

On the Problem of Spurious Non-Linear Effects in Aggregated Scores:

Investigating Differentiation of Cognitive Abilities using Item Level Data

Dylan Molenaar



UNIVERSITY OF AMSTERDAM

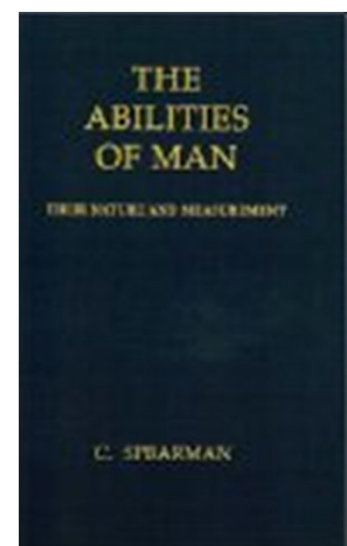
SEM working group, 15th of March 2018, Amsterdam, The Netherlands

THE ABILITIES OF MAN

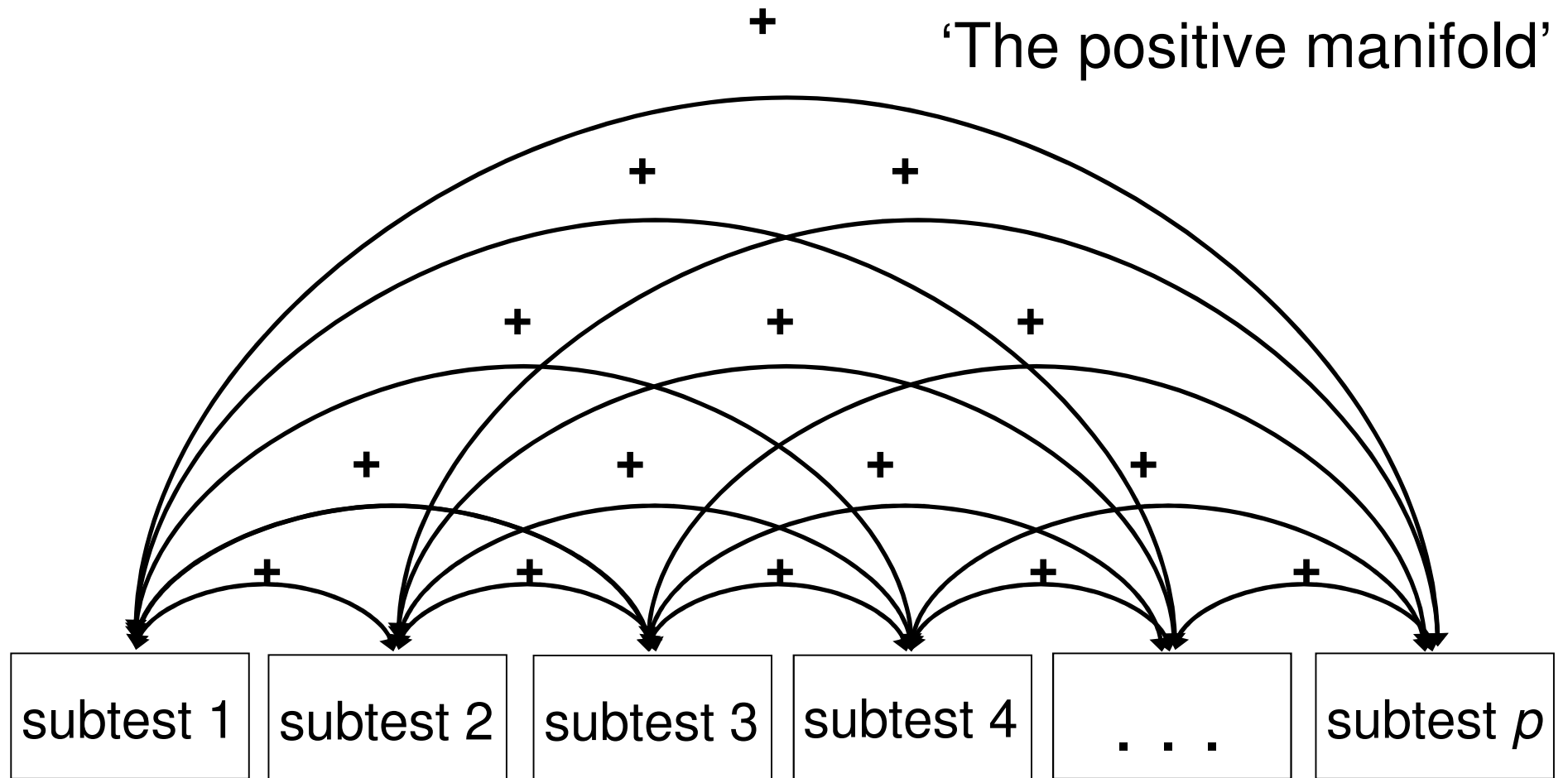
78 NORMAL CHILDREN (Corrected for Attenuation).

