Supplementary material

Content

Methods supplement 2

Results supplement 4

Supplementary references 6

Supplementary Tables 7

Supplementary Table S1. 7

Supplementary Table S2 8

Supplementary Table S3. 9

Supplementary Table S4. 10

Supplementary Figures 11

Supplementary Figure S1 11

Supplementary Figure S2 12

Supplementary Figure S3 13

Methods supplement

A special case of "MIMIC" (multiple indicators multiple causes) and "OTL" (opportunity to learn) models (MacIntosh & Hashim 2002; Muthén 1989; Muthén et al. 1991) expresses a latent overall trait (a factor score distribution) as

$$\eta = \gamma z + \zeta \,, \tag{1}$$

where z is the covariate of sex (e.g. z = 1 for men and 0 for women) and ζ is a normally distributed residual with zero mean and variance ψ, where ζ is independent of z. If the constant γ is not 0, sex has a mean difference effect on the latent ASPD factor η. In addition, sex can have an independent effect β_j on the latent liability to endorse the criterion j (i.e., threshold), denoted y_j^*, with

$$y_j^* = \lambda_j \eta + \beta_j z + \varepsilon_j \,, \tag{2}$$

where λ_j is factor loading for item j and ε_j is normally distributed criterion residual with zero mean that is independent of η and z. The term $\beta_j z$ represents a criterion-specific sex difference in mean. Because we standardize the conditional variance of y_j^* given z to 1, the variance of ε_j is $1 - \lambda_j^2 \psi$.

This model translates to (2-parameter normal ogive) Item Response Theory model with "difficulty" (or criterion severity as an indicator of latent ASPD) and "discriminability" (or criterion sensitivity to latent ASPD) parameters

$$b_j = [(\tau_j - \beta_j z)\lambda_j^{-1} - \mu_\eta]\sigma_{\eta\eta}^{-1/2} \,, \tag{3}$$

$$a_j = \lambda_j (1 - \lambda_j^2 \psi)^{-1/2} \sigma_{\eta\eta}^{1/2} \,, \tag{4}$$

where τ_j is the threshold that y_j^* needs to exceed for the item j to be endorsed, and herein the latent-factor mean μ_η is set to 0 and the latent-factor variance $\sigma_{\eta\eta}$ is set to 1.

A number of observations can be made from these equations. A sex difference in factor loadings (λ_j) obviously has a pervasive effect on the psychometric structure, but an effect of a sex difference in thresholds is more specific and absent for criterion sensitivity (discriminability). The conditional latent criterion severities (difficulties) can be written as

$$(b_j \mid z = 1) = [(\tau_j - \beta_j)\lambda_j^{-1} - \mu_\eta]\sigma_{\eta\eta}^{-1/2}, \tag{5}$$

and

$$(b_j \mid z = 0) = [\tau_j \lambda_j^{-1} - \mu_\eta]\sigma_{\eta\eta}^{-1/2}. \tag{6}$$

Thus, a sex difference in criterion severity, b_j, is captured by the difference β_j in male and female thresholds, $\tau_j^* = \tau_j - \beta_j$ and τ_j, respectively. Testing for the presence of sex difference can take either one of two equivalent approaches: estimating τ_j and β_j and testing the null hypothesis $\beta_j = 0$ or estimating τ_j^* and τ_j and testing the null hypothesis $\tau_j^* = \tau_j$. A similar interpretation can be given for equality between multi-category ordinal thresholds.

Although we define $\sigma_{\eta\eta} = 1$, in a biometric model $\sigma_{\eta\eta} = \sigma_{A\eta} + \sigma_{C\eta} + \sigma_{E\eta}$, with one variance component per each mode of inheritance (genes, shared environment, and non-shared environment). While the value of severity is not affected by sex differences in the ACE partition, its etiologic interpretation is; the difference can be tested in a common pathway biometric model e.g. by fixing the loadings A → F and E → F of Figure 1a across the sexes. Analogously, the specific variances are $1 - \lambda_j^2\psi = 1 - \lambda_j^2(\psi_A + \psi_C + \psi_E)$ where the sum $\psi_A + \psi_C + \psi_E$ is fixed but not the specific values of the summands. Again, a sex difference in specific summands would not imply different sensitivity (discriminability) for a criterion as an indicator of ASPD, but different etiologies for the observed sensitivity [see eq. (4)].

