Supplementary Table 1. Concordance rates of the two peas item in the Two Peas Questionnaire within twin pairs.

Two peas	Zygosity		
	MZ	DZ	unknown
Concordant ordinary resemblance	3.2%	77.9%	28.6%
Concordant two peas in a pod	90.5%	9.5%	62.5%
Discordant (one twin ordinary resemblance, one twin two peas in a pod)	6.3%	12.6%	8.9%

Note. MZ = monozygotic twins. DZ = dizygotic twins. Unknown = twin pairs with no DNA-based zygosity. Two peas = When you were children, were you and your twin as alike as two peas in a pod, or of ordinary family resemblance?

Supplementary Table 2. Model comparisons to identify reference item among twin pairs with

DNA -based zygosity.

	RMSEA [90%CI]	CFI	TLI	χ^2	df	p
Two peas	.067 [.055, .078]	.981	.975	6.190	2	.045
Parents	.067 [.055, .078]	.980	.975	29.688	4	< .001
Relatives	.067 [.055, .078]	.980	.975	23.138	4	< .001
Teachers	.067 [.055, .078]	.980	.975	22.146	4	< .001
Strangers	.067 [.055, .078]	.980	.975	3.307	4	.508

Note. All models are compared against a constrained model where the factor loadings and threshold parameters of all items are constrained to be equal between MZ and DZ twins. In the augmented models, the factor loadings and threshold parameters of the corresponding item are allowed to differ between MZ and DZ twins. To control for Type I error for multiple testing, Bonferroni corrected alpha (.05/5 = .01) was used.

Supplementary Table 3. Goodness-of-fit indices and model comparisons testing differential item functioning (DIF) on each item, with "Strangers" as the reference item.

	RMSEA [90%CI]	CFI	TLI	χ^2	df	p
Two peas	.051 [.037, .064]	.990	.985	2.656	2	.2649
Parents	.062 [.049, .074]	.985	.979	26.149	4	< .001
Relatives	.057 [.044, .070]	.987	.982	20.637	4	< .001
Teachers	.056 [.043, .069]	.988	.982	20.927	4	< .001

Note. To control for Type I error for multiple testing, Bonferroni corrected alpha (.05/4 = .0125) was used.

Supplementary Table 4. Estimated factor loadings and thresholds of the Item Factor Analysis (IFA) models for the self-report zygosity items, separately for MZ and DZ among twin pairs with DNA-based zygosity.

		MZ		DZ	
		Est	(*SE*)	Est	(*SE*)
Loadings					
Two peas	λ_1	.750	(*.06*)	.657	(*.08*)
Parents	λ_2	.485	(*.04*)	.587	(*.09*)
Relatives	λ_3	.688	(*.03*)	.612	(*.06*)
Teachers	λ_4	.680	(*.04*)	.724	(*.05*)
Strangers	λ_5	.655	(*.05*)	.721	(*.05*)
Thresholds					
Two peas	τ_{11}	-1.499	(*.07*)	.995	(*.09*)
Parents	τ_{21}	-.693	(*.05*)	.955	(*.08*)
	τ_{22}	.217	(*.05*)	1.590	(*.10*)
	τ_{23}	1.197	(*.06*)	1.924	(*.14*)
Relatives	τ_{31}	-1.915	(*.09*)	.175	(*.07*)
	τ_{32}	-1.115	(*.06*)	.888	(*.08*)
	τ_{33}	.034	(*.05*)	1.584	(*.11*)
Teachers	τ_{41}	-2.167	(*.10*)	-.034	(*.08*)
	τ_{42}	-1.591	(*.07*)	.527	(*.08*)
	τ_{43}	-.461	(*.05*)	1.242	(*.09*)
Strangers	τ_{51}	-2.483	(*.13*)	-.230	(*.07*)
	τ_{52}	-1.991	(*.09*)	.199	(*.08*)
	τ_{53}	-1.336	(*.06*)	.830	(*.08*)
Mean		0	-	0	-
Variance		1	-	1	-
Model fit					
RMSEA [90%CI]		.061 [.046, .076]		.024 [.000, .061]	
CFI		.981		.999	
TLI		.969		.998	
SRMR		.058		.042	

Note. MZ = monozygotic twins. DZ = dizygotic twins. SE = standard error. RMSEA = Root mean square error approximation. CFI = Comparative Fit Index. TLI = Tucker Lewis Index. SRMR = Standardized Root Mean Square Residual.
Only parameters of one twin are shown here, as all item parameters are constrained to be the same within twin pairs. Means and variances are fixed to 0 and 1, respectively.