	1	2	3	4	5	6	7	8	9	10	11	12
1. Opposites- - -	—	75	78	71	62	64	72	78	57	40	46	33
2. Observation - -	75	—	72	58	60	58	67	56	58	56	52	29
3. Absurdities - -	78	72	—	53	41	44	79	68	41	46	34	29
4. Memory sentences -	71	58	53	—	54	61	54	37	54	55	19	43
5. Crossing o's - -	62	60	41	54	—	73	48	54	38	36	52	35
6. Geometrical figs. -	64	58	44	61	73	—	45	48	30	42	48	35
7. Discrim. length - -	72	67	79	54	48	45	—	56	49	30	31	06
8. Crossing patterns -	78	56	68	37	54	48	56	—	30	21	27	18
9. Memory form - - -	57	58	41	54	38	30	49	30	—	24	31	29
10. Tapping - - - -	40	56	46	55	36	42	30	21	24	—	29	18
11. Strength of grip -	46	52	34	19	52	48	31	27	31	29	—	28
12. Interpret pictures -	33	29	29	43	35	35	06	18	29	18	28	—

Mean = 0.466.

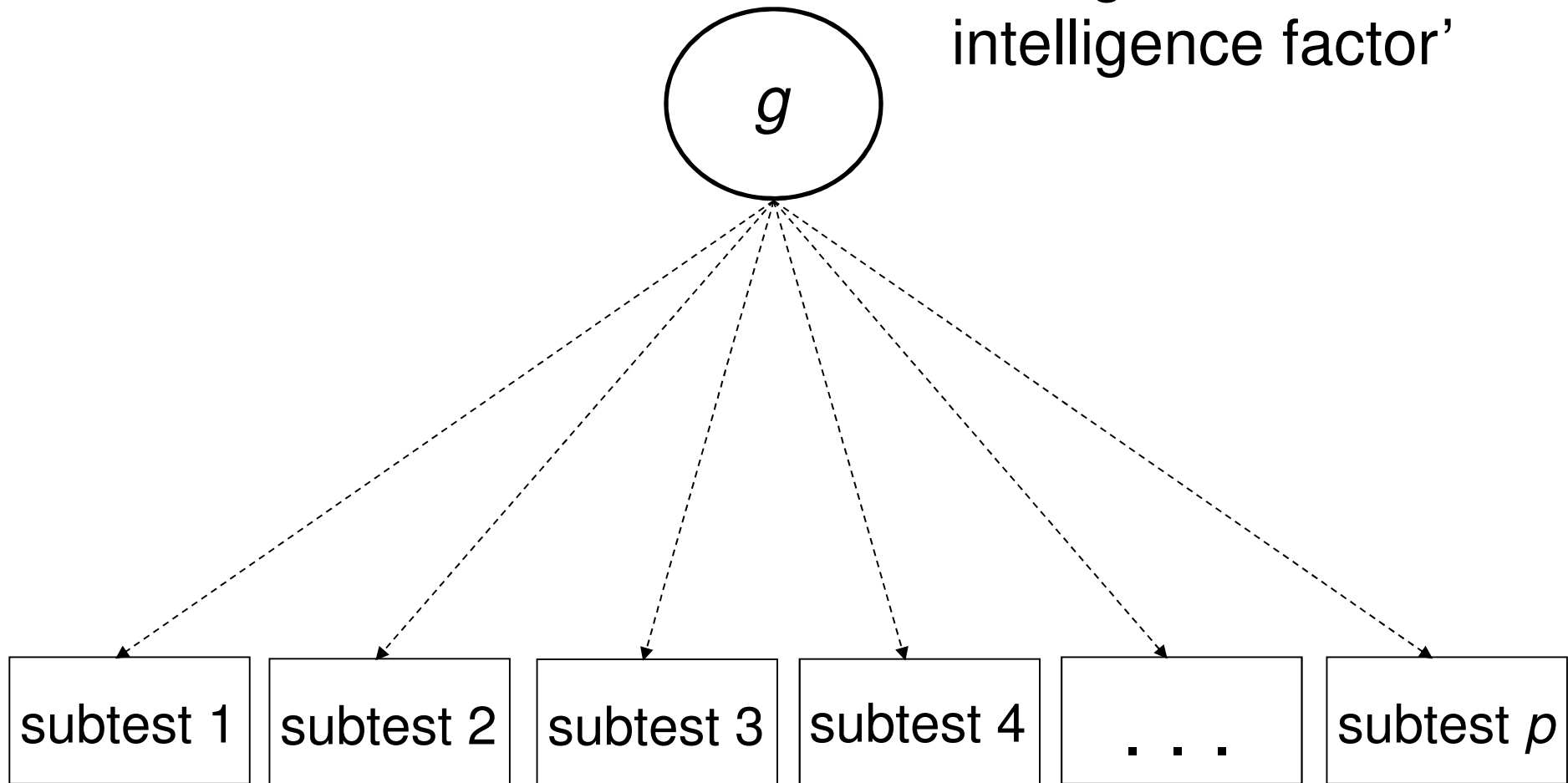


Intelligence research



Intelligence research

'The general intelligence factor'