In the case we find that $\tau_j^* = \tau_j - \beta$ for all j, where β is a constant scalar value, we have (#criteria) × (#ordinal classes – 1) – 1 less parameters to estimate (altogether 13 here), and may express equation (2) as

$$y_j^* = \lambda_j(\eta + \beta z / \lambda_j) + \varepsilon_j \,. \tag{7}$$

If all the factor loadings were equal (i.e., $\lambda_j = \lambda_i$ for all i and j as assumed e.g. in Cronbach's alpha), the scalar translation of thresholds would amount to a constant difference in male and female factor scores. If they are not, the interpretation can nevertheless be close to that. In the main manuscript, we found λ_j values between 0.57 and 0.89 and β was estimated at 0.48. Thus, β/λ_j was between 0.54 and 0.84, implying that the simple interpretation of "no DIF and a factor-score sex difference" has no more inaccuracy than 0.3 standard deviations on the standard normal-variate scale (with s.d. of only 0.11 for β/λ_j). Regardless of the interpretation, it is an empirical question which model fits better, one with $\tau_j^* = \tau_j - \beta$ or one with separately estimated values for each τ_j^* and τ_j.

The Figure S1 shows both the freely estimated and the scalar-translation (restricted) threshold estimates for the best-fit common pathway Model IV. In line with model selection results, there are no significant differences between the estimates.

Results supplement

Although consistency with a previous study required us to exclude the conduct-disorder criterion from our structural analyses, the below figures and tables show the results with Conduct Disorder (CD) included as one of the ASPD criteria. In addition, the results are for the same-sex twins only to complement the analysis of full data (including separate-sex twins)

in the main manuscript. As in the main manuscript, we observed only one phenotypic factor (Figure S2). The obtained factor loadings were very similar to those in the Table 1, with an additional loading of 0.754 for CD (not shown). For completeness, Table S1 shows the Mplus-estimated phenotypic polychoric correlations, and Table S2 shows estimates for cross-twin, cross-trait correlations.

Further, the results on biometric model comparison were largely consistent with the analyses without CD, with the exception that while BIC and SABIC otherwise agreed, they conflicted on the overall best model when CD was included (Table S3). However, the SABIC was practically the same (indecisive) for the two best models, one of which was the same model as in the analyses without CD and the other was the best model from the previous study that used the same independent pathway models. Notice, however, that the previously reported model was favored only in the previously unstudied CD-including data. Although not quite the same, the parameters we estimated for the independent pathway Model V did have some resemblance with the estimates of the previous study (Table S4).

These results suggest that the three models may be difficult to distinguish and more theoretical attention should be given to the principles of model selection in this context, and more empirical attention to an even wider range of models. As for now, we opted for the most robustly selected model, consistently indicated by BIC: the common pathway Model IV. For completeness, Figure S3 presents the parameter estimates for that model in the same-sex twin data only; comparing to Figure 3 in the main manuscript, the estimates are very similar as for the full data.

Supplementary references

MacIntosh R, Hashim S (2003) Variance estimation for converting MIMIC model parameters to IRT parameters in DIF analysis. *Appl Psychol Meas* 27: 372–379.

Muthén BO (1989) Using item-specific instructional information in achievement modeling. *Psychometrika* 54: 385–396.

Muthén BO, Kao C-F, Burstein L (1991) Instructionally sensitive psychometrics: Application of a new IRT-based detection technique to mathematics achievement test items. *J Educ Meas* 28: 1–22.

Olsson, U. (1979) Maximum likelihood estimation of the polychoric correlation coefficient. *Psychometrika* 44, 443-460.

Supplementary Tables

Supplementary Table S1.

Phenotypic polychoric correlations underlying the Mplus factor analyses in same-sex twins

Criterion	#1	#2	#3	#4	#5	#6	#7
1. Not conforming							
2. Deceitfulness	0.489						
3. Impulsivity, or failure to plan	0.467	0.458					
4. Irritability/repeated fights	0.711	0.440	0.444				
5. Reckless disregard	0.485	0.306	0.434	0.46			
6. Irresponsibility	0.517	0.496	0.652	0.532	0.341		
7. Lack of remorse	0.763	0.596	0.496	0.806	0.605	0.454	
8. Conduct disorder	0.652	0.438	0.523	0.529	0.375	0.598	0.648