Supplementary Table 5. Estimated factor loadings and thresholds of the latent class analysis

(LCA) models for the self-report zygosity items among twin pairs without DNA-based zygosity.

		MZ		DZ	
		Est	(*SE*)	Est	(*SE*)
Loadings					
Two peas	λ_1	4.78	(.34)	3.43	(.19)
Parents	λ_2	1.82	(.06)	4.37	(.17)
Relatives	λ_3	2.49	(.06)	5.04	(.20)
Teachers	λ_4	2.87	(.11)	4.33	(.18)
Strangers	λ_5	4.73	(.21)	4.07	(.16)
Thresholds					
Two peas	τ_{11}	-5.0	(.40)	1.37	(.12)
Parents	τ_{21}	-.94	(.06)	1.58	(.10)
	τ_{22}	1.31	(.06)	2.80	(.12)
	τ_{23}	4.18	(.10)	5.23	(.20)
Relatives	τ_{31}	-4.62	(.11)	.26	(.09)
	τ_{32}	-2.19	(.08)	1.41	(.12)
	τ_{33}	.94	(.08)	3.58	(.15)
Teachers	τ_{41}	-6.37	(.14)	.09	(.08)
	τ_{42}	-3.98	(.10)	1.07	(.12)
	τ_{43}	-.61	(.08)	3.10	(.12)
Strangers	τ_{51}	-10.67	(.40)	-.31	(.08)
	τ_{52}	-7.76	(.26)	.44	(.10)
	τ_{53}	-4.80	(.20)	1.90	(.12)
Mean		0	-	0	-
Variance		1	-	1	-

Note. MZ = monozygotic twins. DZ = dizygotic twins. SE = standard error. RMSEA = Root mean square error approximation. CFI = Comparative Fit Index. TLI = Tucker Lewis Index. SRMR = Standardized Root Mean Square Residual.

Supplementary Table 6. Item response probabilities from the item factor analysis (IFA) model

obtained from the sample with DNA-based zygosity.

	Zygosity	Ordinary resemblance	Two peas in a pod		
Two peas	MZ	.07	.93	-	-
	DZ	.84	.16	-	-
		Never confused	Rarely confused	Sometimes confused	Always confused
Parents	MZ	.24	.34	.30	.12
	DZ	.83	.11	.03	.03
Relatives	MZ	.03	.10	.38	.49
	DZ	.57	.24	.13	.06
Teachers	MZ	.01	.04	.27	.68
	DZ	.49	.21	.19	.11
Strangers	MZ	.02	.02	.07	.91
	DZ	41	.17	.22	.20

Note. The probabilities may not add up to 1.00 due to rounding error.

Supplementary Table 7. Item response probabilities from the latent class analysis (LCA) model obtained from the sample without DNA-based zygosity.

	Zygosity	Ordinary resemblance	Two peas in a pod		
Two peas	MZ	.17	.83	-	-
	DZ	.64	.36	-	-
		Never confused	**Rarely confused**	**Sometimes confused**	**Always confused**
Parents	MZ	.35	.35	.25	.05
	DZ	.63	.09	.14	.13
Relatives	MZ	.07	.17	.39	.38
	DZ	.52	.08	.14	.25
Teachers	MZ	.03	.09	.31	.57
	DZ	.51	.08	.16	.25
Strangers	MZ	.02	.05	.11	.83
	DZ	.47	.07	.13	.33