78 NORMAL CHILDREN (Corrected for Attenuation)

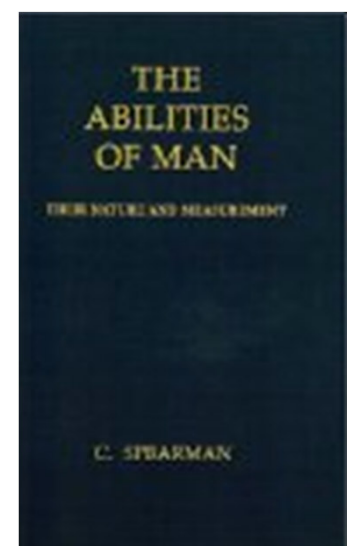
	1	2	3	4	5	6	7	8	9	10	11	12
1. Opposites- - -	—	75	78	71	62	64	72	78	57	40	46	33
2. Observation - -	75	—	72	58	60	58	67	56	58	56	52	29
3. Absurdities - -	78	72	—	53	41	44	79	68	41	46	34	29
4. Memory sentences -	71	58	53	—	54	61	54	37	54	55	19	43
5. Crossing o's - -	62	60	41	54	—	73	48	54	38	36	52	35
6. Geometrical figs. -	64	58	44	61	73	—	45	48	30	42	48	35
7. Discrim. length - -	72	67	79	54	48	45	—	56	49	30	31	06
8. Crossing patterns -	78	56	68	37	54	48	56	—	30	21	27	18
9. Memory form - - -	57	58	41	54	38	30	49	30	—	24	31	29
10. Tapping - - - -	40	56	46	55	36	42	30	21	24	—	29	18
11. Strength of grip -	46	52	34	19	52	48	31	27	31	29	—	28
12. Interpret pictures -	33	29	29	43	35	35	06	18	29	18	28	—

Mean = 0.466.

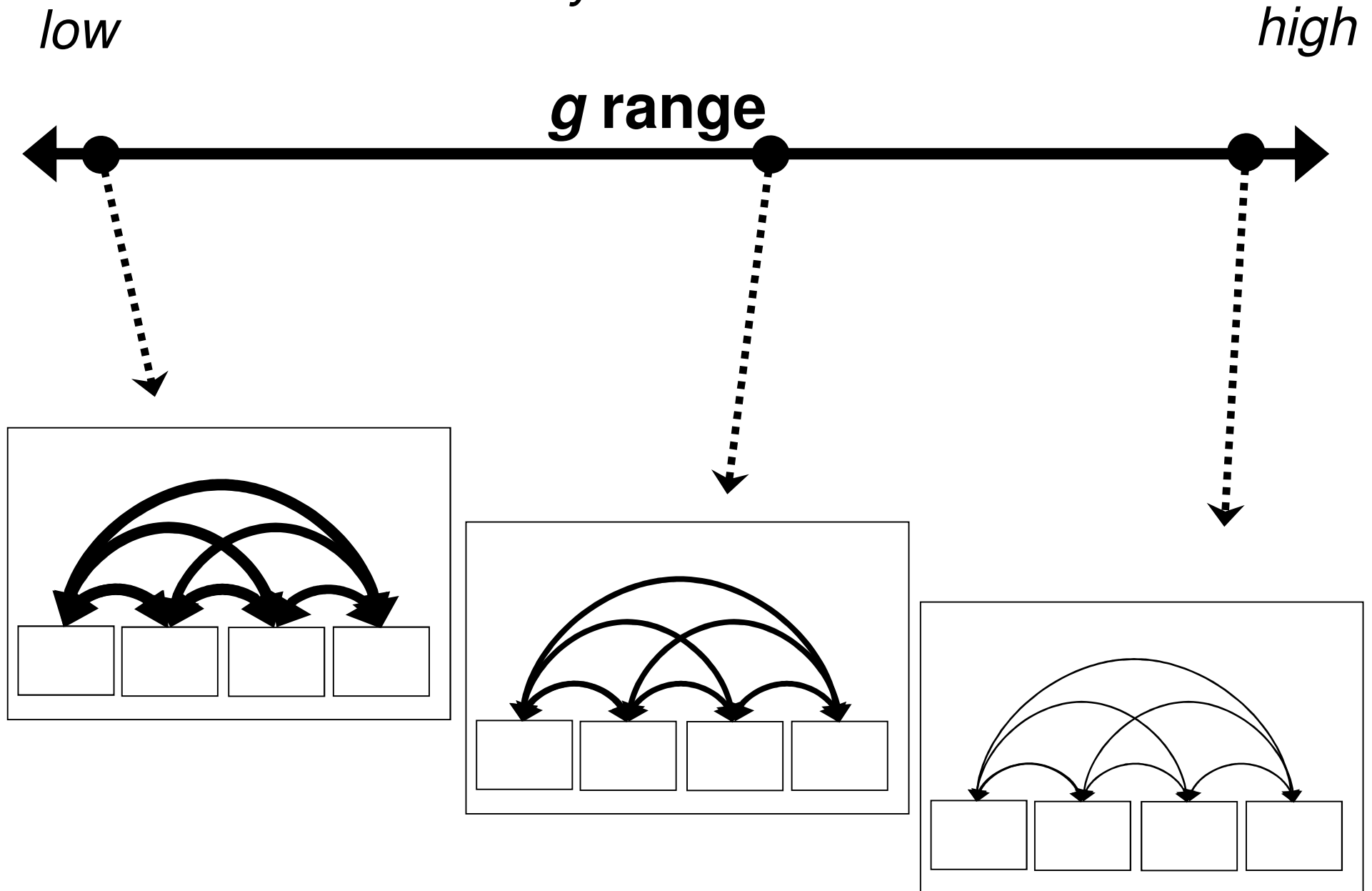
22 DEFECTIVE CHILDREN (Corrected for Attenuation)