Supplementary Table S2

Cross-twin correlation estimates for the DSM-IV Antisocial Personality Disorder criteria

	Twin 1 vs Twin 2 criteria	#1	#2	#3	#4	#5	#6	#7	CD
MZs	1. Not conforming	0.581	0.417	0.216	0.612	0.040	0.403	0.724	0.535
	2. Deceitfulness	0.377	0.454	0.229	0.538	-0.121	0.274	0.453	0.495
	3. Impulsivity, or failure to plan	0.433	0.347	0.29	0.583	0.020	0.456	0.607	0.543
	4. Irritability/repeated fights	0.458	0.367	-0.443	0.706	-0.391	-0.226	0.592	0.326
	5. Reckless disregard	0.441	0.243	-0.340	0.400	0.156	0.190	0.539	0.341
	6. Irresponsibility	0.222	0.248	0.161	0.232	0.120	0.399	0.368	0.447
	7. Lack of remorse	0.550	0.330	0.146	0.558	0.052	0.338	0.544	0.711
	Conduct disorder (CD)	0.444	0.182	0.182	0.380	0.048	0.293	0.501	0.514
		#1	#2	#3	#4	#5	#6	#7	CD
DZs	1. Not conforming	0.236	-0.117	0.277	0.35	-0.116	0.290	0.259	0.124
	2. Deceitfulness	0.079	-0.117	0.321	0.17	-0.640	0.028	-0.007	0.050
	3. Impulsivity, or failure to plan	0.039	0.261	0.316	-0.457	-0.082	0.178	-0.375	0.108
	4. Irritability/repeated fights	0.442	0.203	0.549	0.441	-0.289	0.554	0.417	0.453
	5. Reckless disregard	0.182	0.072	0.273	0.189	-0.154	0.253	0.203	0.221
	6. Irresponsibility	0.141	0.178	0.171	-0.371	-0.022	0.0724	-0.063	0.082
	7. Lack of remorse	0.259	0.144	0.416	0.293	-0.472	0.342	0.215	0.232
	Conduct disorder (CD)	0.409	0.013	0.333	0.372	0.085	0.362	0.344	0.214

Note: Rows represent twin 1 and columns twin 2 (labels 1 and 2 are exchangeable). The values are from the "two-step" approximation procedure as implemented in R package "polycor", version 0.7-9 (Olson, 1979), and are provided strictly to satisfy the curiosity of readers and reviewers. The program adjusts the matrix to make it positive definite. None of the analyses in the manuscript use these values. Instead, full information maximum likelihood estimation on raw data was used in biometric modeling, together with models that incorporated sex differences in the criterion endorsement thresholds.

Supplementary Table S3.

Comparison of biometric models with and without conduct disorder (CD).

Model	Sex effects	Common factors	Specific factors	BIC	BIC (wCD)	SABIC	SABIC (wCD)
IP Models		**#A-#C-#E**	**a-c-e**				
I	Present	1A-1C-1E	a-c-e	-96796.75	-109526.0	4858.308	6638.324
II	Absent	1A-1C-1E	a-c-e	-97001.84	-109765.4	4737.915	6495.659
III	Absent	1A-0C-1E	a-e	**-97076.73**	**-109845.9**	**4707.489**	6466.037
IV	Absent	0A-1C-1E	c-e	-97065.25	-109840.0	4718.976	6471.896
V	Absent	2A-0C-1E	a-e	-97047.17	-109827.9	4717.994	**6461.753***
VI	Absent	1A-0C-2E	a-e	-97052.04	-109822.2	4713.124	6467.482
VII	Absent	3A-0C-1E	a-e	-97026.77	-109761.2	4722.518	6509.454
CP Models		**#(A-C-E)**	**a-c-e**				
I	Present	1-ACE	a-c-e	-96881.94	-109648.0	4856.902	6612.738
II	Absent	1-ACE	a-c-e	-97050.10	-109813.8	4731.551	6495.563
III	Absent	1-CE	c-e	-97096.69	-109864.9	4710.366	6473.044
IV*	Absent	1-AE	a-e	**-97104.65***	**-109875.6***	**4702.414***	**6462.325**
V	Absent	2-ACE	a-c-e	-96989.70	-109752.4	4757.615	6519.446
VI	Absent	2-CE	c-e	-97054.34	-109821.8	4724.743	6485.011
VII	Absent	2-AE	a-e	-97060.89	-109830.5	4718.192	6476.312
VIII	Absent	3-AE	a-e	-96987.85	-109748.3	4756.891	6520.948