Note. The probabilities may not add up to 1.00 due to rounding error.
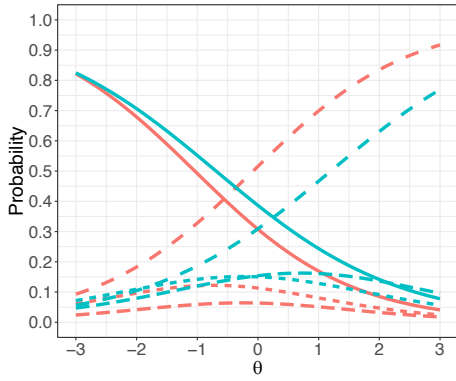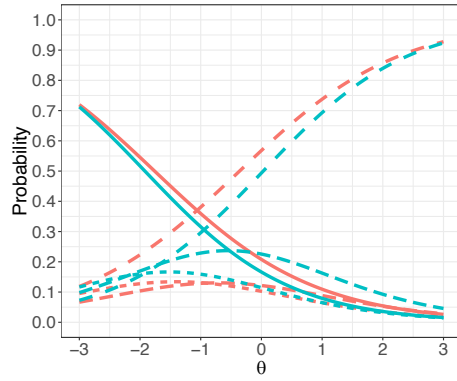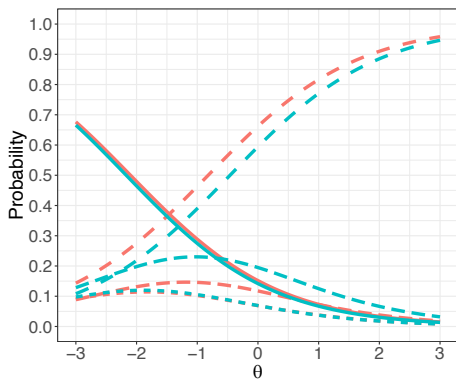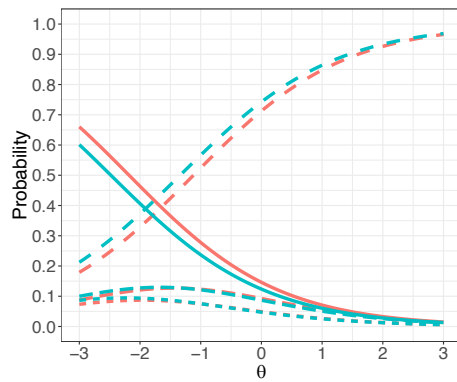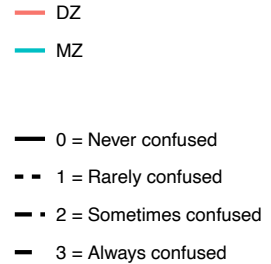
Supplementary Figure 1. Category response curves (CRCs) of four of the self-report zygosity questionnaire items by zygosity among twin pairs with DNA-based zygosity.
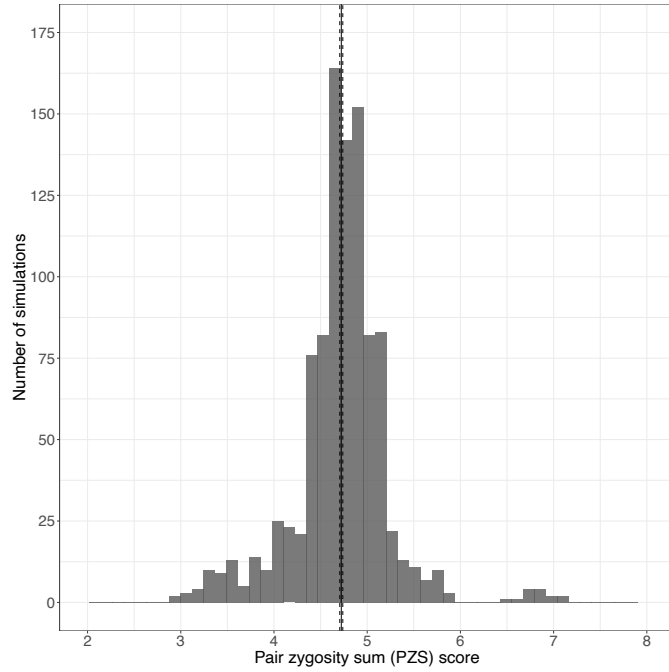
(a) Distribution of optimum cutoff pair zygosity sum (PZS) values.
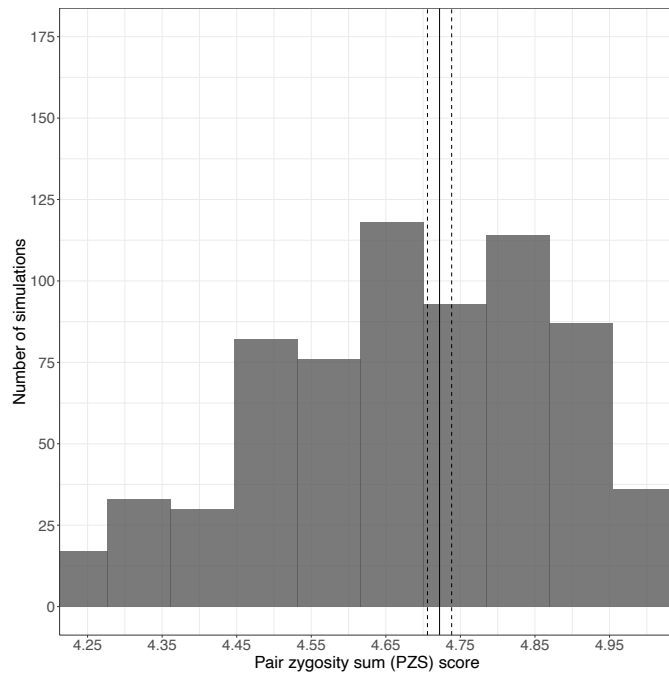
(b) Distribution of optimum cutoff pair zygosity sum (PZS) values, zoomed.