	1	2	3	4	5	6	7	8	9	10	11	12
1. Absurdities - - -	—	100	100	98	97	100	100	100	98	94	94	79
2. Opposites- - - -	100	—	97	95	87	91	85	76	85	87	70	72
3. Crossing patterns -	100	97	—	91	80	88	68	92	74	78	76	67
4. Crossing o's - - -	98	95	91	—	85	77	84	67	76	81	73	55
5. Memory sentences -	97	87	80	85	—	73	90	68	88	65	78	68
6. Observation - - -	100	91	88	77	73	—	76	83	71	86	59	65
7. Memory form - - -	100	85	68	84	90	76	—	65	67	70	77	75
8. Interpret. pictures -	100	76	92	67	68	83	65	—	74	80	80	59
9. Geometrical figs. -	98	85	74	76	88	71	67	74	—	65	60	62
10. Discrim. length - -	94	87	78	81	65	86	70	80	65	—	51	45
11. Tapping - - - -	94	70	76	73	78	59	77	80	60	51	—	61
12. Strength of grip -	79	72	67	55	68	65	75	59	62	45	61	—

Mean = 0.782.



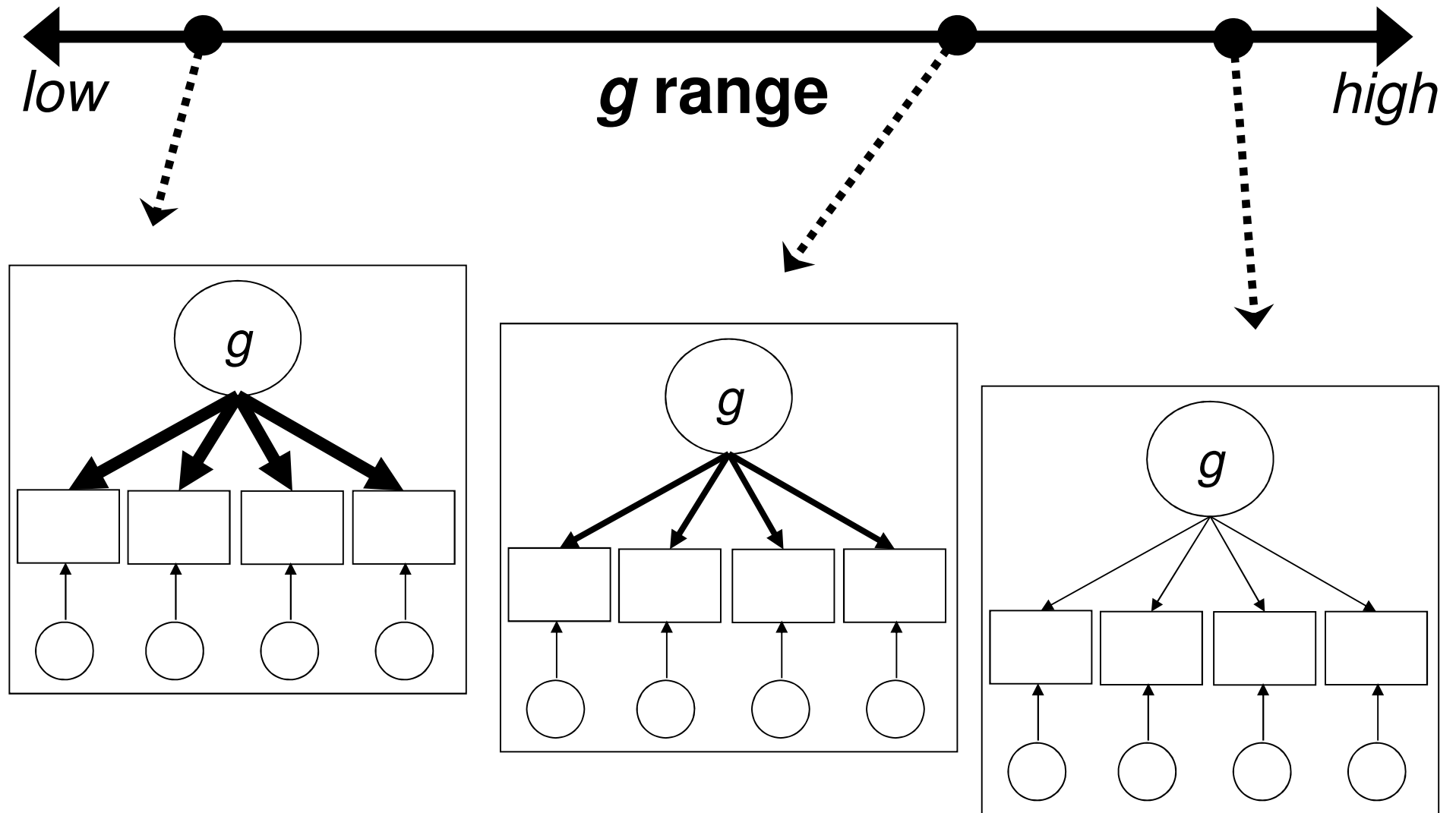
*Spearman's Law of Diminishing Returns, or
'Ability Differentiation'*



Research into ability differentiation

- “First wave”: **Multi-group approaches**
 - Split IQ subtest scores (e.g., Deary et al., 1996; Detterman & Daniels, 1989),
 - Split factor scores (Carlstedt, 2001; Reynolds & Keith, 2007)
 - Use existing groups (e.g., Detterman & Daniels, 1989; Spearman, 1927; te Nijenhuis, & Hartmann, 2006).
- “Second wave”: **Single group approaches**
 - Quadratic effect of g (e.g., Molenaar, Dolan, Wicherts, & Van der Maas, 2010; Murray, Dixon, & Johnson, 2013; Tucker-Drob, 2009)
 - Mixture of strong and weak g -effects (Reynolds, Keith, & Beretvas, 2010)

