0.937	0.038
Strong Measurement Invariance	955/440	235	0.951	0.942	0.037	978/440	235	0.958	0.950	0.033	929/440	235	0.947	0.937	0.039
Strict Measurement Invariance	1072/500	175	0.946	0.943	0.037	1109/500	175	0.953	0.950	0.033	1117/500	175	0.933	0.930	0.041

Note. MLR/df = Maximum Likelihood Robust chi-square/degrees of freedom ratio; NFP = Number of free Parameters; CFI = Comparative fit index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation.
[a]Factor variances and factor covariances are invariant for all three age groups.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM10 displays strong test-retest stability coefficients over five years across method groups in young, middle-aged, and old adults.

Table SM 10

Five year Test-Retest Stability Coefficients on the Manifest Scale Level (With Identical Methods For 2005 to 2009 Or With Method Change)

	Young adults					Middle-aged adults					Older adults				
	N	E	O	A	C	N	E	O	A	C	N	E	O	A	C
Method change[a]	.469	.556	.512	.485	.459	.517	.590	.577	.478	.418	.475	.488	.532	.386	.429
PAPI	.535	.527	.568	.488	.553	.559	.604	.571	.566	.534	.515	.549	.572	.454	.488
CAPI	.581	.533	.502	.485	.524	.533	.600	.564	.534	.506	.578	.570	.589	.490	.552
SELF-Alone	.619	.692	.647	.555	.581	.640	.699	.671	.579	.501	.560	.597	.609	.507	.439
SELF-Not-Alone	.610	.702	.659	.519	.554	.681	.685	.668	.588	.557	.652	.695	.671	.624	.548

Note. Young adults: for Change, *N* = 1,151; for PAPI, *N* = 784; for CAPI, *N* = 617; for SELF-Alone, *N* = 677; for SELF-Not-Alone, *N* = 1,003. Middle-aged adults: for Change, *N* = 1,266; for PAPI, *N* = 1,131; for CAPI, *N* = 1033; for SELF-Alone, *N* = 688; for SELF-Not-Alone, *N* = 1,385. Old adults: for Change, *N* = 760; for PAPI, *N* = 1,128; for CAPI, *N* = 951; for SELF-Alone, *N* = 341; for SELF-Not-Alone, *N* = 544. N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness.

[a]Change condition includes participants that changed between PAPI, CAPI, SELF-Alone, and SELF-Not-Alone.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM11 shows zero-order Cronbach`s Alpha across method groups in the total sample.

Table SM 11

Cronbach's Alpha of Manifest Item Scores of the Five Personality Scales Across Methods (wave 2005)

	N	E	O	A	C
PAPI	.607	.632	.666	.508	.632
CAPI	.603	.666	.607	.498	.597
SELF-Alone	.560	.674	.594	.493	.591
SELF-Not-Alone	.623	.679	.617	.507	.598
CATI	.584	.642	.575	.451	.487
TOTAL SAMPLE	.603	.656	.629	.501	.601

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM12 shows mean levels and standard errors for the manifest scale values in wave 2005. For Openness and Extraversion we found the well documented effect of linear decrease over the lifespan (McCrae et al., 1999; Roberts et al., 2006). Additionally, comparable to the results reported by Donellan and Lucas (2008) and Terracciano et al. (2005), we found an increase of Conscientiousness from young to middle age, but a decrease of Conscientiousness from middle age to old age. Furthermore, in line with previous research (e.g., McCrae et al., 1999; Roberts et al., 2006; Terracciano et al., 2005) we found a linear increase of Agreeableness from young to old adulthood. Results concerning mean differences in Neuroticism are somewhat mixed as some studies suggest a decrease of Neuroticism over the life span (e.g., McCrae et al., 1999), while others show high mean level stability of Neuroticism after the third decade (Roberts et al., 2006). In our study that included only adults older than 20 years, mean levels of Neuroticism slightly increased with age. This effect is consistent with other studies that analyzed the SOEP dataset (Donellan & Lucas, 2008).

Table SM12

Manifest Scale Means of Young, Middle-Aged and Older Adults

	Young adults		Middle-aged adults		Older adults	
	M	SE	M	SE	M	SE
N	3.89	0.015	3.97	0.014	4.05	0.016
E	4.95	0.014	4.83	0.013	4.68	0.015
O	4.61	0.014	4.55	0.013	4.37	0.017
A	5.38	0.012	5.44	0.011	5.55	0.013
C	5.80	0.011	6.03	0.001	5.95	0.012

Note. N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM13 shows results of an Exploratory Factor Analysis of the BFI-S across all three method contexts. This analysis was conducted in order to allow comparability of the instrument when applied in research without applying ESEM for the statistical analysis of raw data. Table SM13 shows that variances accounted for in the EFA are comparable across all three age groups. No differences were observed with regard to variances accounted for by the five constructs when comparing the three assessment contexts.