Supplementary Figure 2. Distribution of optimum cutoff pair zygosity sum (PZS) values across the 1000 cross-validated logistic regression models. The solid line illustrates the average optimum cutoff value at PZS = 4.7, and the dashed lines illustrate the standard error (*SE* = .03).
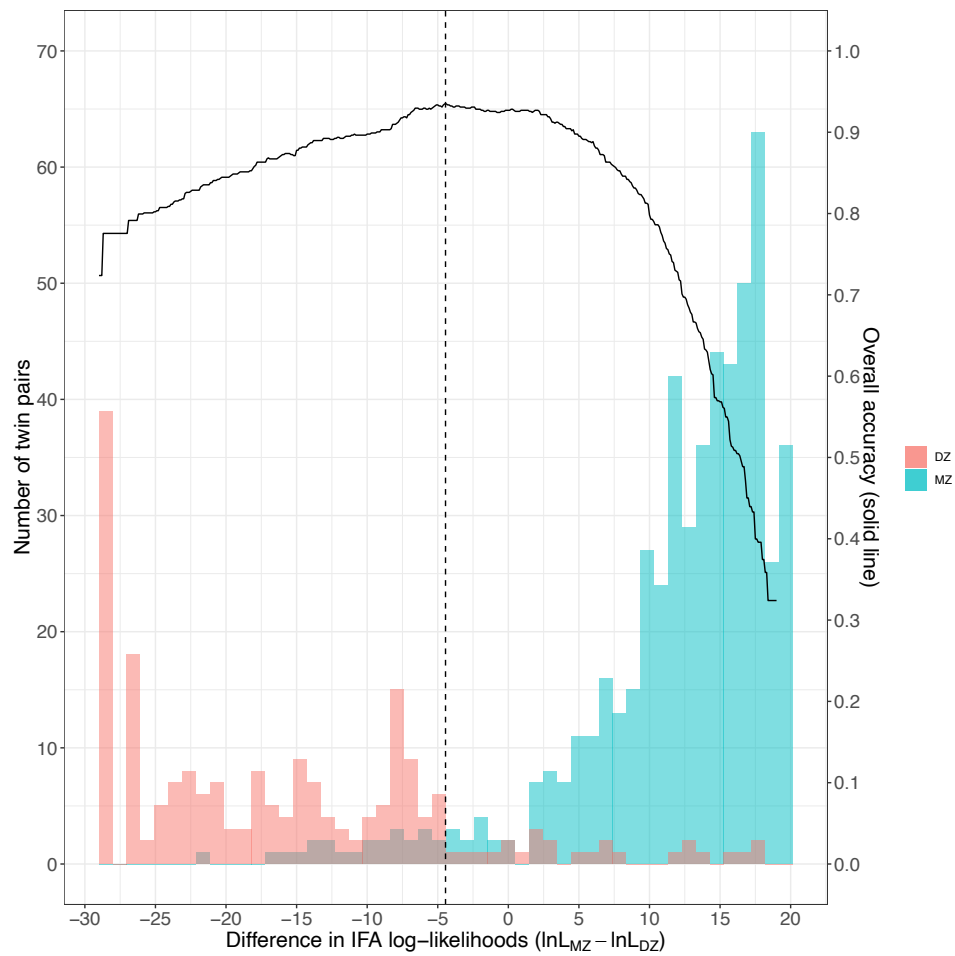
Supplementary Figure 3. Distribution of response probabilities from the item factor analysis (IFA) model