Note: only the same-sex twins were used. The best-fit independent-pathway and common-pathway models are highlighted with bold font. Numbers of factors in a biometric ACE model are given in the form A-C-E with the number of factors in front of the letters. Presence versus absence of criterion-specific effects is indicated by corresponding subset of "a-c-e" components. Absence of "sex effects" refers to male and female factor structures that are constrained to be equal, instead of being freely estimated (sex effects "present").

wCD = analysis run with Conduct Disorder + the other 7 ASPD criteria.

* = The overall best fit

BIC = Bayesian Information Criterion. Lower BIC values indicate more parsimonious (i.e. better) model.

SABIC = Sample-size Adjusted BIC. Lower SABIC values indicate more parsimonious (i.e. better) model.

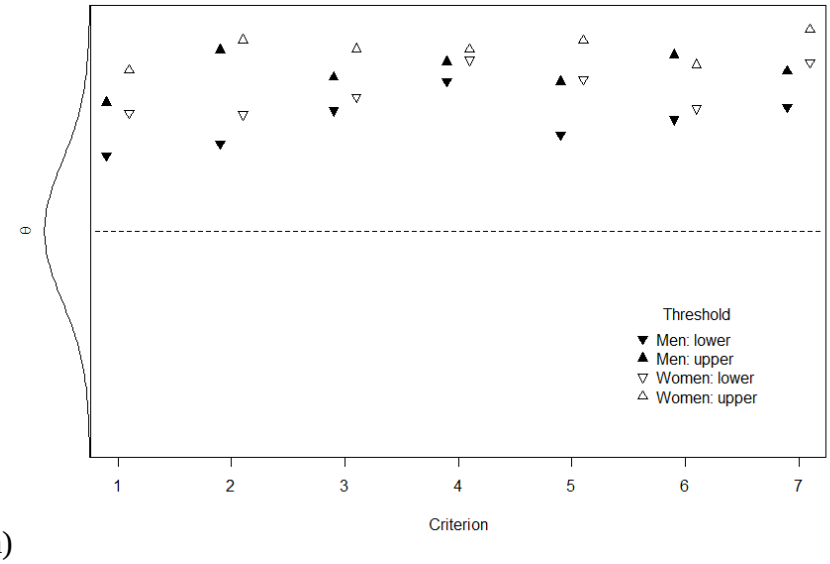
Supplementary Table S4.

Biometric model factor loadings with conduct disorder

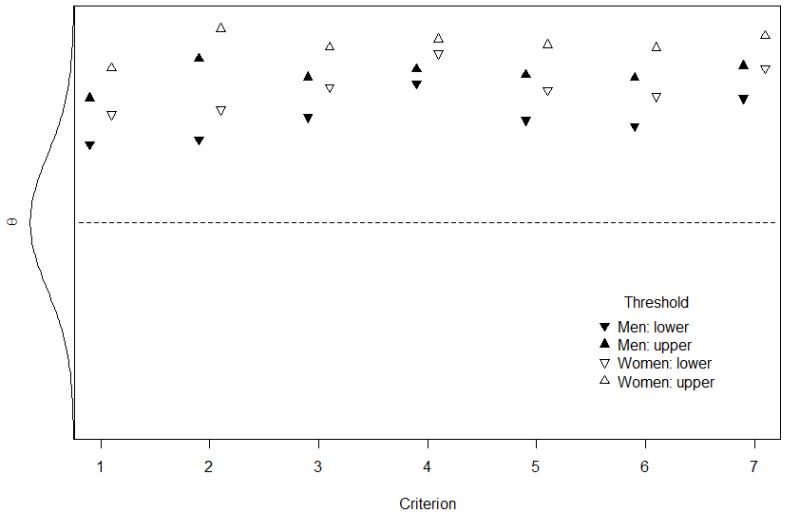
	Common-pathway Model IV			**Independent-pathway Model V**				
	Factor	**Specific effects**		**Factors**				**Specific**
	A1	**A**	**E**	**A1**	**A2**	**E1**	**A**	**E**
1. Not conforming	**0.796**	0.277	0.538	**0.649**	0.286	-0.442	0.0507	0.547
2. Deceitful	**0.586**	0.417	0.694	0.285	0.23	-0.488	0.472	0.637
3. Failure to plan	**0.686**	0.483	0.543	0.147	**0.564**	-0.489	0.348	0.547
4. Irritability/repeated fights	**0.801**	0.582	0.14	**0.629**	0.103	**-0.567**	0.521	0.0393
5. Reckless disregard	0.484	0.426	0.765	0.148	0.081	**-0.620**	0.444	0.625
6. Irresponsibility	**0.735**	0.437	0.519	0.16	**0.664**	-0.498	0.181	0.502
7. Lacks remorse	**0.876**	0.00001	0.483	**0.66**	0.145	**-0.629**	0.00001	0.384
8. Conduct disorder	**0.745**	0.288	0.602	0.472	**0.511**	-0.329	0.00001	0.639