Table SM13

Explained Variance (Percent) For FACE-CATI-SELF[a] in an EFA With VARIMAX Rotation

	Young adults	Middle-aged adults	Older adults[a]
N	7.77	7.83	7.40
E	9.04	7.56	6.96
O	9.09	10.00	12.31
A	7.12	7.44	7.93
C	9.56	9.06	8.28

Note. N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness.
[a]For older adults: only highly educated participants of the CATI study were included.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM14 shows test-retest correlated uniqunesses between the same item measured 2005 and 2009.

Table SM 14

Five-Year Test-Retest Correlated Uniquenesses between the same item measured 2005 and 2009 .

	Young adults			Middle-aged adults			Older adults		
I see Myself as Someone Who ...	CHANGE	FACE	SELF	CHANGE	FACE	SELF	CHANGE	FACE	SELF
Worries a lot	.277	.332	.395	.349	.341	.427	.281	.358	.433
Gets nervous easily	.176	.087	.186	.106	.121	.056	-.087	.023	-.118
Remains calm in tense situations[a]	.218	.138	.298	.296	.198	.311	.247	.170	.314
Is talkative	.177	.211	.347	.245	.212	.349	.223	.241	.165
Is outgoing, sociable	.249	.286	.306	.236	.254	.325	.293	.196	.289
Is reserved[a]	.128	.262	.289	.147	.232	.268	.207	.200	.281
Is original, comes up with new ideas	.217	.127	.256	.146	.244	.337	.270	.123	.221
Values artistic, aesthetic experiences	.396	.420	.540	.379	.397	.509	.399	.402	.510
Has an active imagination	.211	.157	.433	.354	.259	.382	.225	.243	.326
Is sometimes rude to others[a]	.296	.369	.377	.337	.374	.333	.287	.347	.455
Has a forgiving nature	.286	.254	.379	.322	.293	.418	.162	.255	.390
Is considerate and kind to almost everyone	-.011	-.103	-.003	-.264	-.082	-.040	-.132	-.051	.040
Does a thorough job	.081	.001	.155	.197	.112	.171	.159	.084	.286
Tends to be lazy[a]	.295	.356	.412	.276	.311	.340	.286	.321	.344
Does things efficiently	.064	.214	.173	.149	.203	.194	.109	.159	.152

Note. Young adults: Change, *N* = 709; FACE, *N* = 1,720; SELF, *N* = 1,803. Middle-aged adults: Change, *N* = 763; FACE, *N* = 2,589; SELF, *N* = 2,151. Older adults: Change, *N* = 351; FACE, *N* = 2,459; SELF, *N* = 914.

[a]Item was recoded (inversed).

Tables SM15 and Tables SM16 report on measurement invariance testing across method conditions within three age-groups within one multi-group comparision (age-by-method). In this step of additional analyses, we repeated measurement invariance testing across method groups (FACE, CATI, SELF) within age groups (young, middle-aged, old) in one multi-group comparison across nine (age-by-method) groups to parallel the results for the method invariance testing within age groups. This multi-age-multi-method comparison simultaneously tests measurement invariance over age groups and methods groups (see Tables SM15 and SM16 below).

First, we tested configural invariance over all nine groups. Again, configural invariance was not supported in the Old age/CATI group. Therefore, we included only highly educated older participants in CATI in the following analyses. Table SM15 summarizes the goodness-of-fit statistics for the invariance models in the nine groups.

Table SM15

Summary of Goodness-of-Fit Statistics for Measurement Invariance for Age By Method[a]

	AGE (young, middle-aged, old) x METHOD (FACE, CATI, SELF)[a]				
Model	MLR/df	Nfp	CFI	TLI	RMSEA
Configural Inv.	1923/360	855	.969	.919	.044
Weak Invariance	2701/760	455	.962	.952	.034
Strong Invariance	4055/840	375	.936	.928	.041
Strong Invariance between Method Groups	3138/820	395	.954	.947	.035
Partial Strong Invariance	3211/830	385	.953	.946	.036
Strict Invariance	4156/950	265	.936	.937	.039
Strict Invariance between Method Groups	3617/920	295	.947	.945	.036
Partial Strict Invariance	3783/944	271	.944	.944	.036
Factor Means Inv.	4976/984	231	.921	.924	.042

