

Aanvullende gegevens normen, betrouwbaarheid en validiteit van de COTAPP

28 september 2020

Speciale dank gaat uit naar dr. Lieke Voncken (Universiteit Tilburg) en prof. dr. Marieke Timmerman (Rijksuniversiteit Groningen) voor hun onmisbare en grote bijdrage aan deze bijlage.

Inhoudsopgave

1	Normen	3
1.1	Equivalentie van de steekproefomvang in continue en klassieke norming	3
1.2	Modelpassing van de gebruikte normeringsmodellen	12
1.2.1	Basisrapportage	16
1.2.2	Specialistische rapportage	44
1.3	Representativiteitsgegevens per leeftijdsgroep	100
1.4	Gegevens over verschillen tussen subgroepen	105
1.4.1	Subgroepen naar migratieachtergrond	106
1.4.2	Subgroepen naar geslacht	110
1.4.3	Subgroepen naar Totaal IQ	115
1.5	Informatie over de meetnauwkeurigheid	119
1.6	Gegevens over de scoreverdelingen	120
2	Betrouwbaarheid	127
2.1	Parallellie van testhelften	127
2.2	Split-half betrouwbaarheid op basis van leeftijdsgroepen	136
2.3	Test-hertest betrouwbaarheid	141
2.4	Vergelijking huidige betrouwbaarheidsgegevens met die van de handleiding	144
3	Criteriumvaliditeit	146
3.1	Representativiteit steekproeven o.b.v. achtergrondgegevens	146
4	Referenties	148

1 Normen

1.1 Equivalentie van de steekproefomvang in continue en klassieke norming

Om te onderzoeken in hoeverre de gebruikte steekproefomvang voldoende groot is voor alle leeftijdsgroepen en testonderdelen is een vergelijking gemaakt tussen de precisie onder continue normering en de precisie onder klassieke normering per leeftijdsgroep. De precisie is vergeleken op basis van de geschatte standaardfout van het steekproefgemiddelde.

Voor de klassieke normering is de standaardfout per groep j per testonderdeel t (SE_{tj}^{clas}) berekend op basis van de geobserveerde gepoolde standaarddeviatie van het testonderdeel in de normeringssteekproef (SD_t^p) en de steekproefomvang (n), als volgt:

$$SE_{tj}^{clas} = \frac{SD_t^p}{\sqrt{n}}$$

De standaardfout onder klassieke normering is berekend voor $n = 200$, $n = 300$ en $n = 400$. Volgens de COTAN richtlijnen is een steekproefomvang van $n = 200$ en $n = 300$ respectievelijk voldoende en goed voor tests voor relatief minder belangrijke beslissingen op individueel niveau; een steekproefomvang van $n = 300$ en $n = 400$ is respectievelijk voldoende en goed voor tests voor belangrijke beslissingen op individueel niveau.

De standaardfout onder lineair continue normeren is berekend volgens de methode beschreven door Bechger, Hemker en Maris (2009). De volgende formule is toegepast:

$$SE_{tj}^{con} = \left(\frac{1}{SS_{tj}} \left(SD_t^p \left(\frac{\sum_{j=1}^J j^2 n_{tj}}{\sum_{j=1}^J n_{tj}} + j^2 - 2jM \right) \right) \right)^{\frac{1}{2}}$$

met SE_{tj}^{con} de standaardfout voor testonderdeel t van groep j onder lineair continue normeren, M gelijk aan de gemiddelde groep, n_{tj} de geobserveerde steekproefomvang voor testonderdeel t van groep j , en $SS_{tj} = (j - M)^2 n_{tj}$ de bijbehorende kwadratensom.¹

Voor de berekening van deze geschatte standaardfout per testonderdeel is gebruik gemaakt van de gepoolde standaarddeviatie en de steekproefomvangen per leeftijdsgroep, zoals geobserveerd in de normeringssteekproef. Deze gepoolde standaarddeviaties en steekproefomvangen per leeftijdsgroep en testonderdeel zijn weergegeven in tabel 1.

De laagste leeftijdsgroep (5;6-5;11) en de hoogste leeftijdsgroep (13;0-13;6) vallen buiten het leeftijdsbereik van de test. Deze twee leeftijdsgroepen zijn meegenomen om de schattingsprecisie te vergroten voor de buitenste leeftijdsgroepen binnen het leeftijdsbereik van de test. Bij de vergelijking van de standaardfouten onder continue en klassieke normering dienen de laagste en hoogste leeftijdsgroepen natuurlijk buiten beschouwing gelaten te worden.

In tabel 2 staan per testonderdeel de gepoolde standaarddeviaties, de geschatte standaardfout per leeftijdsgroep onder lineair continue normering, en de geschatte standaardfout onder klassieke normering voor $n = 200$, $n = 300$ en $n = 400$.

¹Deze vergelijking is de gecorrigeerde versie van vergelijking (4) van Bechger, Hemker en Maris; die betreft namelijk de gekwadrateerde standaardfout, niet de standaardfout zelf

Tabel 1: Steekproefomvang per leeftijdsgroep en testonderdeel

	5;6-5;11	6;0-6;11	7;0-7;11	8;0-8;11	9;0-9;11	10;0-10;11	11;0-11;11	12;0-12;11	13;0-13;6
BASISRAPPORTAGE									
Verwerkingsnelheid									
Reactiesnelheid	22	121	173	138	155	148	166	77	32
Variabiliteit in reactiesnelheid	22	121	173	138	155	148	166	77	32
Reactiesnelheid tgv cognitieve belasting	22	121	173	138	155	148	166	77	32
Versnellingscapaciteit in reactiesnelheid	22	121	173	138	155	148	166	77	32
Variabiliteit in beslistijd	22	121	173	138	155	148	166	77	32
Aandacht									
Volgehouden aandacht	22	121	173	138	155	148	166	77	32
Arousalregulatie	22	121	173	138	155	148	166	77	32
Onoplettendheid	22	121	173	138	155	148	166	77	32
Executieve controle									
Slordigheid	22	121	173	138	155	148	166	77	32
Interferentiegevoeligheid	22	121	173	138	155	148	166	77	32
Wachttijdafweging	22	121	173	138	155	148	166	77	32
(Werk)geheugen									
Reactiesnelheid eenvoudig	22	121	173	138	155	148	166	77	32
Fouten eenvoudig	22	121	173	138	155	148	166	77	32
Reactiesnelheid complex	13	102	154	126	152	142	163	77	32
Fouten complex	13	102	154	126	152	142	163	77	32
Reactiesnelheid zeer complex	9	63	104	91	122	119	143	70	31
Fouten zeer complex	9	63	104	91	122	119	143	70	31
Leersnelheid									
Leersnelheid	22	121	173	138	155	148	166	77	32
Gedrag tijdens de taak									

Benodigde ondersteuning	22	121	173	138	155	147	166	77	32
Verbaliseren	22	121	173	138	155	147	166	77	31
Bewegelijkheid	22	121	173	138	155	147	166	77	32

SPECIALISTISCHE RAPPORTAGE

Mediane reactiesnelheid

Blok 1	22	121	173	138	155	148	166	77	32
Blok 2	22	121	173	138	155	148	166	77	32
Blok 3	22	121	173	138	155	148	166	77	32
Blok 4 compatibele trials	22	121	173	138	155	148	166	77	32
Blok 4 incompatibele trials	22	121	173	138	155	148	166	77	32
Blok 5 1-back	22	121	173	138	155	148	166	77	32
Blok 5 2-back	13	102	154	126	152	142	163	77	32
Blok 5 2-back PI	9	63	104	91	122	119	143	70	31
Blok 6	22	121	173	138	155	148	166	77	32
Blok 7 deel 1	22	121	173	138	155	148	166	77	32
Blok 7 deel 2	22	121	173	138	155	148	166	77	32

Intra-individuele coëfficiënt van variatie

Blok 1	22	121	173	138	155	148	166	77	32
Blok 2	22	121	173	138	155	148	166	77	32
Blok 3	22	121	173	138	155	148	166	77	32
Blok 4 compatibele trials	22	121	173	138	155	148	166	77	32
Blok 4 incompatibele trials	22	121	173	138	155	148	166	77	32
Blok 5 1-back	22	121	173	138	155	148	166	77	32
Blok 5 2-back	13	102	154	126	152	142	163	77	32
Blok 5 2-back PI	9	63	104	91	122	119	143	70	31
Blok 6	22	121	173	138	155	148	166	77	32
Blok 7 deel 1	22	121	173	138	155	148	166	77	32

Blok 7 deel 2	22	121	173	138	155	148	166	77	32	
Aantal fouten										
Blok 2	22	121	173	138	155	148	166	77	32	
Blok 3	22	121	173	138	155	148	166	77	32	
Blok 4 compatibele trials	22	121	173	138	155	148	166	77	32	
Blok 4 incompatibele trials	22	121	173	138	155	148	166	77	32	
Blok 5 1-back	22	121	173	138	155	148	166	77	32	
Blok 5 2-back	13	102	154	126	152	142	163	77	32	
Blok 5 2-back PI	9	63	104	91	122	119	143	70	31	
Blok 7 deel 1	22	121	173	138	155	148	166	77	32	
Blok 7 deel 2	22	121	173	138	155	148	166	77	32	
Contrastparameters reactiesnelheid										
Blok 2 minus Blok 1	22	121	173	138	155	148	166	77	32	
Blok 3 minus Blok 2	22	121	173	138	155	148	166	77	32	
Blok 4 minus Blok 2	22	121	173	138	155	148	166	77	32	
Blok 7 minus Blok 2	22	121	173	138	155	148	166	77	32	
Blok 4 incompatibele trials minus compatibele trials	22	121	173	138	155	148	166	77	32	
Blok 7 deel 2 minus Blok 7 deel 1	22	121	173	138	155	148	166	77	32	
Contrastparameters fouten										
Blok 3 minus Blok 2	22	121	173	138	155	148	166	77	32	
Blok 4 minus Blok 2	22	121	173	138	155	148	166	77	32	
Blok 7 minus Blok 2	22	121	173	138	155	148	166	77	32	
Blok 4 incompatibele trials minus compatibele trials	22	121	173	138	155	148	166	77	32	
Blok 7 deel 2 minus Blok 7 deel 1	22	121	173	138	155	148	166	77	32	
Blokoeverkoepelende afwijkende responsen										
Aantal premature responsen	22	121	173	138	155	148	166	77	32	
Aantal irrelevante knopdrukken	22	121	173	138	155	148	166	77	32	
Aantal extra responsen	22	121	173	138	155	148	166	77	32	

Aantal responscompetities	22	121	173	138	155	148	166	77	32
---------------------------	----	-----	-----	-----	-----	-----	-----	----	----

Uit tabel 2 blijkt dat de standaardfout onder lineair continue normeren voor alle leeftijdsgroepen binnen het leeftijdsbereik van alle testonderdelen lager is dan onder de klassieke normering met steekproefomvangen per groep van $n = 200$. Voor $n = 300$ is de standaardfout hoger voor 6 testonderdelen voor de laagste leeftijdsgroep binnen het leeftijdsbereik van de test (i.e., 6;0-6;11) en voor 3 testonderdelen voor de hoogste leeftijdsgroep (i.e., 12;0-12;11). Voor $n = 400$ zijn van alle testonderdelen de buitenste leeftijdsgroepen binnen het leeftijdsbereik van de test (i.e., 6;0-6;11 en 12;0-12;11) de standaardfouten onder lineair continue normeren hoger dan onder klassieke normering, en voor 3 testonderdelen voor de leeftijdsgroep 7;0-7;11. Volgens COTAN is dit acceptabel, zolang de mate van verslechtering bij de extreme groepen gering is, en de winst voor de middengroepen substantieel. Dit is voor alle COTAPP testonderdelen het geval, want de standaardfouten in de middengroepen zijn laag tot erg laag in vergelijking met de standaardfouten onder klassiek normeren. Samenvattend geven de resultaten aan dat de gebruikte steekproefomvang onder lineair continue normeren overeenkomt met het oordeel 'goed', als de COTAPP wordt ingezet als een test voor relatief minder belangrijke beslissingen op individueel niveau; en 'voldoende' tot 'goed', als de COTAPP wordt ingezet als een test voor belangrijke beslissingen op individueel niveau. Aangezien de COTAPP nooit als enige instrument gebruikt wordt om belangrijke beslissingen te nemen (aan-/afwezigheid ontwikkelingsstoornis, aan-/afwezigheid van onderpresteren of over-/ondervraging), maar altijd gecombineerd wordt met anamnestiche informatie en observaties binnen een proces van Bayesiaanse gevolgtrekking (i.e. klinisch redeneren), is de gebruikte steekproefomvang goed.

Bovenstaande standaardfouten zijn berekend voor een lineair continue normeringsmodel, onder aanname van normaal verdeelde residuen (i.e., het standaard lineaire regressiemodel). De gefitte modellen zijn GAMLSS modellen (Rigby, & Stasinopoulos, 2005). Deze zijn flexibeler dan het standaard lineaire regressiemodel. Deze flexibiliteit is er op twee manieren, namelijk in de relatie tussen leeftijd en de model parameters, en in de verdeling van de scores conditioneel op de leeftijd. De relatie is op niet-lineaire manier gemodelleerd, en de verdeling van de scores is gemodelleerd met de Box-Cox Power Exponential (BCPE)-verdeling, de Beta Binomiale (BB)-verdeling of, elk eenmalig, de multinomiale verdeling of normaalverdeling. Daarnaast is bij een aantal subtests de binomiale verdeling toepast om de kans op het maken van een bepaalde subtest te modelleren.

Het was nodig om deze flexibele modellen toe te passen om een goede modelfit te verkrijgen. Aan de aannames van het lineaire continue normeringsmodel is daarmee niet voldaan. Het is onbekend hoe de verhouding is tussen de schattingsnauwkeurigheid van het lineaire continue normeringsmodel en schattingsnauwkeurigheid onder de door ons toegepaste modellen.

Dit is zo omdat er tot op heden geen eenvoudige en eenduidige methode beschikbaar is om de schattingsnauwkeurigheid van de GAMLSS modellen in de context van continue normering te bepalen (zie bijv. Timmerman, Voncken & Albers, 2020). De bovenstaande aanpak van een vergelijking op basis van de geschatte standaardfout van het steekproefgemiddelde kan niet worden toegepast. Dat wil zeggen, voor GAMLSS modellen kunnen er weliswaar standaardfouten van de modelparameters (zoals μ, σ, ν, τ) berekend worden, maar die standaardfouten geven zelf weinig informatie in de context van continue normering. Het gaat hierbij immers om de nauwkeurigheid van de distributieschatting als geheel (of, anders gezegd, de percentielen (of z

scores), conditioneel op leeftijd), en niet om de nauwkeurigheid van de modelparameters zelf. Onder het lineaire model is dit wel mogelijk, omdat daarbij alleen de parameter μ afhangt van de leeftijd. Bij GAMLSS modellen daarentegen wordt de nauwkeurigheid van de distributieschatting bepaald door meerdere parameters, met ieder eigen standaardfouten, die ieder voor zich onvoldoende informatief zijn voor de schattingsnauwkeurigheid van de distributie.

Hierbij is het ook van belang om op te merken dat het klassieke model voor een kleine aantal leeftijdsgroepen van subtests weliswaar kleinere standaardfouten gelden, maar dat dit standaardfouten zijn onder vergaande aannames. Schending van die aannames leidt tot systematische fouten, die niet uitgedrukt zijn in de grootte van de standaardfouten. De bij de COTAPP gebruikte GAMLSS gebaseerde continue normeringsmodellen hebben vanwege de flexibiliteit minder aannames (terwijl overfitting voldoende wordt tegengegaan (zie Timmerman, Voncken en Albers, 2020), en dus ook een kleiner risico van grote systematische fouten. De gehanteerde GAMLSS modellen leiden dus niet alleen -in de meeste situaties- tot een kleinere standaardfout, maar geven ook nog een kleiner risico op andere fouten. Op basis hiervan, en het feit dat de standaardfouten onder lineair continue normeren voor alle testonderdelen en leeftijdsgroepen duidelijk lager zijn dan onder klassiek normeren voor $N = 300$ en $N = 200$, en voor $N = 400$ alleen voor buitenste leeftijdsgroepen binnen het leeftijdsbereik van het testonderdeel hoger, zijn we van mening dat de normgrootte minimaal equivalent is aan klassiek normeren volgens de COTAN eisen.

Tabel 2: Standaardfouten onder continue en klassiek normeren per leeftijdsgroep en testonderdeel

Leeftijdsgroep - steekproefomvang	Pooled SD					SE Continue					SE Klassiek		
	5;6-5;11	6;0-6;11	7;0-7;11	8;0-8;11	9;0-9;11	10;0-10;11	11;0-11;11	12;0-12;11	13;0-13;6	n = 200	n = 300	n = 400	
BASISRAPPORTAGE													
Verwerkingssnelheid													
Reactiesnelheid	5.00	0.34	0.27	0.21	0.17	0.16	0.18	0.22	0.28	0.35	0.35	0.29	0.25
Variabiliteit in reactiesnelheid	2.74	0.18	0.15	0.12	0.09	0.09	0.10	0.12	0.16	0.19	0.19	0.16	0.14
Reactiesnelheid ten gevolge van cognitieve belasting	195.06	13.14	10.58	8.32	6.65	6.08	6.89	8.71	11.04	13.63	13.79	11.26	9.75
Versnellingscapaciteit in reactiesnelheid	100.14	6.74	5.43	4.27	3.41	3.12	3.54	4.47	5.67	7.00	7.08	5.78	5.01
Variabiliteit in beslistijd	0.12	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01
Aandacht													
Volgehouden aandacht	181.31	12.21	9.84	7.73	6.18	5.65	6.41	8.09	10.26	12.67	12.82	10.47	9.07
Arousalregulatie	196.35	13.22	10.65	8.37	6.69	6.12	6.94	8.77	11.12	13.72	13.88	11.34	9.82
Onoplettendheid	6.78	0.46	0.37	0.29	0.23	0.21	0.24	0.30	0.38	0.47	0.48	0.39	0.34
Executieve controle													
Slordigheid	1.16	0.08	0.06	0.05	0.04	0.04	0.04	0.05	0.07	0.08	0.08	0.07	0.06
Interferentiegevoeligheid	139.58	9.40	7.57	5.95	4.76	4.35	4.93	6.23	7.90	9.76	9.87	8.06	6.98
Wachttijdafweging	2.12	0.14	0.11	0.09	0.07	0.07	0.07	0.09	0.12	0.15	0.15	0.12	0.11
(Werk)geheugen													
Reactiesnelheid eenvoudig	313.10	21.09	16.99	13.35	10.67	9.76	11.06	13.98	17.73	21.88	22.14	18.08	15.65
Fouten eenvoudig	1.99	0.13	0.11	0.08	0.07	0.06	0.07	0.09	0.11	0.14	0.14	0.11	0.10
Reactiesnelheid complex	434.57	31.50	25.43	19.98	15.78	14.02	15.55	19.62	25.01	31.05	30.73	25.09	21.73
Fouten complex	2.05	0.15	0.12	0.09	0.07	0.07	0.07	0.09	0.12	0.15	0.14	0.12	0.10
Reactiesnelheid zeer complex	450.57	39.04	31.73	25.03	19.57	16.60	17.46	21.69	27.80	34.81	31.86	26.01	22.53
Fouten zeer complex	1.98	0.17	0.14	0.11	0.09	0.07	0.08	0.10	0.12	0.15	0.14	0.11	0.10
Leersnelheid													
Leersnelheid	27.13	1.83	1.47	1.16	0.92	0.85	0.96	1.21	1.54	1.90	1.92	1.57	1.36
Gedrag tijdens de taak													
Benodigde ondersteuning	1.71	0.11	0.09	0.07	0.06	0.05	0.06	0.08	0.10	0.12	0.12	0.10	0.09
Verbalisatie	2.46	0.17	0.13	0.10	0.08	0.08	0.09	0.11	0.14	0.17	0.17	0.14	0.12
Beweeglijkheid	3.10	0.21	0.17	0.13	0.11	0.10	0.11	0.14	0.18	0.22	0.22	0.18	0.15
SPECIALISTISCHE RAPPORTAGE													
Mediane reactiesnelheid													
Blok 1	217.74	14.66	11.81	9.29	7.42	6.79	7.69	9.72	12.33	15.22	15.40	12.57	10.89
Blok 2	155.02	10.44	8.41	6.61	5.28	4.83	5.48	6.92	8.78	10.83	10.96	8.95	7.75
Blok 3	163.34	11.00	8.86	6.97	5.57	5.09	5.77	7.29	9.25	11.41	11.55	9.43	8.17
Blok 4 compatibele trials	190.77	12.85	10.35	8.14	6.50	5.94	6.74	8.52	10.80	13.33	13.49	11.01	9.54
Blok 4 incompatibele trials	227.74	15.34	12.36	9.71	7.76	7.10	8.05	10.17	12.89	15.92	16.10	13.15	11.39
Blok 5 1-back	313.10	21.09	16.99	13.35	10.67	9.76	11.06	13.98	17.73	21.88	22.14	18.08	15.65
Blok 5 2-back	434.57	31.50	25.43	19.98	15.78	14.02	15.55	19.62	25.01	31.05	30.73	25.09	21.73
Blok 5 2-back PI	450.57	39.04	31.73	25.03	19.57	16.60	17.46	21.69	27.80	34.81	31.86	26.01	22.53
Blok 6	661.01	44.52	35.86	28.19	22.52	20.60	23.35	29.51	37.42	46.20	46.74	38.16	33.05

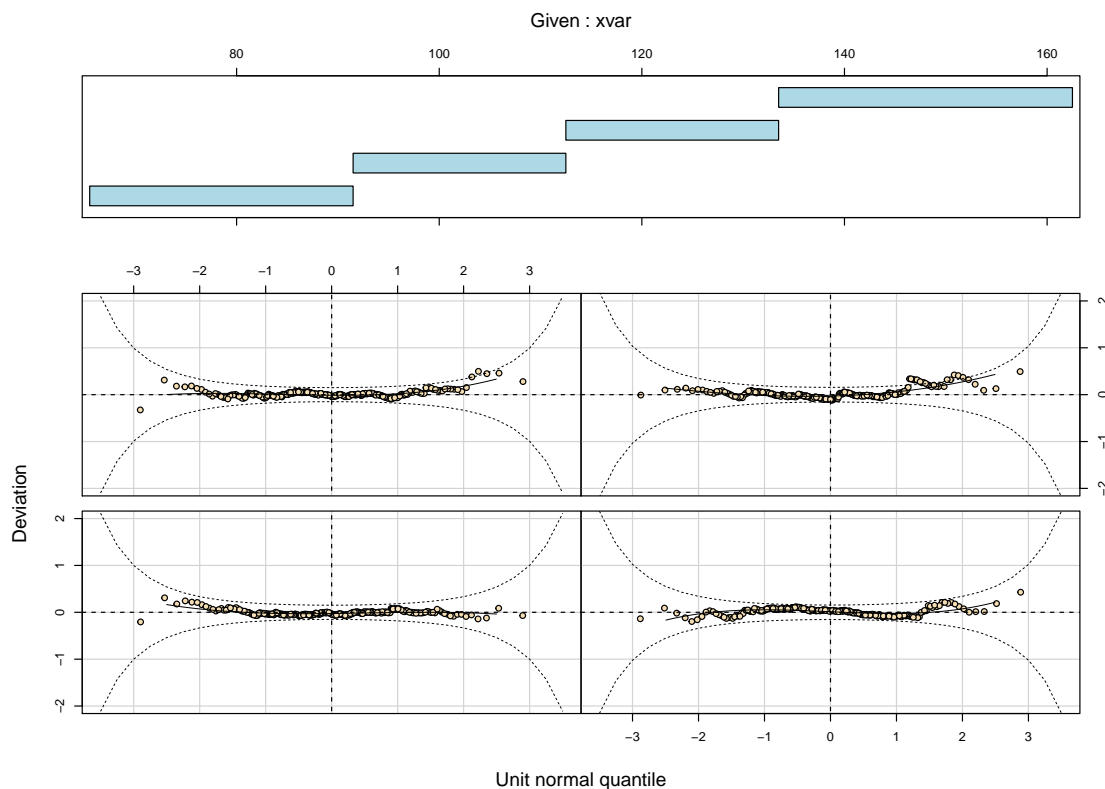
Leeftijdsgroep - steekproefomvang	Pooled SD			SE Continue							SE Klassiek		
	5;6-5;11	6;0-6;11	7;0-7;11	8;0-8;11	9;0-9;11	10;0-10;11	11;0-11;11	12;0-12;11	13;0-13;6	$n = 200$	$n = 300$	$n = 400$	
Blok 7 deel 1	237.35	15.99	12.88	10.12	8.09	7.40	8.39	10.60	13.44	16.59	16.78	13.70	11.87
Blok 7 deel 2	318.26	21.43	17.27	13.57	10.85	9.92	11.24	14.21	18.02	22.24	22.50	18.38	15.91
Intra-individuele coefficient van variatie													
Blok 1	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Blok 2	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blok 3	0.05	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blok 4 compatibele trials	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blok 4 incompatibele trials	0.06	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Blok 5 1-back	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Blok 5 2-back	0.08	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Blok 5 2-back PI	0.09	0.01	0.01	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.00	0.00
Blok 6	0.12	0.01	0.01	0.00	0.00	0.00	0.00	0.01	0.01	0.01	0.01	0.01	0.01
Blok 7 deel 1	0.07	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Blok 7 deel 2	0.09	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.01	0.00	0.00
Aantal fouten													
Blok 2	2.44	0.16	0.13	0.10	0.08	0.08	0.09	0.11	0.14	0.17	0.17	0.14	0.12
Blok 3	2.91	0.20	0.16	0.12	0.10	0.09	0.10	0.13	0.16	0.20	0.21	0.17	0.15
Blok 4 compatibele trials	1.62	0.11	0.09	0.07	0.06	0.05	0.06	0.07	0.09	0.11	0.11	0.09	0.08
Blok 4 incompatibele trials	2.68	0.18	0.15	0.11	0.09	0.08	0.09	0.12	0.15	0.19	0.19	0.15	0.13
Blok 5 1-back	1.99	0.13	0.11	0.08	0.07	0.06	0.07	0.09	0.11	0.14	0.14	0.11	0.10
Blok 5 2-back	2.05	0.15	0.12	0.09	0.07	0.07	0.07	0.09	0.12	0.15	0.14	0.12	0.10
Blok 5 2-back PI	1.98	0.17	0.14	0.11	0.09	0.07	0.08	0.10	0.12	0.15	0.14	0.11	0.10
Blok 7 deel 1	1.87	0.13	0.10	0.08	0.06	0.06	0.07	0.08	0.11	0.13	0.13	0.11	0.09
Blok 7 deel 2	1.30	0.09	0.07	0.06	0.04	0.04	0.05	0.06	0.07	0.09	0.09	0.07	0.06
Contrastparameters reactiesnelheid													
Blok 2 minus Blok 1	195.06	13.14	10.58	8.32	6.65	6.08	6.89	8.71	11.04	13.63	13.79	11.26	9.75
Blok 4 minus Blok 2	100.14	6.74	5.43	4.27	3.41	3.12	3.54	4.47	5.67	7.00	7.08	5.78	5.01
Blok 7 minus Blok 2	181.31	12.21	9.84	7.73	6.18	5.65	6.41	8.09	10.26	12.67	12.82	10.47	9.07
Blok 4 incompatibele trials minus compatibele trials	86.41	5.82	4.69	3.69	2.94	2.69	3.05	3.86	4.89	6.04	6.11	4.99	4.32
Blok 7 deel 2 minus Blok 7 deel 1	196.35	13.22	10.65	8.37	6.69	6.12	6.94	8.77	11.12	13.72	13.88	11.34	9.82
Contrastparameters fouten													
Blok 3 minus blok 2	3.11	0.21	0.17	0.13	0.11	0.10	0.11	0.14	0.18	0.22	0.22	0.18	0.16
Blok 4 minus Blok 2	3.48	0.23	0.19	0.15	0.12	0.11	0.12	0.16	0.20	0.24	0.25	0.20	0.17
Blok 7 minus Blok 2	2.57	0.17	0.14	0.11	0.09	0.08	0.09	0.11	0.15	0.18	0.18	0.15	0.13
Blok 4 incompatibele trials minus compatibele trials	2.43	0.16	0.13	0.10	0.08	0.08	0.09	0.11	0.14	0.17	0.17	0.14	0.12
Blok 7 deel 2 minus Blok 7 deel 1	1.53	0.10	0.08	0.07	0.05	0.05	0.05	0.07	0.09	0.11	0.11	0.09	0.08
Blokoverkoepelende afwijkende responsen													
Aantal premature responsen	1.88	0.13	0.10	0.08	0.06	0.06	0.07	0.08	0.11	0.13	0.13	0.11	0.09
Aantal irrelevante knopdrukken	1.44	0.10	0.08	0.06	0.05	0.04	0.05	0.06	0.08	0.10	0.10	0.08	0.07
Aantal extra responsen	27.13	1.83	1.47	1.16	0.92	0.85	0.96	1.21	1.54	1.90	1.92	1.57	1.36
Aantal responscompetities	2.64	0.18	0.14	0.11	0.09	0.08	0.09	0.12	0.15	0.18	0.19	0.15	0.13

1.2 Modelpassing van de gebruikte normeringsmodellen

In deze sectie wordt informatie gegeven over de passing van de GAMLSS eindmodellen. Dit gebeurt door de wormplot, en door de centielcurveplot (voor een continue uitkomstvariabele) of de z -curveplot (voor een discrete uitkomstvariabele) van ieder gefit model weer te geven. Hieronder geven we een uitgebreide toelichting op deze plots voor één testonderdeel.

In Figuur 1 is de wormplot van 'Mediane reactiesnelheid Blok 1' (Specialistische rapportage) te zien. Een wormplot is een serie van detrended Q-Q plots, per subrange van de leeftijd. De worm plot visualiseert hoe goed het model de steekproefdata past. De blauwe strepen boven de worm plots geven de leeftijdsrange in elke detrended Q-Q plot weer; deze zijn rijgewijs geordend van onder-links naar boven-rechts. In iedere Q-Q plot wordt ook een 95% betrouwbaarheidsband geplott; onder het model is de verwachting dat 95% van de observaties binnen de 95% betrouwbaarheidsband ligt. Een wormplot met alle observaties erg dicht bij de horizontale lijn kan wijzen op een model dat lijdt aan overfitting - en dat is geen goed idee. Voor meer uitleg over en illustraties van dit fenomeen verwijzen we naar Timmerman, Voncken & Albers (2020).

Figuur 1: Wormplot van het normeringsmodel voor 'Mediane reactiesnelheid Blok 1'

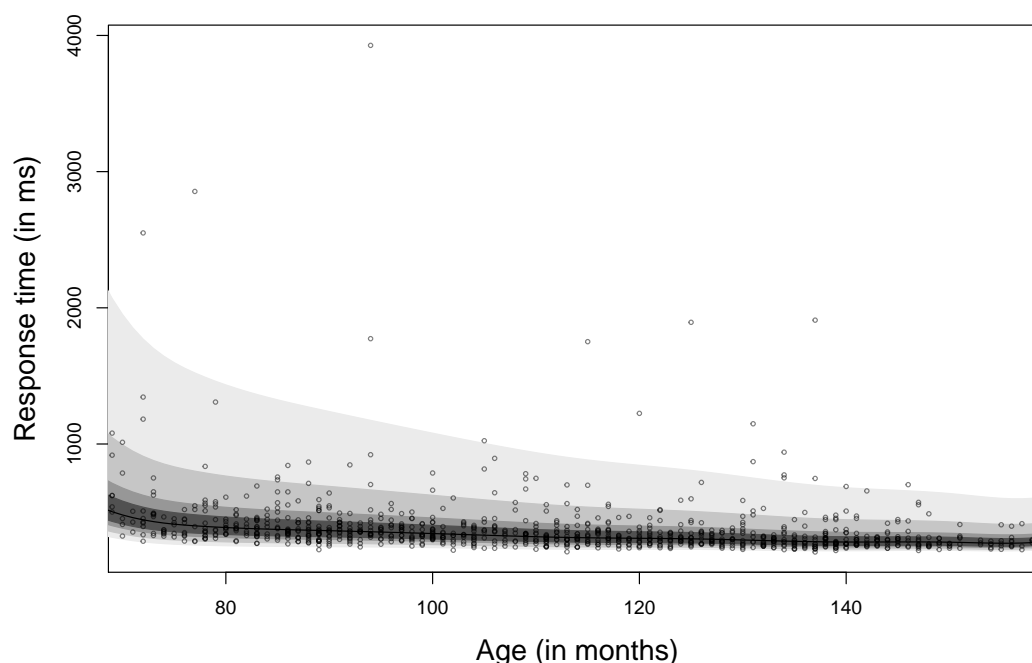


De wormplot van het model voor 'Mediane reactiesnelheid Blok 1' (Specialistische rapportage) in Figuur 1 laat zien dat de meeste observaties (i.e., punten) binnen de 95% betrouwbaarheidsbanden liggen. Enige afwijking is te zien voor de oudste leeftijdsgroep (plot rechts-boven), waarbij relatief veel positieve afwijkingen voor lijken te komen - dus iets hogere geobserveerde scores dan het model voorspelt. De wormplot suggereert al met al een goede passing van het model over de gehele leeftijdsrange, en biedt daarmee ondersteuning voor het gekozen model en

daarmee de normen.

In Figuur 2 is de centielcurveplot van 'Mediane reactiesnelheid Blok 1' te zien. Hierin zijn de centielcurves voor de centielen 1, 5, 10, 25, 50, 75, 90, 95 en 99, en de geobserveerde normeringssteekproefscores weergegeven als een functie van de leeftijd. De middelste lijn geeft de mediaan (i.e., 50ste centiel) weer; de grenzen van de middelste, donkerste band geven Q1 (i.e., 25ste centiel) en Q3 (i.e., 75ste centiel) weer; de grenzen van de volgende, één-na-donkerste band geven de grenzen van het 10de centiel en het 90ste centiel weer; etcetera.

Figuur 2: Centielcurveplot van het normeringsmodel voor 'Mediane reactiesnelheid Blok 1'



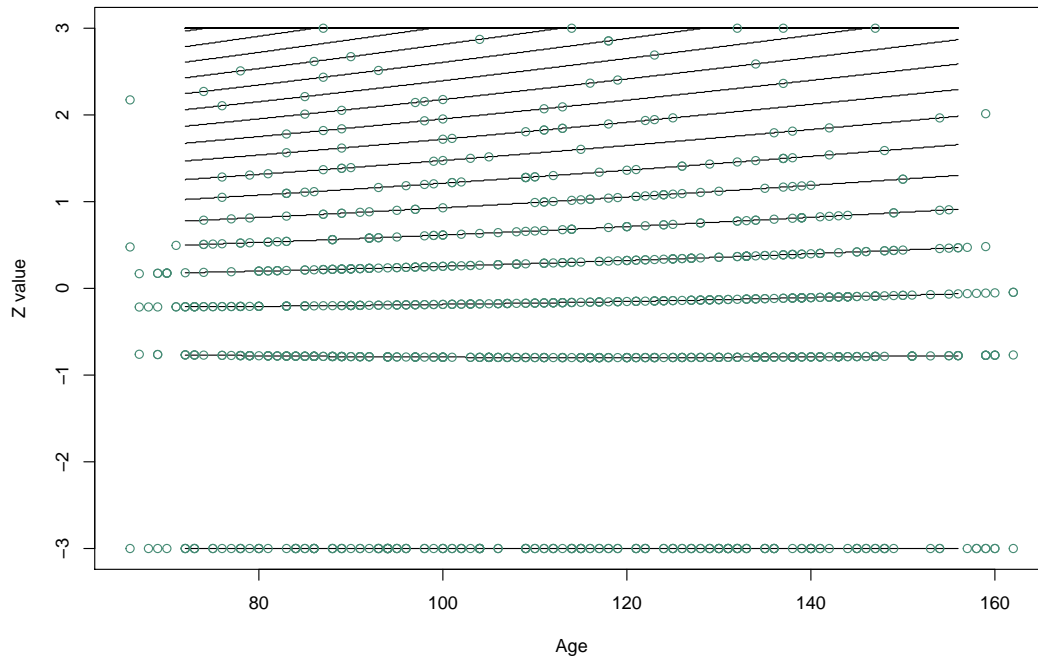
In Figuur 3 is een z -curveplot te zien. Een z -curveplot toont de model geïmpliceerde (genormaliseerde) z -curves per mogelijke ruwe score van de test en de geobserveerde normeringssteekproefscores, beide als een functie van de leeftijd.

Een centielcurveplot en een z -curveplot laten zien

- in hoeverre het model, en daarmee de normen in overeenstemming zijn met theoretische verwachtingen; en
- in welke regio's de normen goed ondersteund worden door empirische data.

Ad a) Voor het grootste deel van de COTAPP testonderdelen geldt dat op theoretische gronden verwacht mag worden dat de verdeling van de ruwe scores met toenemende leeftijd lager wordt (dat wil zeggen: een betere prestatie aangeeft). Verder geldt voor alle testonderdelen dat de relatie tussen de ruwe scoreverdeling en de leeftijd redelijk glad is (dat wil zeggen: niet al te grote sprongen laat zien). Deze testeigenschappen zijn zichtbaar in centielcurveplots en z -curveplots.

Figuur 3: z -curveplot van het normeringsmodel voor 'Aantal fouten Blok 2'



Ad b) De bolletjes in de centielcurveplot geven de geobserveerde scores in de normeringssteekproef weer. Op basis van deze scores is het betreffende GAMLSS model gefit. In de figuur is te zien of de centielcurves dan wel de z -curves over de gehele leeftijdsrange van de test ondersteund worden.

De centielcurveplot van het model voor 'Mediane reactiesnelheid Blok 1' in figuur 2 laat zien dat de centielen monotoon niet-stijgend zijn met de leeftijd (Age), en dat deze gladde curves hebben. Dit is in lijn met de theoretische verwachtingen. Verder is te zien dat het model ondersteund wordt door empirische data over de gehele leeftijdsrange. Deze centielcurveplot biedt daarmee ondersteuning voor het gekozen model en daarmee de normen.

De z -curveplot van het model voor 'Aantal fouten Blok 2' (Specialistische rapportage) in figuur 3 laat zien dat de z -waarden per mogelijke geobserveerde score monotoon niet-stijgend zijn met de leeftijd (Age) - voor iedere mogelijke score, en dat deze gladde curves hebben. Dit is in lijn met de theoretische verwachtingen. Verder is te zien dat het model ondersteund wordt door empirische data over de gehele leeftijdsrange. Deze z -curveplot biedt daarmee ondersteuning voor het gekozen model en daarmee de normen.

Merk op dat in de z -curveplot in figuur 3 ook te zien is dat er maar drie negatieve z -curves voorkomen in de normen (i.e., -3, rond -.8, en rond -.2) van 'Aantal fouten Blok 2'. Dit komt doordat de ruwe uitkomstvariabele discreet, scheef verdeeld is. Het grootste deel van de kinderen scoort heel goed op dit testonderdeel. Dit is geheel niet bezwaarlijk voor het gebruik van de normen, omdat alleen relatief slechte scores (i.e., hoge z -waarden) een klinische betekenis hebben, en niet de scores in het normale scoringsbereik.

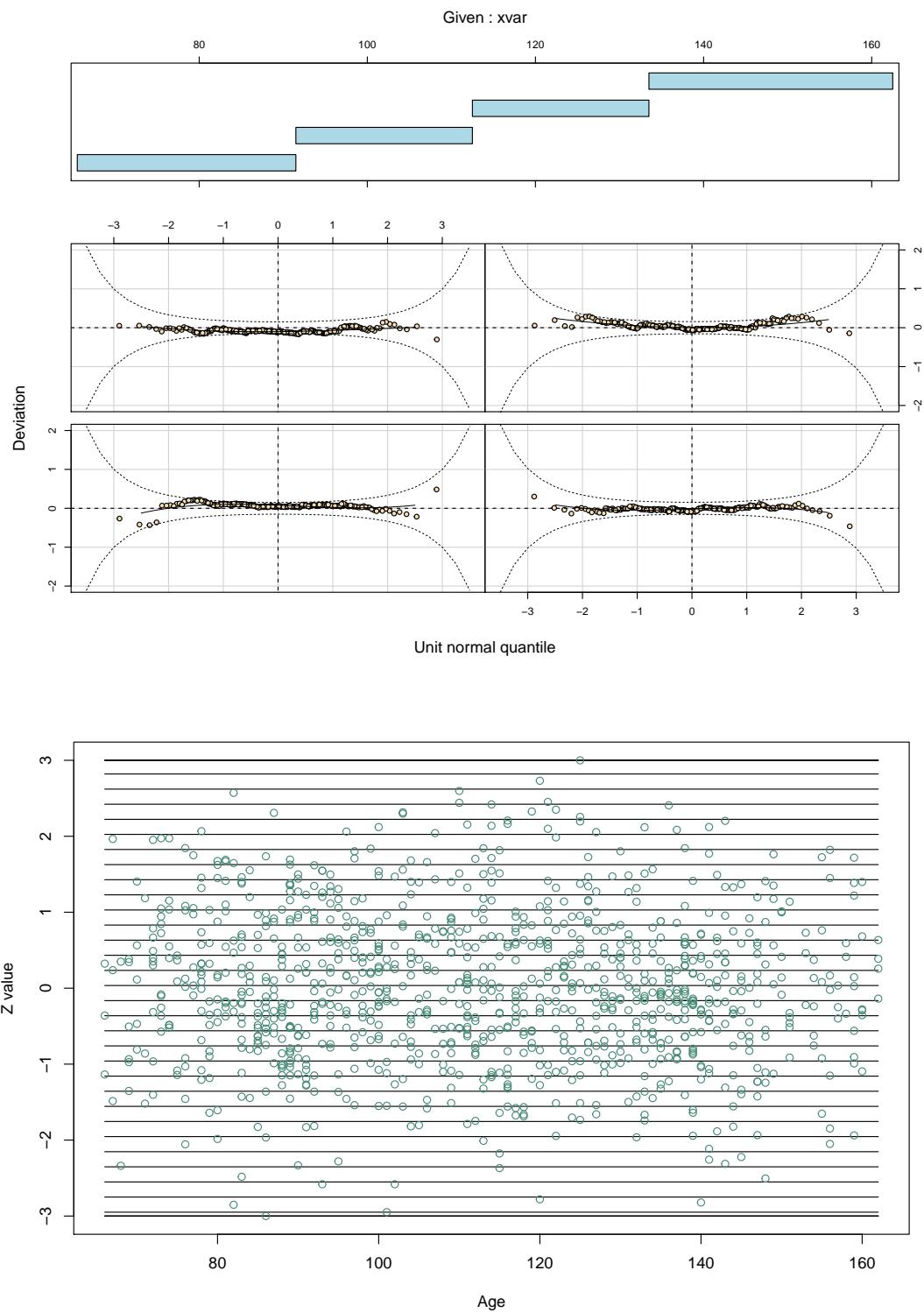
Op de volgende pagina's staan de wormplots en centielcurveplots dan wel z -curveplots voor alle testonderdelen weergegeven. Voor zover er voor een testonderdeel een effect van geslacht was (en dus aparte normen per geslacht beschikbaar zijn) zijn er aparte centielcurveplots dan wel z -curveplots gegeven per sekse. Er is één wormplot weergegeven, omdat het GAMLSS model gefit is op de gehele normeringssteekproef (met in dergelijke gevallen leeftijd en geslacht als predictoren). Waar we dat nodig achten, is een korte toelichting gegeven.