among twin pairs with DNA-based zygosity. The solid line illustrates the overall accuracy rate (true MZ and DZ pairs correctly classified as MZ and DZ, respectively). Dashed line indicates the optimum cutoff value at ΔlnL_{IFA} = -4.45.
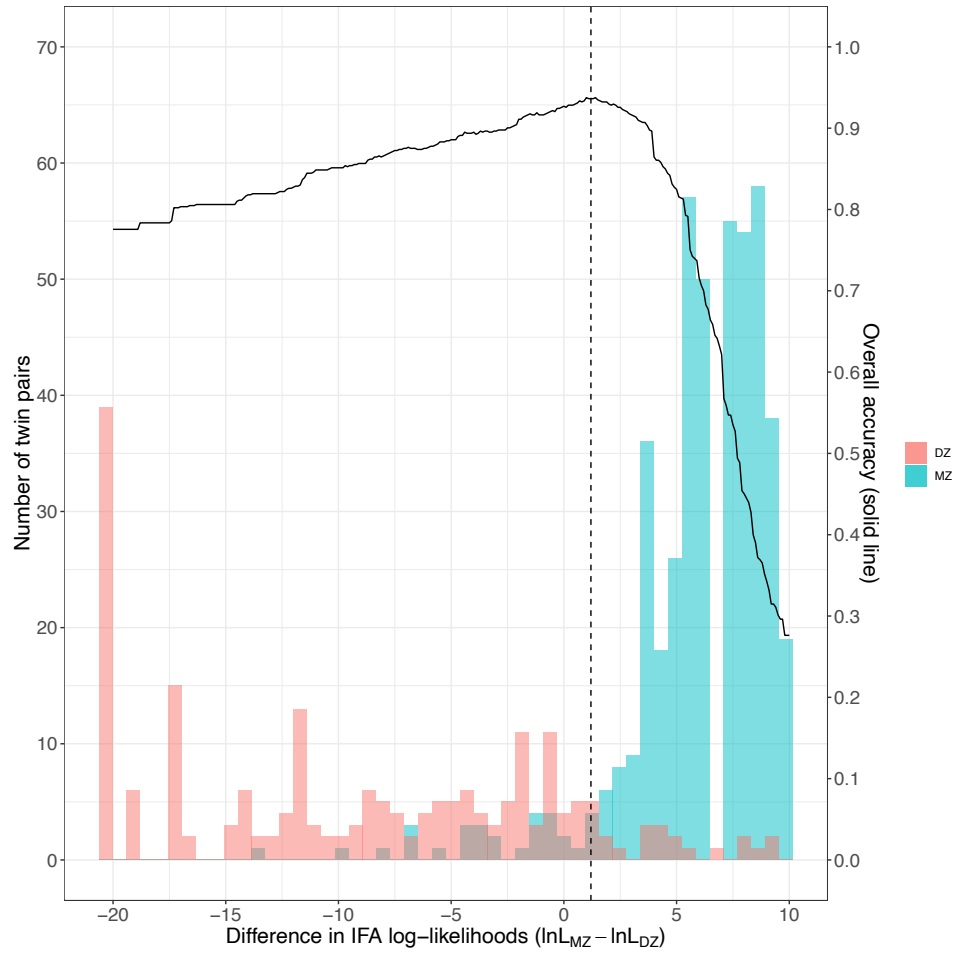
Supplementary Figure 4. Distribution of response probabilities from the latent class analysis (LCA) model among twin pairs with DNA-based zygosity. The solid