Single group approaches



Tucker-Drob (2009), Molenaar et al., (2010), Murray et al., (2013)

Single group approaches

Possible confounds which should be taken into account:

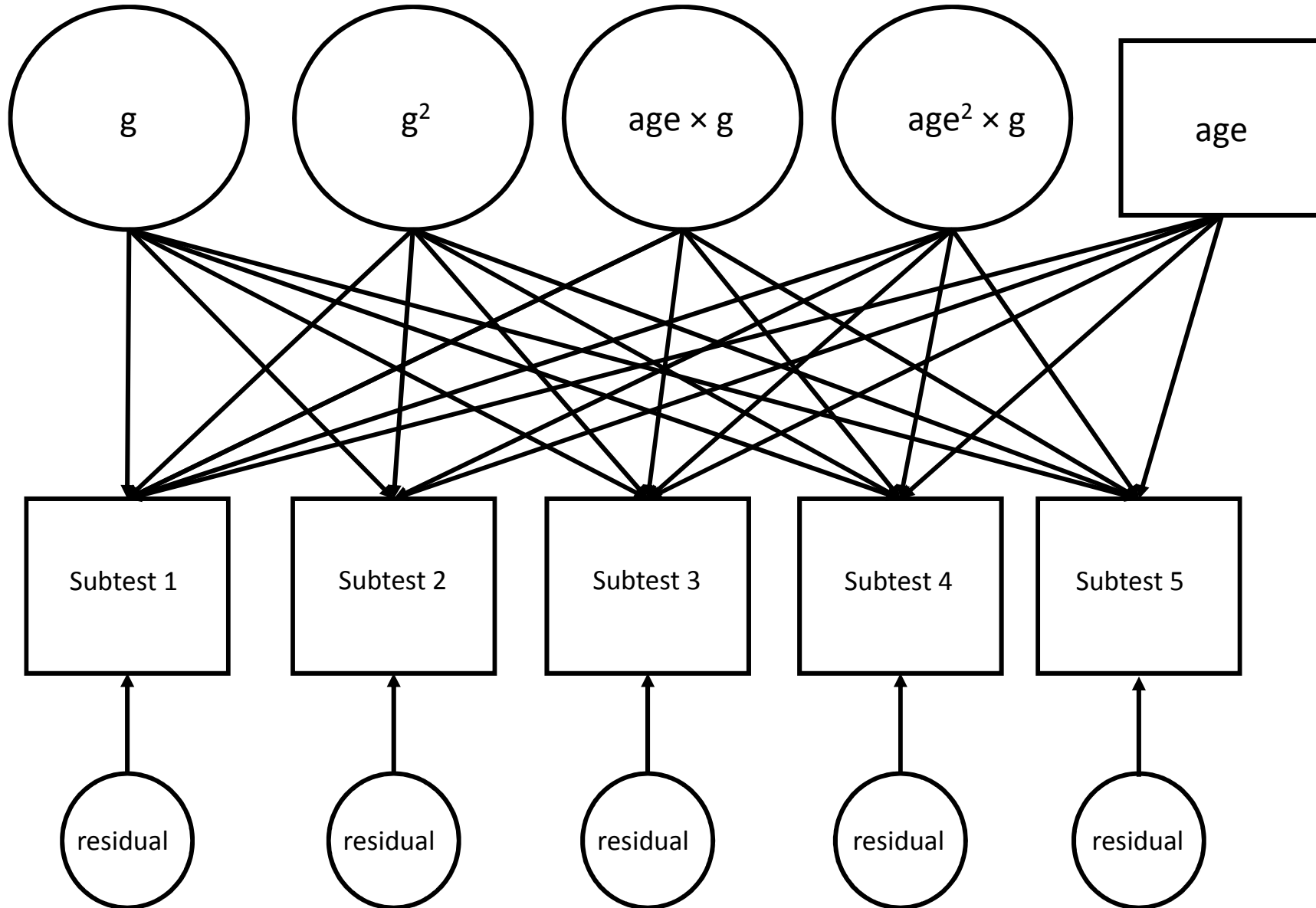
Age differentiation:

Strength of g decreases across age

Age differentiation-dedifferentiation:

Strength of g decreases across age but later increases

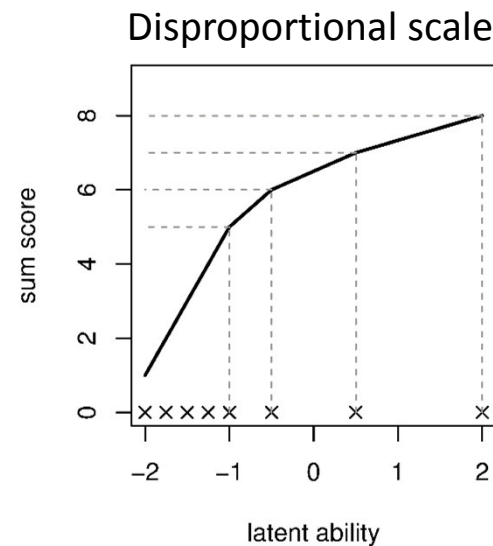
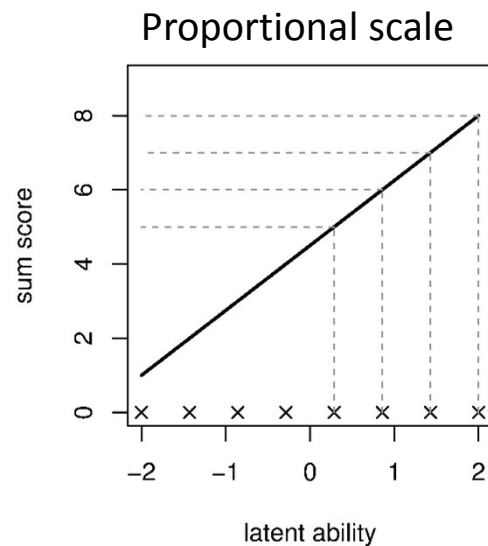
Resulting model



Challenge:

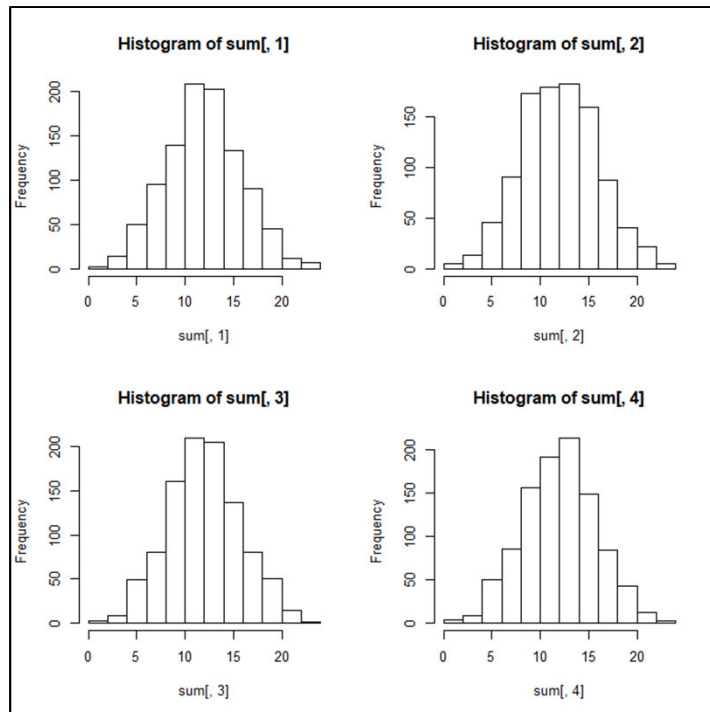
How to compute scale scores

Raw composite scores are unsuitable

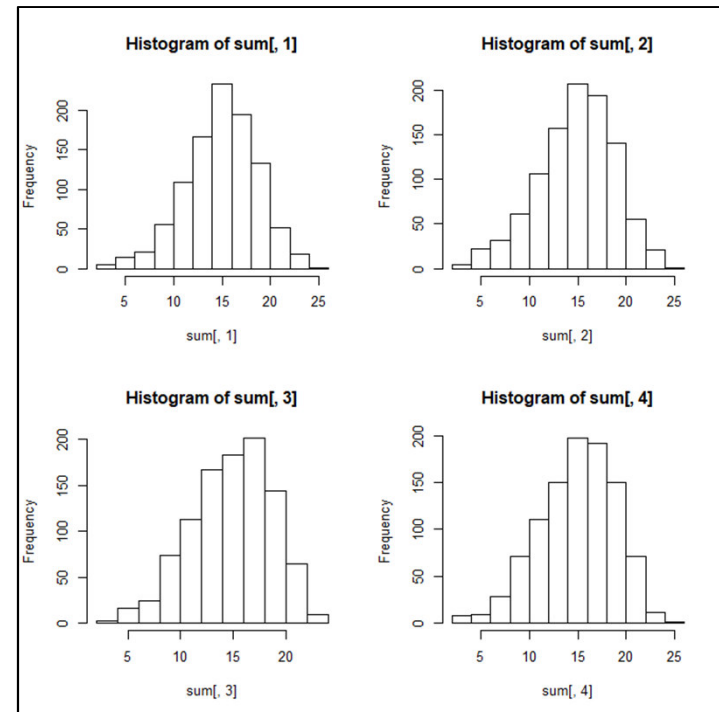


Challenge: How to compute scale scores

Proportional scale



Disproportional scale



Challenge:

How to compute scale scores

False positive rates

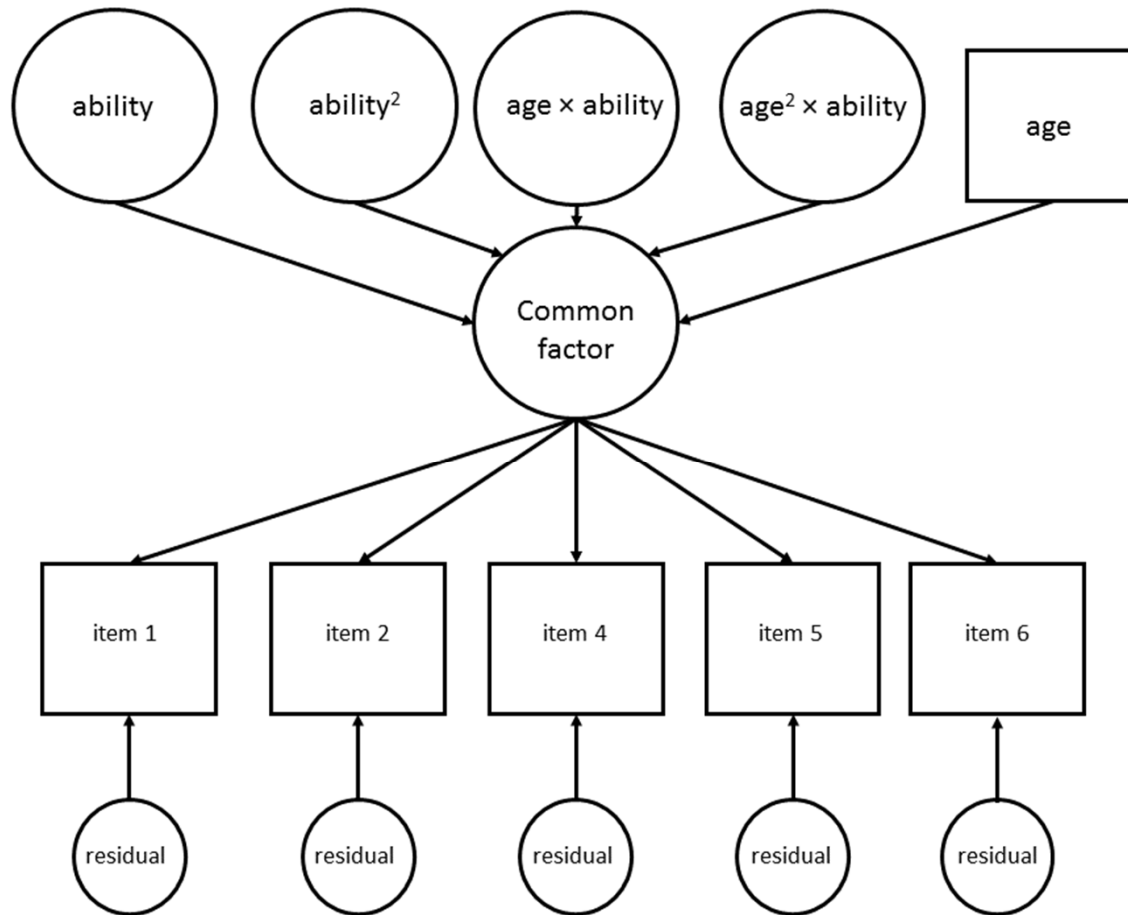
Scale	Measure	Subtest	ability ²	age × ability	age ² × ability
Proportional	Raw sum scores	1	0.09	0.04	0.10
		2	0.03	0.03	0.07
		3	0.06	0.05	0.06
		4	0.05	0.07	0.09

Challenge:

How to compute scale scores

- Use factor scores (Tucker-Drob, 2009)
 - Neglects uncertainty in factor score estimates
- Adjust residuals (Molenaar et al., 2010; Murray et al., 2011)
 - Requires a parametric form for the adjustment

Alternative solution: An item level model



Alternative solution: An item level model

Detection rates rates

False Positive

True Positive

Data			Hit rates		
<i>ability</i> ²	<i>age</i> ×	<i>age</i> ² ×	<i>ability</i> ²	<i>age</i> ×	<i>age</i> ² ×
	<i>ability</i>	<i>ability</i>	<i>ability</i>	<i>ability</i>	
0	0	0	0.03	0.04	0.05
0	0	1	0.06	0.05	0.89
0	1	0	0.02	0.97	0.04
0	1	1	0.04	0.91	0.88
1	0	0	0.70	0.02	0.04
1	0	1	0.74	0.03	0.92
1	1	0	0.78	0.92	0.04
1	1	1	0.81	0.93	0.89

Application

- Standardization data of the Hungarian WAIS IV
 - 17 subtests
 - 1,112 subjects with ages between 16 and 90
- For 4 subtests, the first item was omitted
- For 2 subtests, the last item was omitted
- Measurement invariance holds across 13 age groups
 - Grouping based on sampling design of standardization sample

Results

Coloring on basis of
0.01 significance level

Subtest	ability²	age × ability	age² × ability
BD: Block Design ^a			
SI: Similarities			
DSf: Digit Span forwards ^{a,b}			
DSb: Digit Span backwards			
DSi: Digit Span inverse			
MR: Matrix Reasoning			
VO: Vocabulary			
AR: Arithmetic ^a			
SS: Symbol Search			
VP: Visual Puzzles			
IN: Information ^a			
LN: Letter-Number seq. ^b			
FW: Figure Weights			
CA: Cancellation			
CM: Comprehension			
PC: Picture Completion.			
CO: Coding			