De wormplots en centielcurveplots voor alle uitkomstparameters laten zonder uitzondering een gelijkaardig beeld zien van een goede modelpassing, en modellen in overeenstemming met de theoretische verwachtingen, en ondersteund door empirische data in de gehele leeftijdsrange. Daarmee bieden deze plots ondersteuning voor alle gekozen modellen, en daarmee de normen van alle uitkomstparameters.

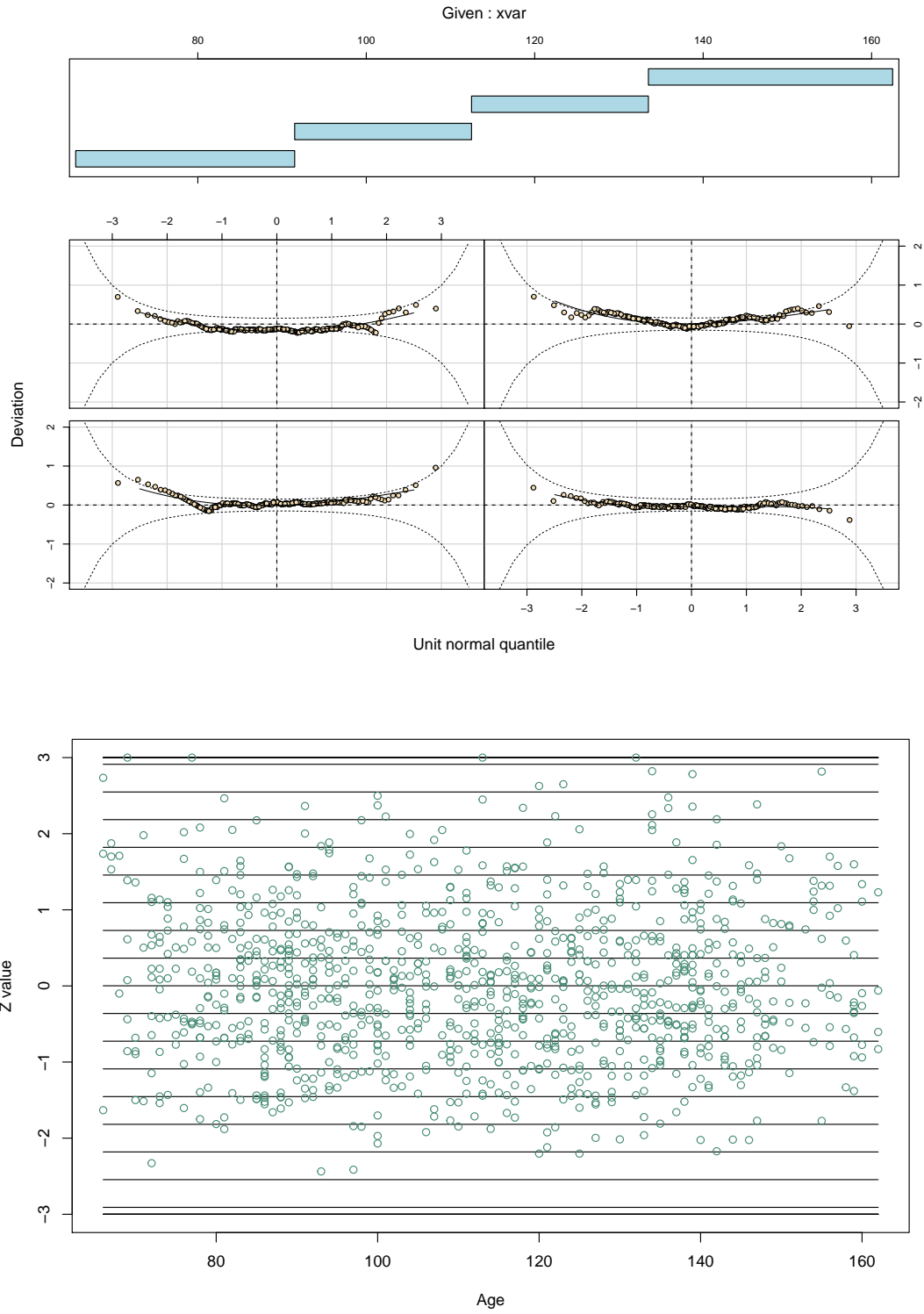
1.2.1 Basisrapportage

VERWERKINGSSNELHEID

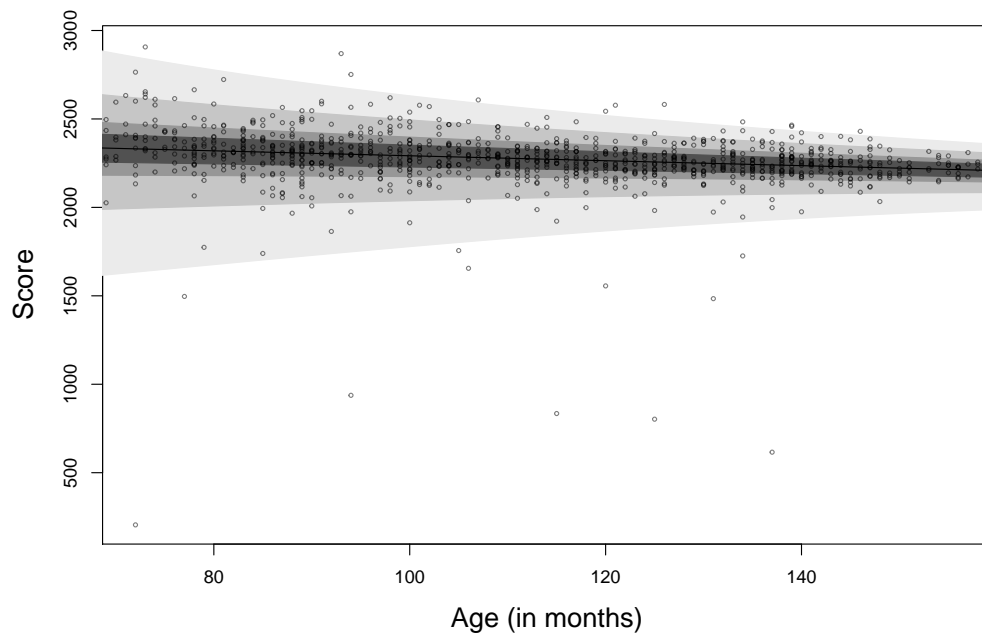
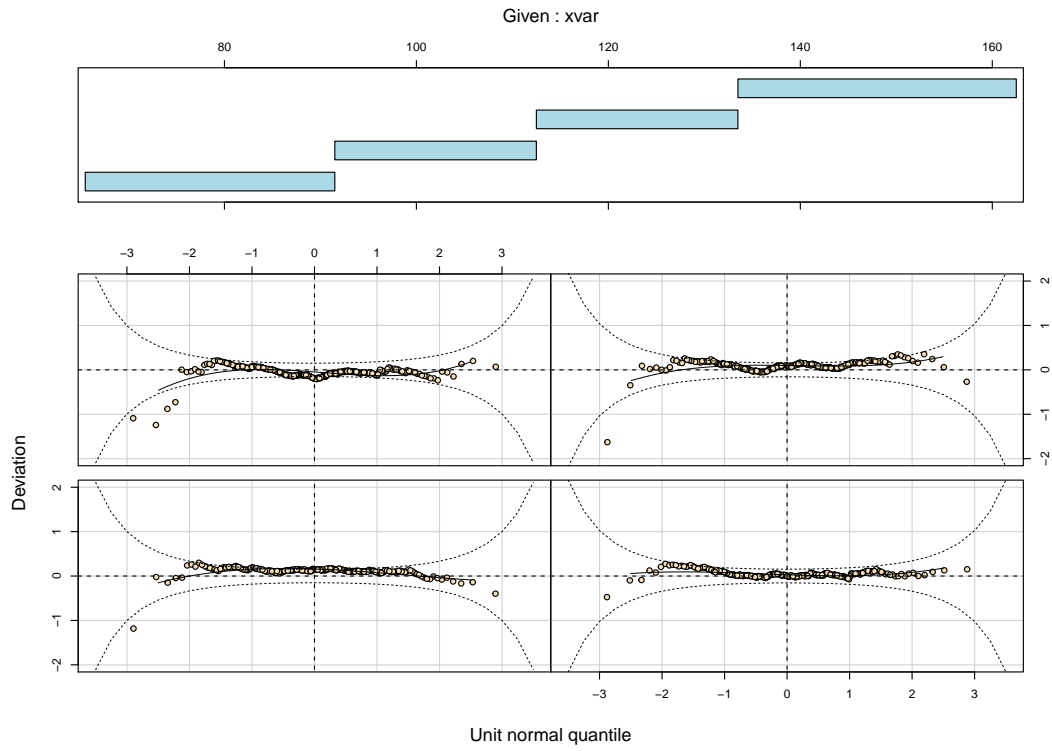
Reactiesnelheid



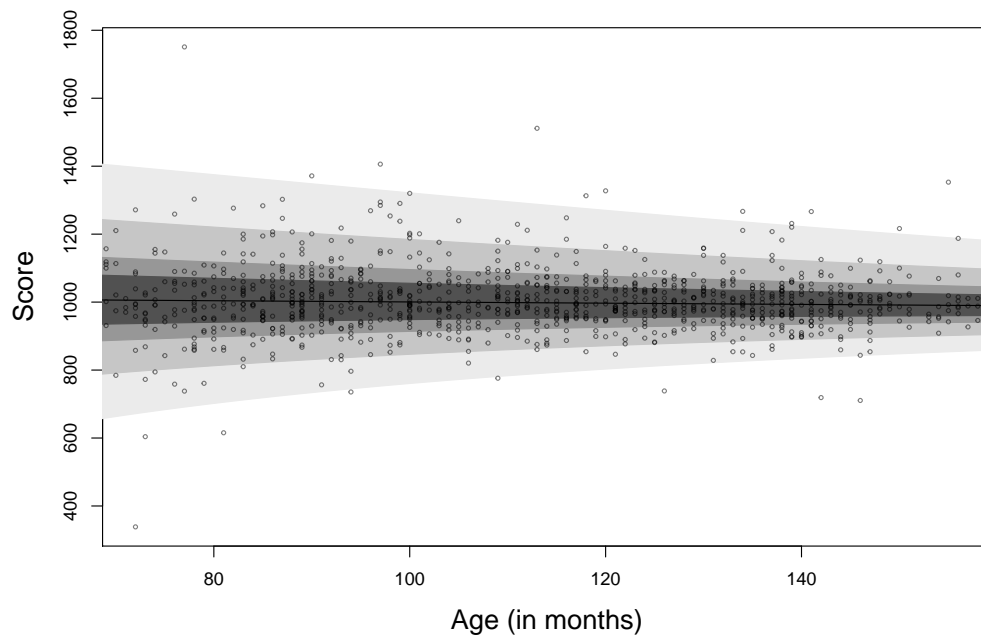
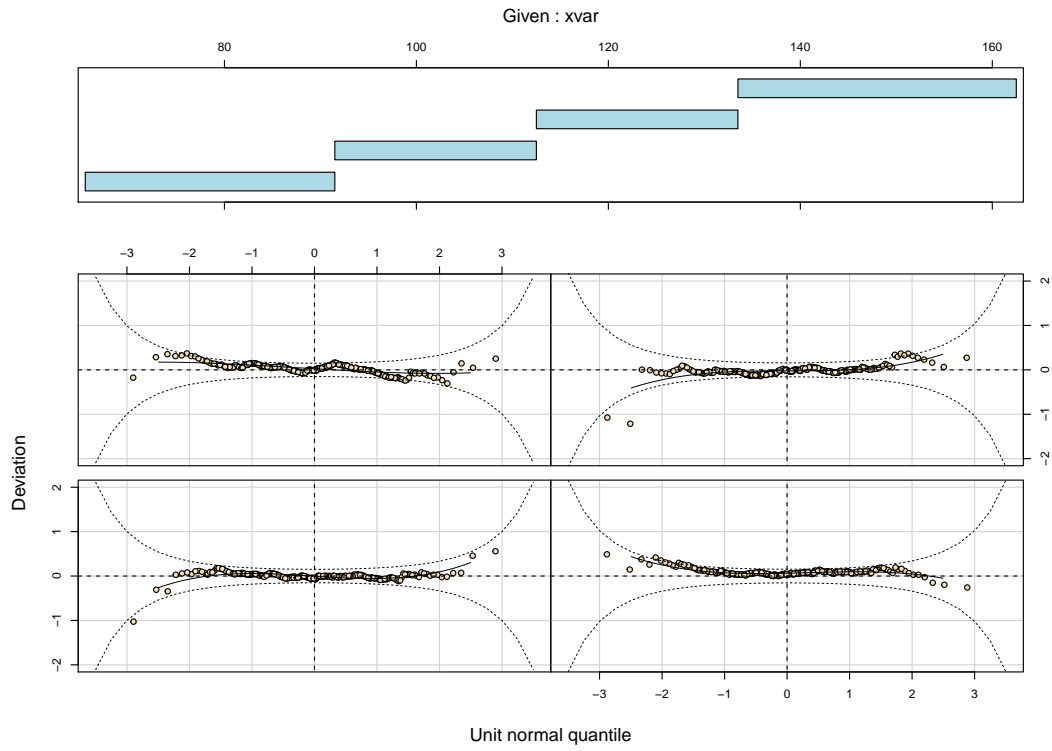
Variabiliteit in reactiesnelheid



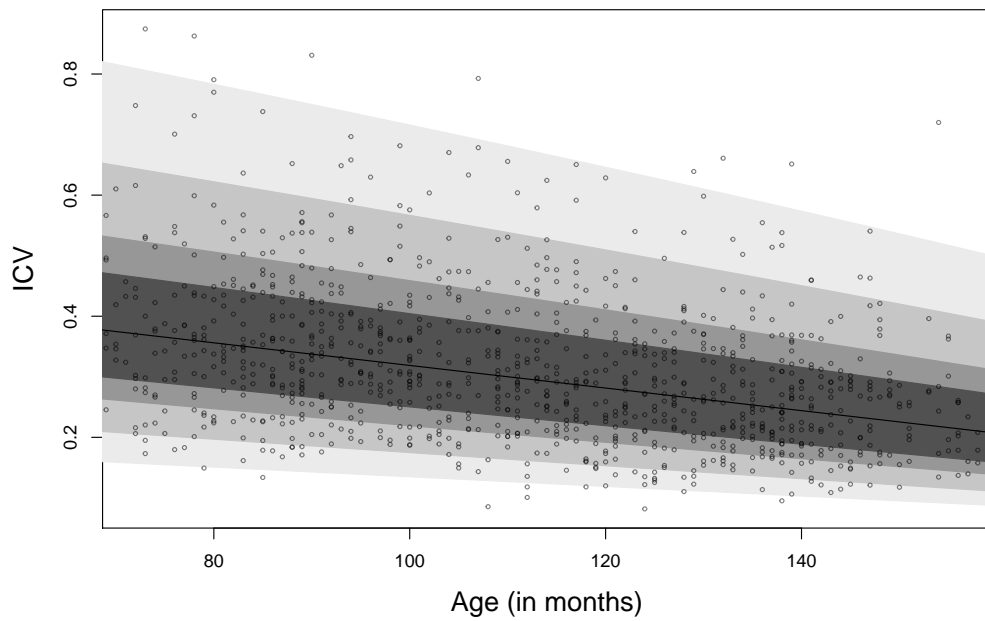
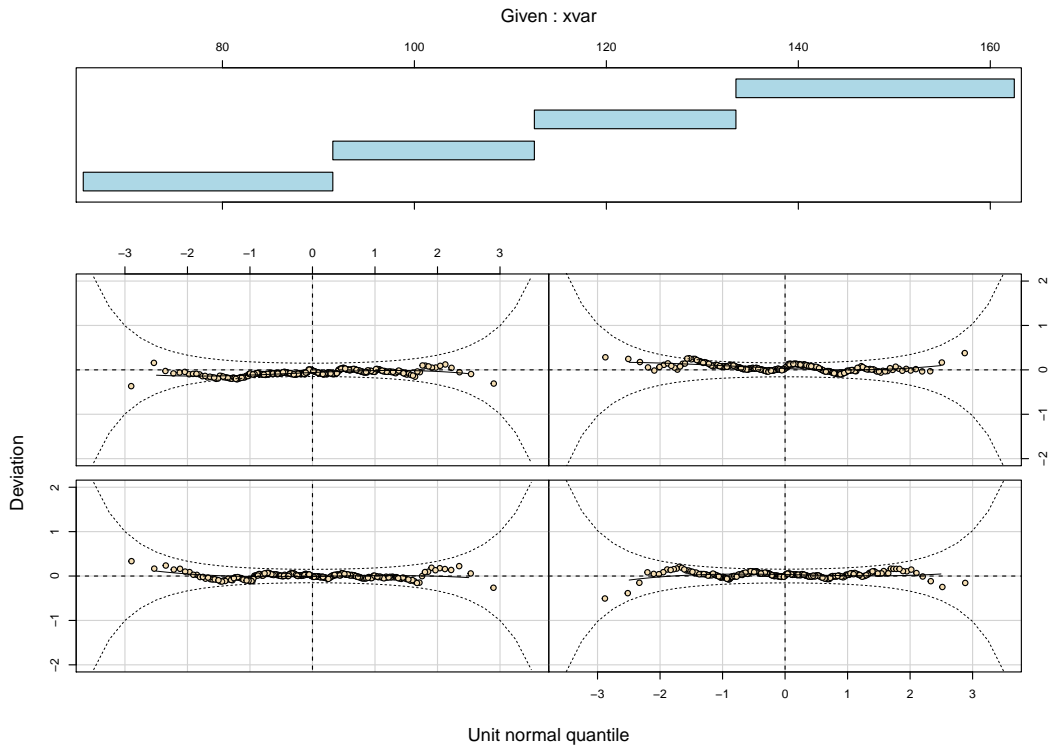
Reactiesnelheid ten gevolge van cognitieve belasting



Versnellingscapaciteit in reactiesnelheid

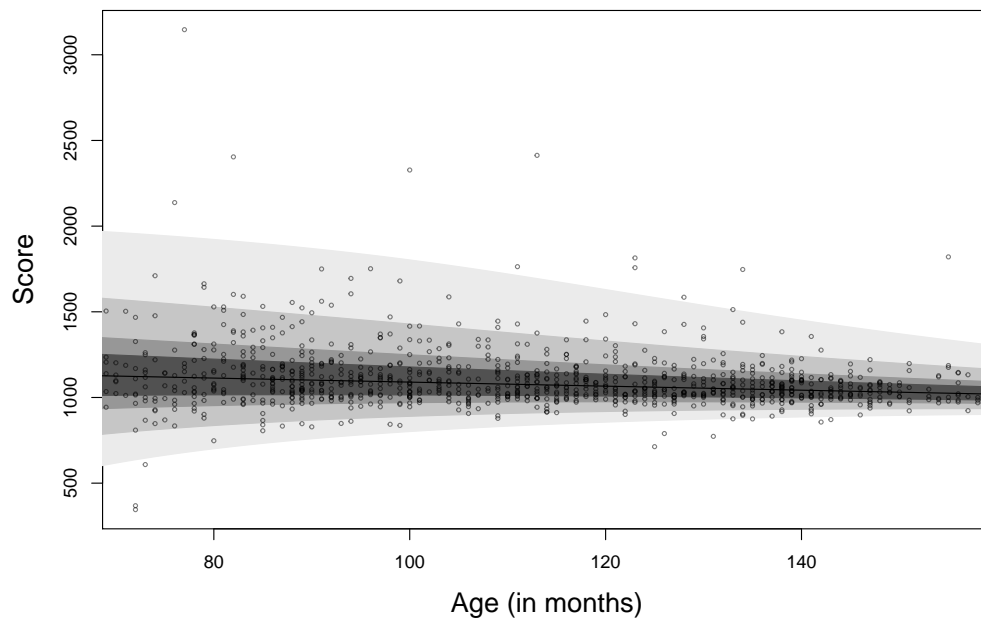
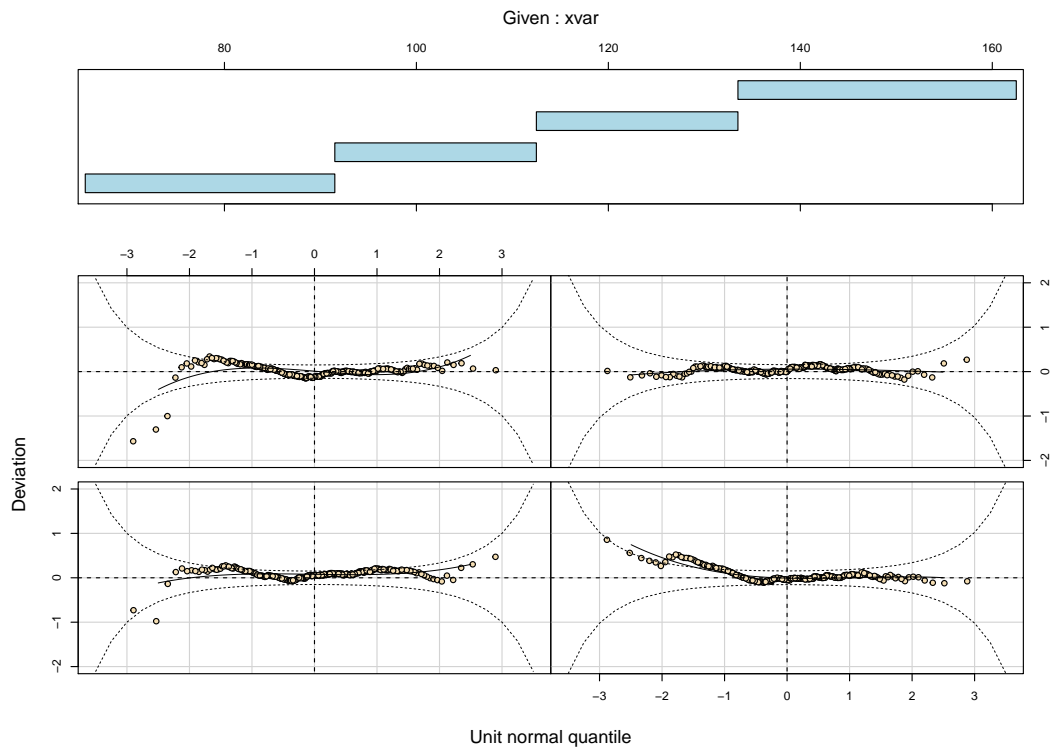


Variabiliteit in beslistijd

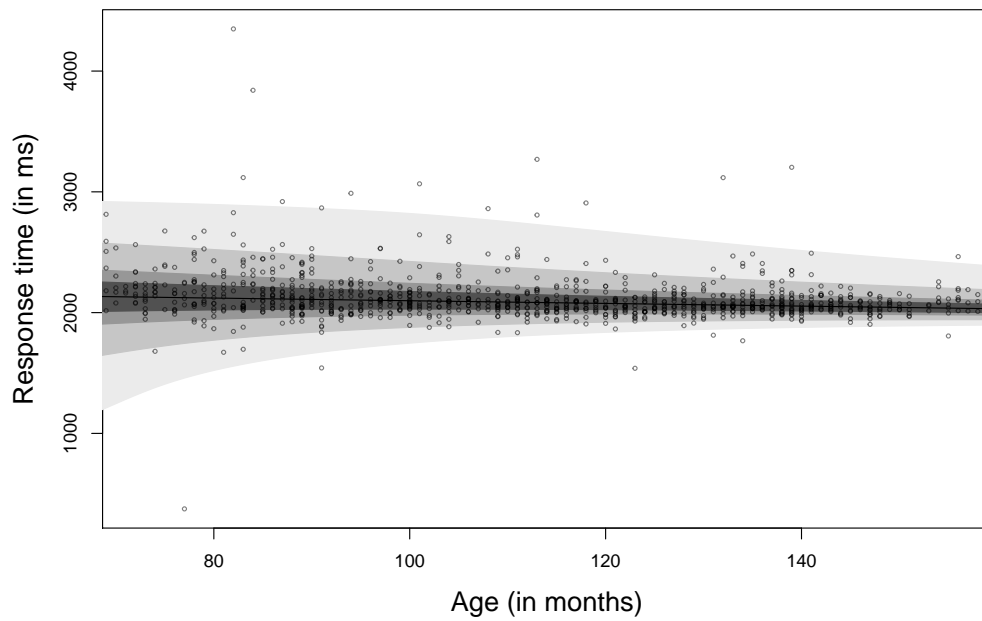
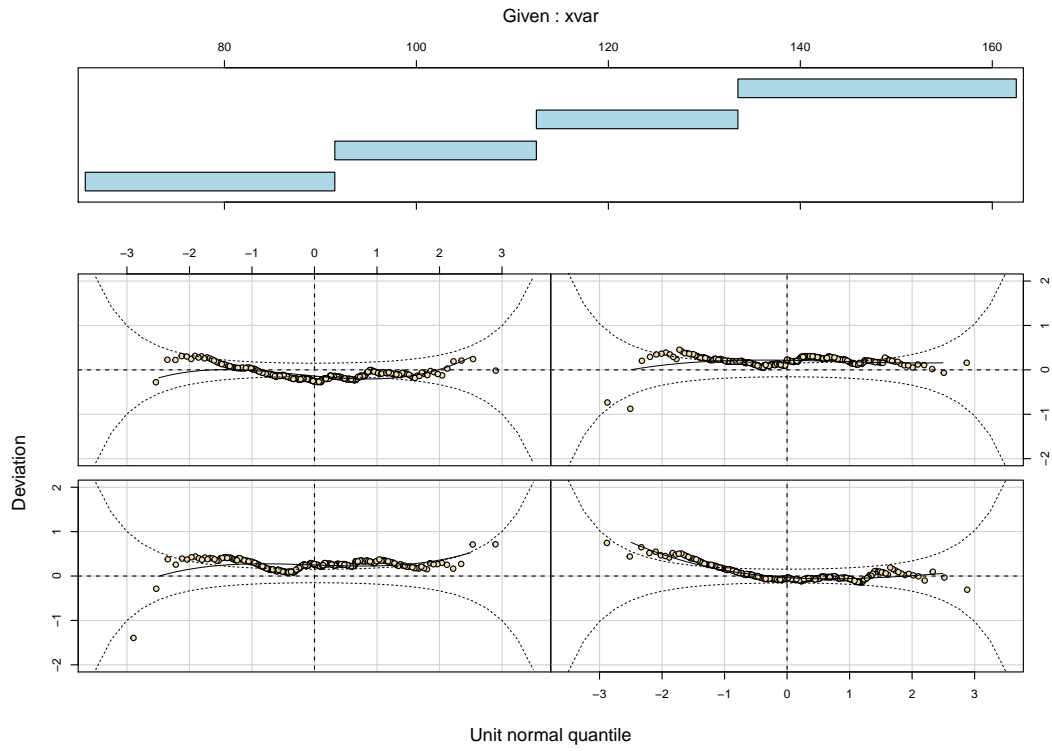


AANDACHT

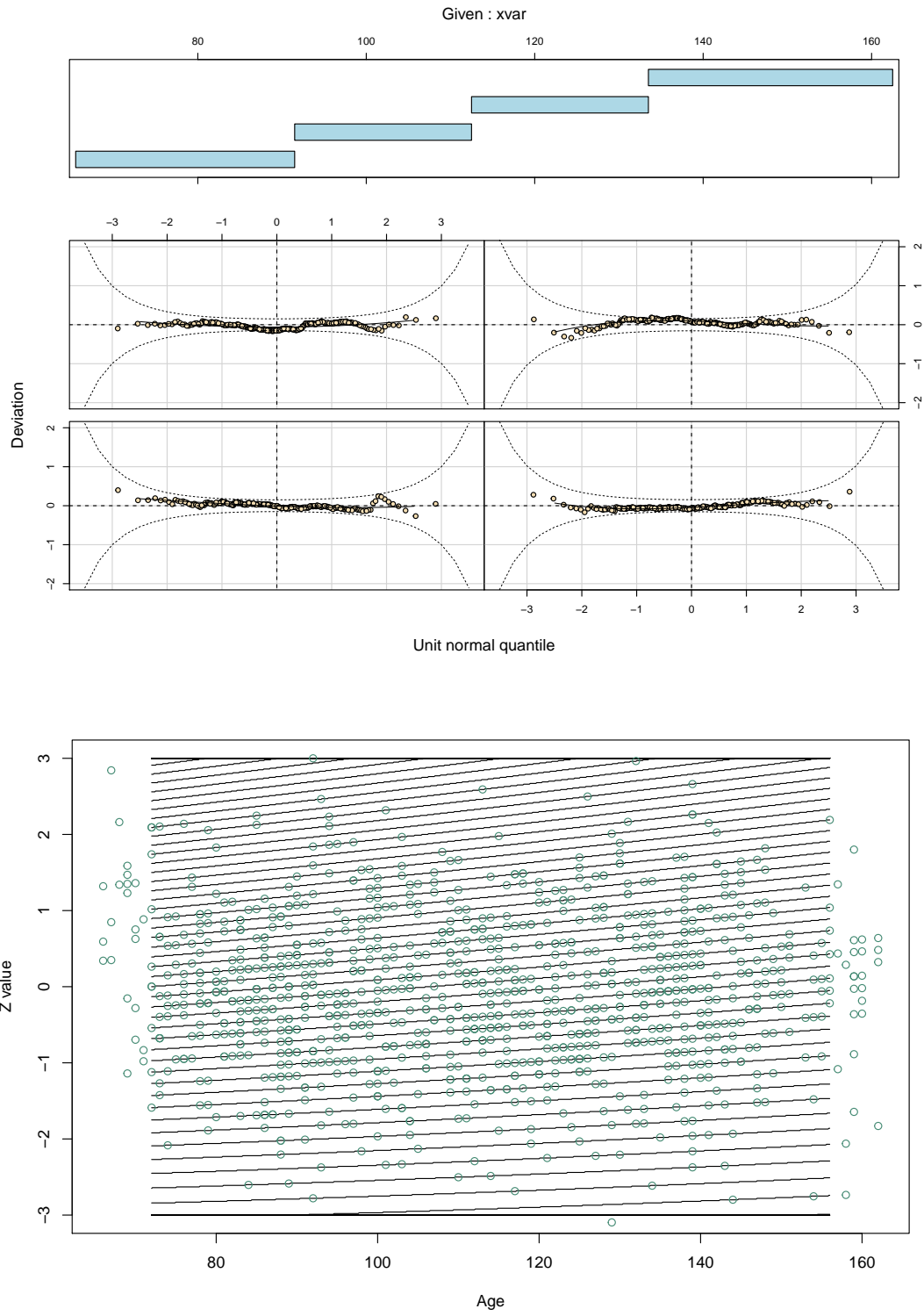
Volgehouden aandacht



Arousalregulatie

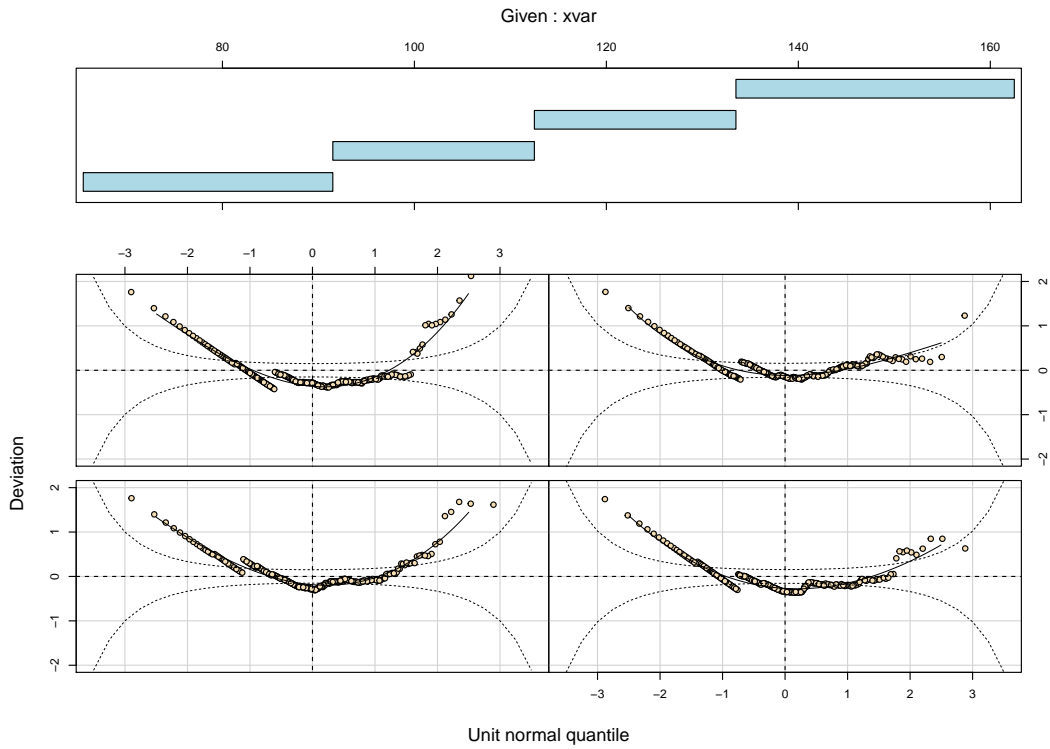


Onoplettendheid

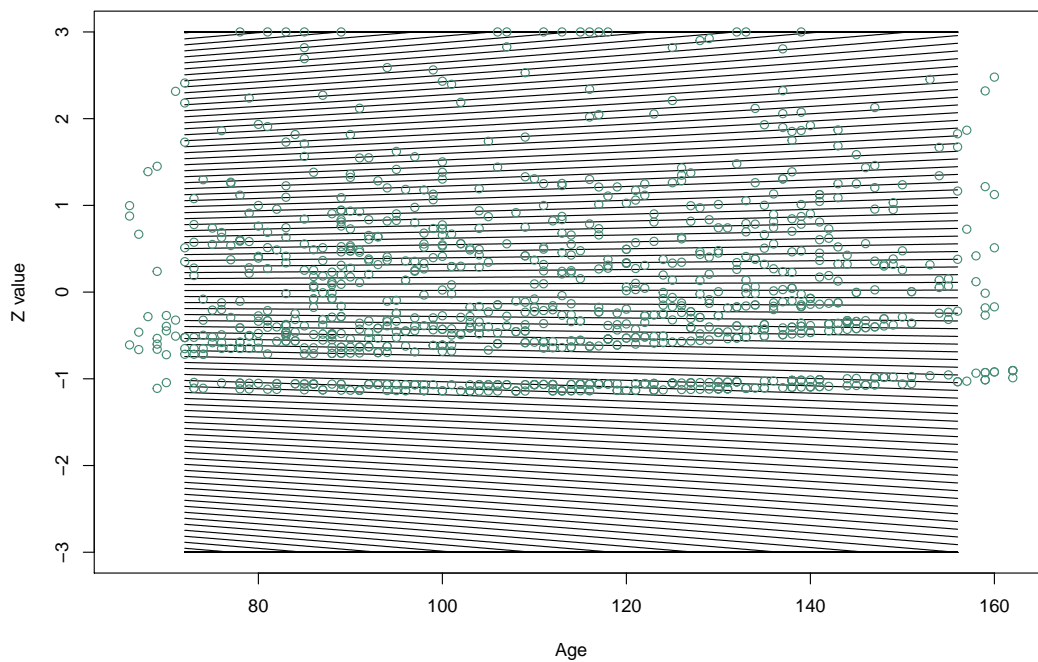


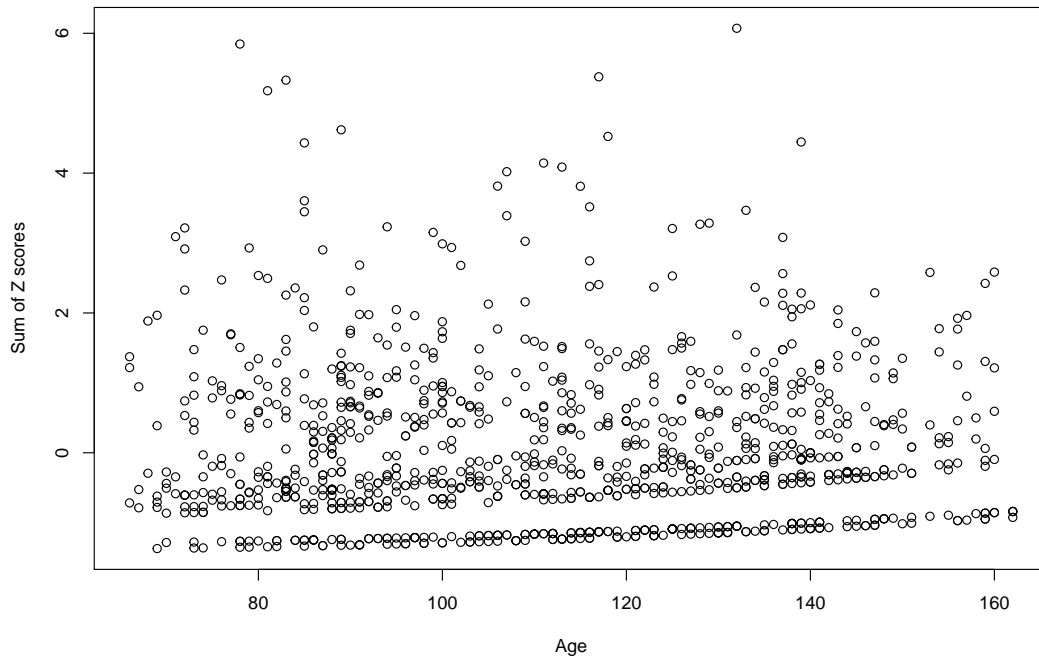
EXECUTIEVE CONTROLE

Slordigheid

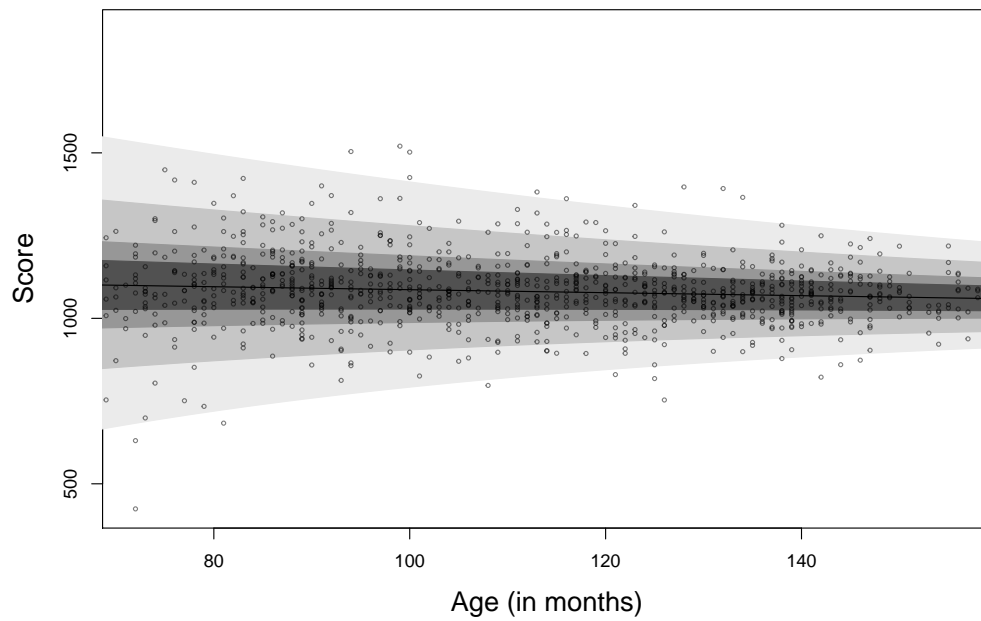
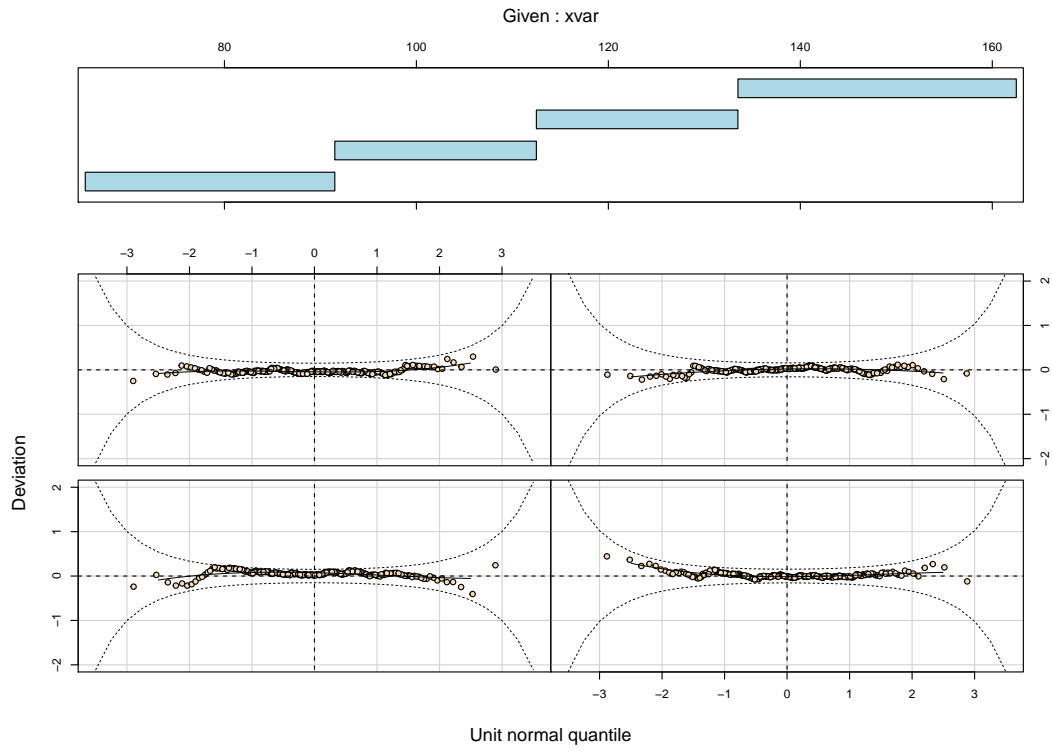


In de z -score plot is te zien dat er geen genormeerde z -scores -1.5 zijn. Dit is te danken aan het plafondeffect van de subtest. Dit heeft geen effect op de klinische interpretatie, omdat hiervoor alleen de relatief slechte scores van belang zijn.