Note. Total Sample: for FACE, *N* = 11,266; for CATI, *N* = 1,178; for SELF, *N* = 8,085. Young adults: for FACE, *N* = 3,164; for CATI, *N* = 527; for SELF, *N* = 3,128. Middle-aged adults: for FACE, *N* = 4,131; for CATI, *N* = 426; for SELF, *N* = 3,425. Old adults: for FACE, *N* = 3,971; for CATI, *N* = 150; for SELF, N = 1,532. MLR/df = Maximum Likelihood Robust chi-square / degrees of freedom ratio; *Nfp* = Number of free parameters; CFI = Comparative fit index; TLI = Tucker-Lewis-Index; RMSEA = Root Mean Square Error of Approximation.
[a]For old adults: only highly educated participants of the CATI study were included.

As Table SM15 shows, fit indices showed good model fit of the configural model and the weak measurement invariance model. However, the difference in CFI values for the strong measurement invariance model as compared to the weak measurement invariance model was above .01 (.936 vs. .962). Also the TLI (.928 vs. .952) and RMSEA (.041 vs. .034) values changed. To rule out the possibility that this lack of strong measurement invariance is due to differential item functioning between method groups we repeated strong measurement invariance testing with constrained item intercepts over method groups and freed item intercepts over age groups. We obtained improved model fit for this strong invariance model between method groups. This finding parallels the results for the within age group testing. However, this multi-age-multi-method comparison pointed to differential item functioning of the BFI-S with regard to age groups (cf. Marsh et al., under review).

In order to directly test whether differential item functioning was due to age or method groups we tested the following partial strong invariance model: The partial invariance model had all 15 item intercepts fixed across method groups, 10 item intercepts fixed across age groups and 5 item intercepts (based on modification indices) freed across age groups. When allowing item intercepts of these 5 items to vary across age group (but not across method group) partial strong invariance was supported. Differences in fit indices between the strict measurement invariance model and the partial strong invariance model revealed a CFI change slightly above .01 (.936 vs. .953). Also, TLI (.937 vs. .946) and RMSEA (.039 vs. .036) differed slightly. However, testing strict measurement invariance only across method groups confirms the assumption of equal factor loadings, equal item intercepts, and equal item uniqueness across method groups. However, this multi-age-multi-method comparison pointed to differences in item uniquenesses of the BFI-S with regard to age groups (cf. Marsh et al., under review). To directly test whether different item uniquenesses were prevalent between age or method groups we tested the following partial strict invariance model: The partial invariance model had all 15 item uniquenesses fixed across method groups, 12 item uniquenesses fixed across age groups and 3 item uniquenesses (based on modification indices) freed across age groups. After we allowed item uniquenesses of these 3 items to vary across age group (but not across method groups) partial strict invariance was supported. In a last step, we tested for invariance of factor means. However, comparison of fit indices between partial strict invariance model and the model with constrained factor means pointed to variance of factor means. In conclusion, as shown in Table SM16, the data show expected age differences on factor means with regard to the BFI-S.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

Table SM16:

Standardized Differences of Factor Means between FACE-CATI-SELF for the Partial Strict Measurement Invariance Model for Young, Middle-aged, and Older adults[a]

	Young adults			Middle-aged adults			Older adults		
	FACE	CATI	SELF	FACE	CATI	SELF	FACE	CATI	SELF
N	0	.081	.176	.092	.217	.256	.240	.271	.433
E	0	.010	-.075	-.081	-.072	-.207	-.150	.028	-.273
O	0	.370	-.044	-.188	.374	-.247	-.453	.338	-.371
A	0	-.195	-.196	.079	-.070	-.096	.351	.006	-.032
C	0	-.126	-.198	.190	-.012	.029	-.015	-.098	-.183

Note. N = Neuroticism; E = Extraversion; O = Openness; A = Agreeableness; C = Conscientiousness.
[a]For old adults: only highly educated participants of the CATI study were included.