line illustrates the overall accuracy rate (true MZ and DZ pairs correctly classified as MZ and DZ, respectively). Dashed line indicates the optimum cutoff value at ΔlnL_{LCA} = 1.2.
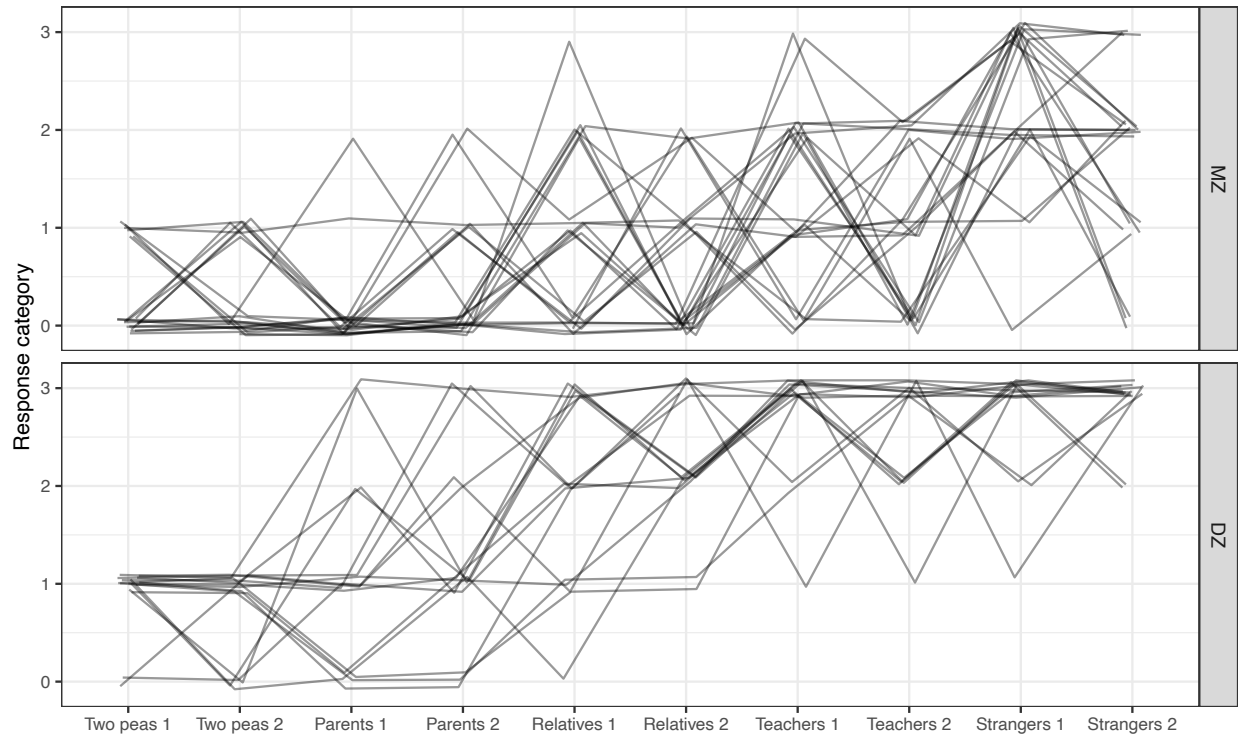
Figure 5. Responses of twin pairs with DNA-based zygosity (23 MZ pairs, 16 DZ pairs) who were consistently misclassified by all three classification methods. Each line reflects the responses of one twin pair.