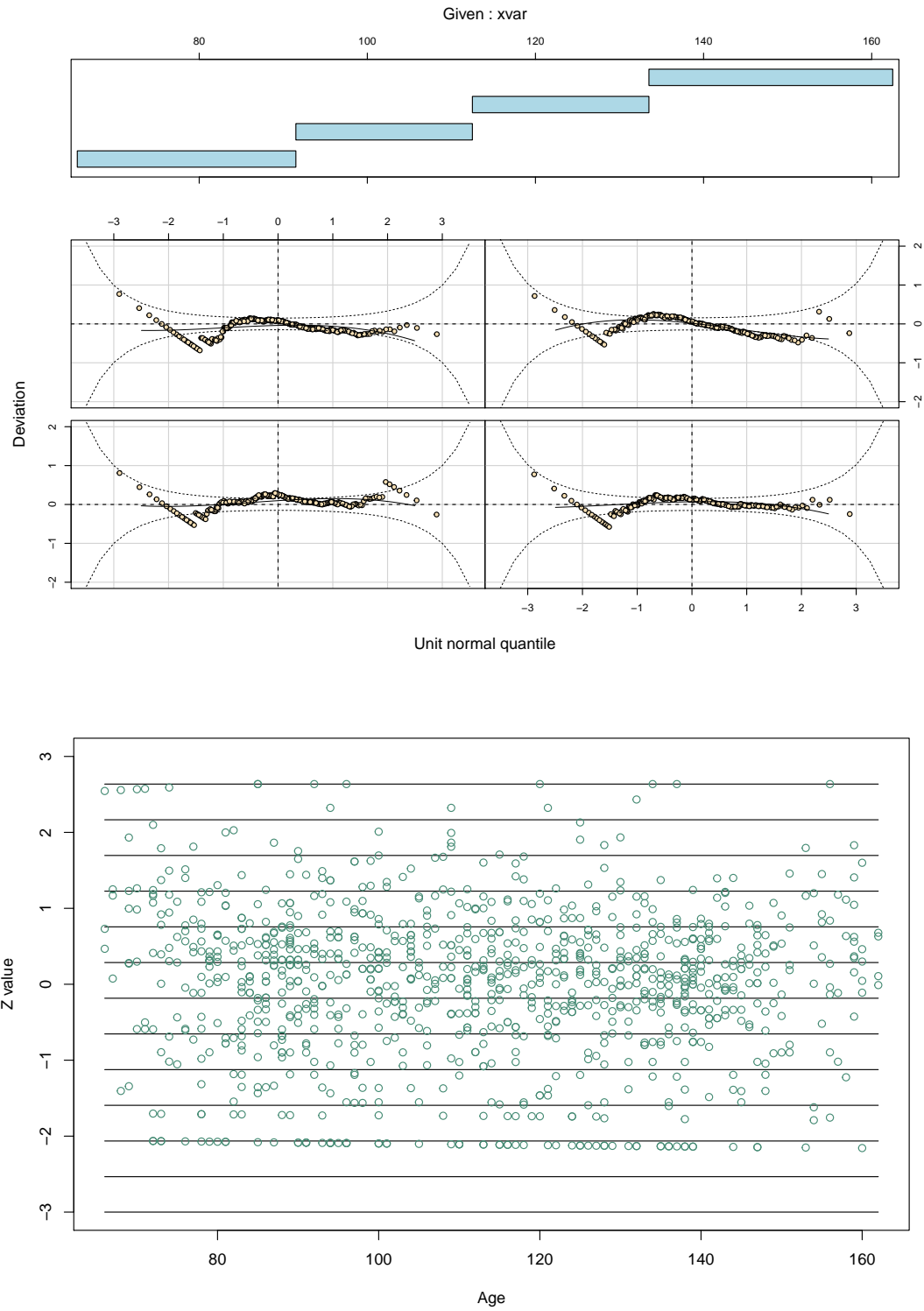




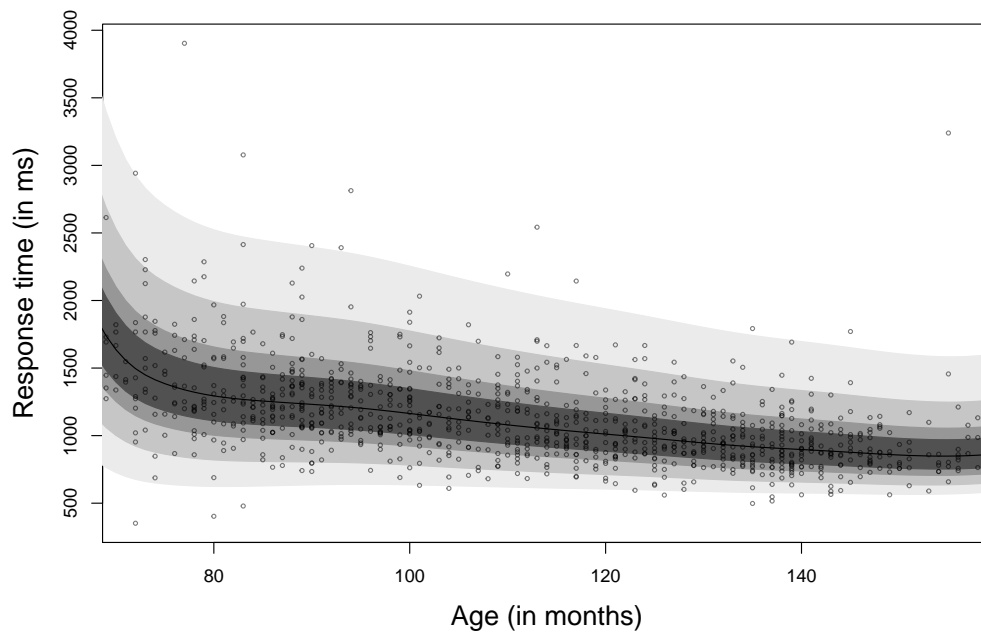
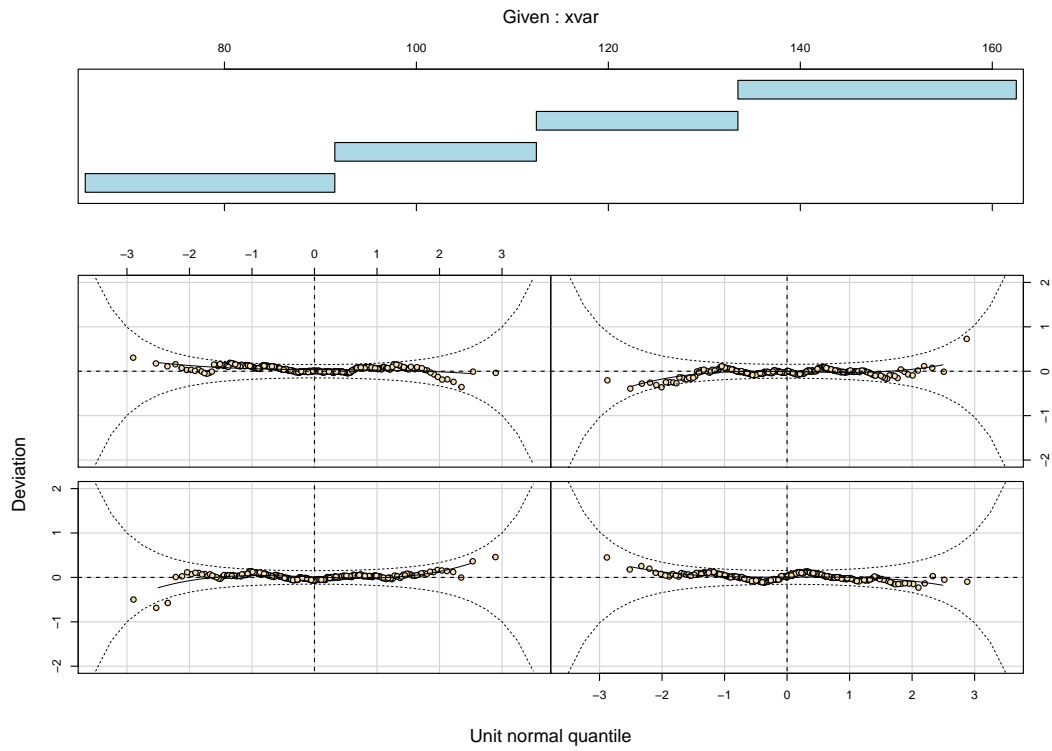
Interferentiegevoeligheid



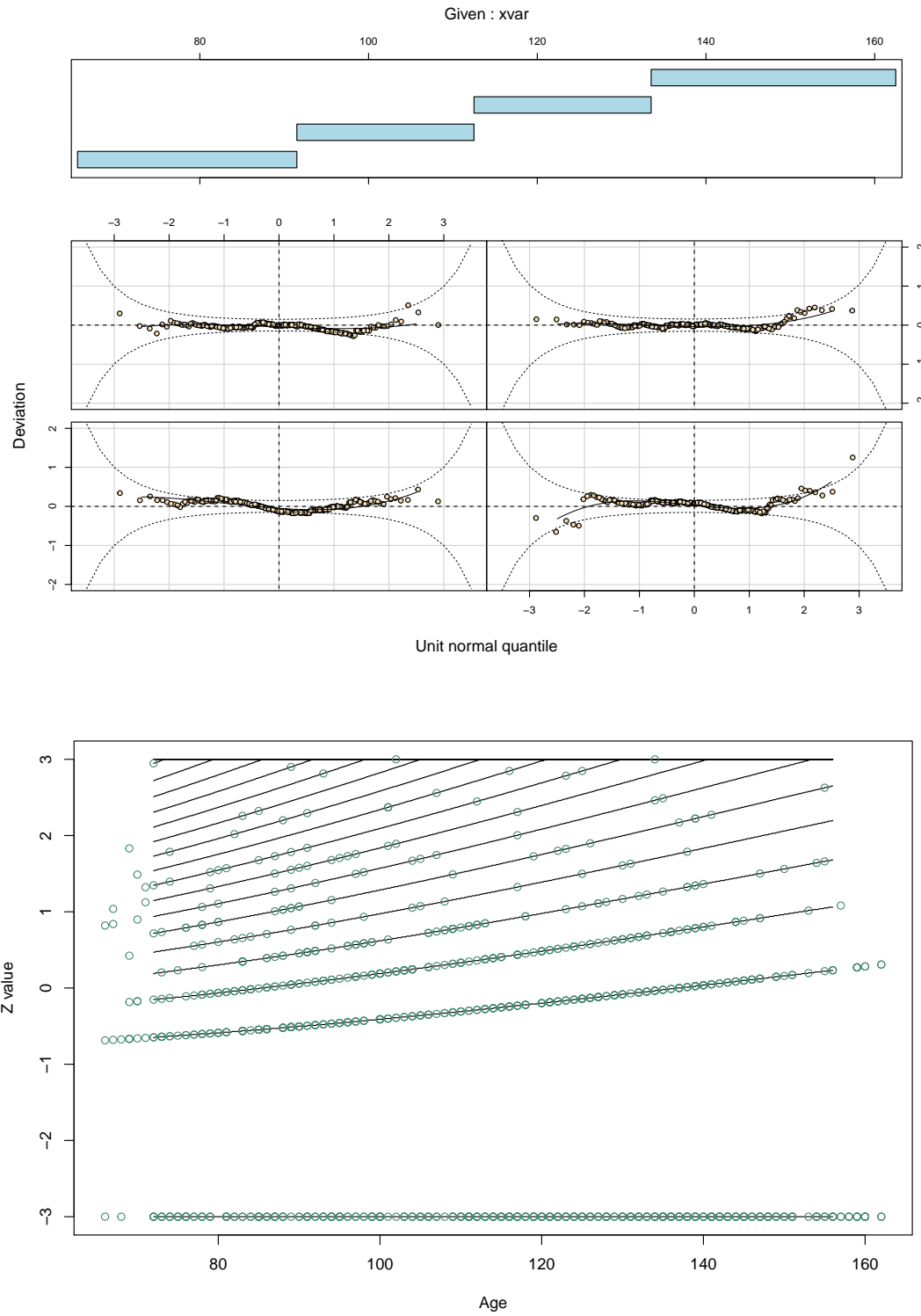
Wachttijdafweging



(WERK)GEHEUGEN
Reactiesnelheid eenvoudig

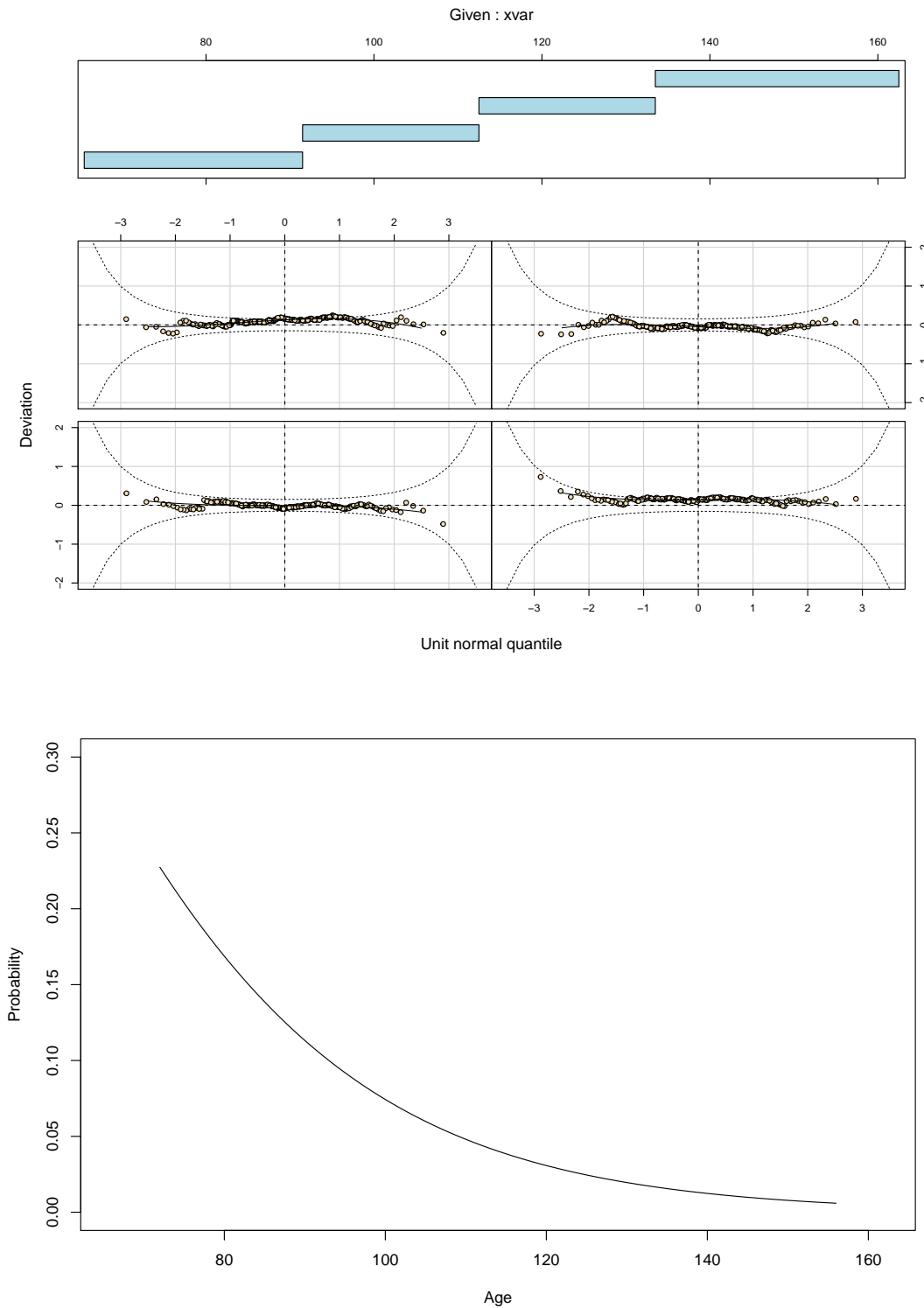


Fouten eenvoudig



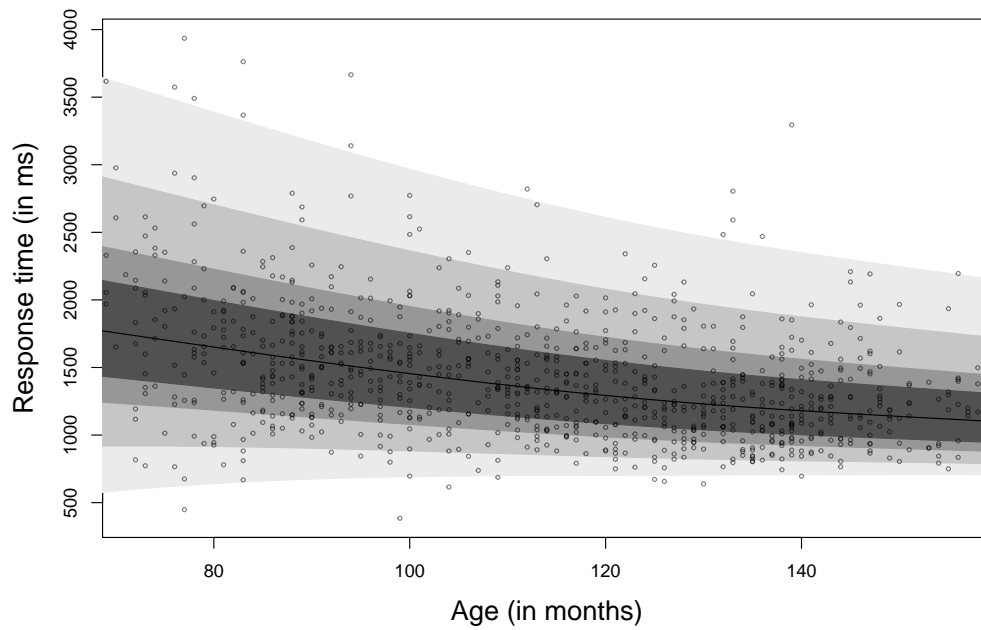
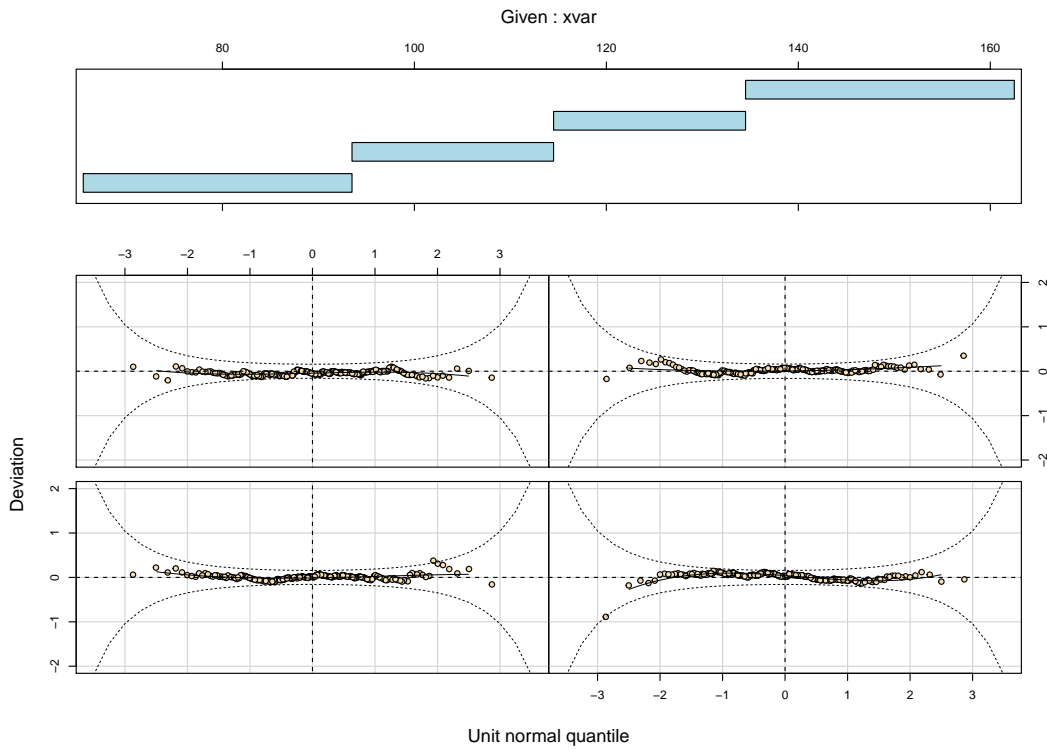
Reactiesnelheid complex

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.



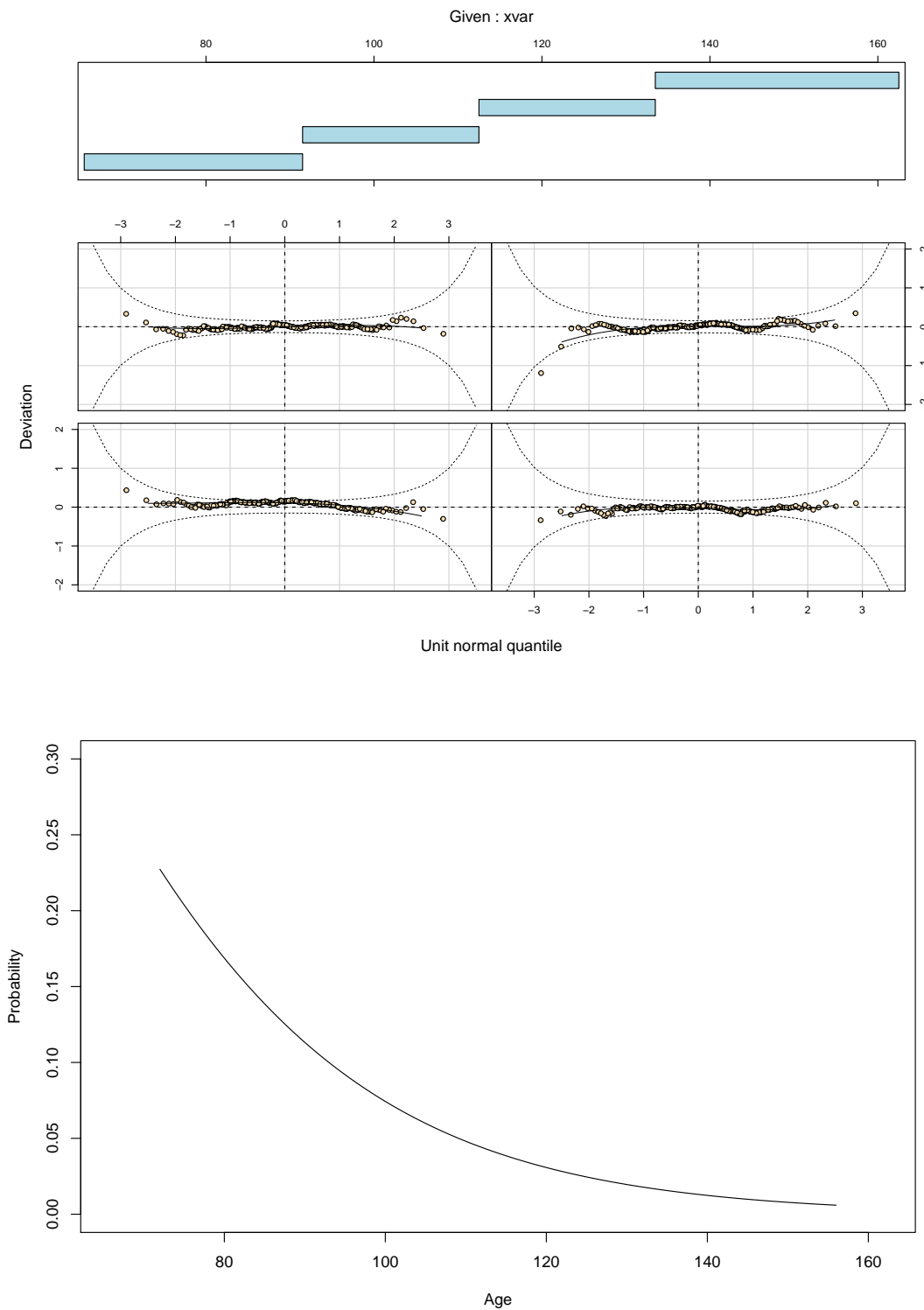
Reactiesnelheid complex (vervolg)

Deze wormplot en centielcurveplot behoren bij het BCPE model voor de Respons tijd.

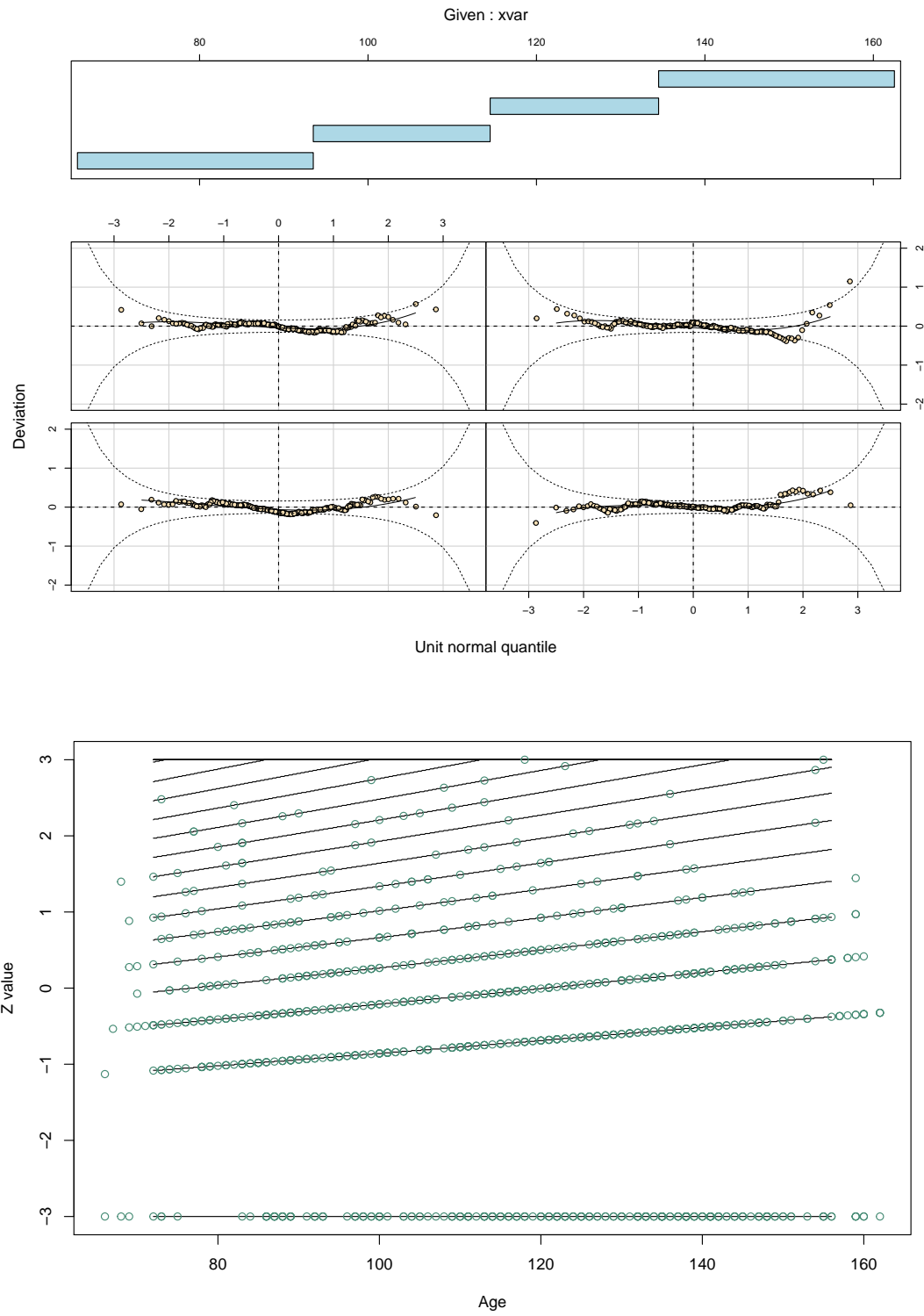


Fouten complex

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.

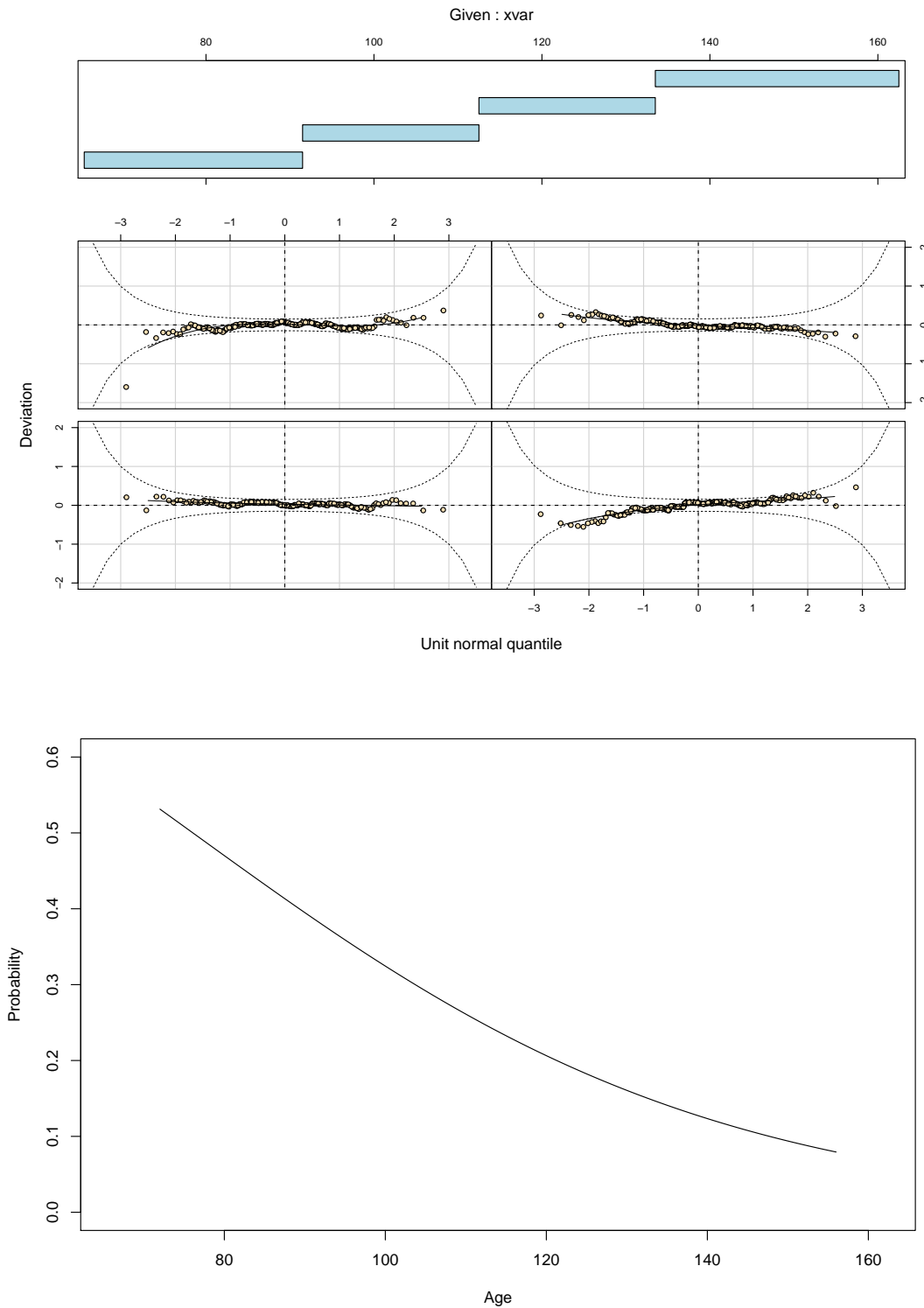


Fouten complex (vervolg)



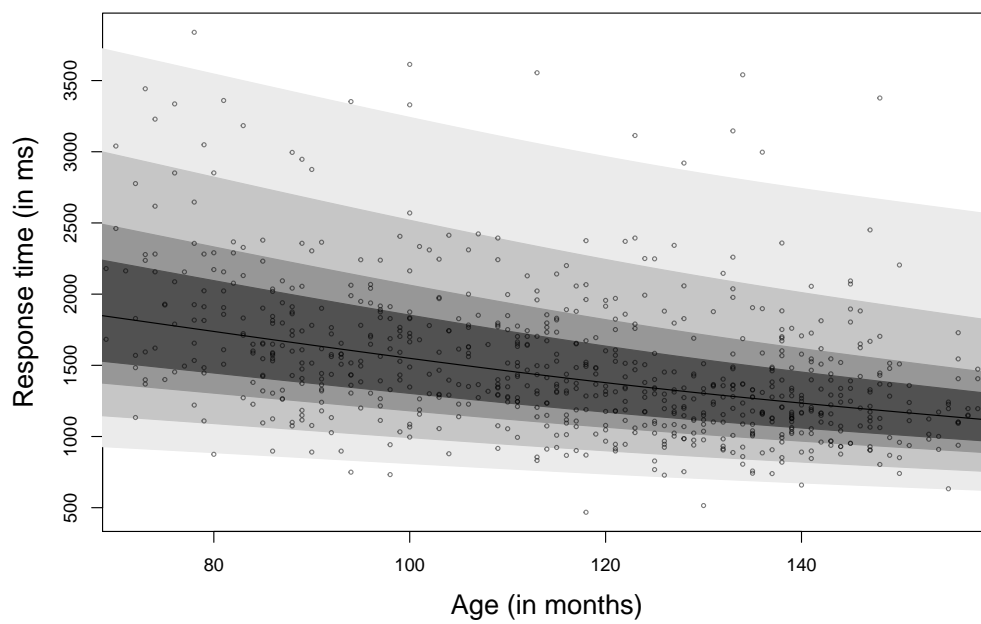
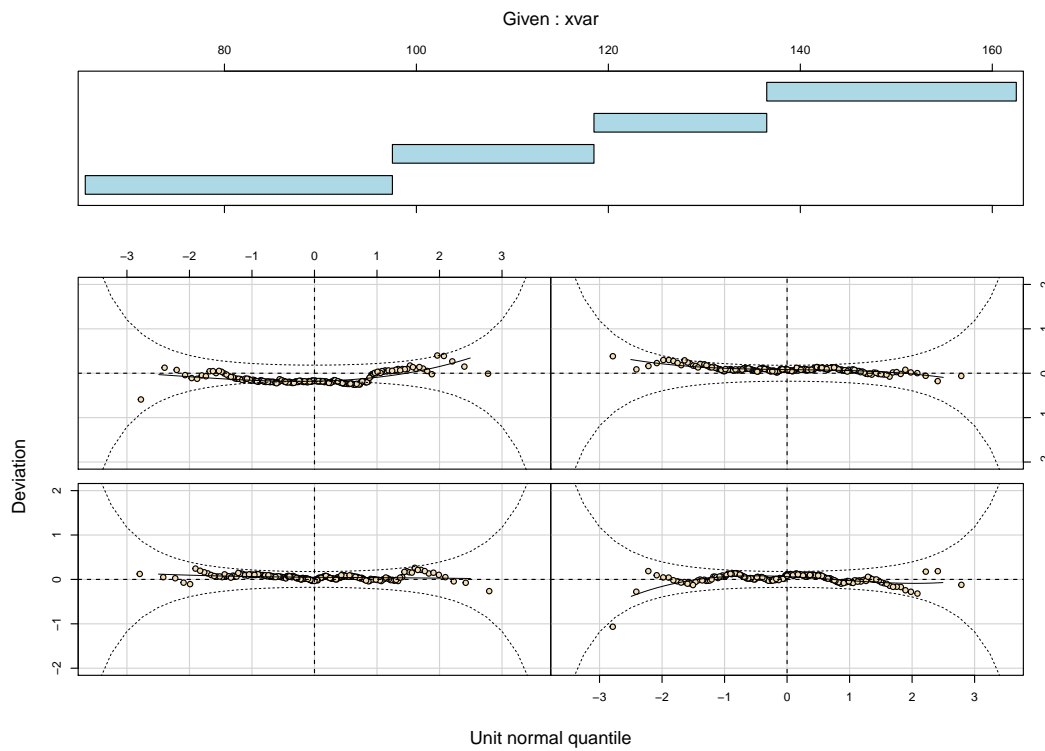
Reactiesnelheid zeer complex

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.



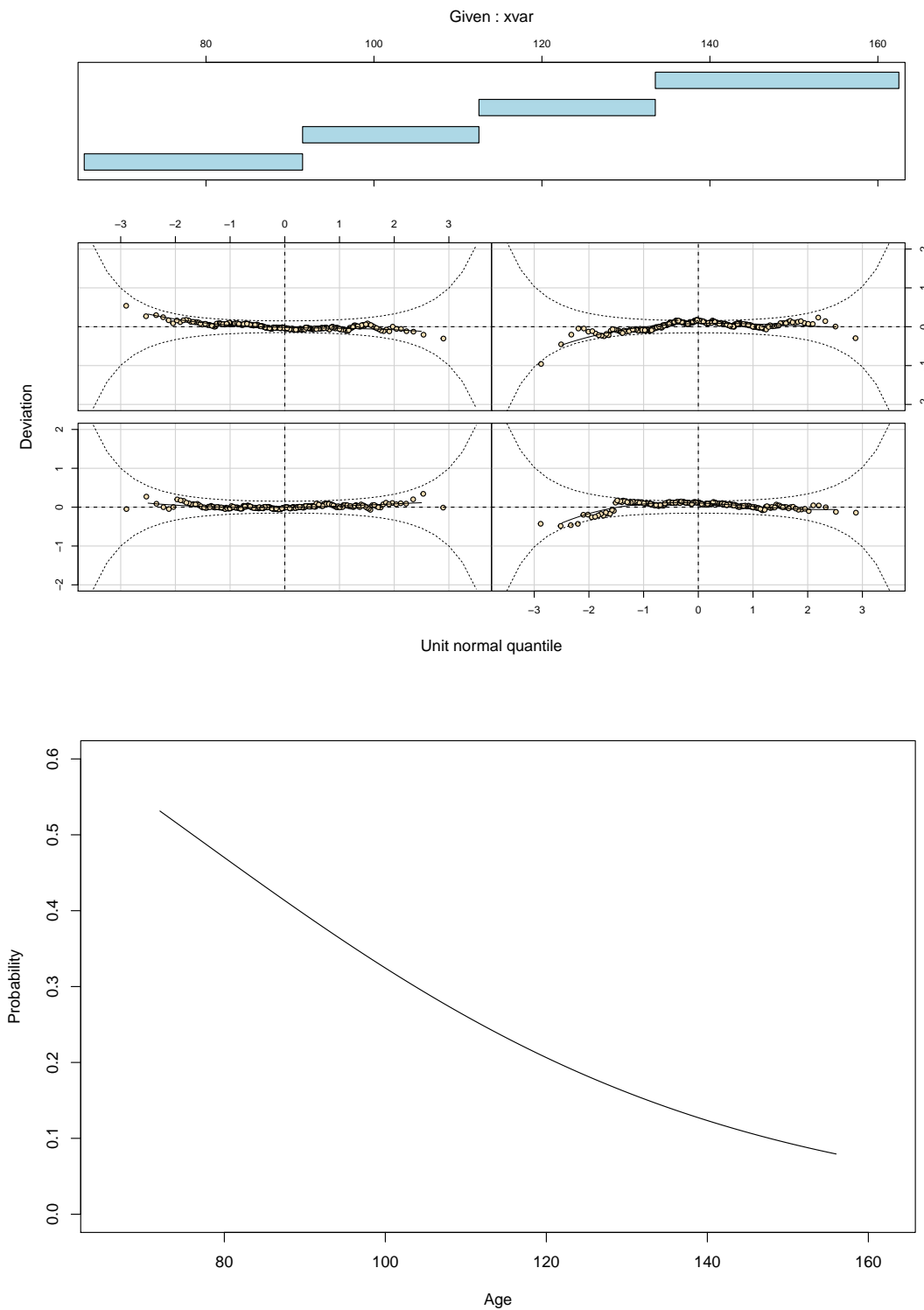
Reactiesnelheid zeer complex (vervolg)

Deze wormplot en centielcurveplot behoren bij het BCPE model voor de Respons tijd.

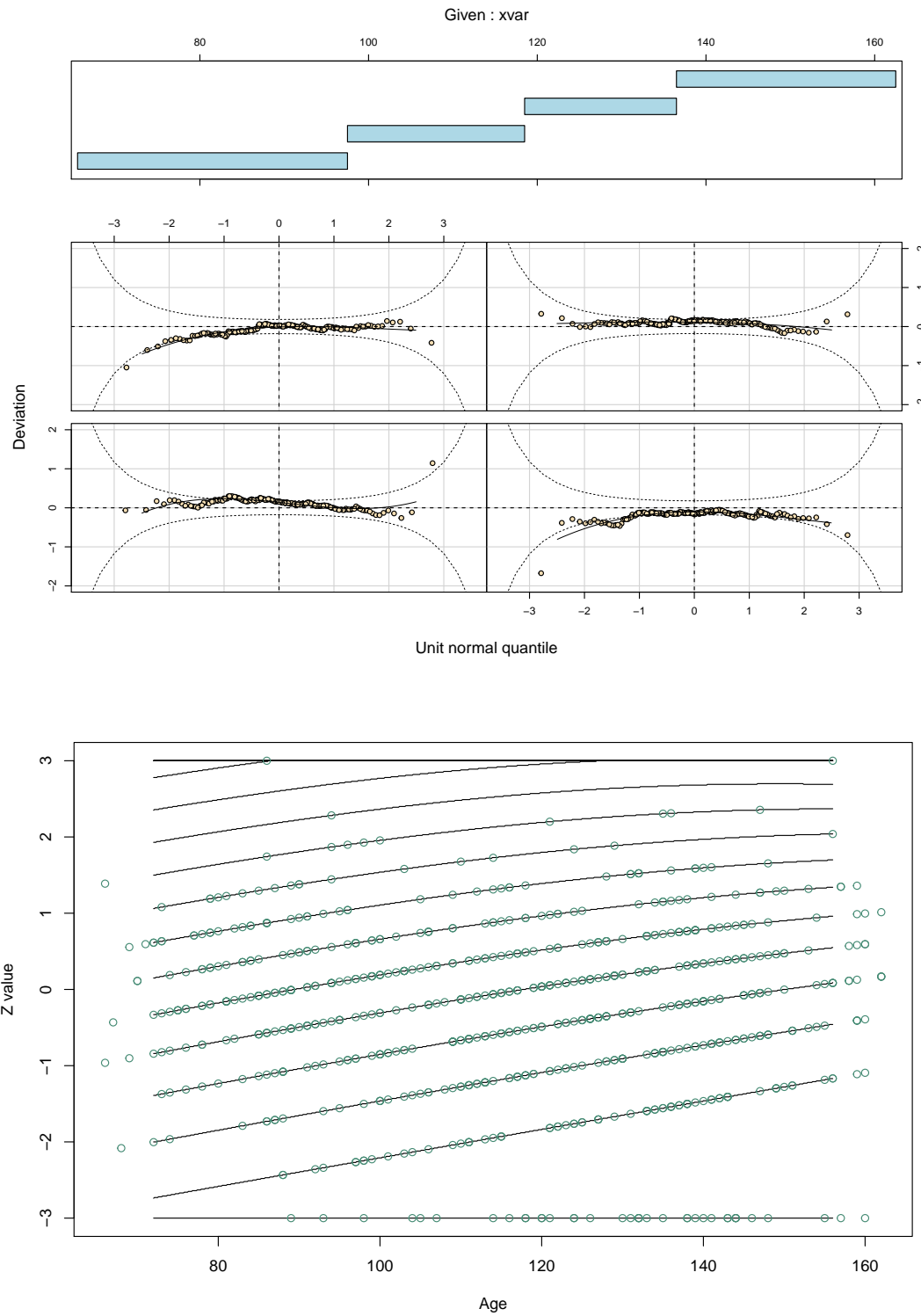


Fouten zeer complex

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.