Note: Factor loadings with absolute value > 0.5 are highlighted with bold font. Varimax-rotation was applied to the genetic factors A1 and A2. In general, “A” stands for additive-genetic effects and “E” for (non-shared) environmental effects.

Supplementary Figures

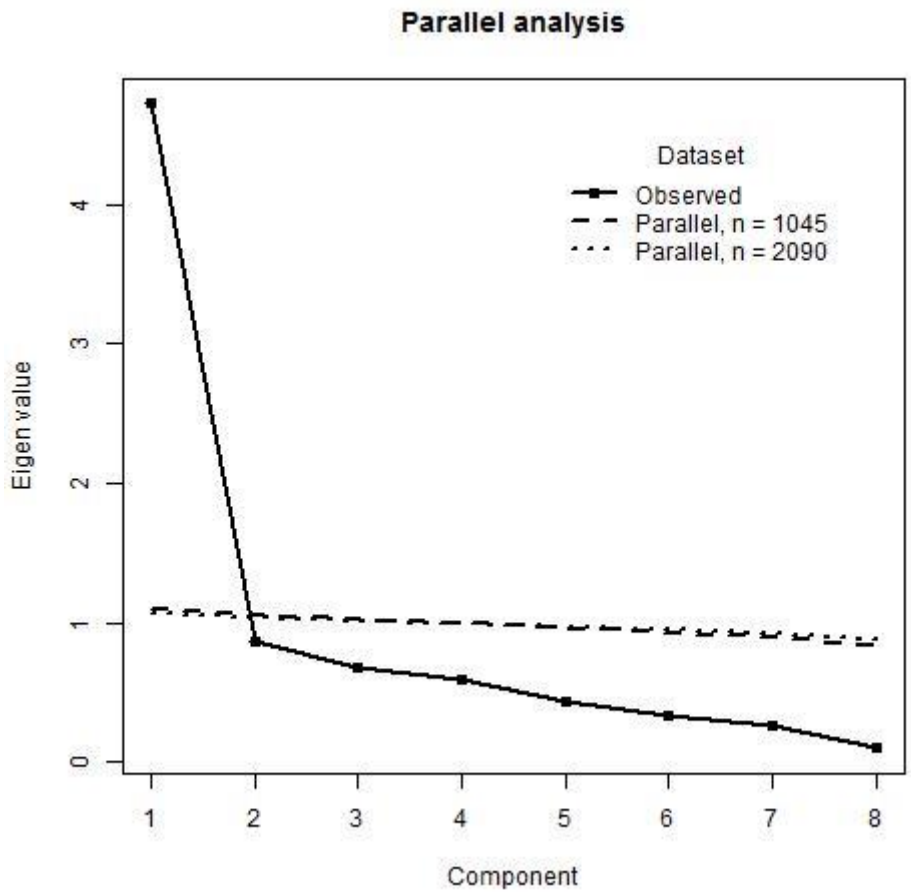


a)