Note. For two peas, response category 0 = ordinary resemblance, 1 = two peas in a pod. For parents, relatives, teachers, and strangers, response category 0 = never confused, 1 = rarely confused, 2 = sometimes confused, 3 = always confused. Numbers after each item indicate responses from twin 1 or twin 2 (arbitrarily assigned).
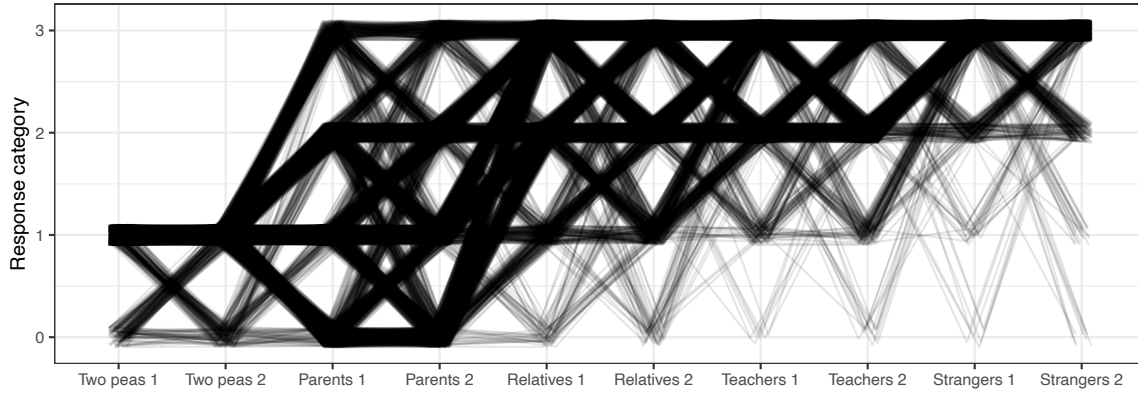
(a) Twins (*n* = 4212 pairs) consistently classified as MZ pairs.
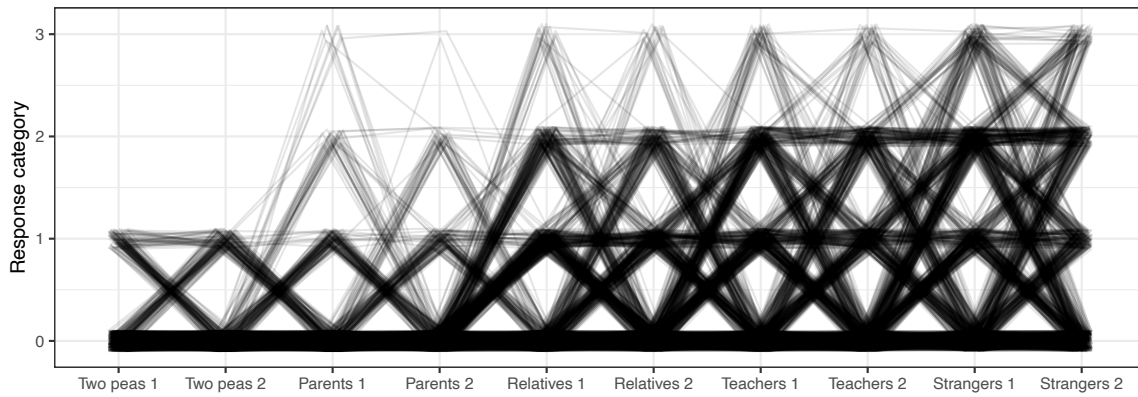
(b) Twins (*n* = 1991 pairs) consistently classified as DZ pairs.
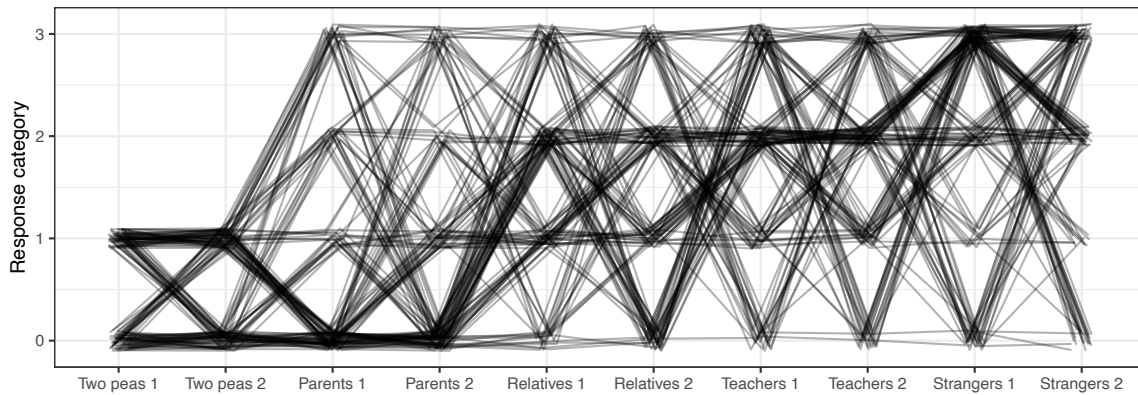
(c) Twins (*n* = 164 pairs) inconsistently classified by the three methods.