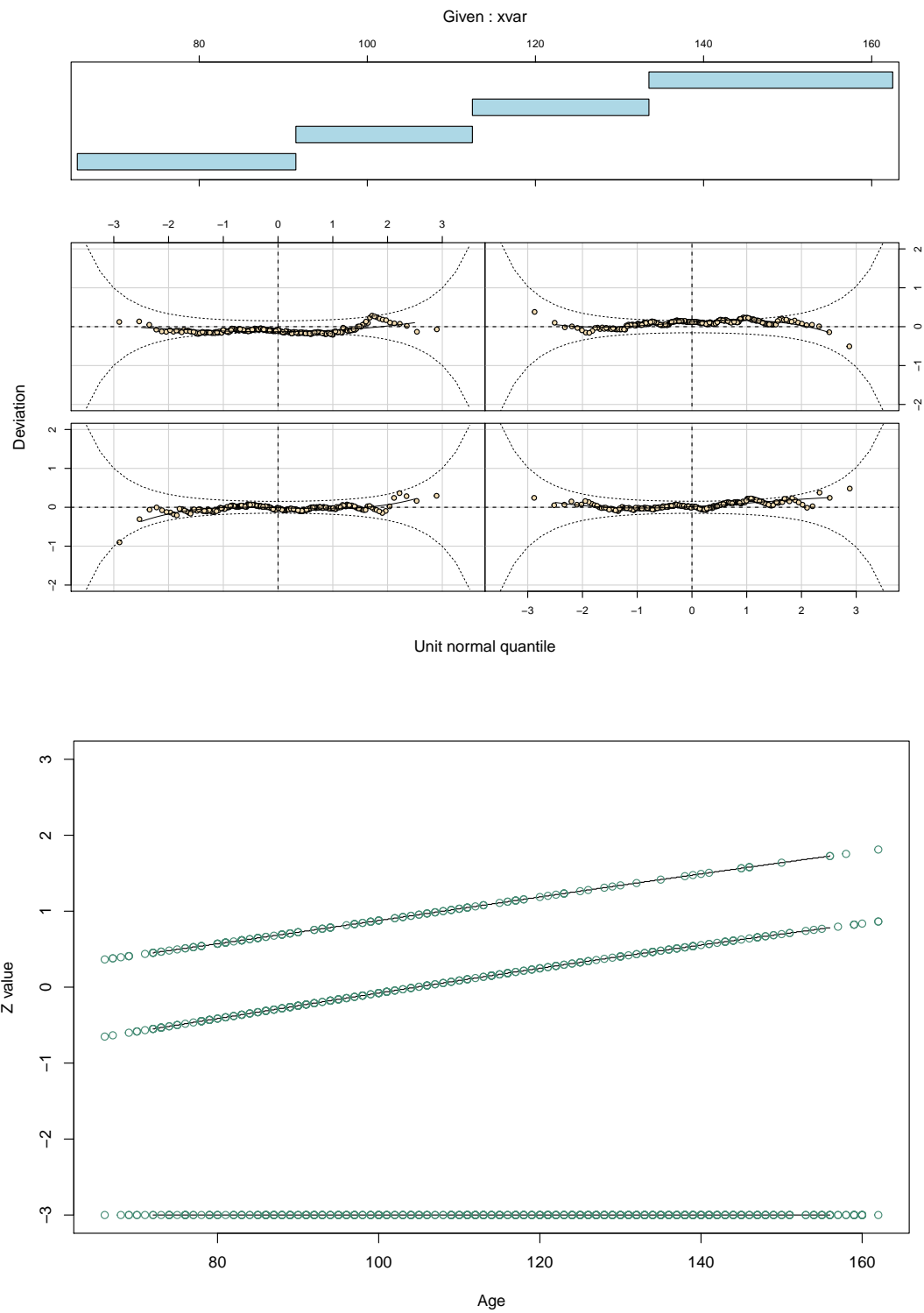


Fouten zeer complex (vervolg)



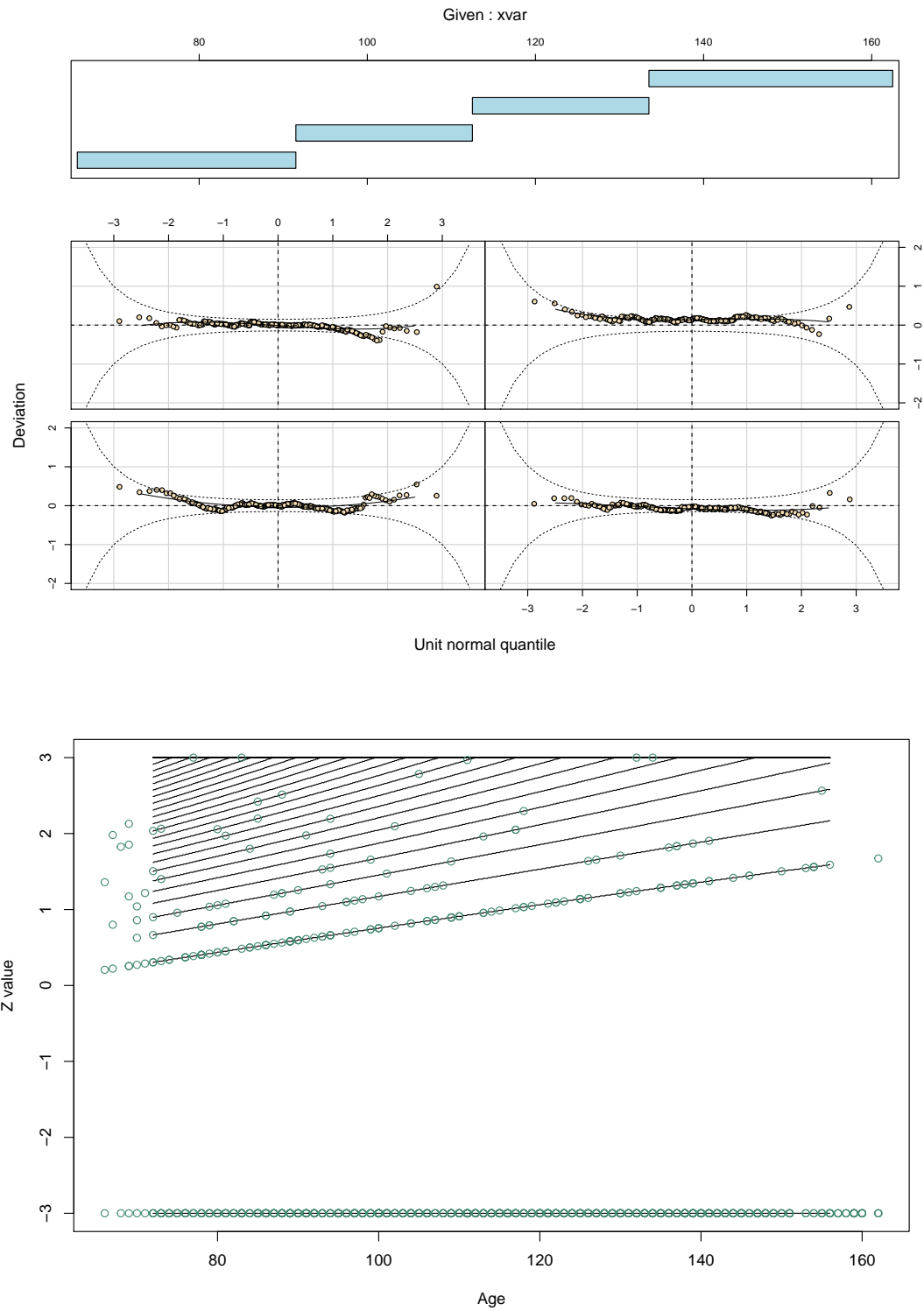
LEERSNELHEID

Leersnelheid

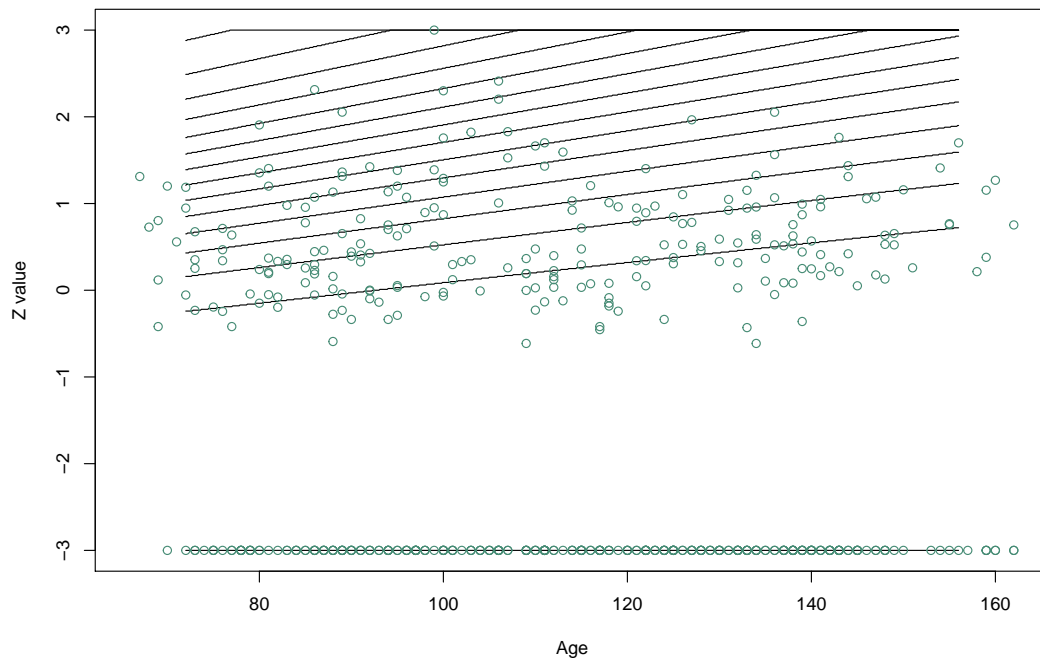
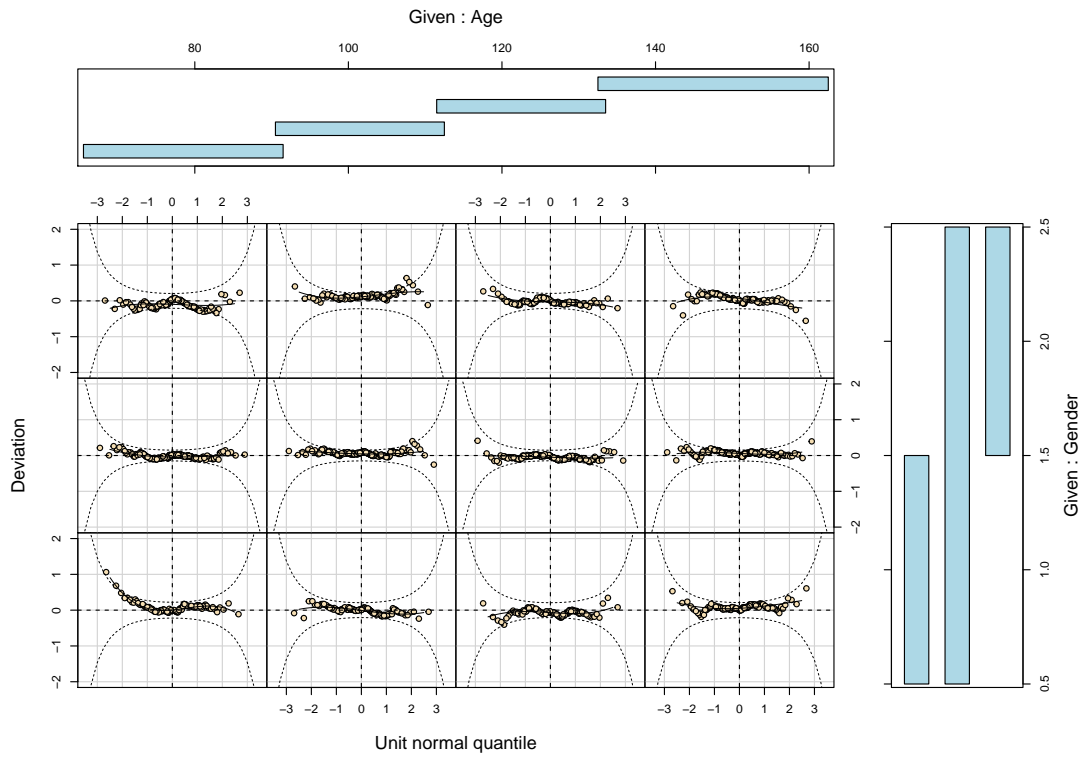


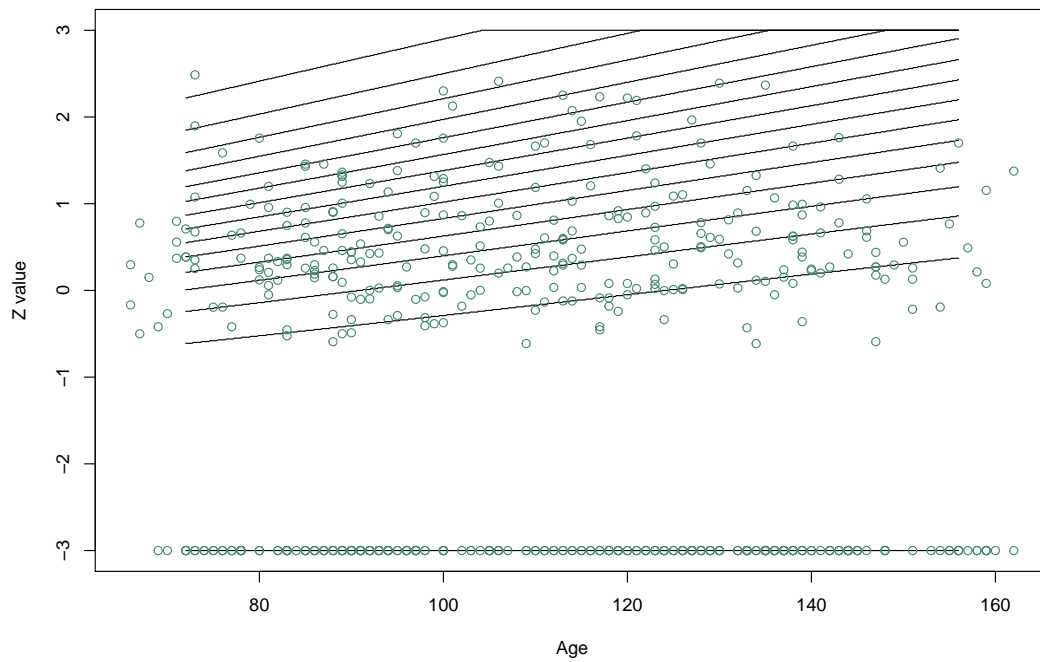
GEDRAG TIJDENS DE TAAK

Benodigde ondersteuning

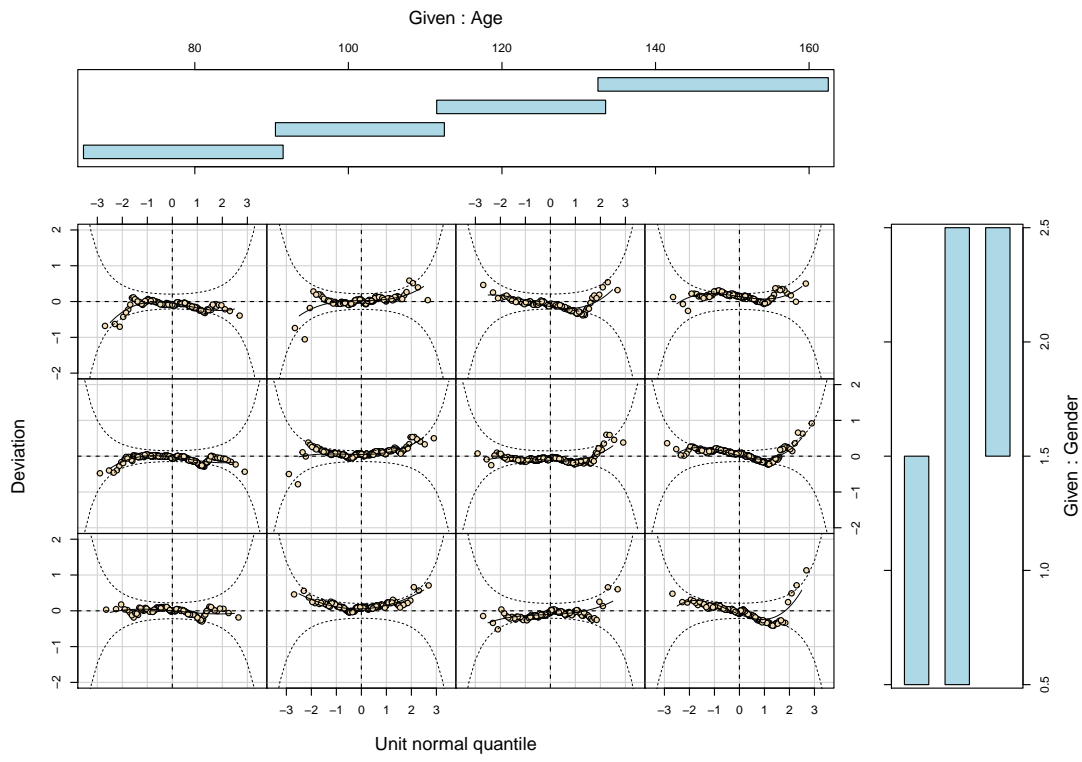


Verbaliseren

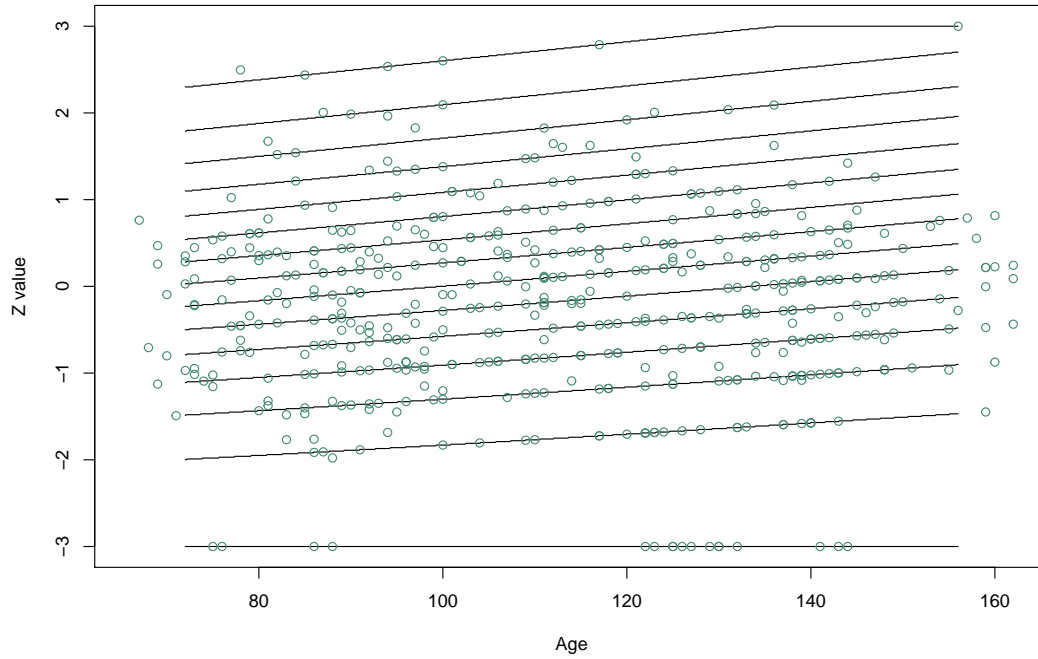




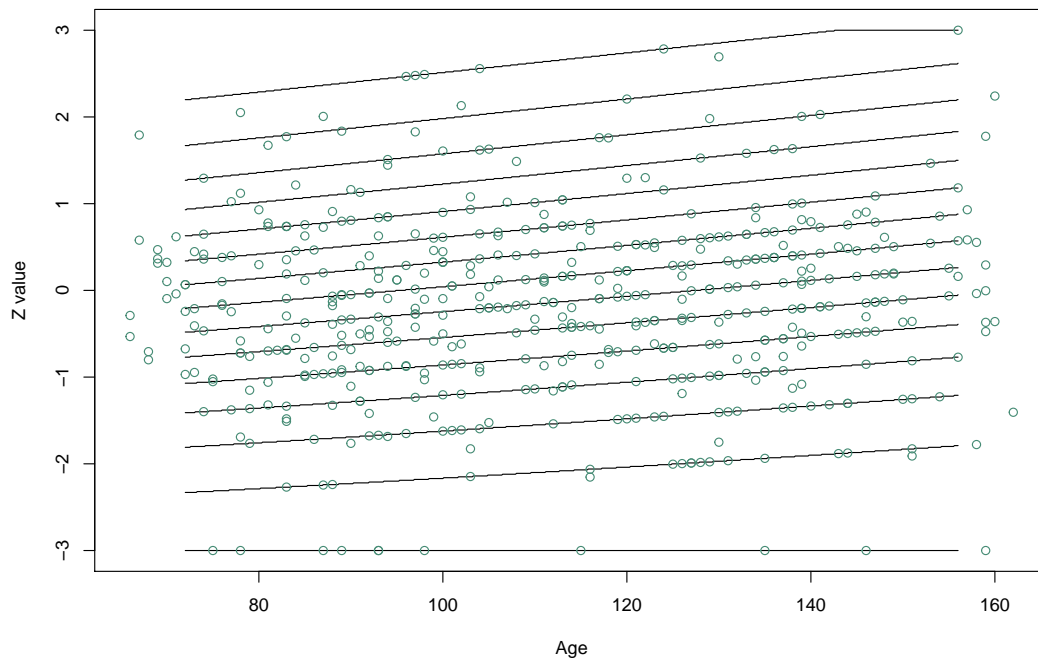
Beweeglijkheid



z-curveplot van Bewegelijkheid voor de meisjes



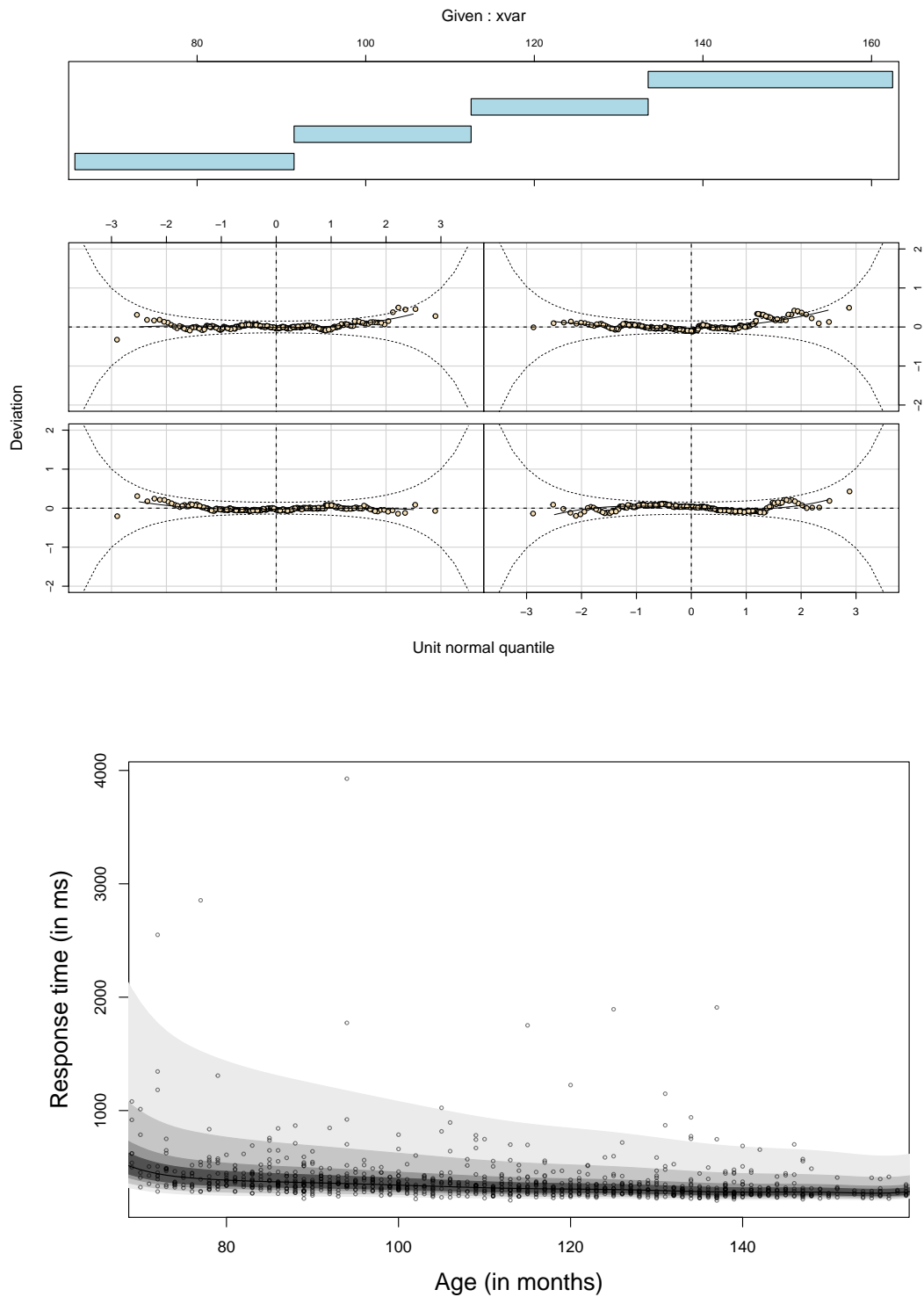
z-curveplot van Bewegelijkheid voor de jongens



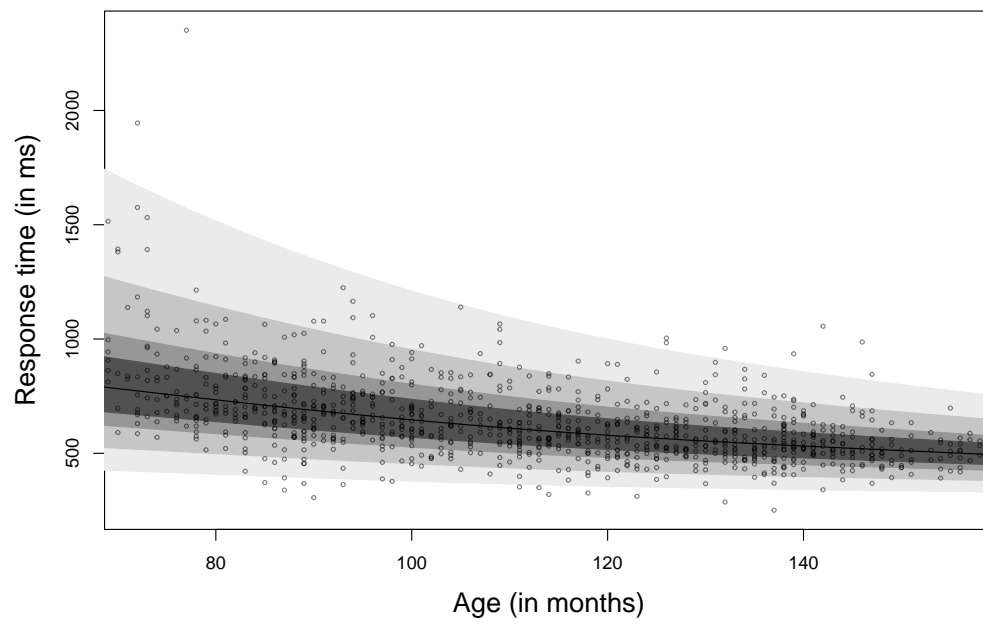
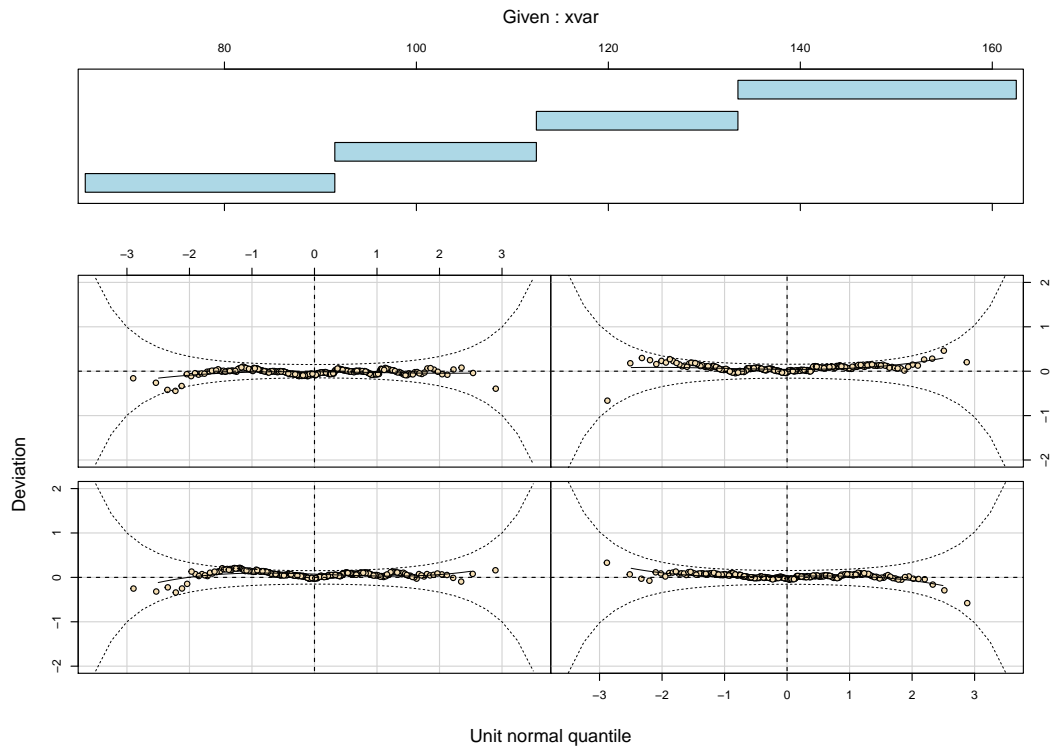
1.2.2 Specialistische rapportage

MEDIANE REACTIESNELHEID

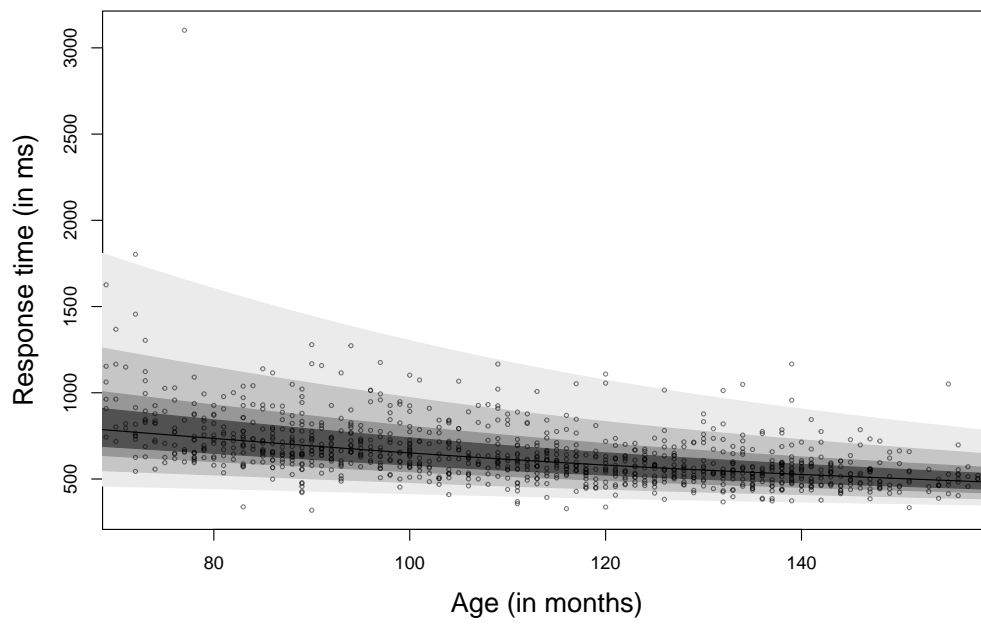
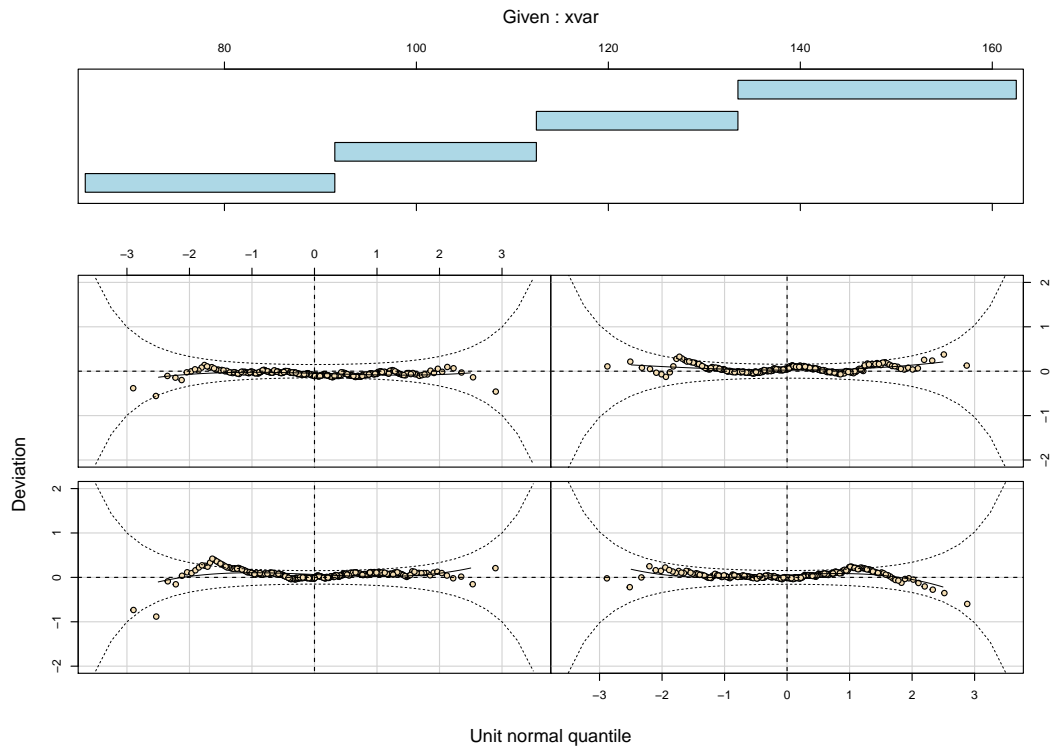
Blok 1



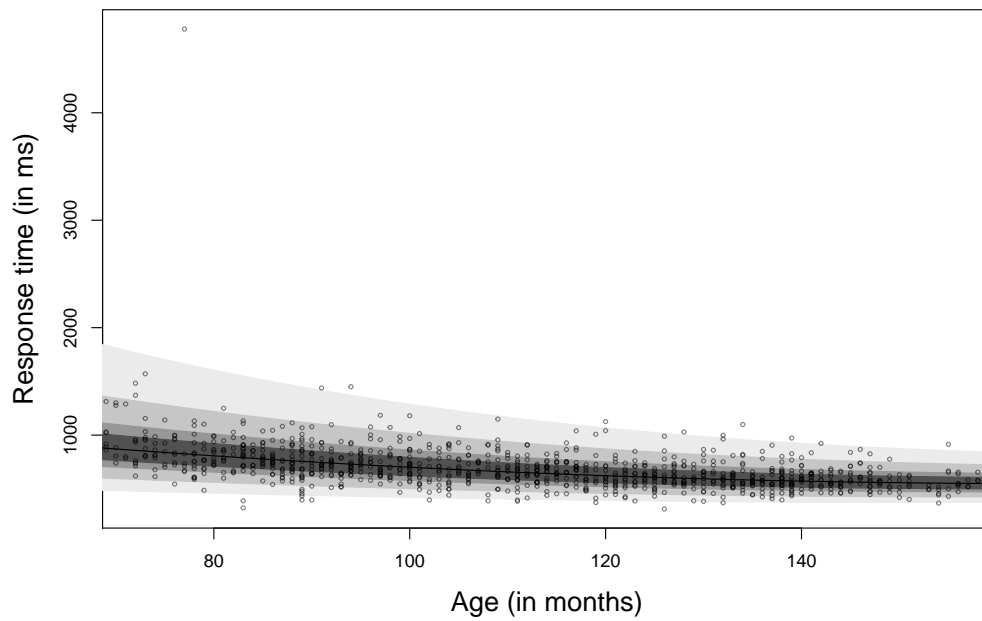
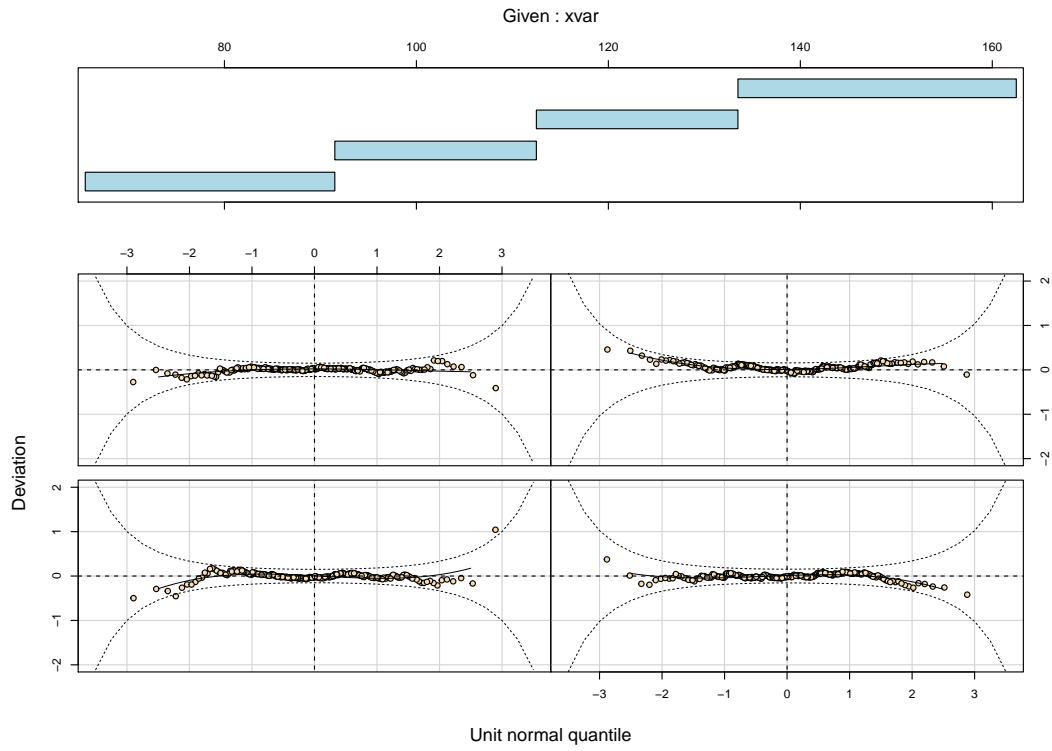
Blok 2



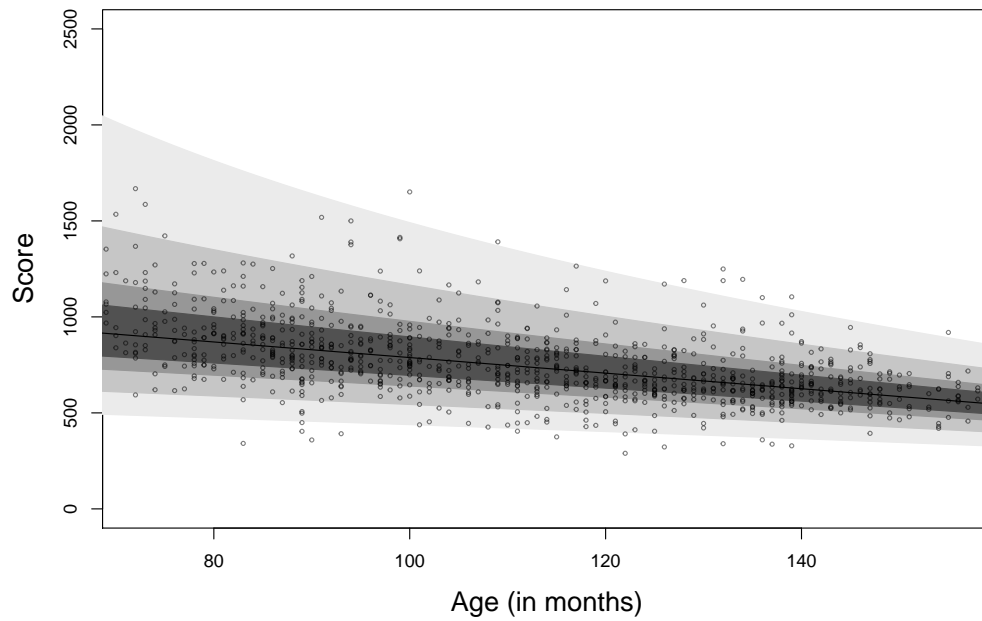
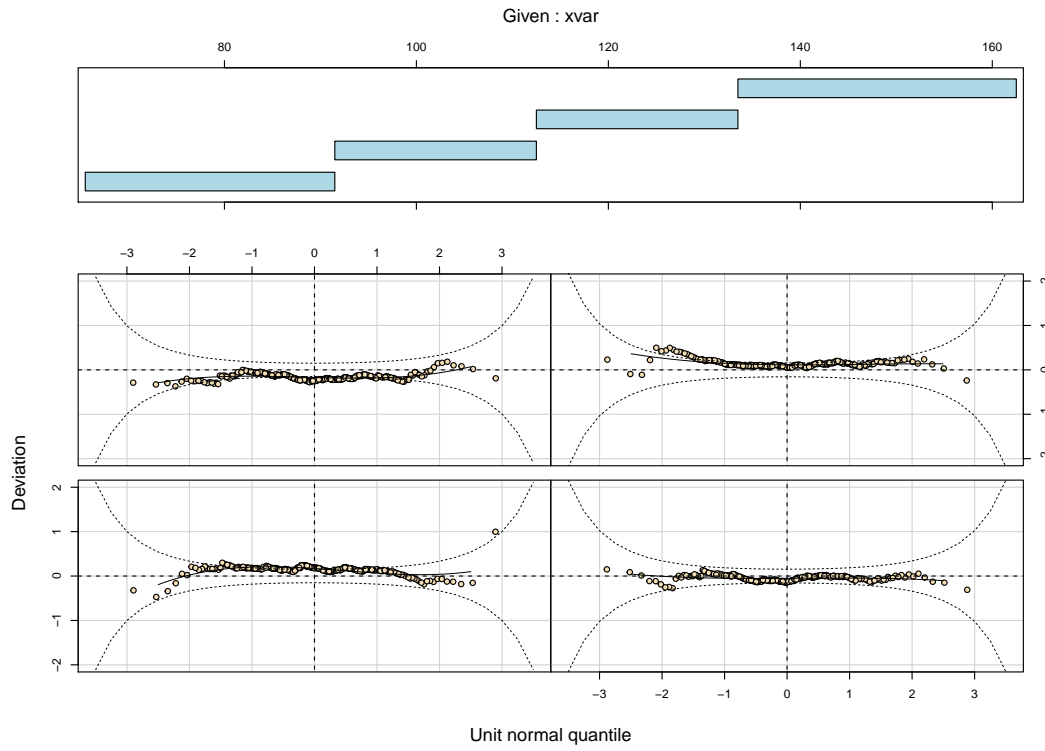
Blok 3