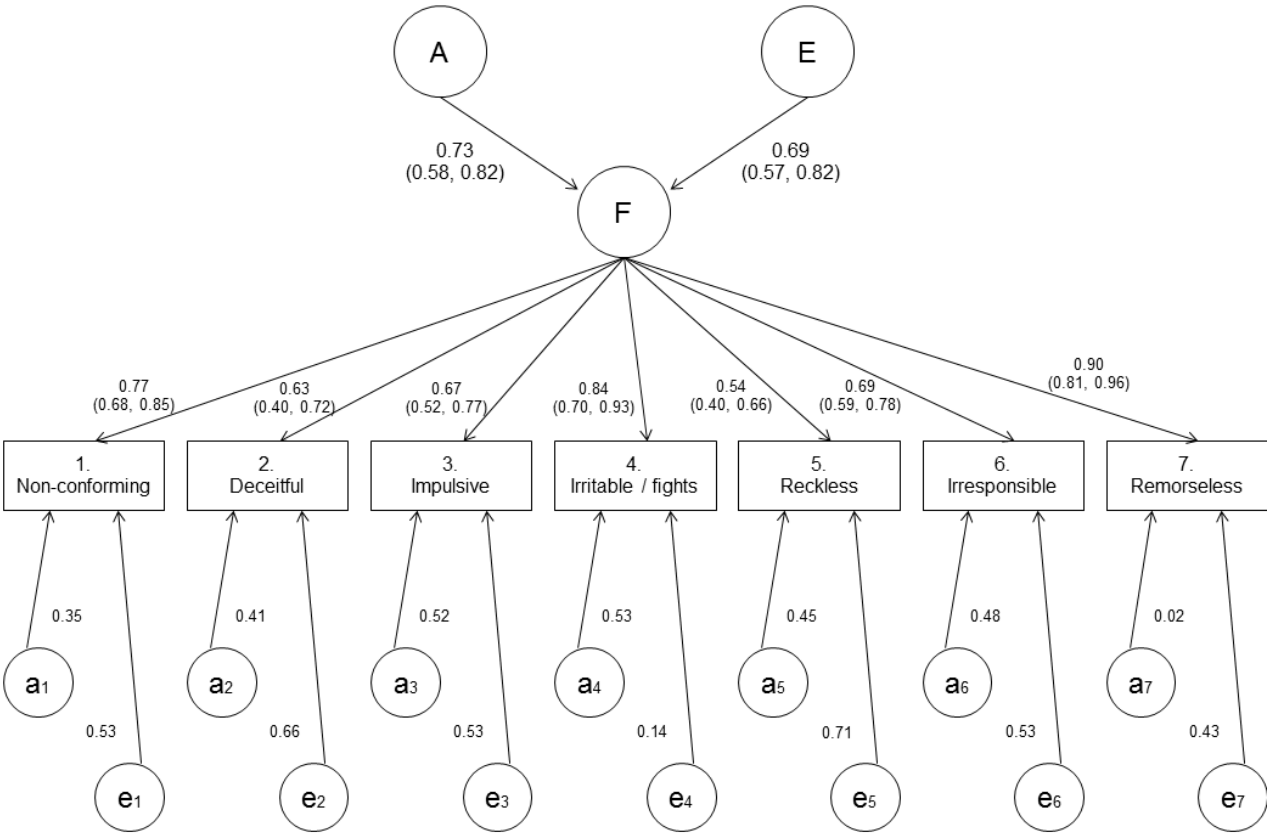


b)

Supplementary Figure S1. Threshold estimates for the ASPD criteria. **a)** Freely estimated thresholds: an ordinal endorsement occurs when the unit-variance, normally distributed, liability (θ) for the criterion exceeds a threshold separately estimated for each criterion and each sex (values are estimates from the common pathway Model IV of Table 2). **b)** Scalar-translation thresholds: instead of unconstrained estimates for both sexes, women's estimates can be a scalar shift, a translation, of men's thresholds for added model parsimony (lower number of estimable parameters); values are for the same model as in the panel **a**, but with the additional constraint.



Supplementary Figure S2. *Scree plot and parallel analysis test for ASPD criteria, including conduct disorder.* The solid line shows the eigenvalues of the weighted-least squares mean- and variance-adjusted polychoric correlation matrix, whereas the dashed (simulated sample size n = 1045) and the dotted (simulated sample size n = 2090) lines indicate 5th percentile values across 1000 replications in parallel analyses using uncorrelated criteria. Scree-plotted observed eigenvalues above the parallel-analyses lines represent structure (i.e., factors) over and above sampling variance. The two parallel analyses lines simply indicate that both perfect correlation (minimum information) and no correlation (maximum information) between the twins would nevertheless lead to the same conclusion



Supplementary Figure S3. *Path diagram and parameter estimates of the best-fitting common-pathway biometric model.* The same model as in Figure 3, but estimated from the same-sex twins only instead of the full data.