Figure 6. Responses of twin pairs without DNA-based zygosity. Each line reflects the responses of one twin pair.

Note. For two peas, response category 0 = ordinary resemblance, 1 = two peas in a pod. For parents, relatives, teachers, and strangers, response category 0 = never confused, 1 = rarely confused, 2 = sometimes confused, 3 = always confused. Numbers after each item indicate responses from twin 1 or twin 2 (arbitrarily assigned).

$r_{consistent}$ = .99 [.99, .99]
$r_{inconsistent}$ = .92 [.90, .94]
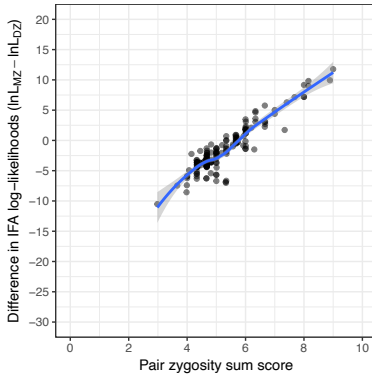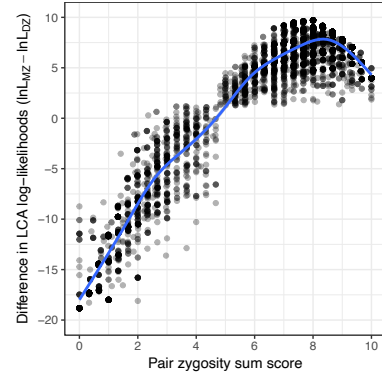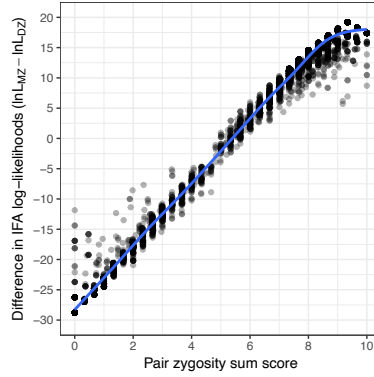
$r_{consistent}$ = .94 [.94, .94]
$r_{inconsistent}$ = -.24 [-.38, -.09]
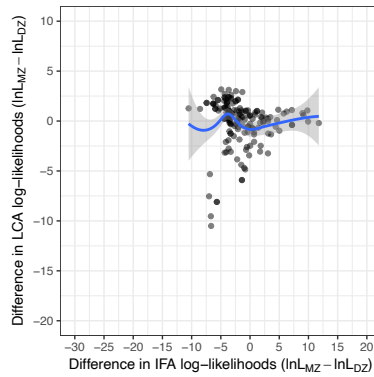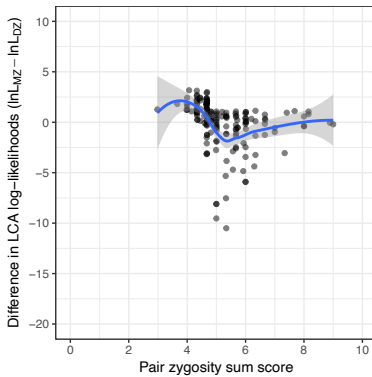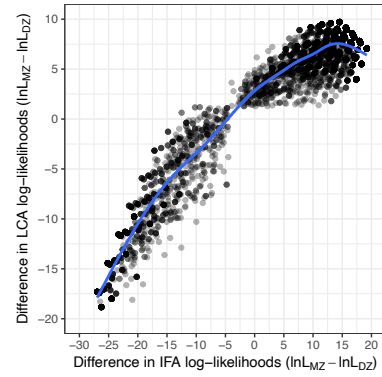
$r_{consistent}$ = .96 [.96, .96]
$r_{inconsistent}$ = 0 [-.15, .16]

Figure 7. Correlations among estimates from the pair zygosity sum (PZS) score, item factor analysis (IFA) model, and latent class analysis (LCA) model among twin pairs without DNA-based zygosity. Twin pairs who were consistently classified as MZ and DZ pairs were illustrated on the top diagonal, and those who were inconsistently classified were illustrated on the bottom diagonal.

$r_{consistent}$ and $r_{inconsistent}$ represent Pearson's correlations for pairs consistently and inconsistently classified, respectively. 95% confidence intervals were shown in brackets.