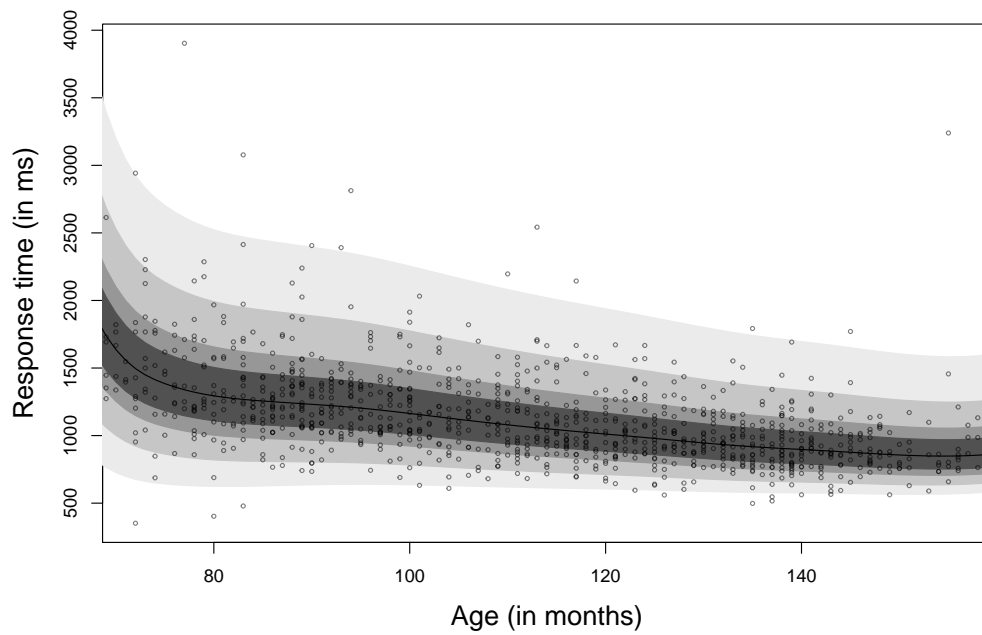
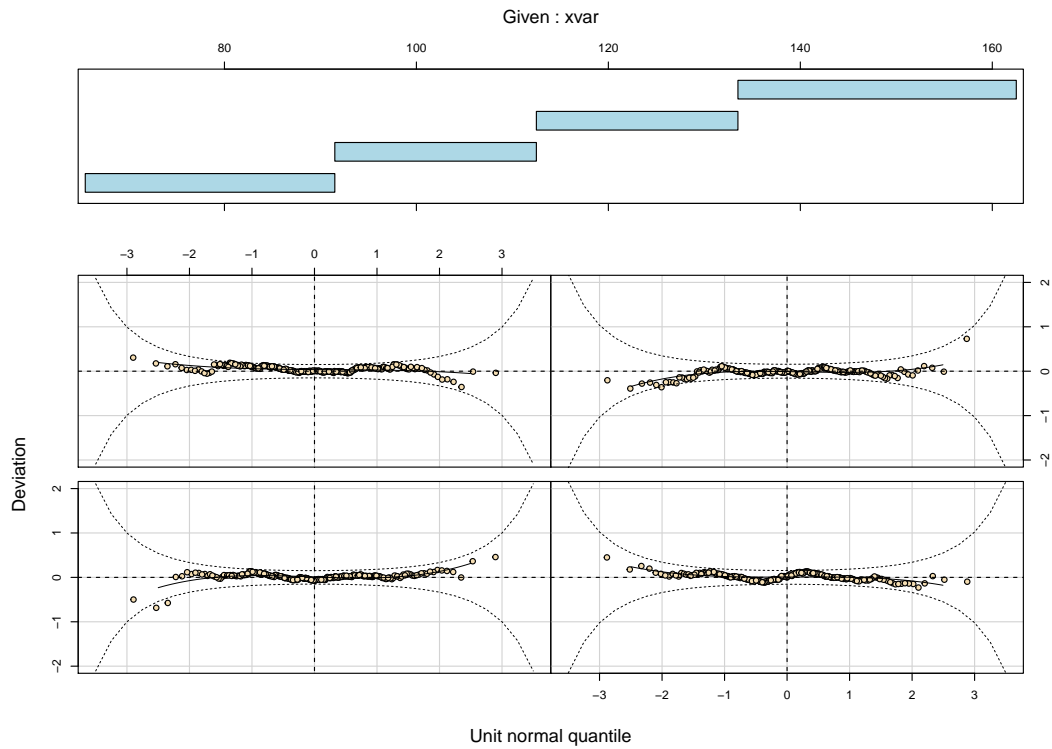
Blok 4 compatibele trials



Blok 4 incompatibele trials

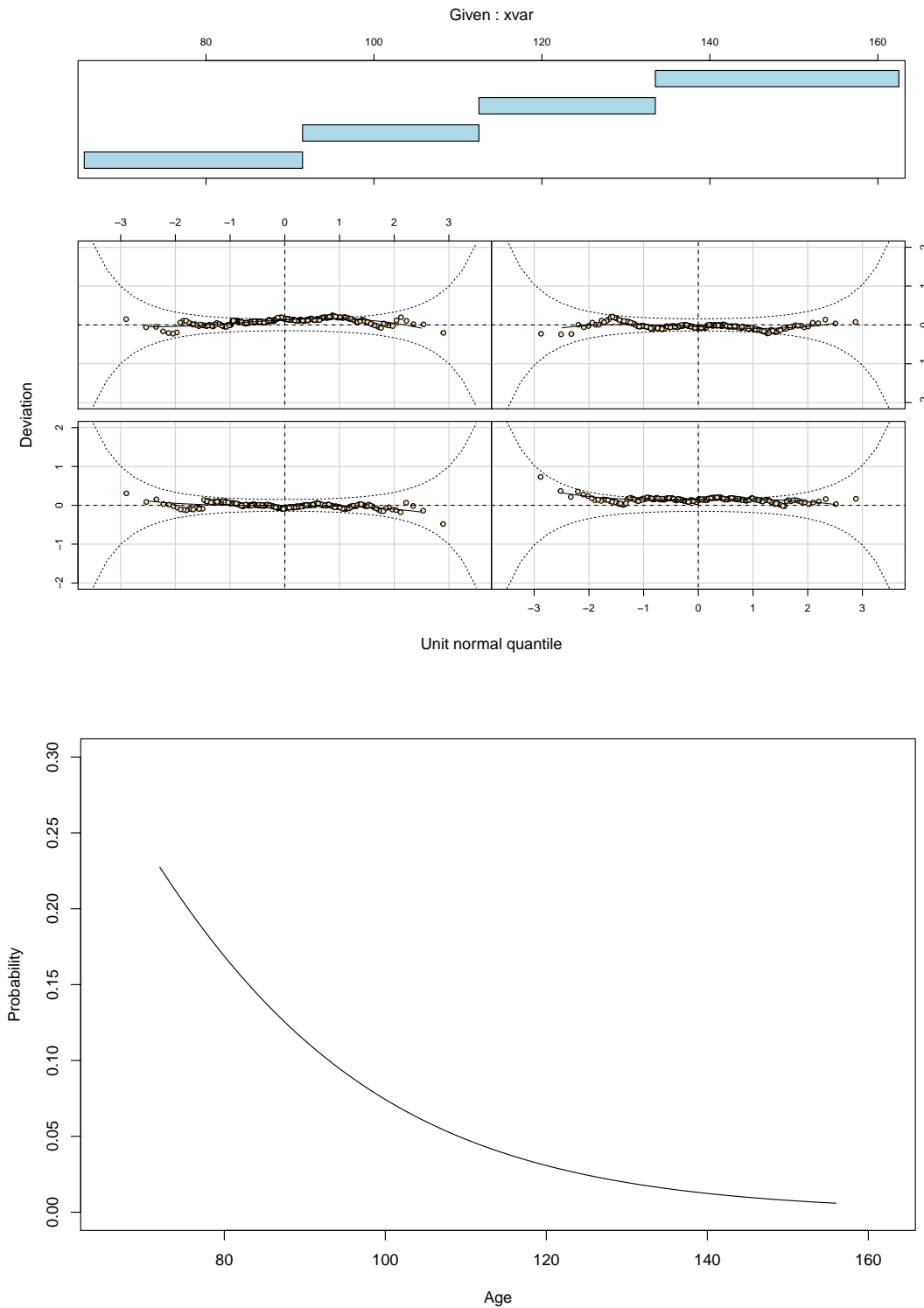


Blok 5 1-back



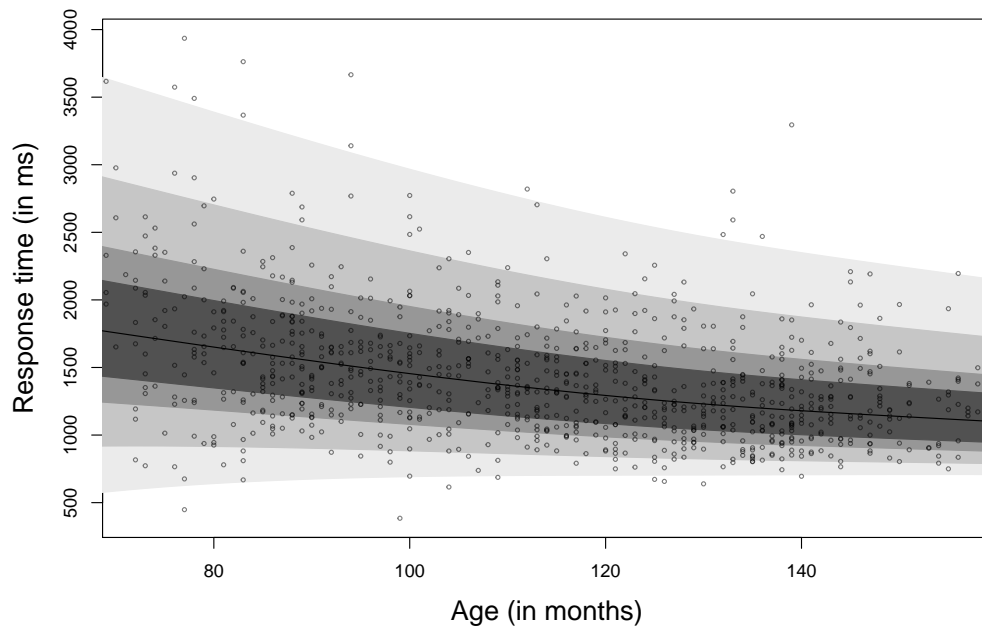
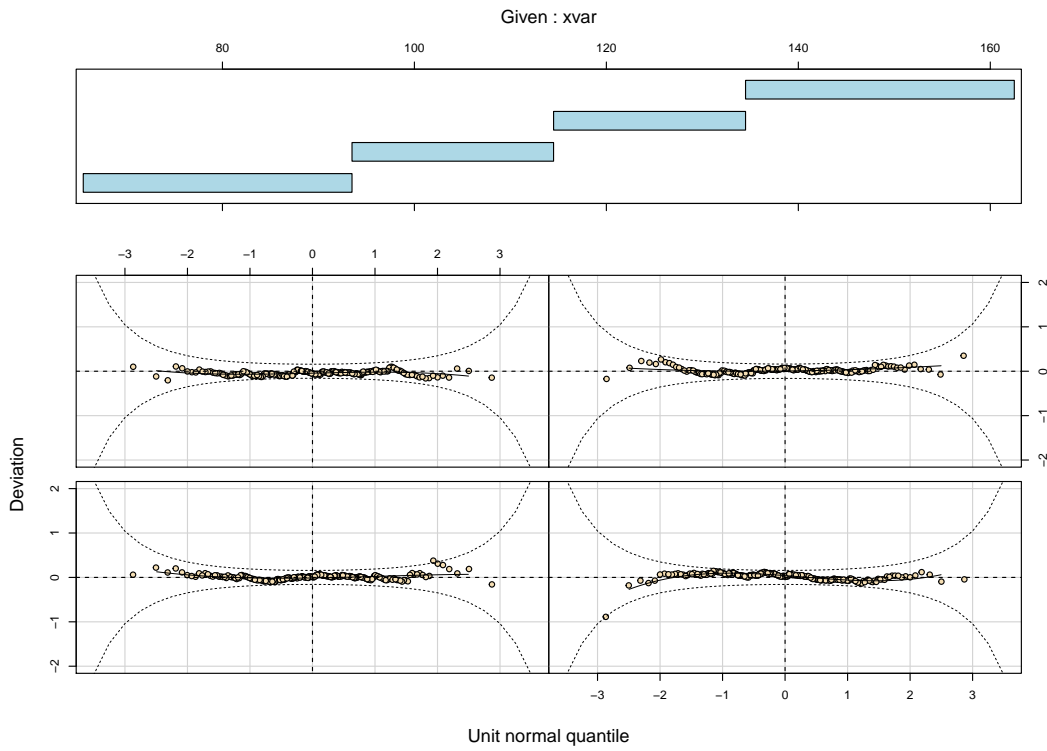
Blok 5 2-back

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.



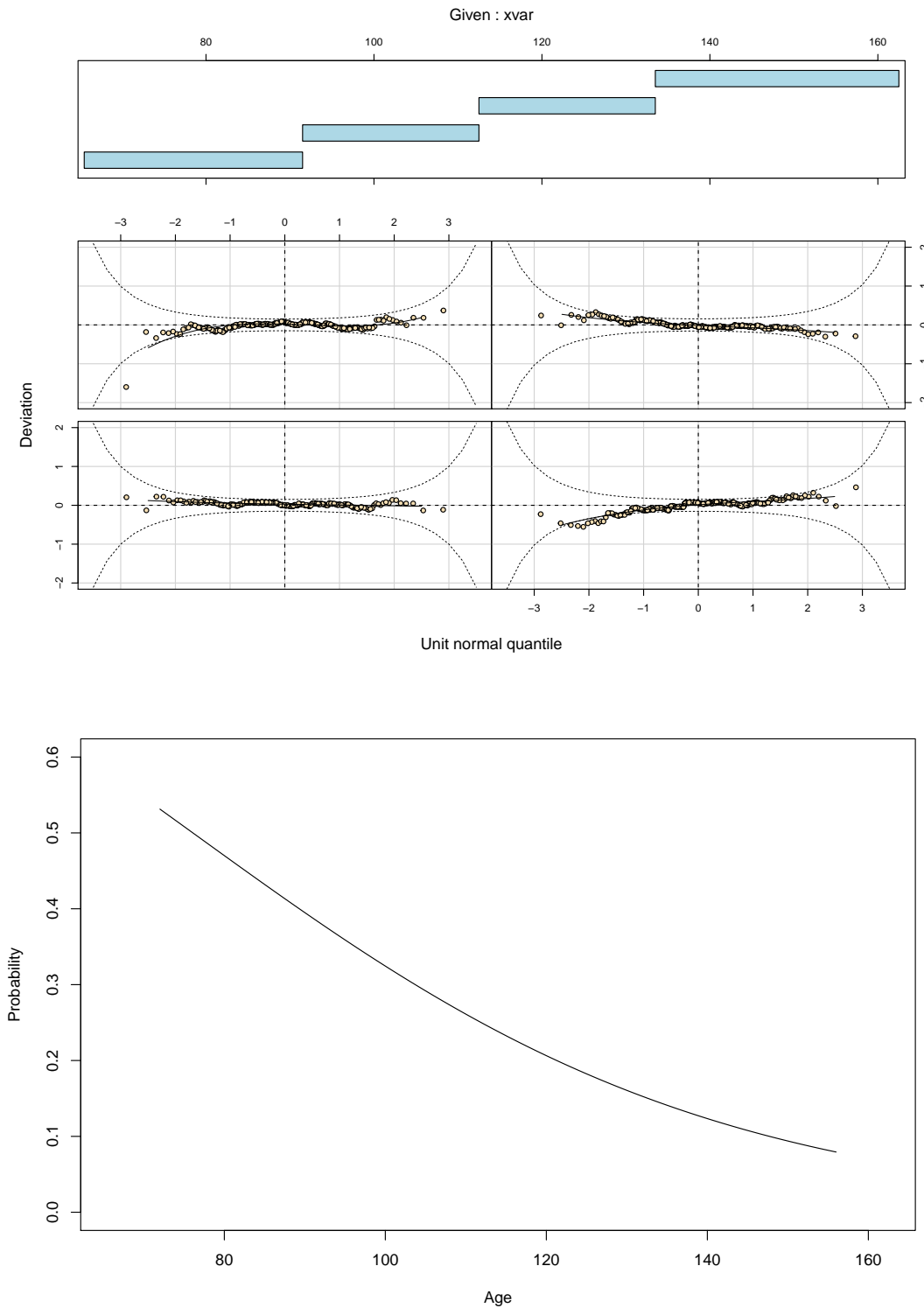
Blok 5 2-back (vervolg)

Deze wormplot en centielcurveplot behoren bij het BCPE model voor de Respons tijd.



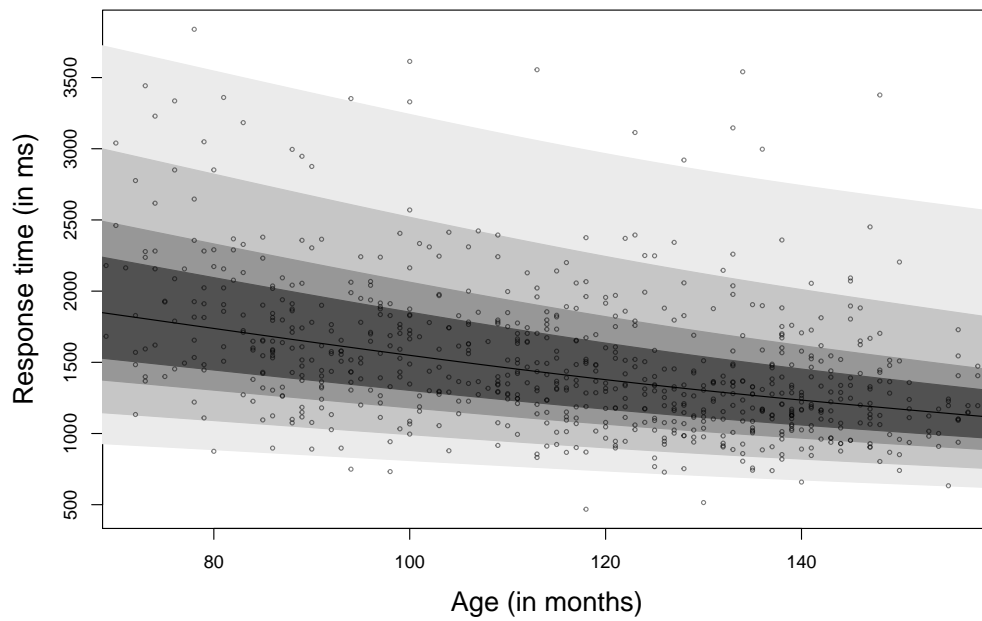
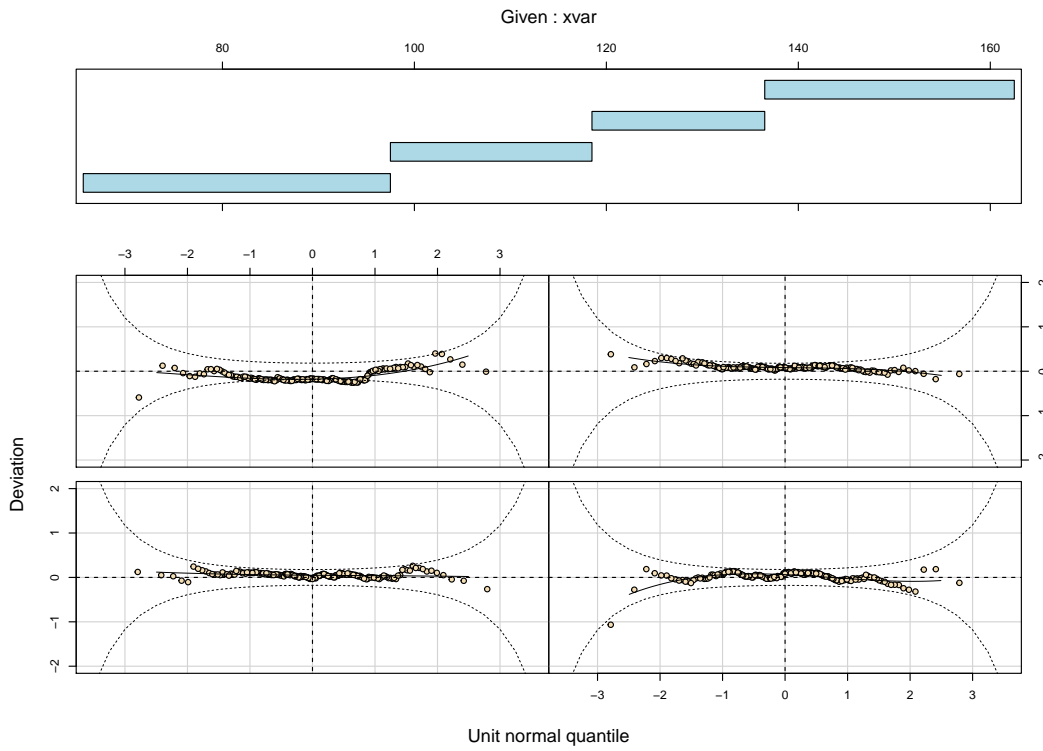
Blok 5 2-back PI

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.

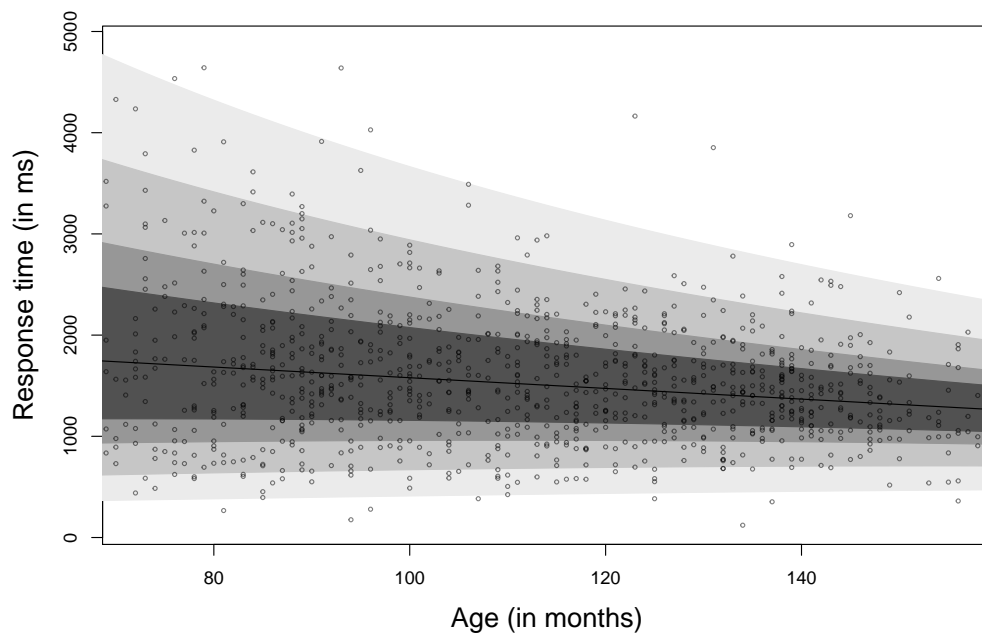
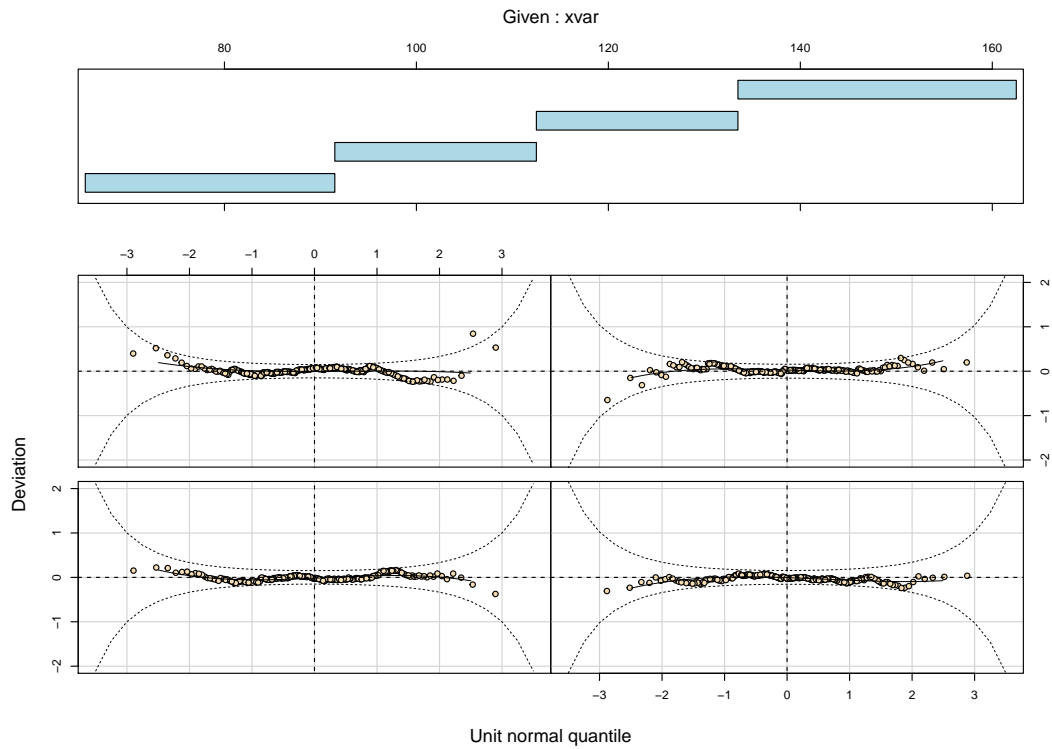


Blok 5 2-back PI (vervolg)

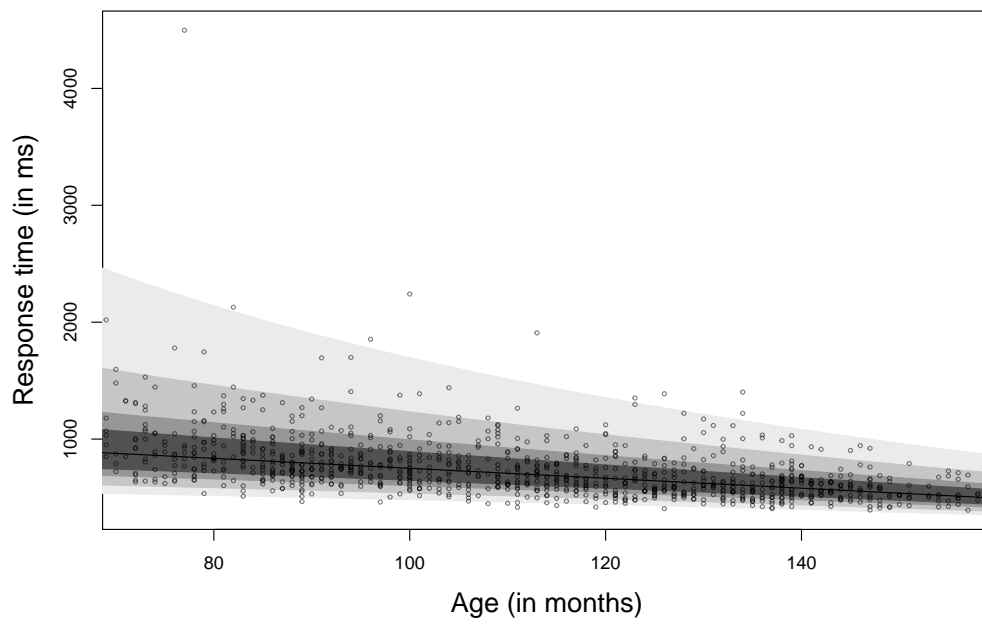
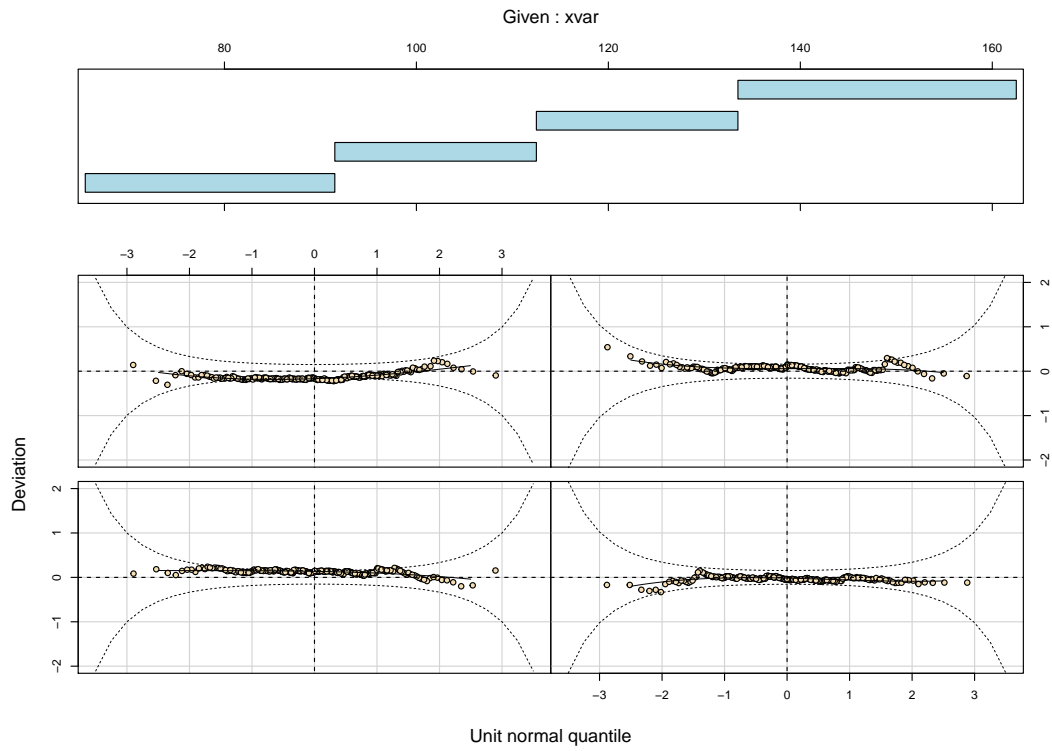
Deze wormplot en centielcurveplot behoren bij het BCPE model voor de Respons tijd.



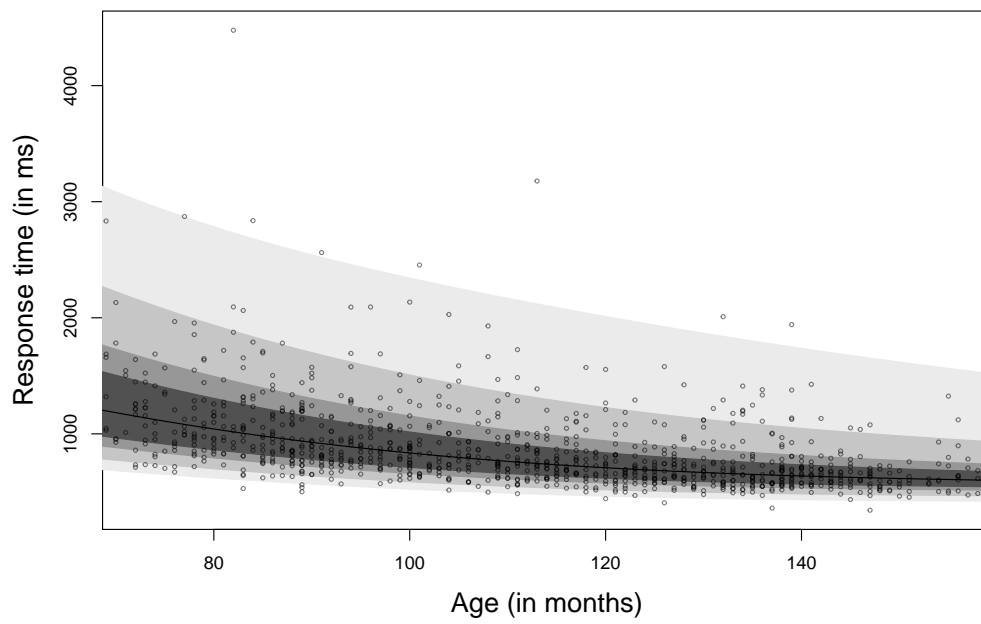
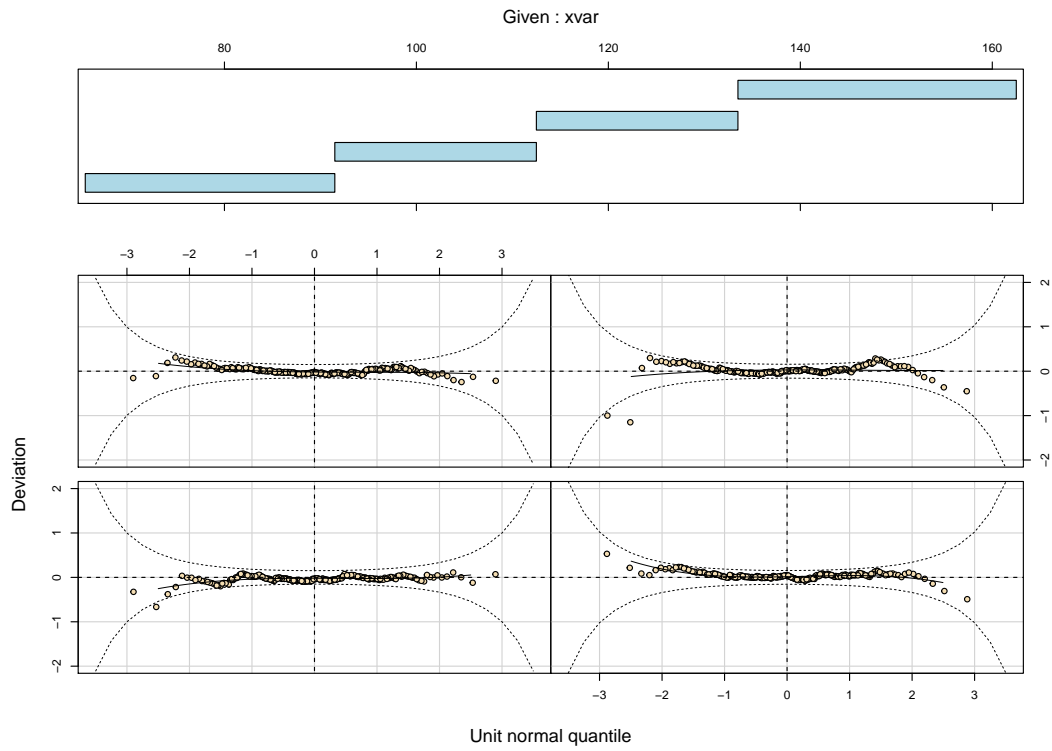
Blok 6



Blok 7 deel 1

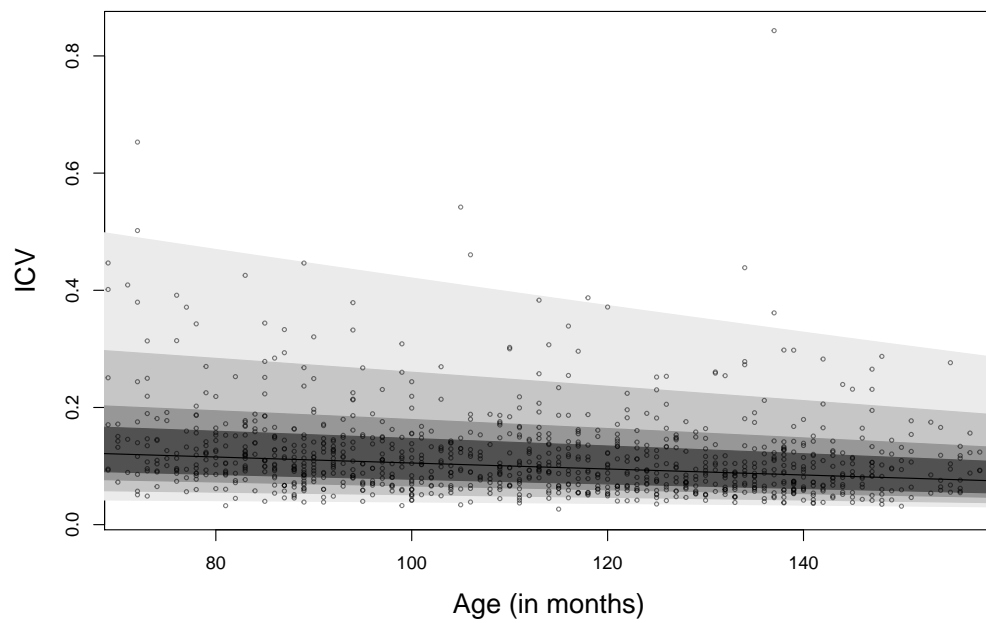
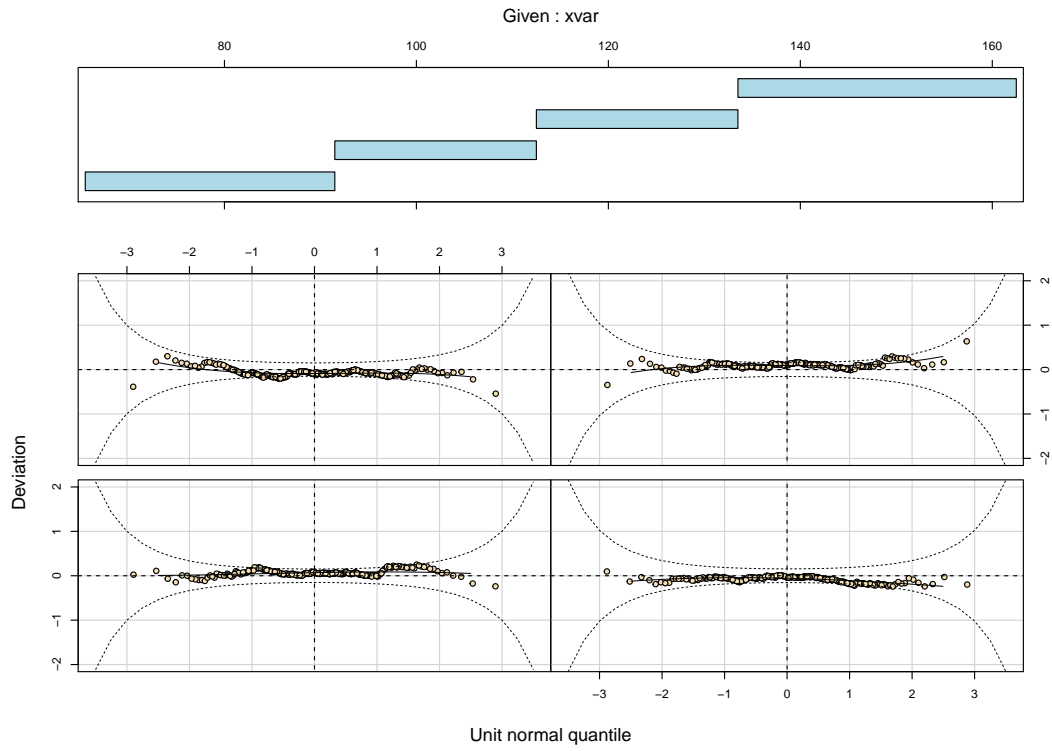