Note. ^a: first item omitted; ^b: last item omitted

Results

Coloring on basis of
0.01 significance level

Subtest	ability ²			age × ability	age ² × ability
	<i>Est</i>	<i>SE</i>	<i>Z</i>		
BD: Block Design ^a	-0.23	0.02	-9.42		
SI: Similarities	-0.03	0.03	-0.86		
DSf: Digit Span forwards ^{a,b}	-0.19	0.01	-13.36		
DSb: Digit Span backwards	-0.17	0.01	-19.33		
DSi: Digit Span inverse	-0.20	0.00	-66.33		
MR: Matrix Reasoning	-0.15	0.03	-5.63		
VO: Vocabulary	-0.06	0.02	-2.75		
AR: Arithmetic ^a	-0.16	0.02	-7.57		
SS: Symbol Search	-	-	-		
VP: Visual Puzzles	-0.20	0.03	-7.69		
IN: Information ^a	-0.004	0.04	-0.11		
LN: Letter-Number seq. ^b	-0.18	0.02	-7.95		
FW: Figure Weights	-0.19	0.02	-9.25		
CA: Cancellation	-0.04	0.02	-1.81		
CM: Comprehension	-	-	-		
PC: Picture Completion.	-0.08	0.03	-2.52		
CO: Coding	-	-	-		

Note. ^a: first item omitted; ^b: last item omitted

Results

Coloring on basis of
0.01 significance level

Subtest	ability ²			age × ability			age ² × ability
	<i>Est</i>	<i>SE</i>	<i>Z</i>	<i>Est</i>	<i>SE</i>	<i>Z</i>	
BD: Block Design ^a	-0.23	0.02	-9.42	-0.09	0.06	-1.66	
SI: Similarities	-0.03	0.03	-0.86	0.12	0.04	2.76	
DSf: Digit Span forwards ^{a,b}	-0.19	0.01	-13.36	0.05	0.03	1.44	
DSb: Digit Span backwards	-0.17	0.01	-19.33	0.01	0.01	0.80	
DSi: Digit Span inverse	-0.20	0.00	-66.33	0.00	0.00	0.33	
MR: Matrix Reasoning	-0.15	0.03	-5.63	0.01	0.05	0.16	
VO: Vocabulary	-0.06	0.02	-2.75	0.09	0.08	1.13	
AR: Arithmetic ^a	-0.16	0.02	-7.57	0.01	0.03	0.32	
SS: Symbol Search	-	-	-	-	-	-	
VP: Visual Puzzles	-0.20	0.03	-7.69	-0.05	0.05	-1.13	
IN: Information ^a	-0.004	0.04	-0.11	0.24	0.08	2.87	
LN: Letter-Number seq. ^b	-0.18	0.02	-7.95	-0.04	0.04	-0.83	
FW: Figure Weights	-0.19	0.02	-9.25	-0.01	0.06	-0.19	
CA: Cancellation	-0.04	0.02	-1.81	0.06	0.03	1.65	
CM: Comprehension	-	-	-	-	-	-	
PC: Picture Completion.	-0.08	0.03	-2.52	0.12	0.06	2.05	
CO: Coding	-	-	-	-	-	-	

Note. ^a: first item omitted; ^b: last item omitted

Results

Coloring on basis of
0.01 significance level

Subtest	ability ²			age × ability			age ² × ability		
	<i>Est</i>	<i>SE</i>	<i>Z</i>	<i>Est</i>	<i>SE</i>	<i>Z</i>	<i>Est</i>	<i>SE</i>	<i>Z</i>
BD: Block Design ^a	-0.23	0.02	-9.42	-0.09	0.06	-1.66	-0.06	0.05	-1.11
SI: Similarities	-0.03	0.03	-0.86	0.12	0.04	2.76	0.07	0.04	1.86
DSf: Digit Span forwards ^{a,b}	-0.19	0.01	-13.36	0.05	0.03	1.44	0.04	0.08	0.48
DSb: Digit Span backwards	-0.17	0.01	-19.33	0.01	0.01	0.80	0.01	0.02	0.73
DSi: Digit Span inverse	-0.20	0.00	-66.33	0.00	0.00	0.33	0.01	0.00	2.00
MR: Matrix Reasoning	-0.15	0.03	-5.63	0.01	0.05	0.16	-0.12	0.04	-3.11
VO: Vocabulary	-0.06	0.02	-2.75	0.09	0.08	1.13	-0.05	0.05	-1.04
AR: Arithmetic ^a	-0.16	0.02	-7.57	0.01	0.03	0.32	-0.04	0.03	-1.31
SS: Symbol Search	-	-	-	-	-	-	-	-	-
VP: Visual Puzzles	-0.20	0.03	-7.69	-0.05	0.05	-1.13	-0.11	0.04	-3.11
IN: Information ^a	-0.004	0.04	-0.11	0.24	0.08	2.87	-0.06	0.04	-1.49
LN: Letter-Number seq. ^b	-0.18	0.02	-7.95	-0.04	0.04	-0.83	0.01	0.04	0.12
FW: Figure Weights	-0.19	0.02	-9.25	-0.01	0.06	-0.19	0.03	0.05	0.65
CA: Cancelation	-0.04	0.02	-1.81	0.06	0.03	1.65	-0.02	0.03	-0.64
CM: Comprehension	-	-	-	-	-	-	-	-	-
PC: Picture Completion.	-0.08	0.03	-2.52	0.12	0.06	2.05	-0.03	0.05	-0.67
CO: Coding	-	-	-	-	-	-	-	-	-

Note. ^a: first item omitted; ^b: last item omitted

Discussion

- Item level analysis:
 - Importance:
 - Exclude the possibility that results are due to scaling
 - Ensure that interaction effects found on aggregated data are also detectable on item level
 - But:
 - No distinction between differentiation across g and across lower-order abilities
 - Ideally: Full item level factor model →
 - Remaining issue: Non-normality of the factor distribution

Full item level factor model