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

M*plus* Input Files for Measurement Invariance Testing

```
title:          FACE_CATI_SELF (configural invariace);
data:           file is facecatiself_mplus.dat;
variable:       names are v1 v2 v4 v5 v6 v8 v9 v10 v11 v13 v14 v3in v7in v12in v15in method;
                grouping is method (0 = FACE 1 = CATI 2 = SELF);
analysis:       Type = general;
                Estimator = MLR;
                Information = observed;
                Rotation = quartimin(oblique);
model:          f1-f5 by v1-v15in (*1);
                [f1-f5@0];
model cati:     f1-f5 by v1-v15in (*1);
                [v1-v15in];
model self:     f1-f5 by v1-v15in (*1);
                [v1-v15in];
output:         standardized sampstat residual modindices tech1 tech4;
```

BEHAVIOR RESEARCH METHODS, 2011. doi: 10.3758/s13428-011-0066-z

```
title:          FACE_CATI_SELF (weak measurement invariace);
data:           file is facecatiself_mplus.dat;
variable:       names are v1 v2 v4 v5 v6 v8 v9 v10 v11 v13 v14 v3in v7in v12in v15in method;
                grouping is method (0 = FACE 1 = CATI 2 = SELF);
analysis:       Type = general;
                Estimator = MLR;
                Information = observed;
                Rotation = quartimin(oblique);
model:          f1-f5 by v1-v15in (*1);
                [f1-f5@0];
model cati:     [v1-v15in];
model self:     [v1-v15in];
output:         standardized sampstat residual modindices tech1 tech4;

title:          FACE_CATI_SELF (strong measurement invariace);
data:           file is facecatiself_mplus.dat;
variable:       names are v1 v2 v4 v5 v6 v8 v9 v10 v11 v13 v14 v3in v7in v12in v15in method;
                grouping is method (0 = FACE 1 = CATI 2 = SELF);
analysis:       Type = general;
                 Estimator = MLR;
                Information = observed;
                Rotation = quartimin(oblique);
model:          f1-f5 by v1-v15in (*1);
output:         standardized sampstat residual modindices tech1 tech4;
```

```
title:      FACE_CATI_SELF (strict measurement invariace);
data:       file is facecatiself_mplus.dat;
variable:   names are v1 v2 v4 v5 v6 v8 v9 v10 v11 v13 v14 v3in v7in v12in v15in method;
            grouping is method (0 = FACE 1 = CATI 2 = SELF);
analysis:   Type = general;
            Estimator = MLR;
            Information = observed;
            Rotation = quartimin(oblique);
model:      f1-f5 by v1-v15in (*1);
            v1 (2); v2 (3); v4 (4); v5 (5); v6 (6); v8 (7); v9 (8); v10 (9); v11 (10); v13 (11); v14 (12);
            v3in (13); v7in (14); v12in (15); v15in (16);
output:     standardized sampstat residual modindices tech1 tech4;
```

```
title       FACE_CATI_SELF (strict measurement invariace with fixed factor means: factor means
            invariance);
data:       file is facecatiself_mplus.dat;
variable:   names are v1 v2 v4 v5 v6 v8 v9 v10 v11 v13 v14 v3in v7in v12in v15in method;
            grouping is method (0 = FACE 1 = CATI 2 = SELF);
analysis:   Type = general;
            Estimator = MLR;
            Information = observed;
            Rotation = quartimin(oblique);
model:      f1-f5 by v1-v15in (*1);
            v1 (2); v2 (3); v4 (4); v5 (5); v6 (6); v8 (7); v9 (8); v10 (9); v11 (10); v13 (11); v14 (12);
            v3in (13); v7in (14); v12in (15); v15in (16);
            [f1-f5@0];
output:     standardized sampstat residual modindices tech1 tech4;
```

References (for SIM only)

Donnellan, M. B., & Lucas, R. E. (2008). Age differences in the Big Five across the life span: Evidence from two national samples. *Psychology and Aging, 23,* 558-566.

Marsh, H. W., Lüdtke, O., Muthén, B., Asparouhov, T., Morin, A. J. S., Trautwein, U., & Nagengast, B. (2010). A new look at the big-five factor structure through exploratory structural equation modeling. *Psychological Assessment, 22.* 471-491.

Marsh, H. W., Nagengast, B., & Morin, A. J. S. (under review). Measurement invariance of big-five factors over the lifespan: ESEM tests of gender, age, plasticity, maturity, and La Dolce Vita effects. *Manuscript under review.*

McCrae, R. R., Costa, P. T.,Jr., Pedroso de Lima, M., Simoes, A., Ostendorf, F., Angleitner, A., et al. (1999). Age differences in personality across the adult life span: Parallels in five cultures. *Developmental Psychology, 35,* 466–477.

Roberts, B. W., Walton, K. E., & Viechtbauer, W. (2006). Patterns of mean-level change in personality traits across the life course: A meta-analysis of longitudinal studies. *Psychological Bulletin, 132,* 1-25.

Terracciano, A., McCrae, R. R., Brant, L. J., & Costa, P. T. Jr. (2005). Hierarchical linear modeling analyses of the NEO-PI-R scales in the Baltimore Longitudinal Study of Aging. *Psychology and Aging, 20,* 493–506.