Blok 7 deel 2

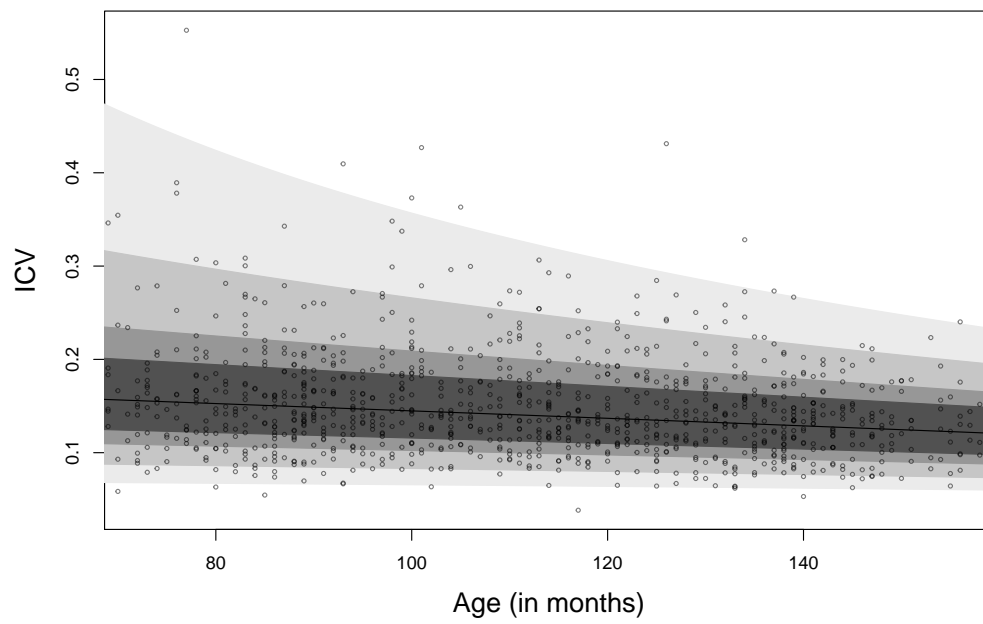
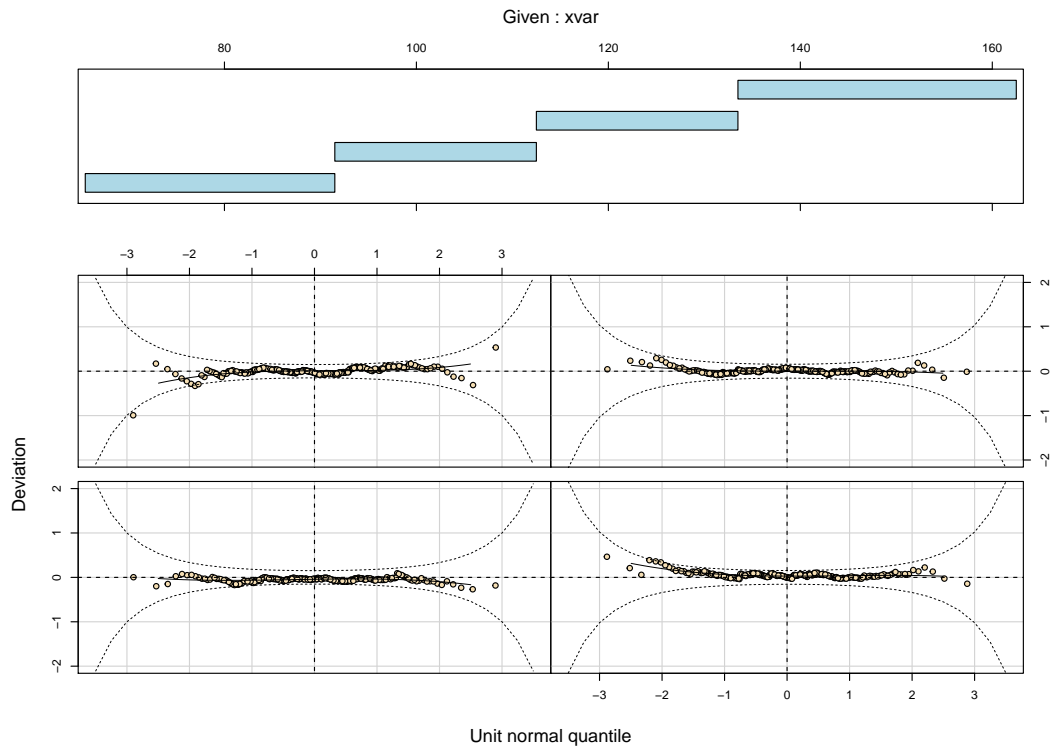


INTRA-INDIVIDUELE COËFFICIËNT VAN VARIATIE

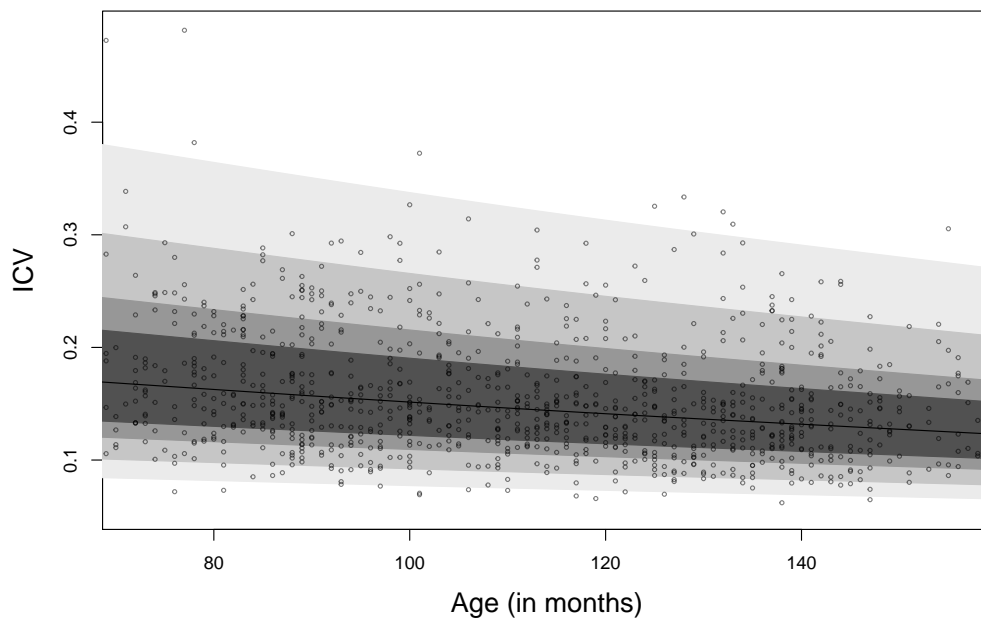
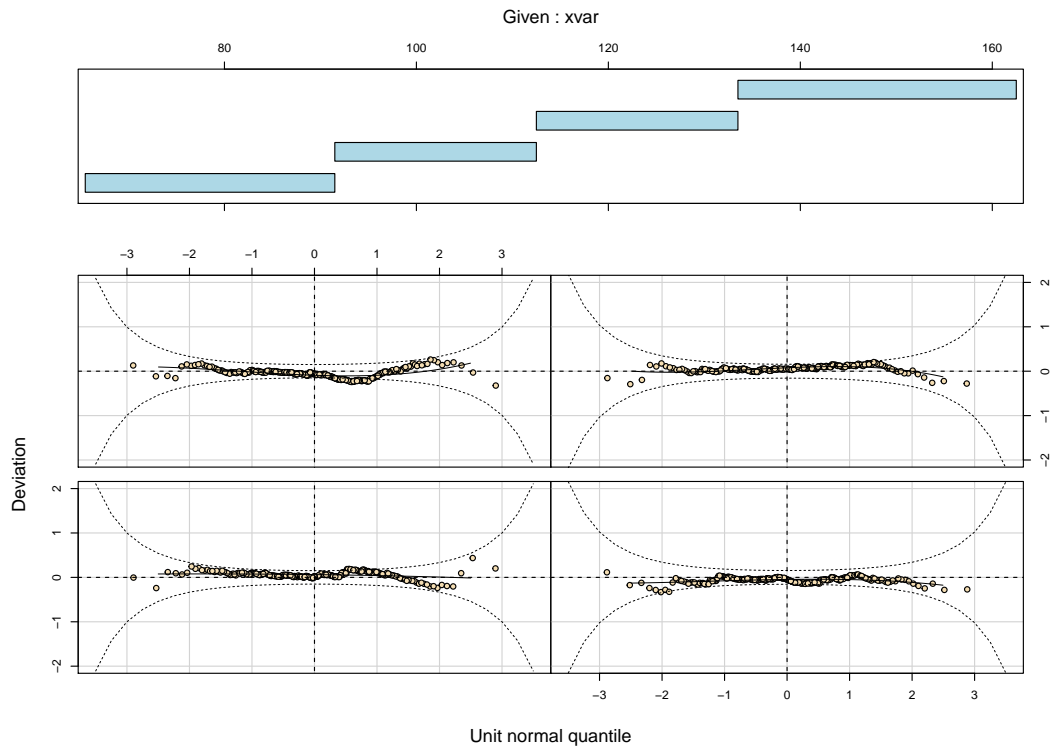
Blok 1



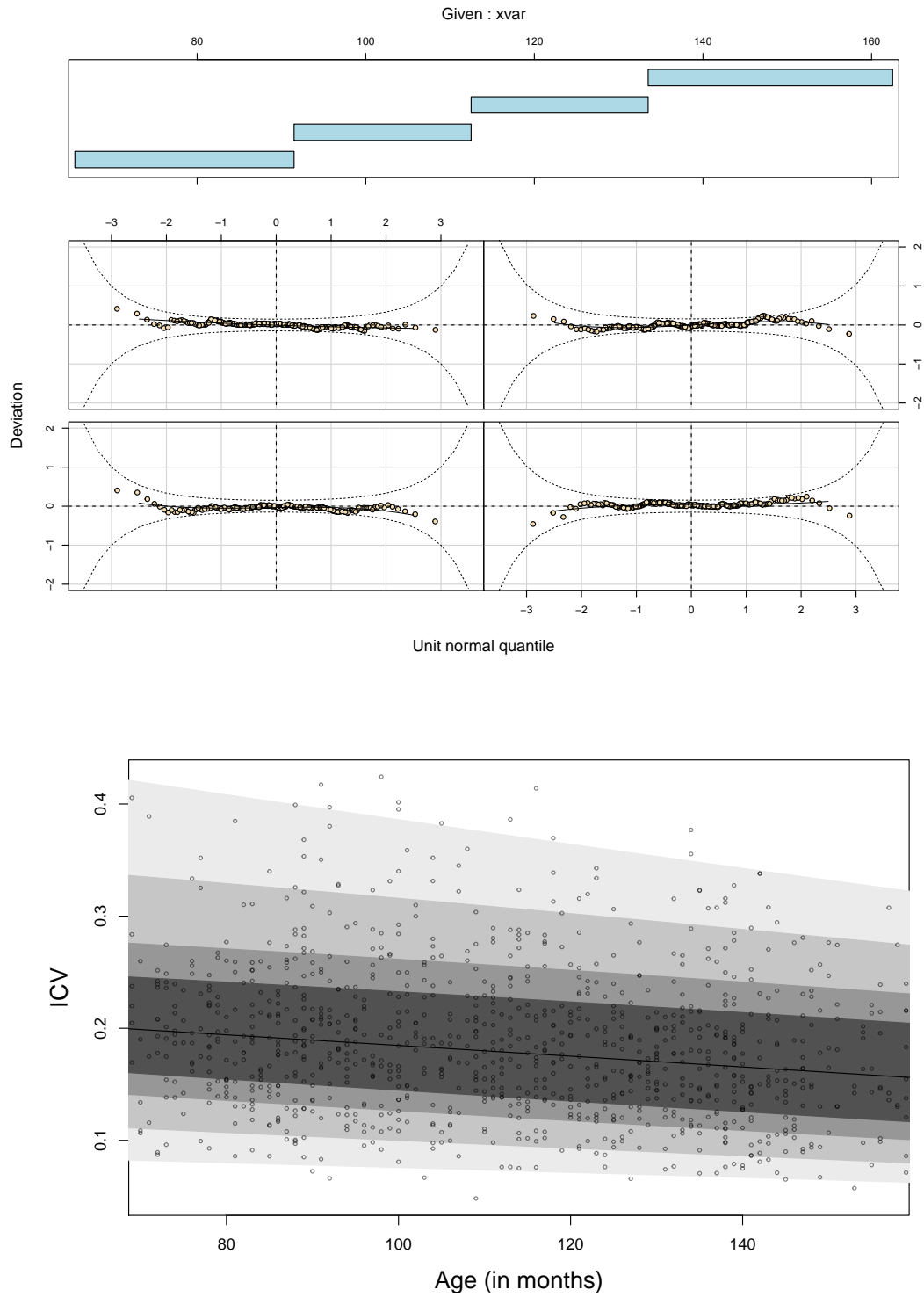
Blok 2



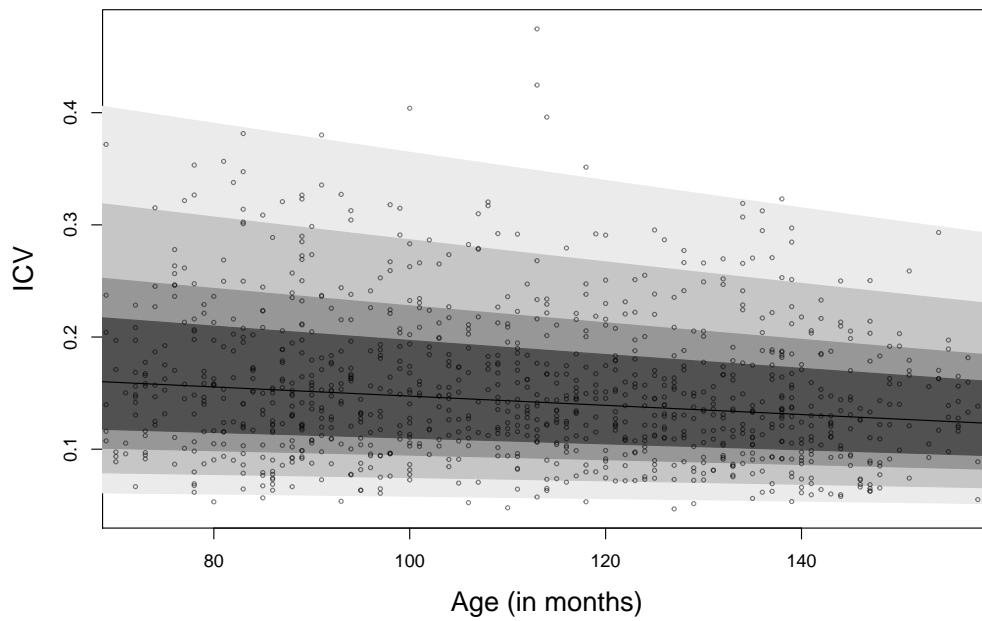
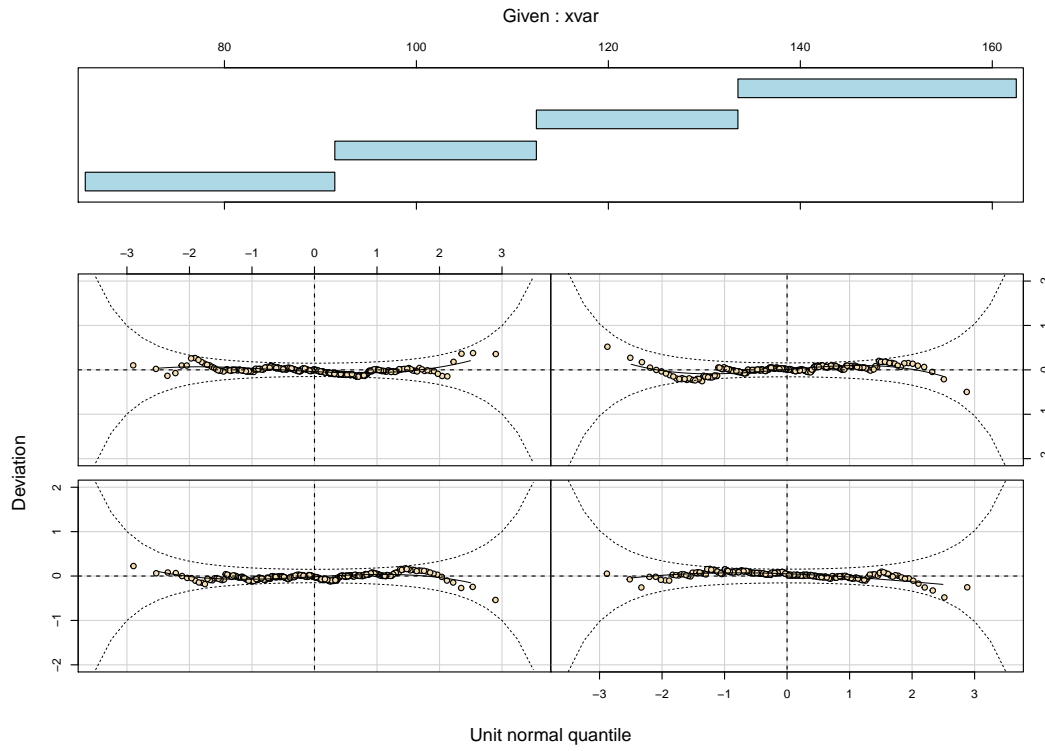
Blok 3



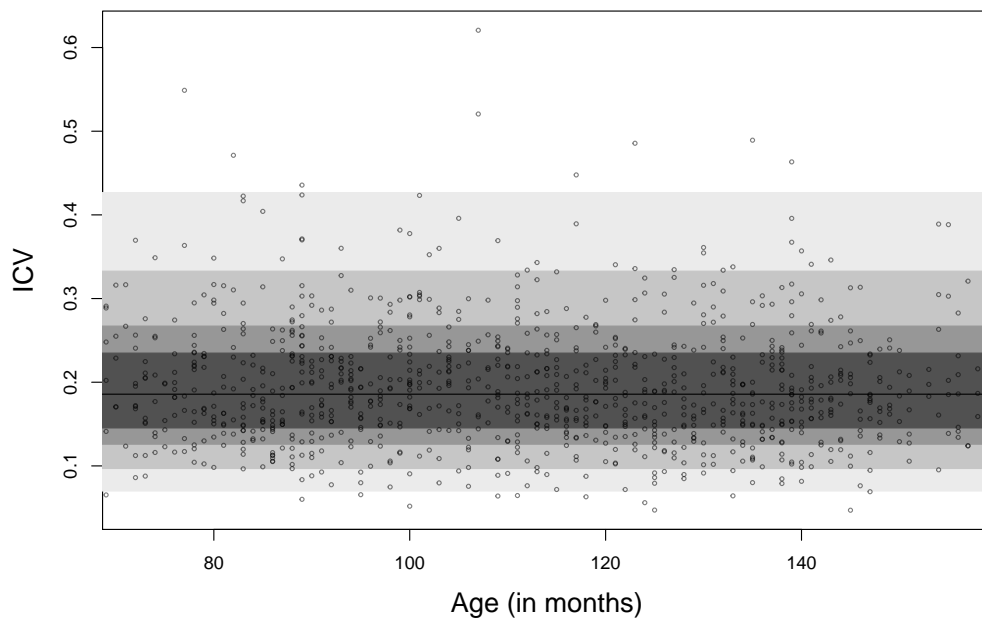
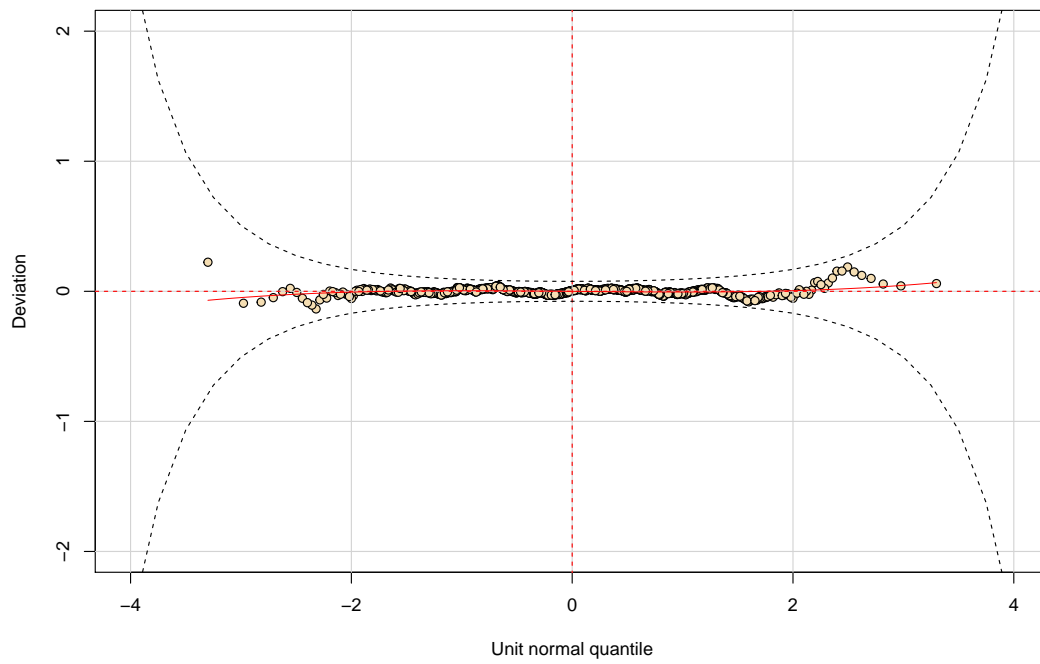
Blok 4 compatibele trials



Blok 4 incompatibele trials

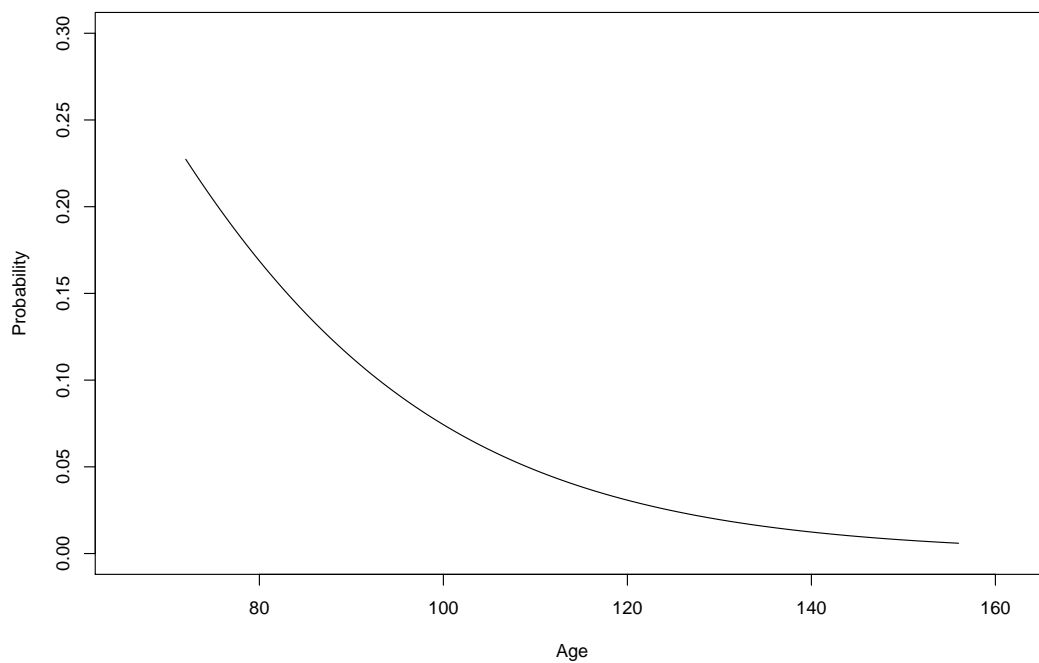
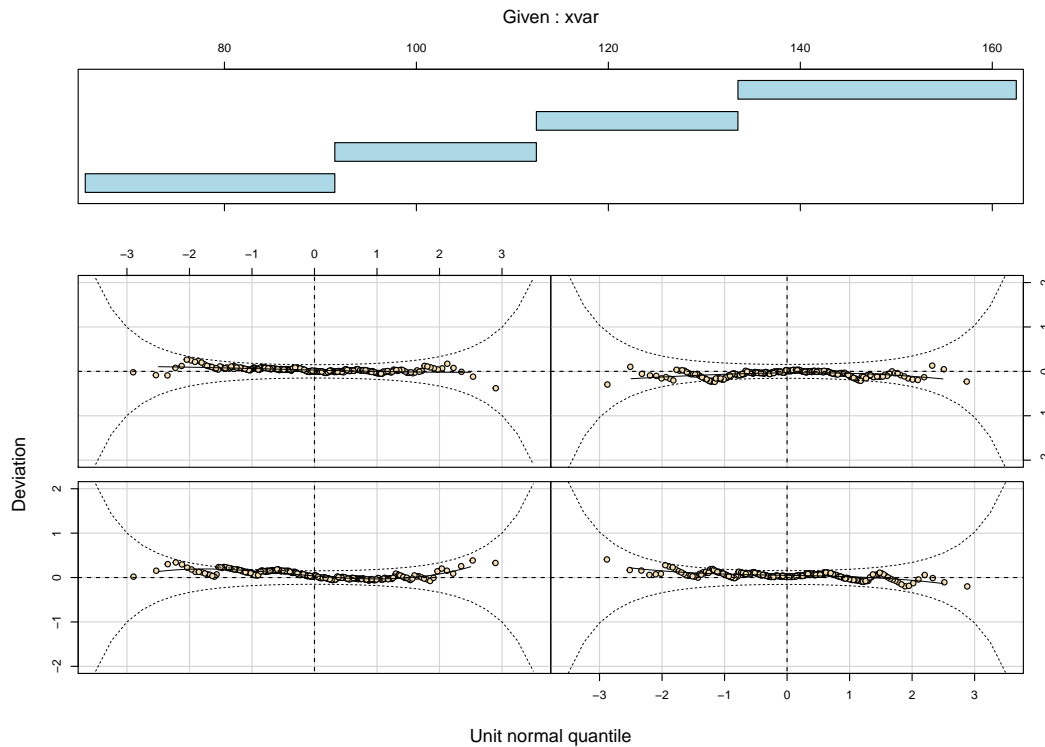


Blok 5 1-back

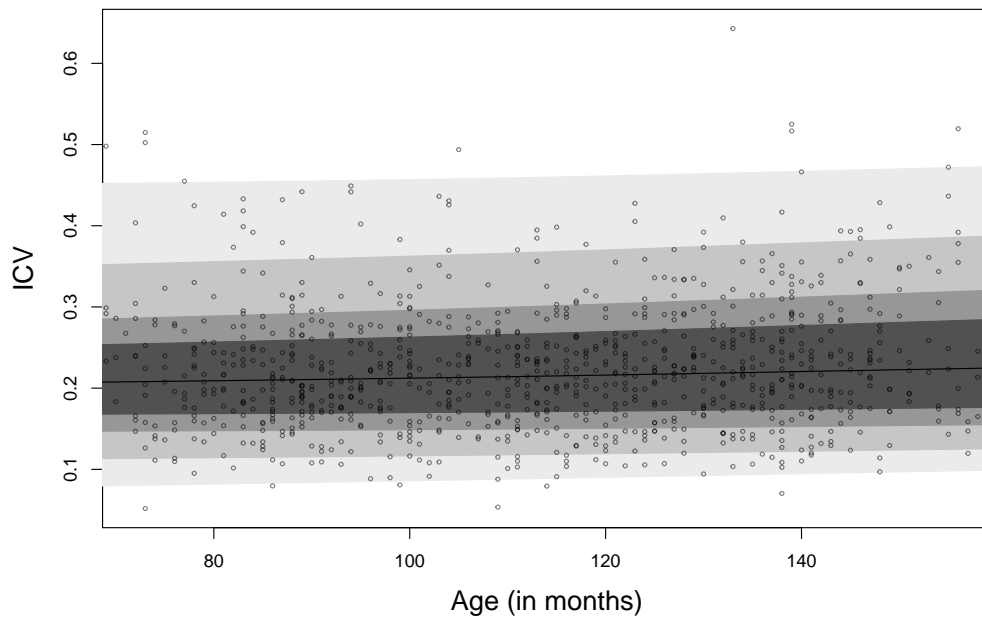
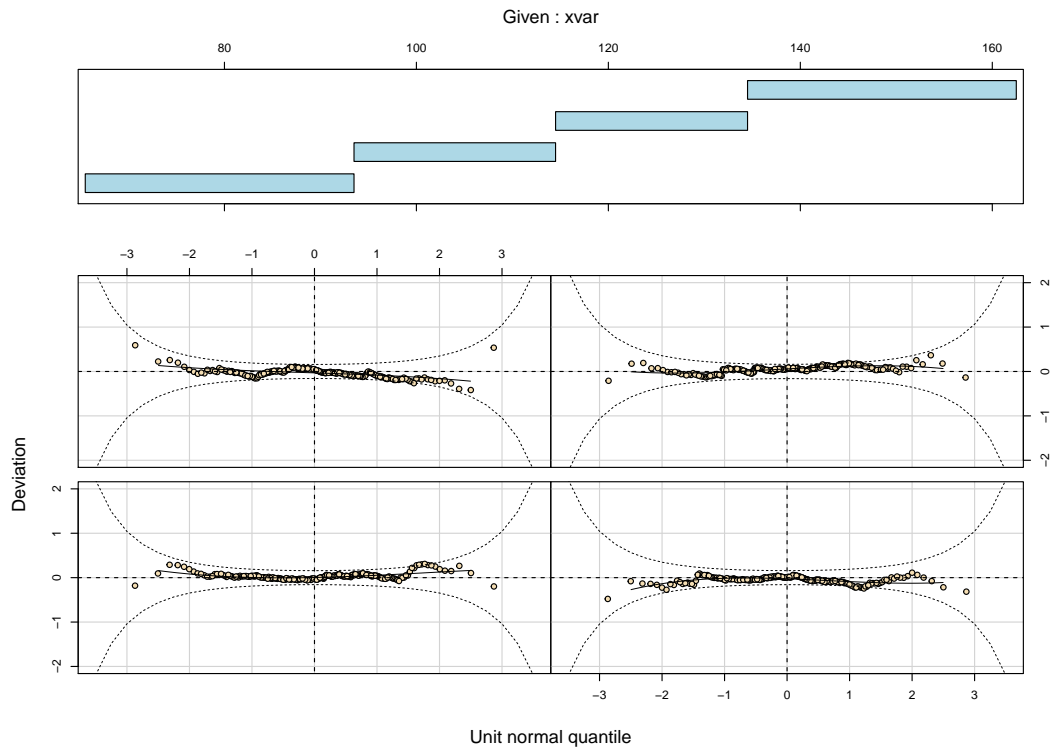


Blok 5 2-back

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.

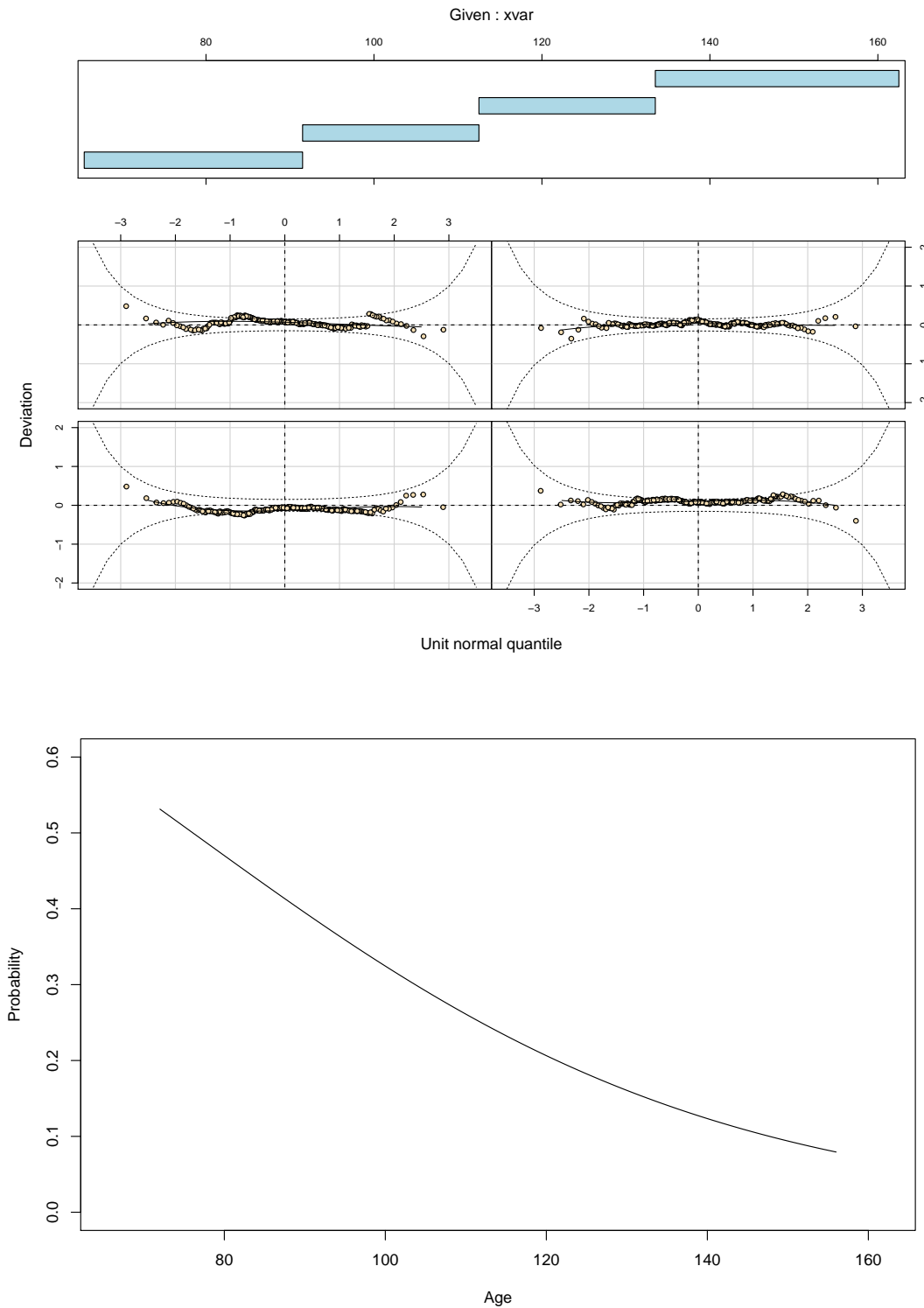


Blok 5 2-back (vervolg)

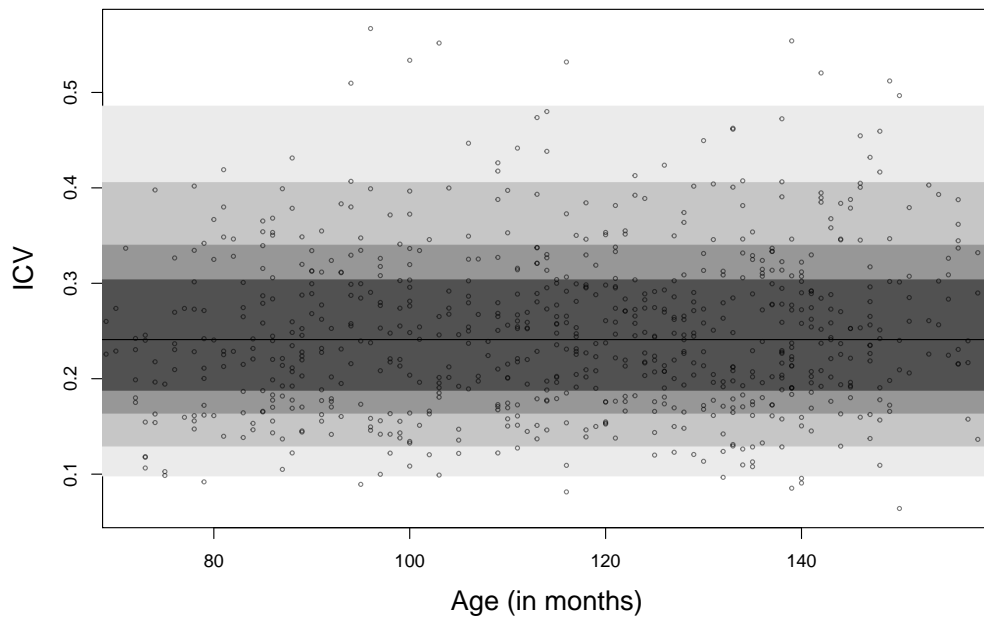
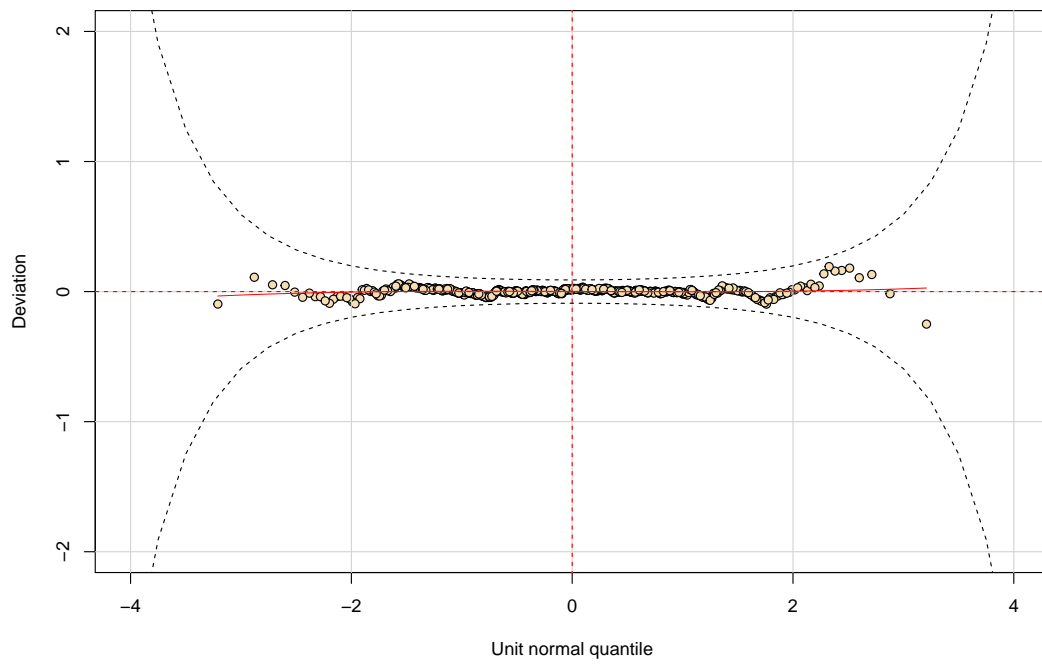


Blok 5 2-back PI

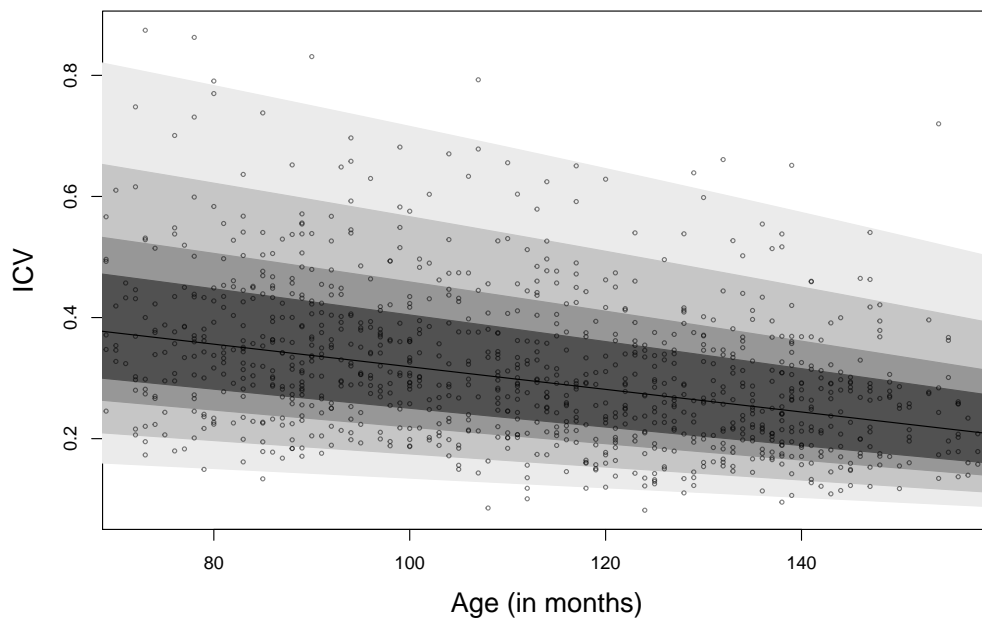
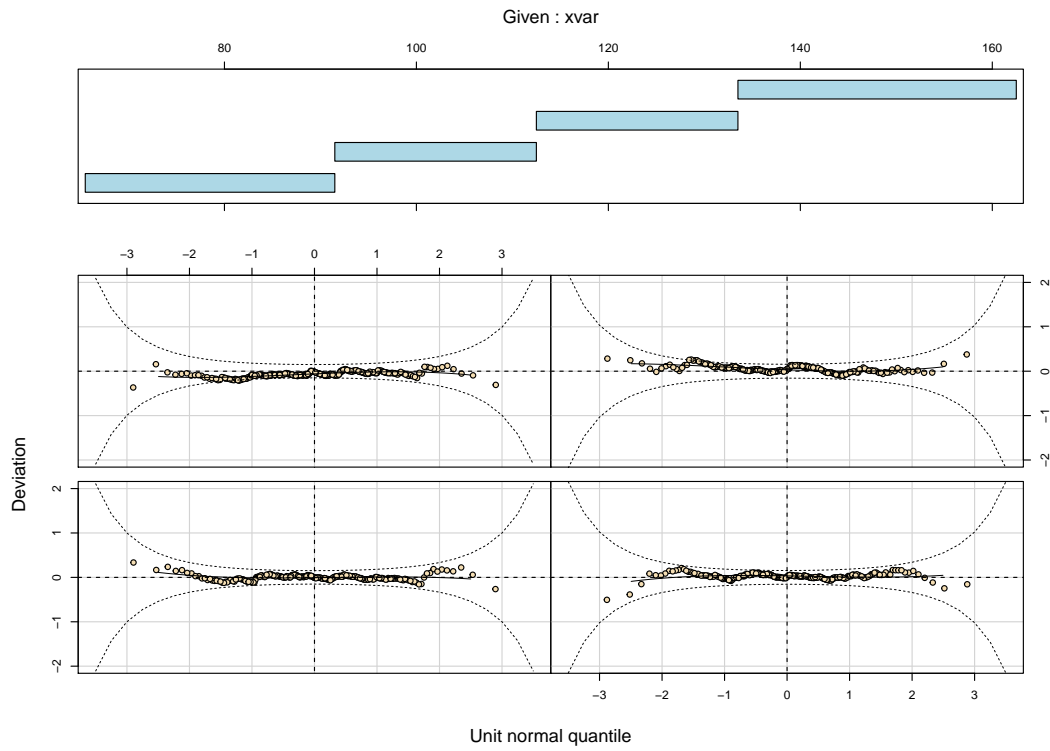
Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.



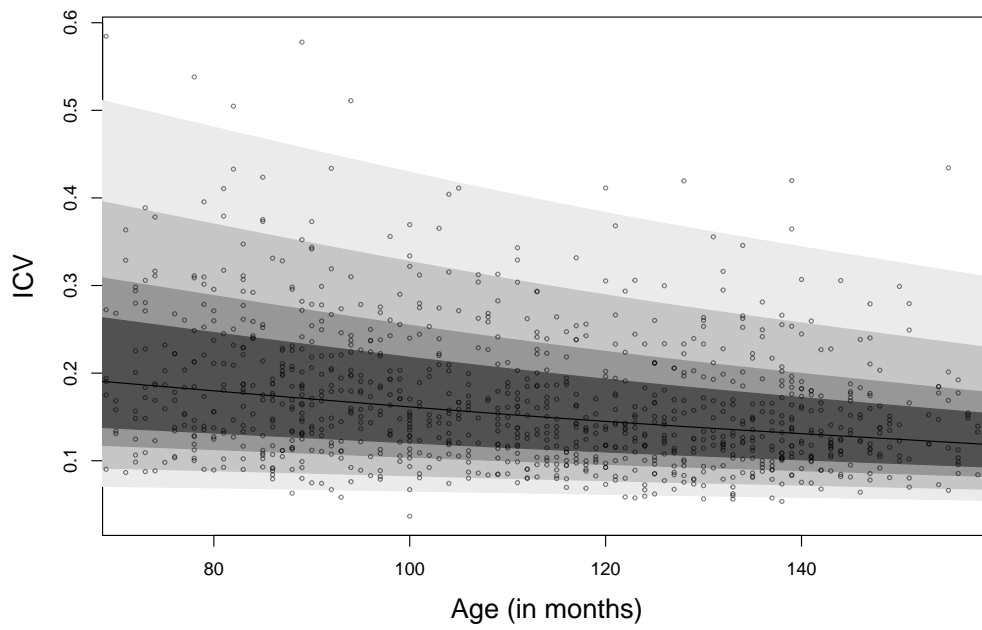
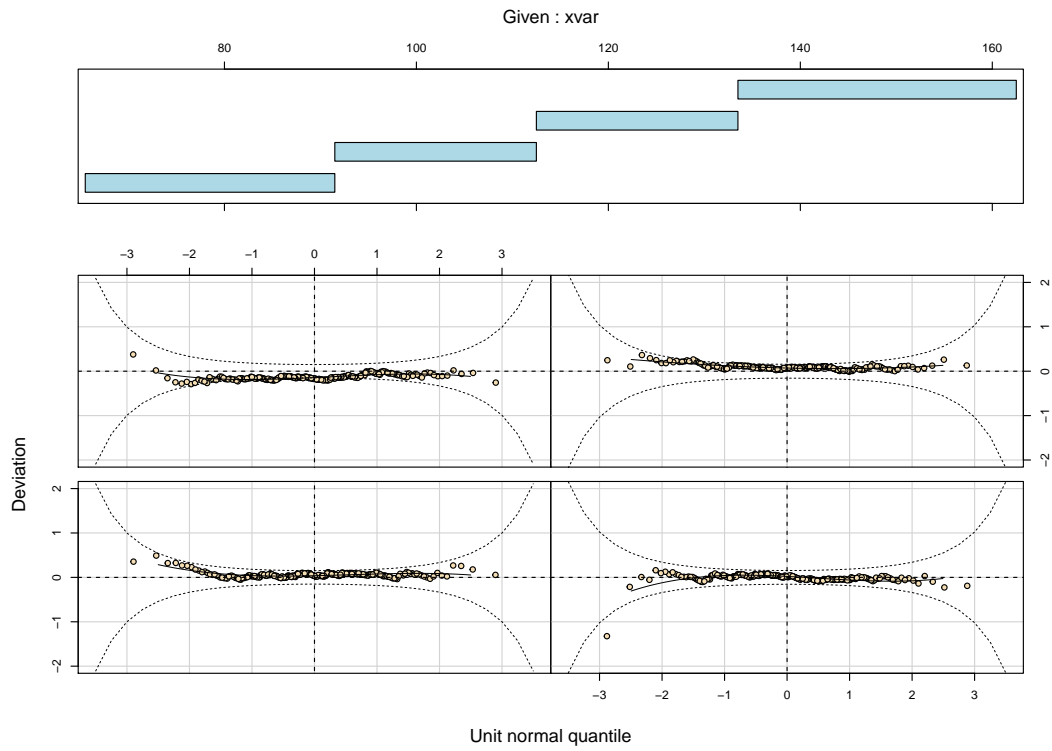
Blok 5 2-back PI (vervolg)



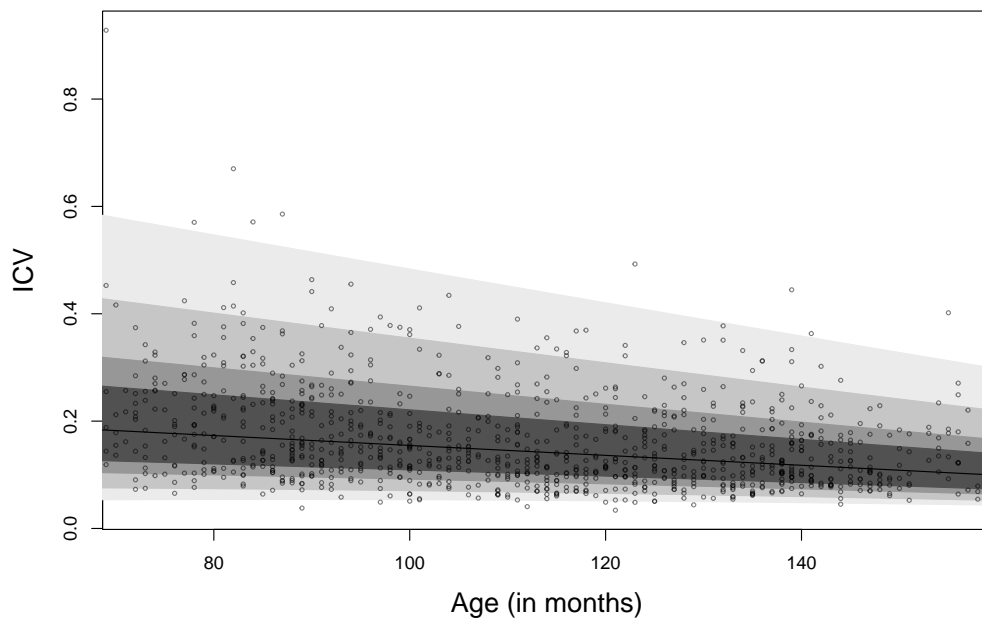
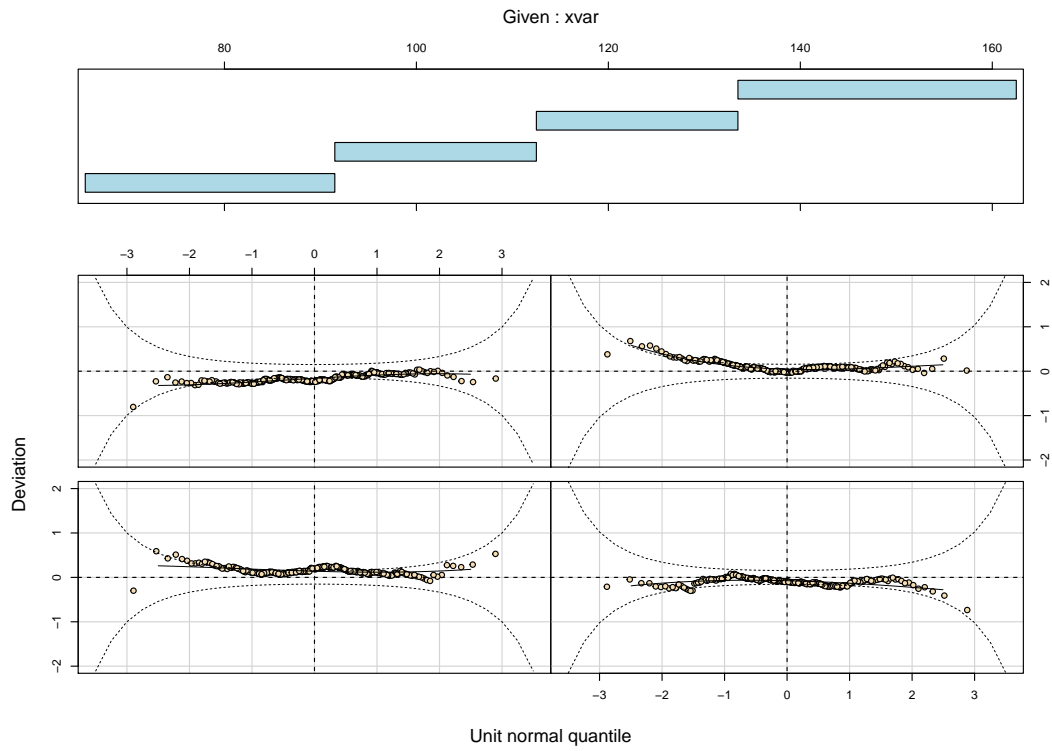
Blok 6



Blok 7 deel 1

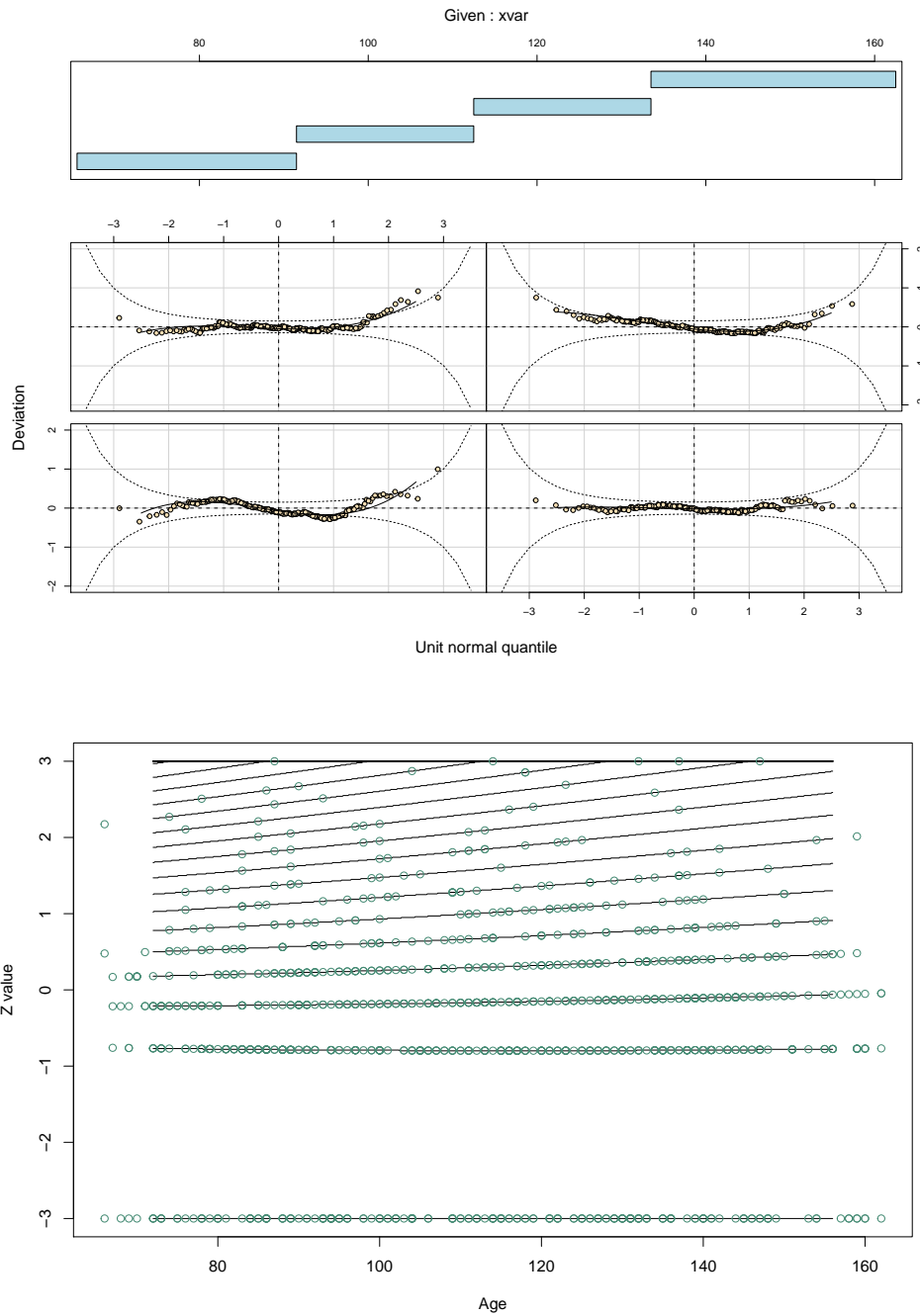


Blok 7 deel 2



AANTAL FOUTEN

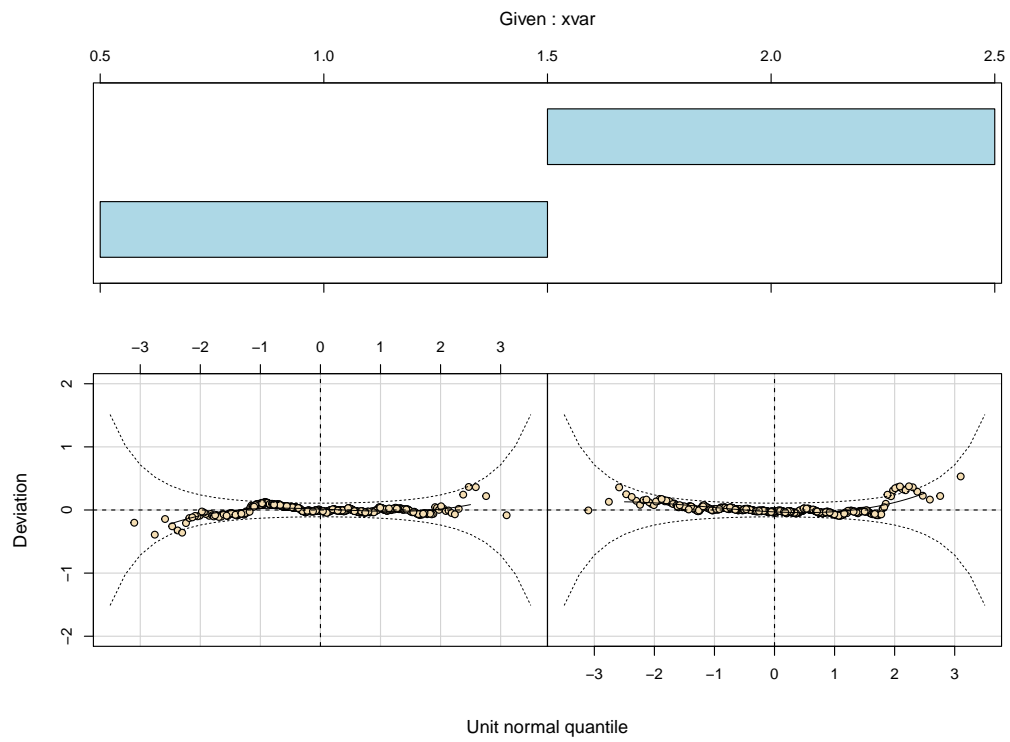
Blok 2



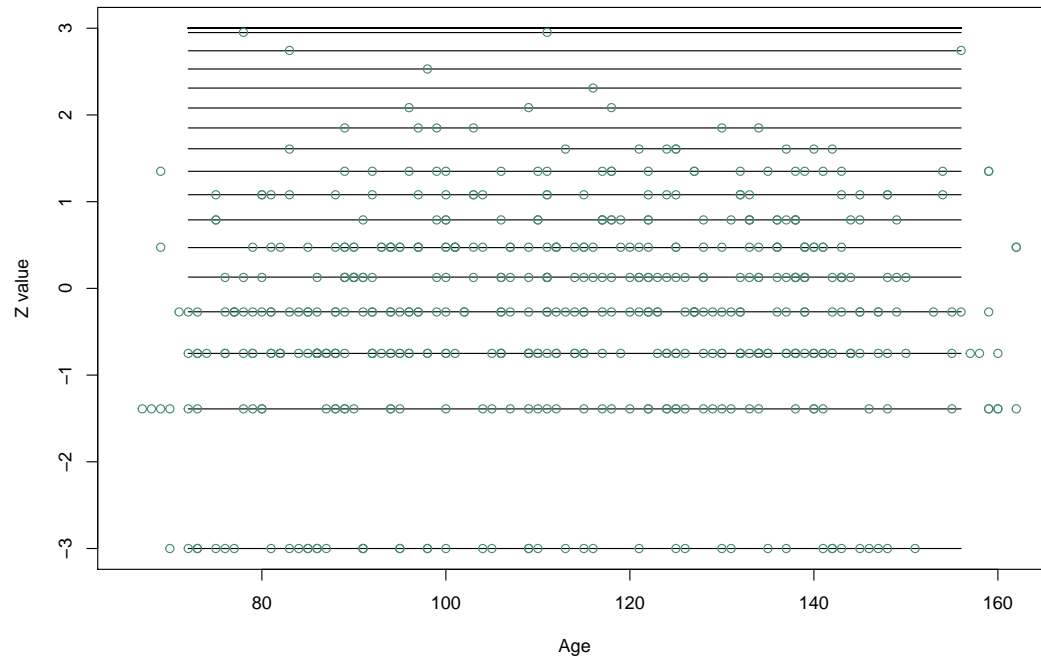
In de z -curveplot, en dus ook in de normen, komen maar drie negatieve z -curves voor. Dit komt doordat de ruwe uitkomstvariabele discreet, scheef verdeeld is. Het grootste deel van de kinderen scoort heel goed op dit testonderdeel. Dit is geheel niet bezwaarlijk voor het gebruik van de normen, omdat alleen relatief slechte scores (i.e., hoge z -waarden) een klinische betekenis hebben, en niet de scores in het normale scoringsbereik. Dit fenomeen is ook zichtbaar in andere testonderdelen, zoals bijvoorbeeld in Fouten Blok 2.

Blok 3

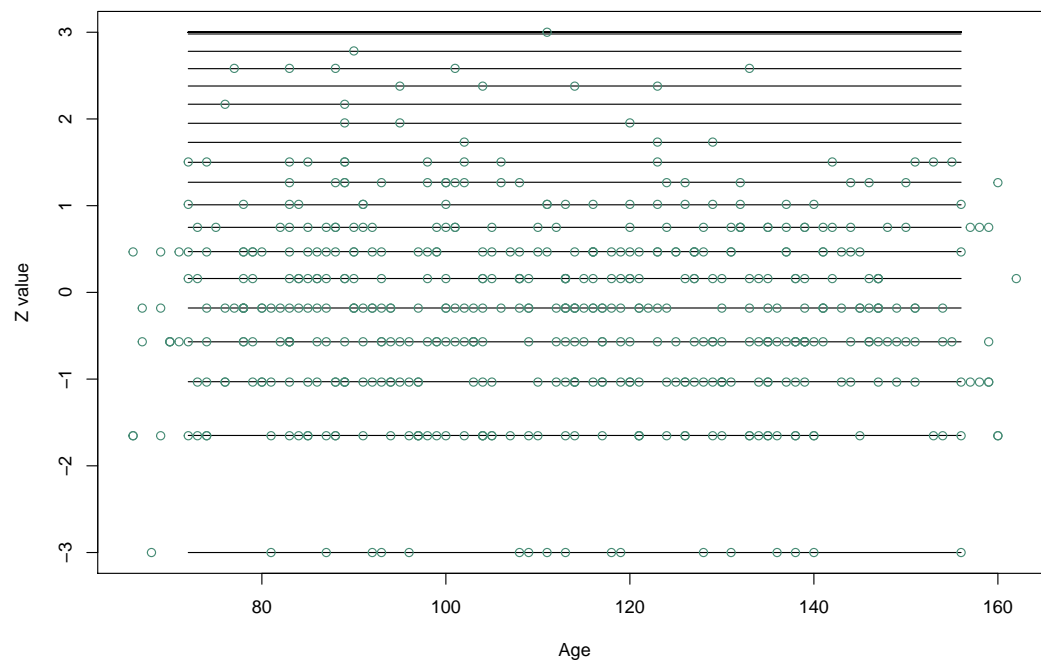
Bij Blok 3 was er alleen een effect van geslacht, en niet van leeftijd. De linker wormplot is voor de jongens, de rechter voor de meisjes.



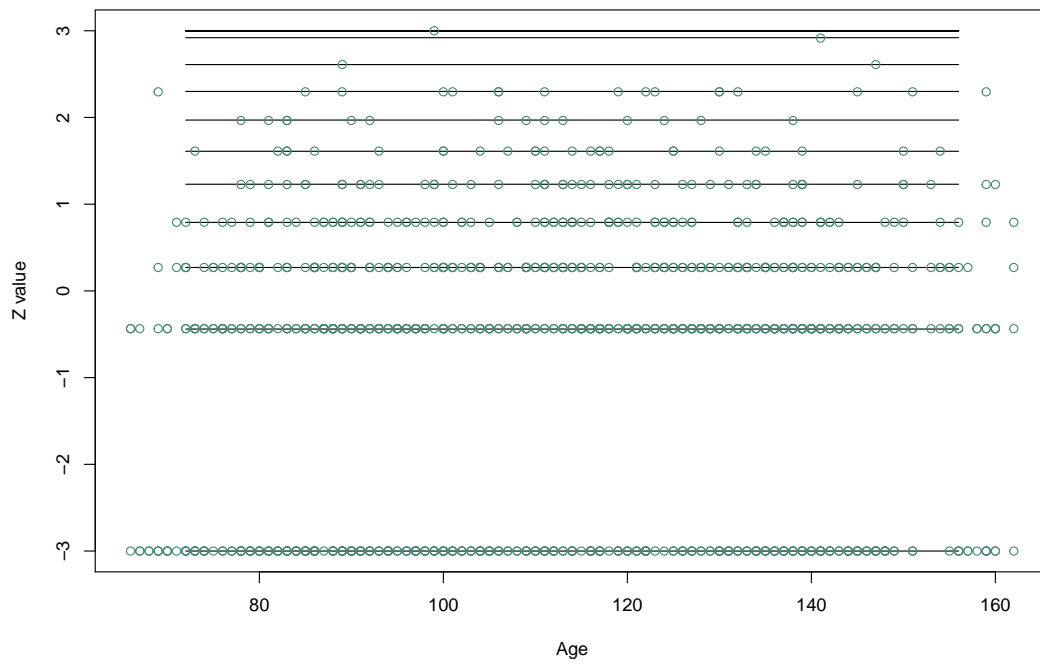
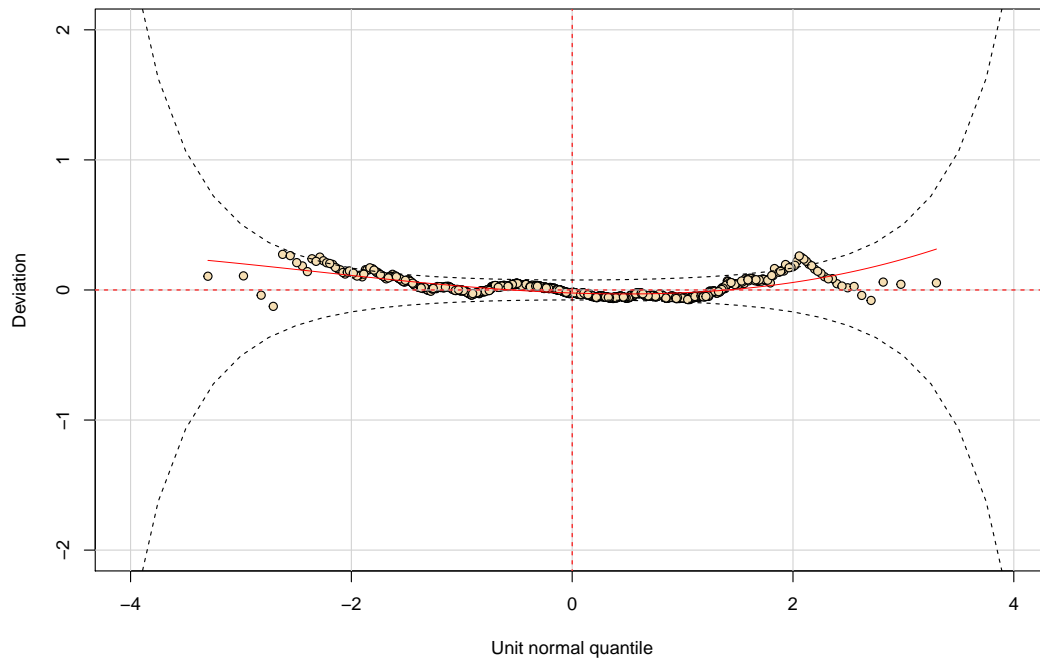
z-curveplot van Blok 3 voor de meisjes



z-curveplot van Blok 3 voor de jongens

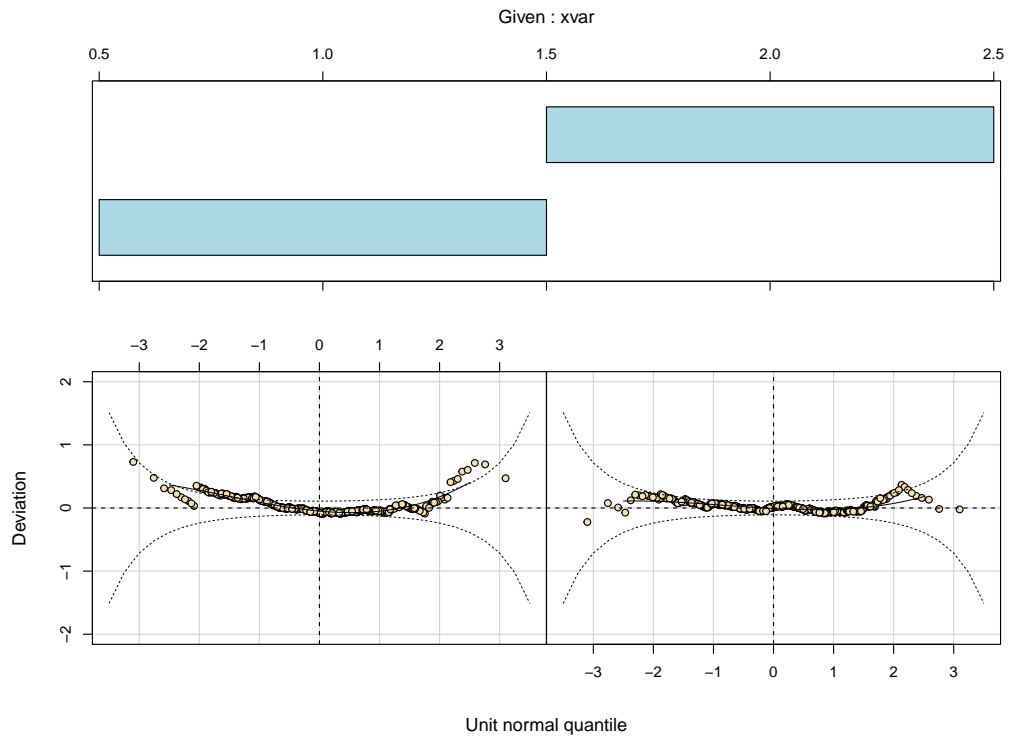


Blok 4 compatibele trials

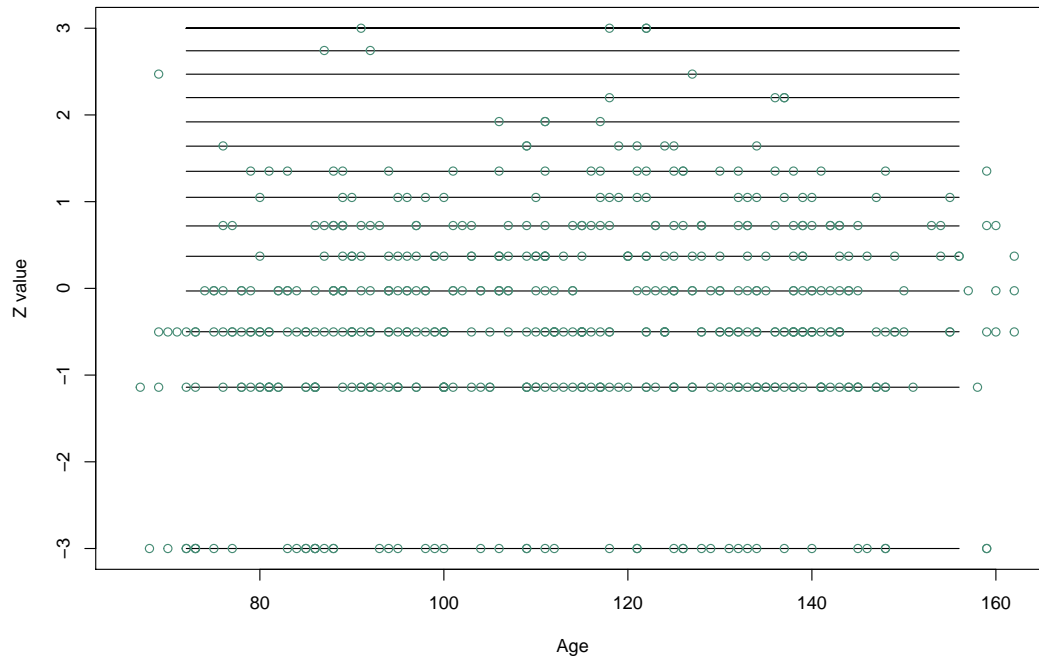


Blok 4 incompatibele trials

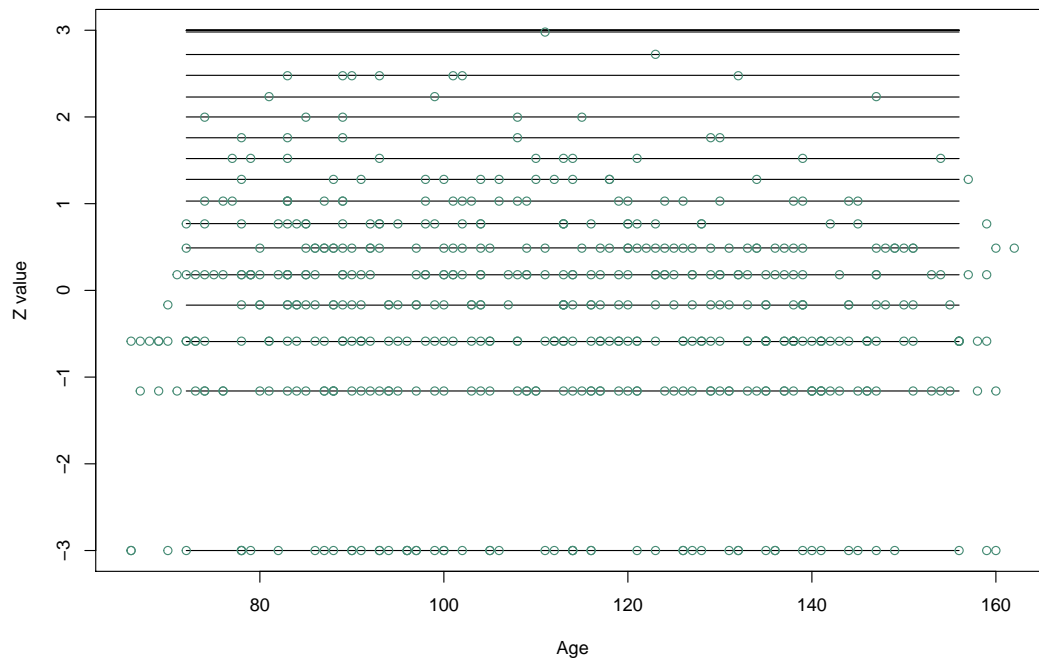
De scores op deze variabele, en daarmee de normen, zijn afhankelijk van sekse, en niet van leeftijd. De linker wormplot is voor de meisjes, de rechter voor de jongens.



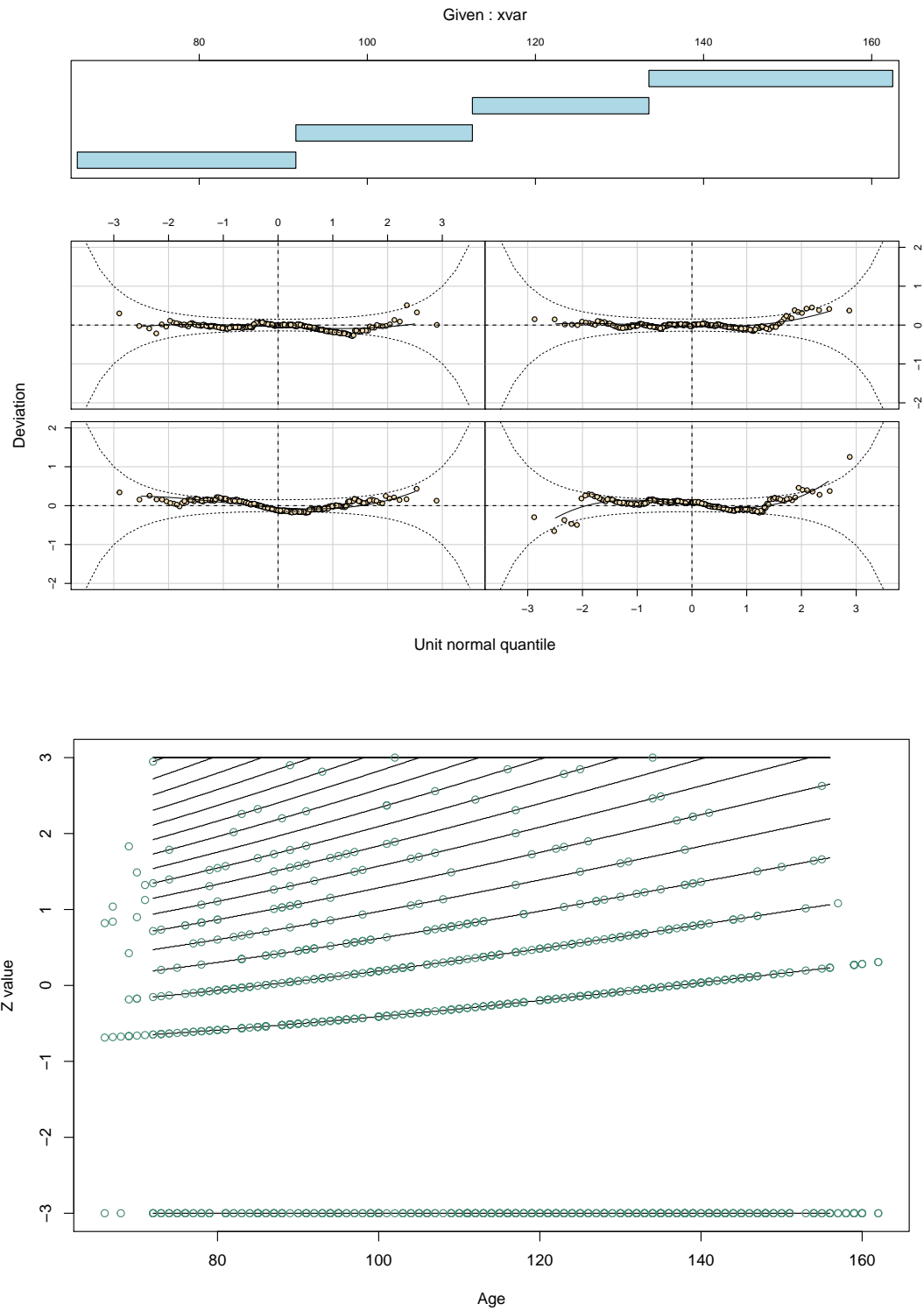
z-curveplot van Blok 4 incompatibele trials voor de meisjes



z-curveplot van Blok 4 incompatibele trials voor de jongens

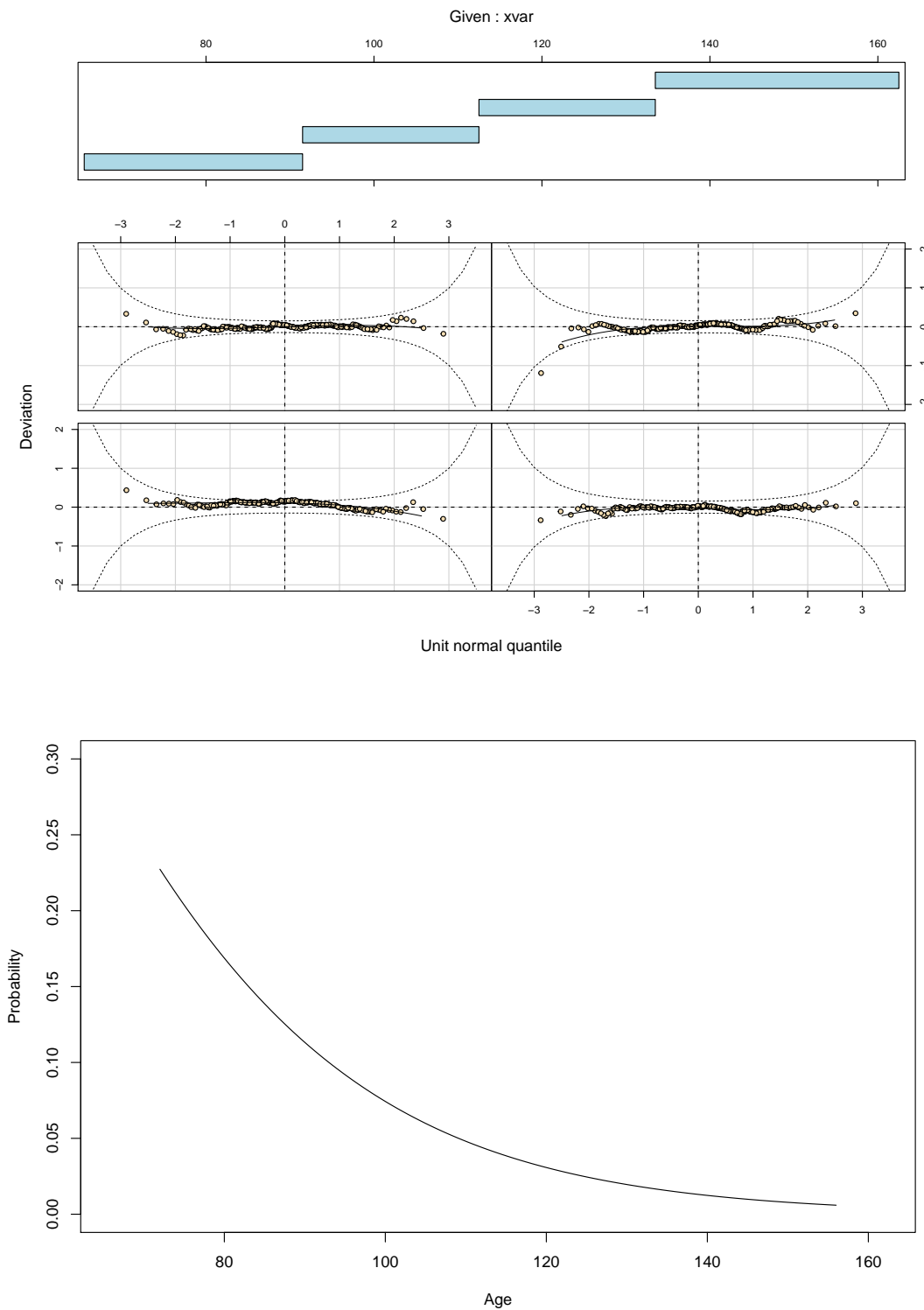


Blok 5 1-back

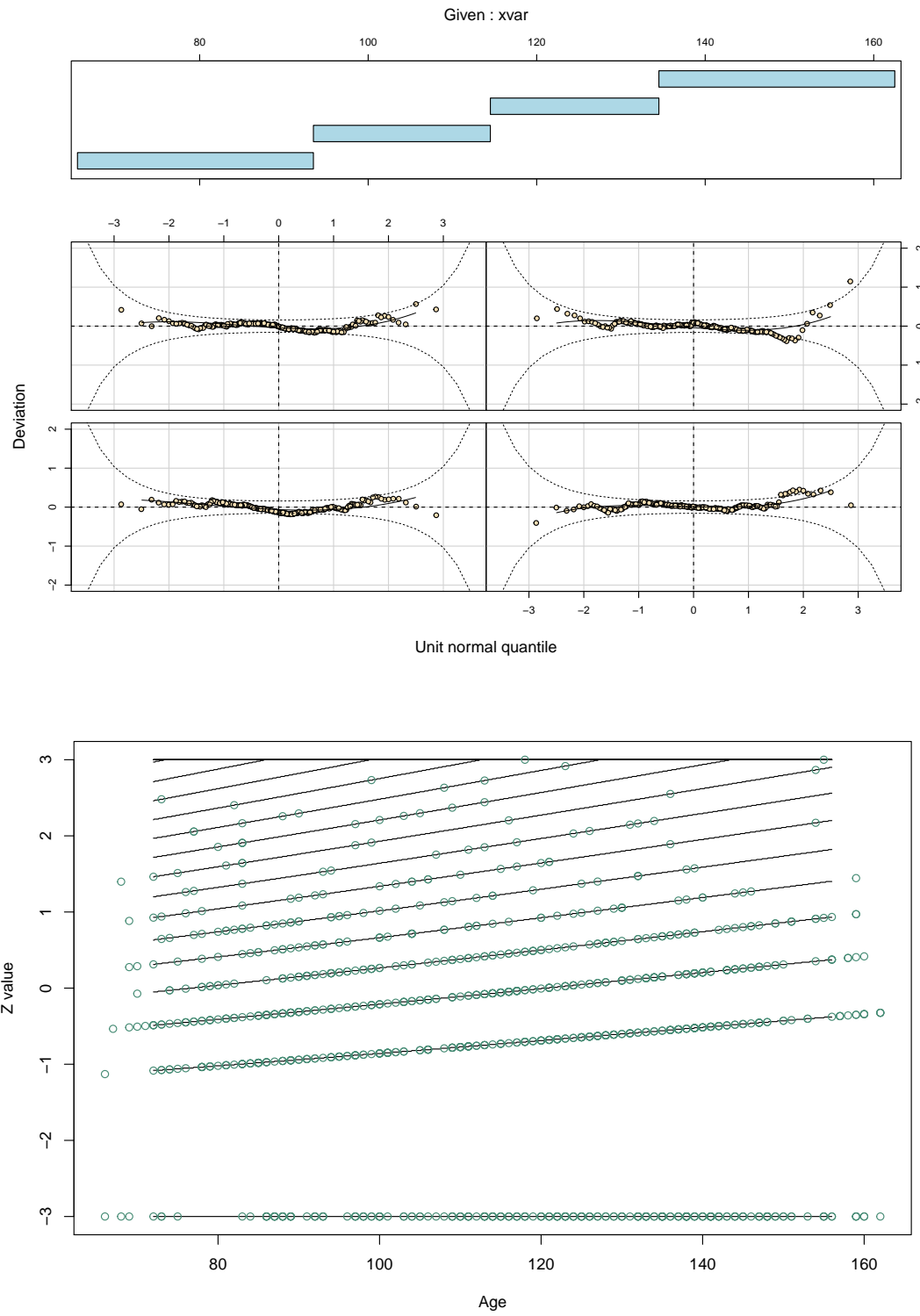


Blok 5 2-back

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.

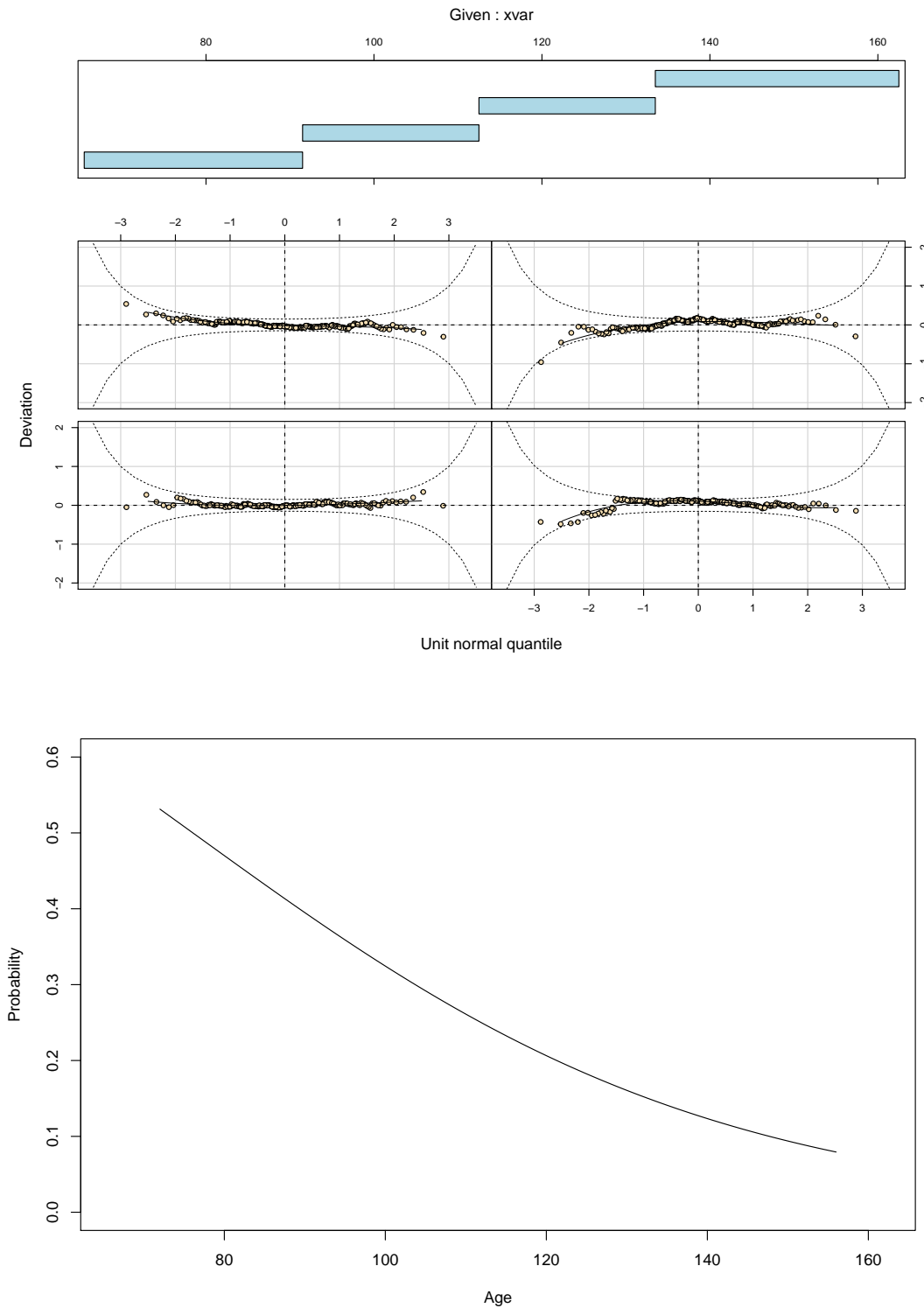


Blok 5 2-back (vervolg)

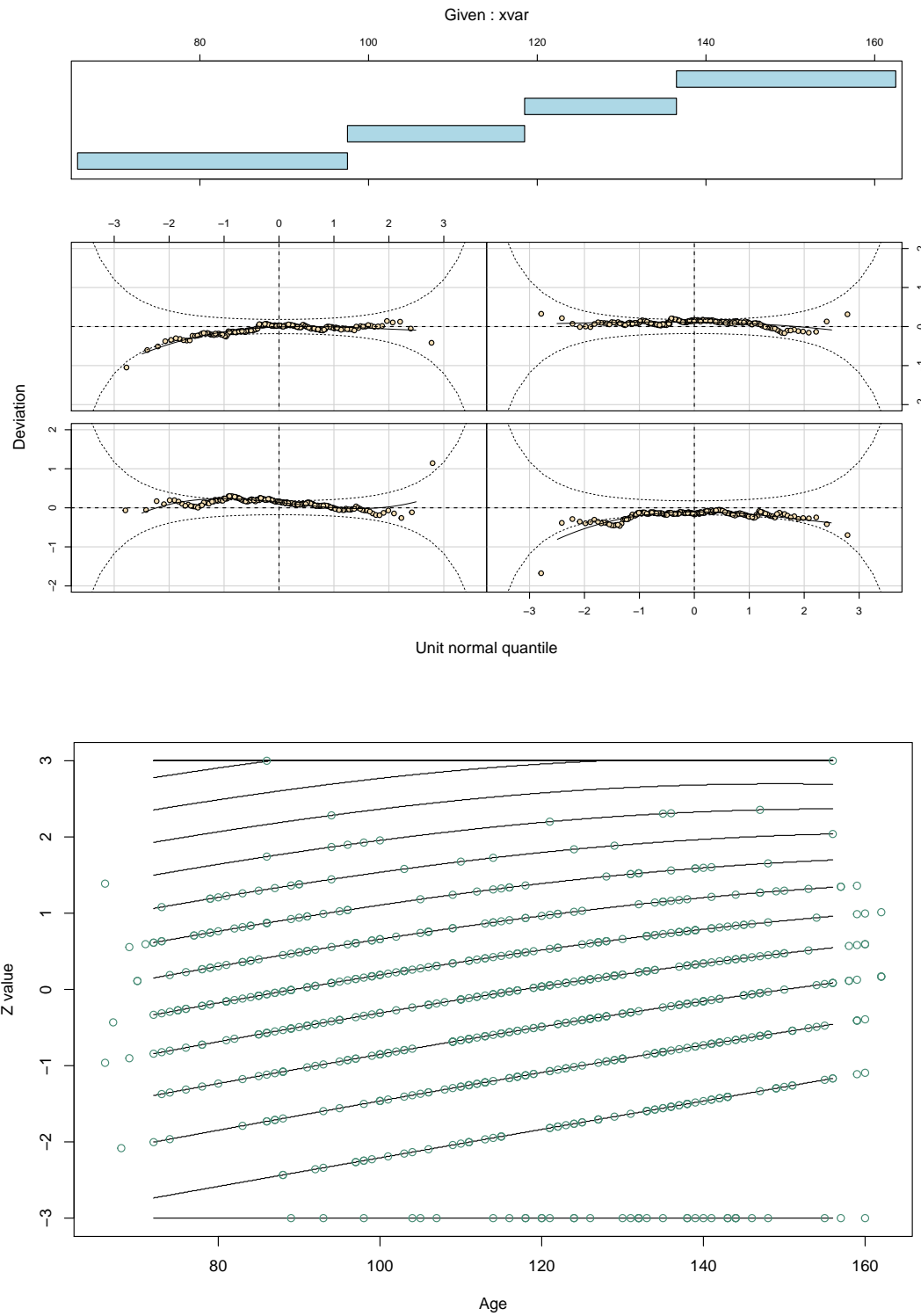


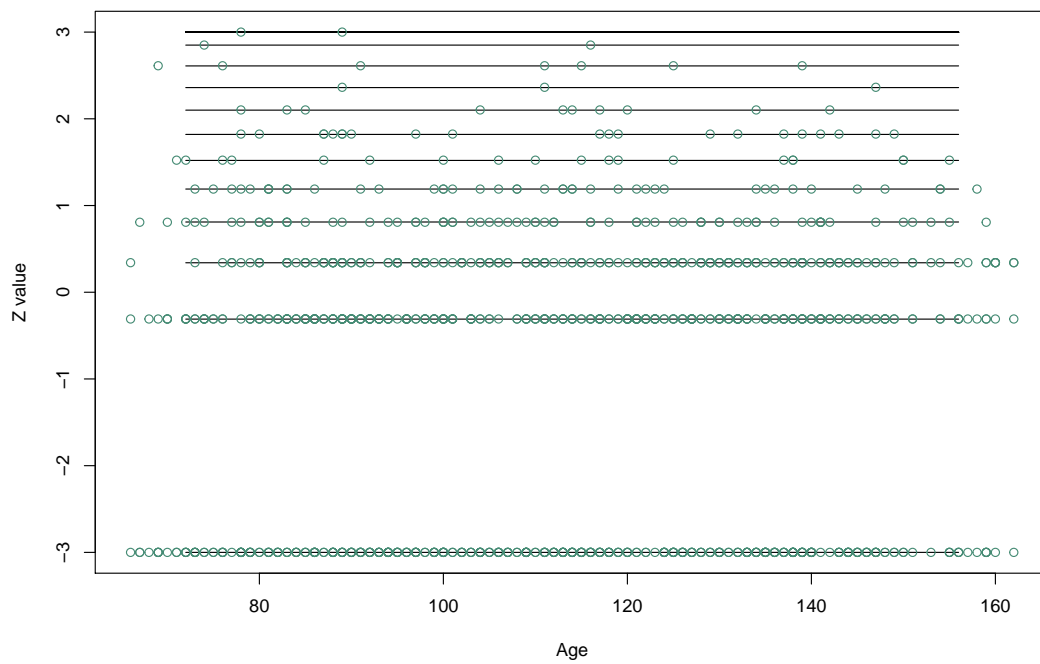
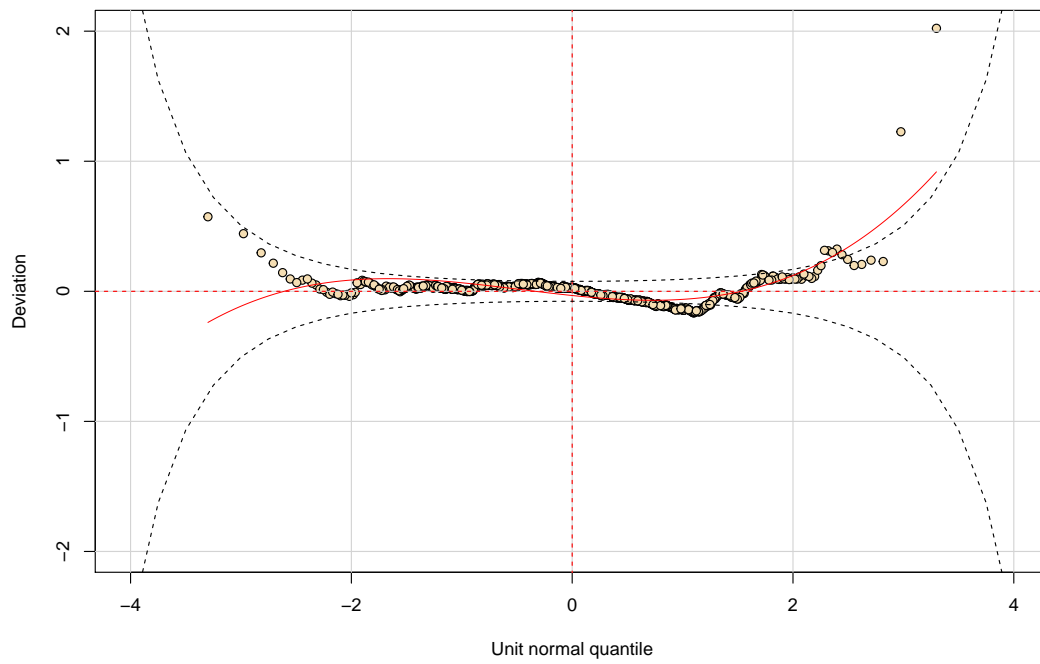
Blok 5 2-back PI

Deze wormplot behoort bij het binomiale model, dat de kans op het **niet** maken van de vraag beschrijft. Omdat dit een binaire variabele betreft is de bijbehorende centielcurveplot niet inzichtelijk. Daarom is deze achterwege gelaten. In plaats daarvan is een plot toegevoegd die de kans op het niet maken van de vraag als een functie van de leeftijd weergeeft.

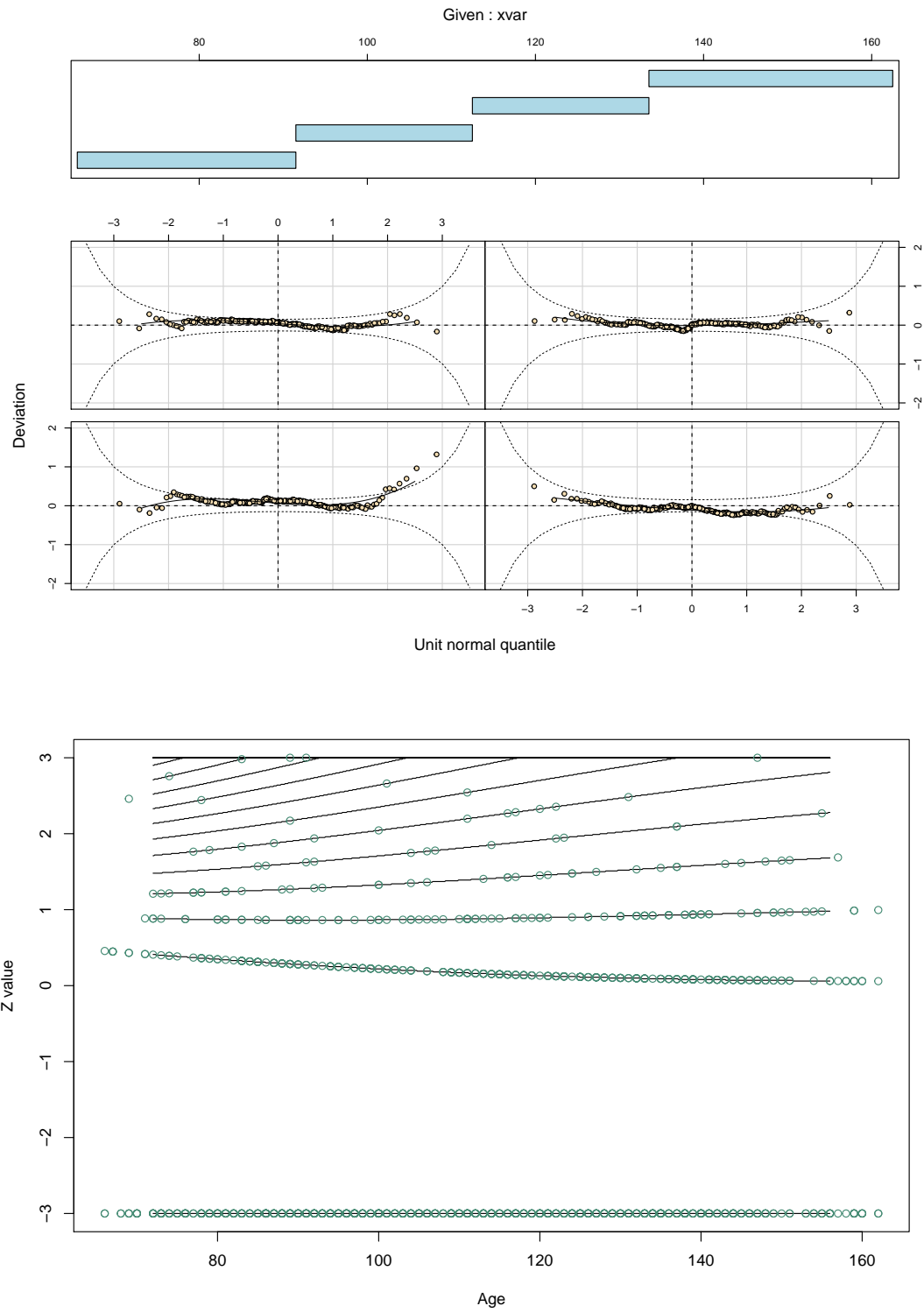


Blok 5 2-back PI (vervolg)



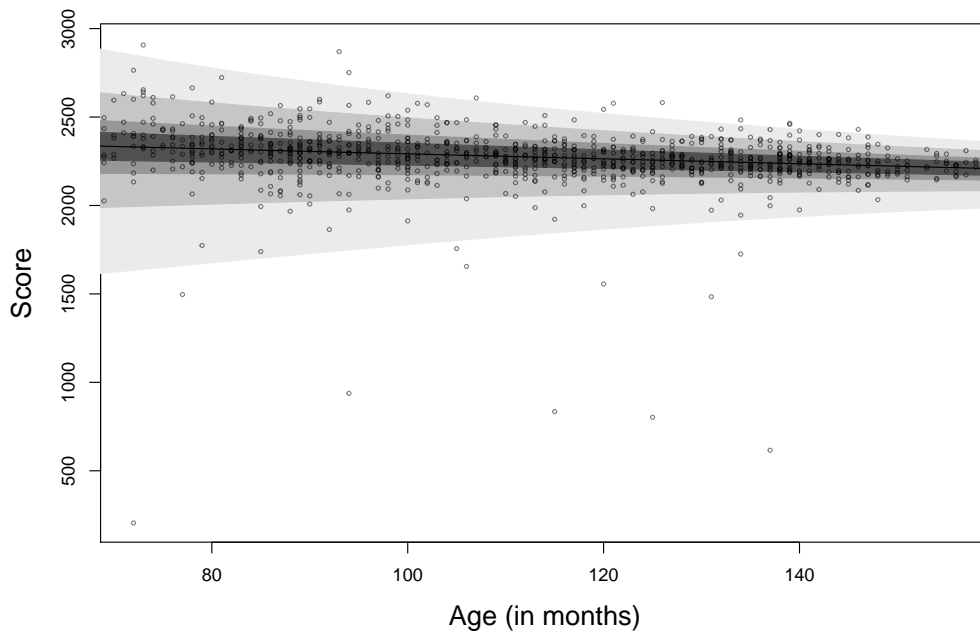
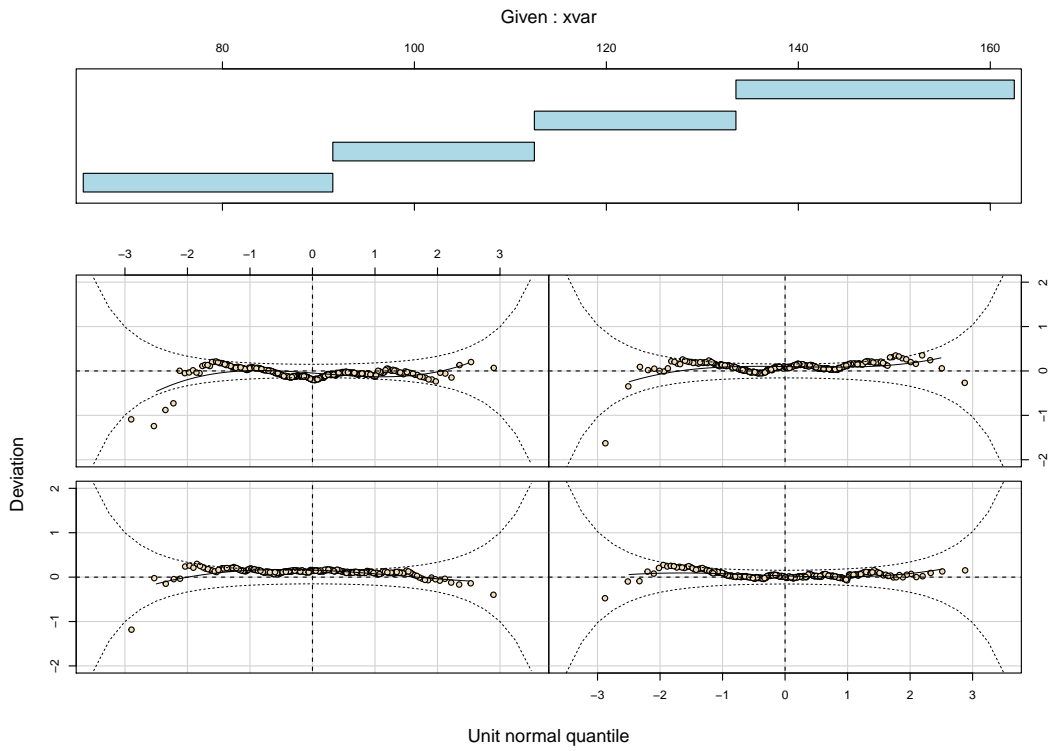


Blok 7 deel 2

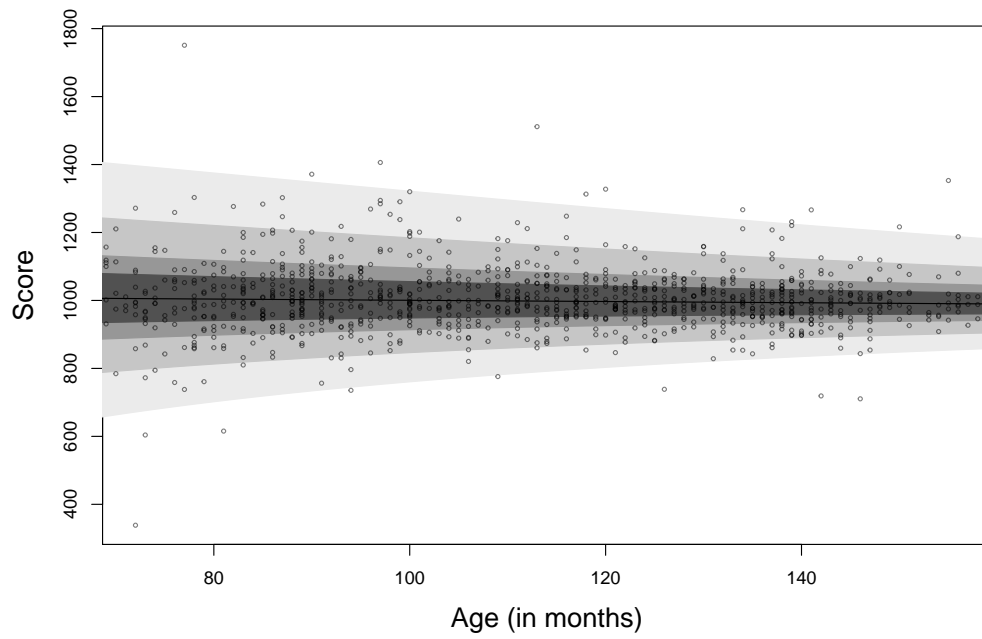
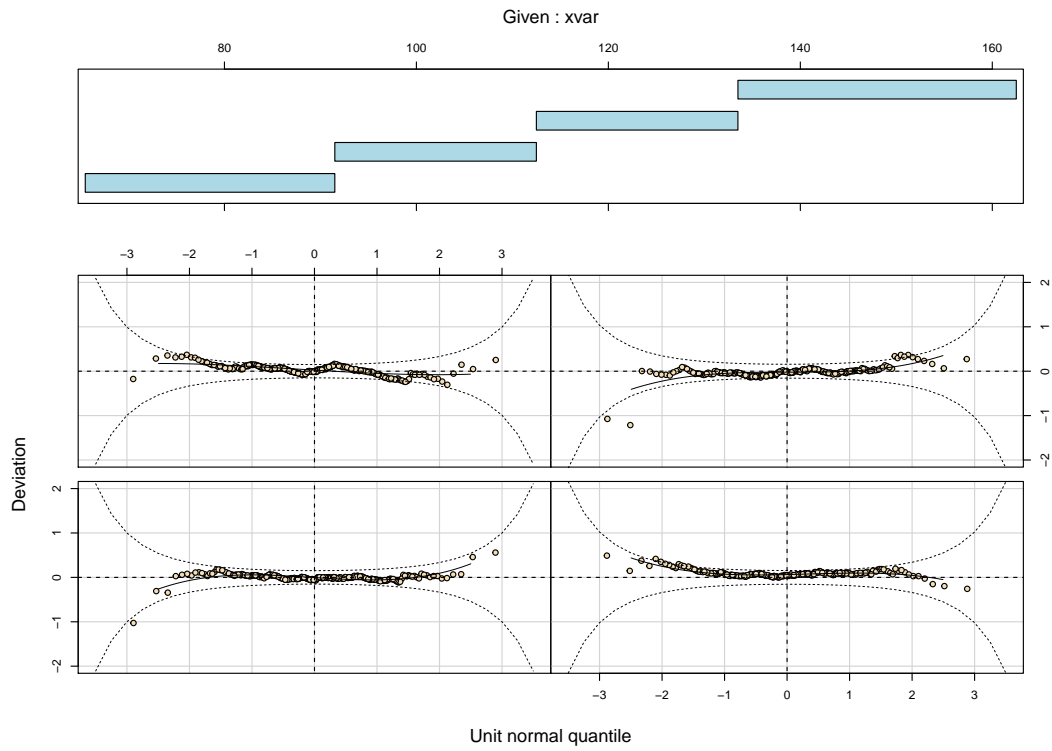


CONTRASTPARAMETERS SNELHEID

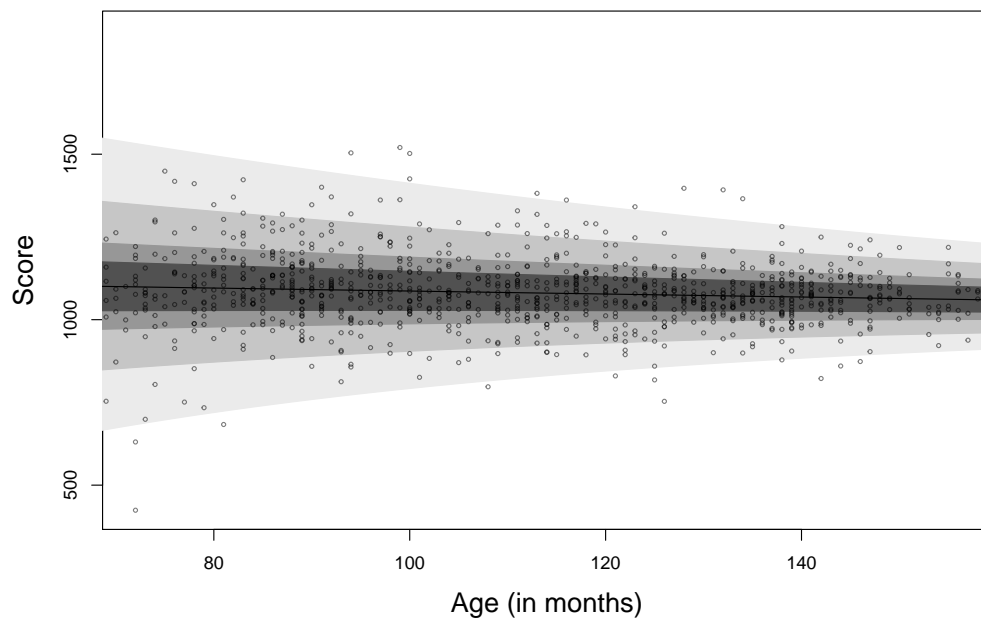
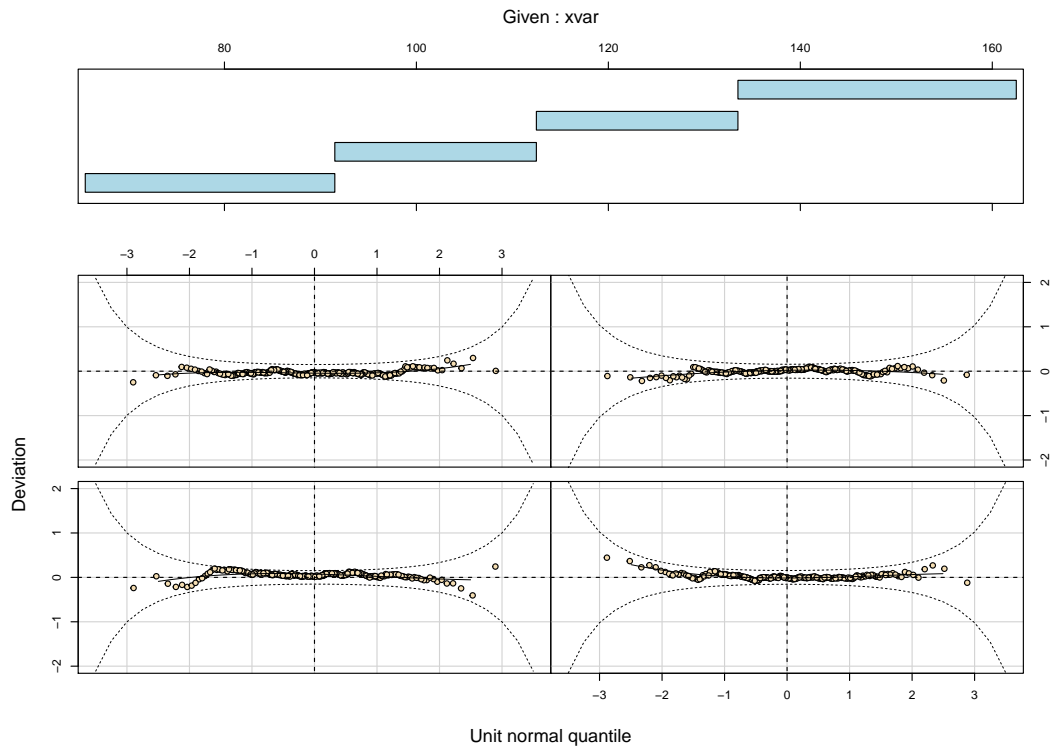
Blok 2 minus Blok 1



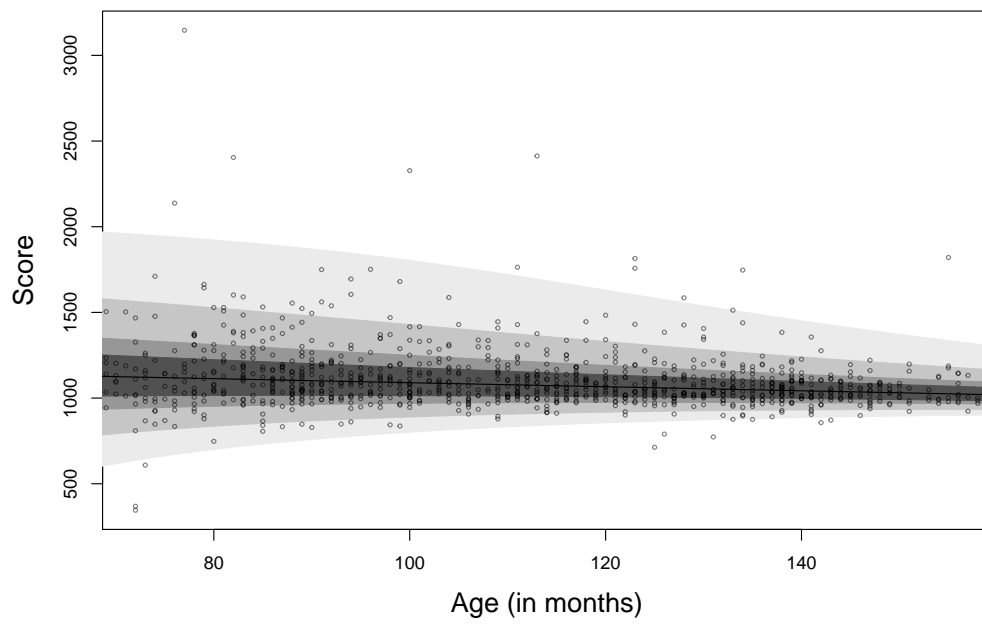
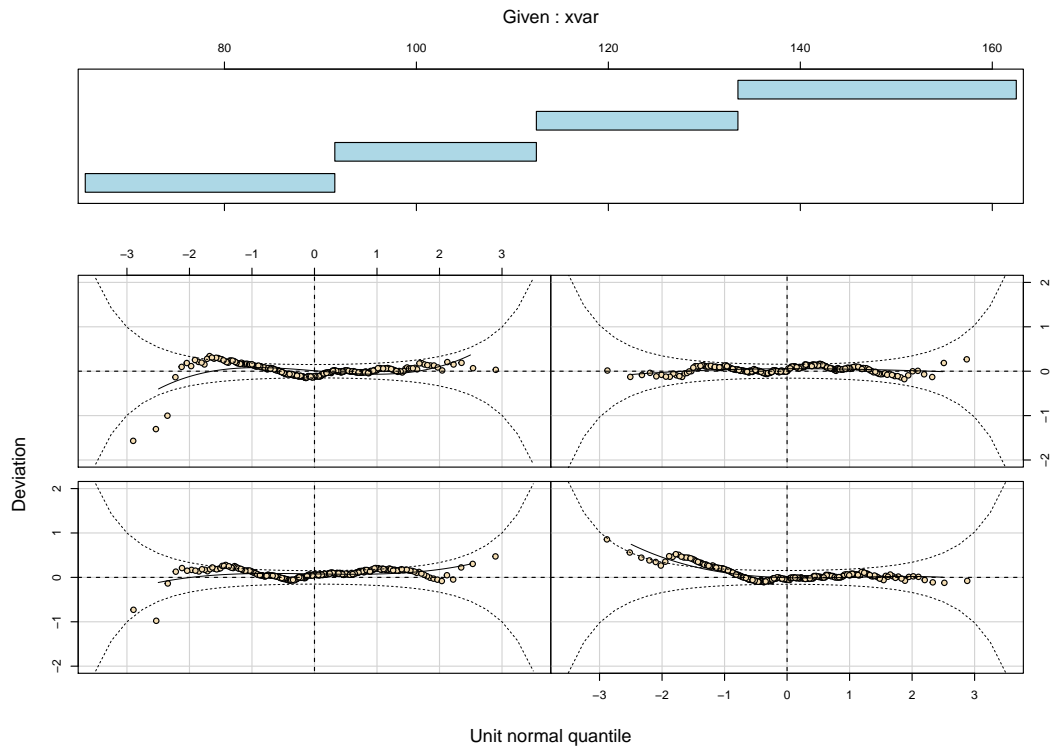
Blok 3 minus Blok 2



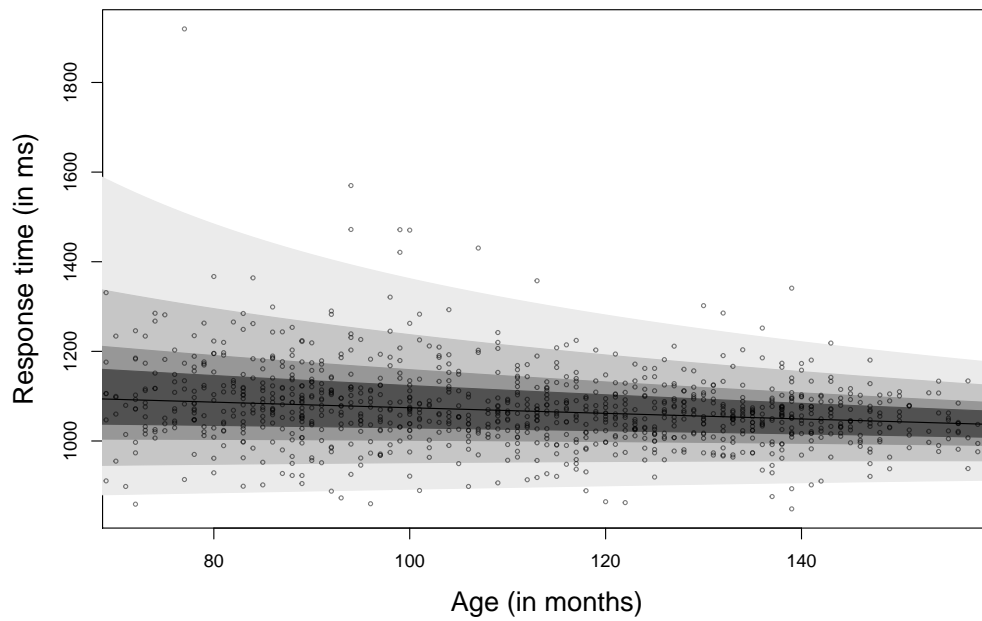
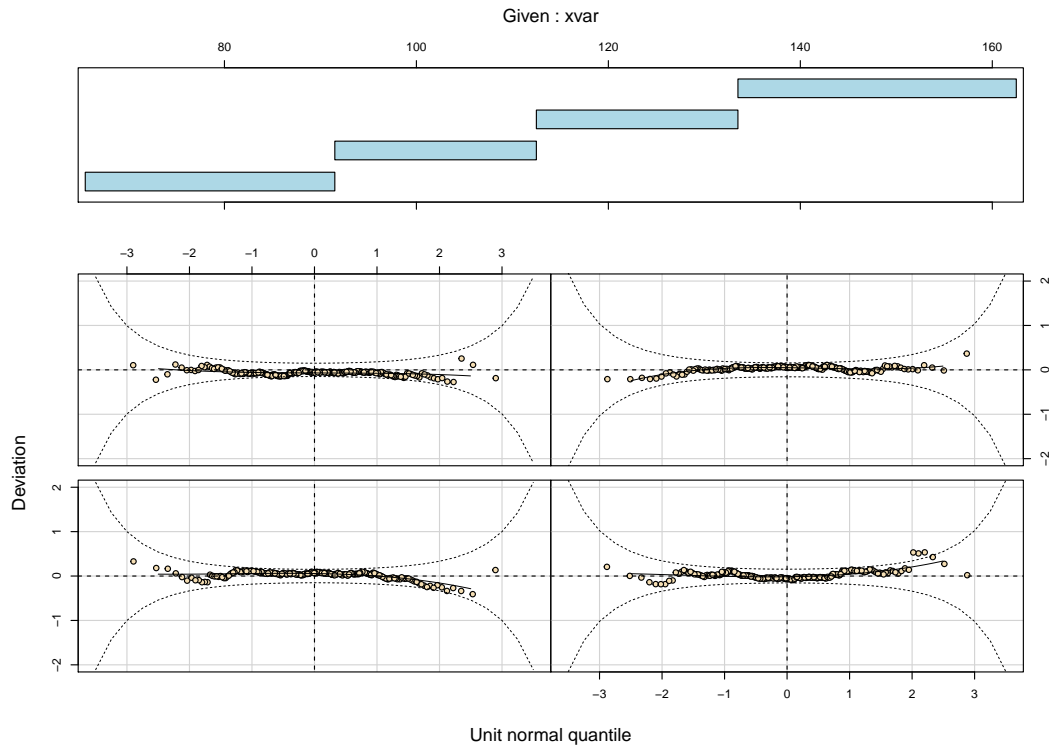
Blok 4 minus Blok 2



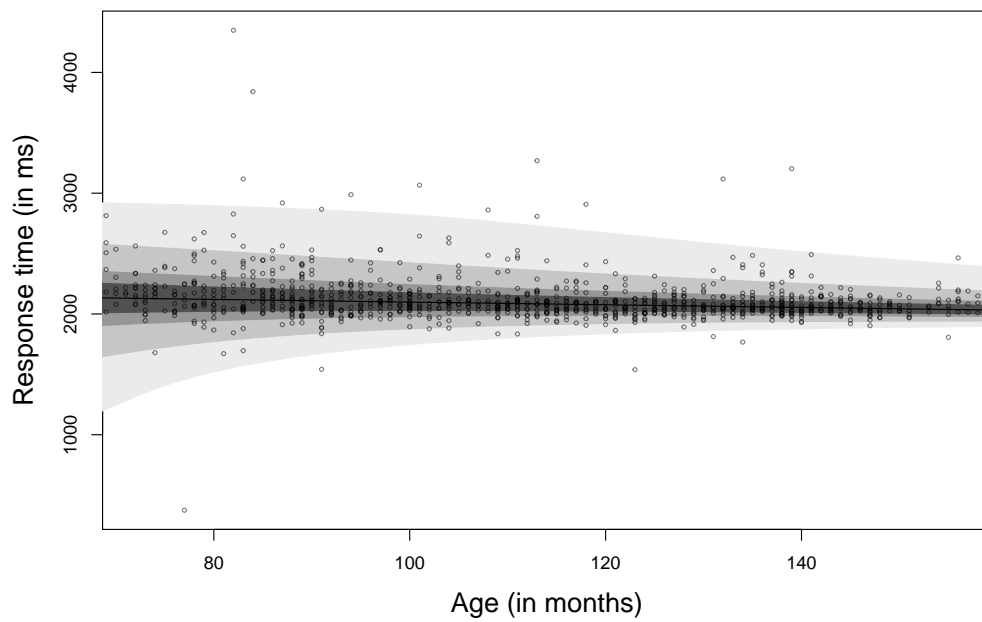
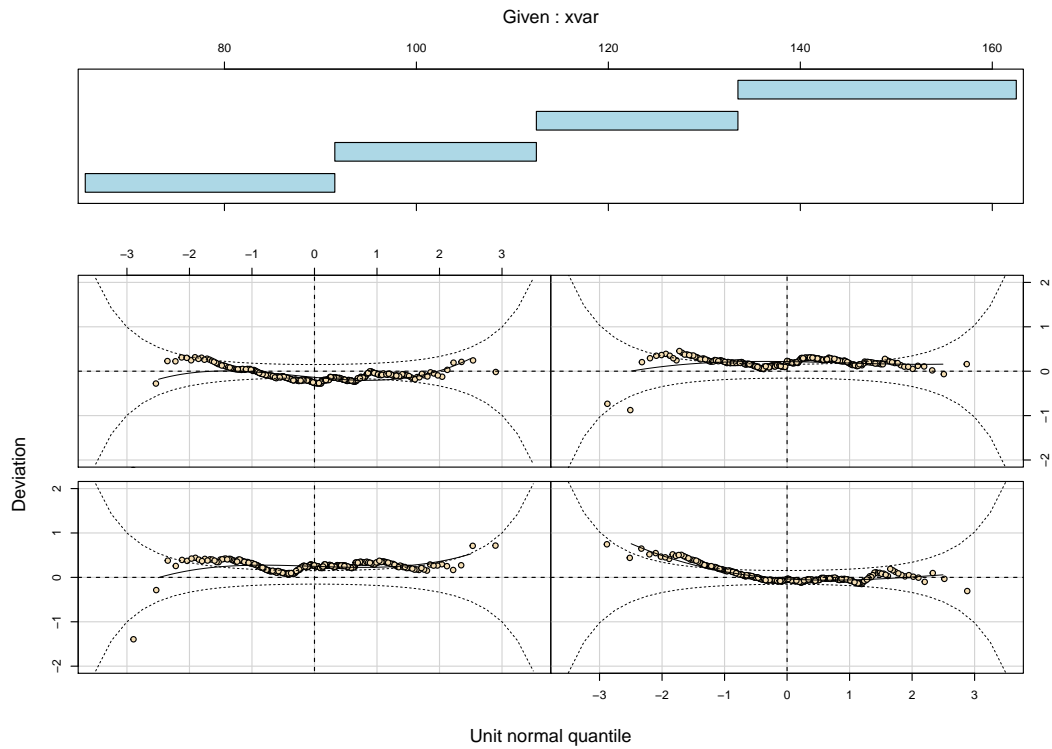
Blok 7 minus Blok 2



Blok 4 incompatible trials minus compatible trials

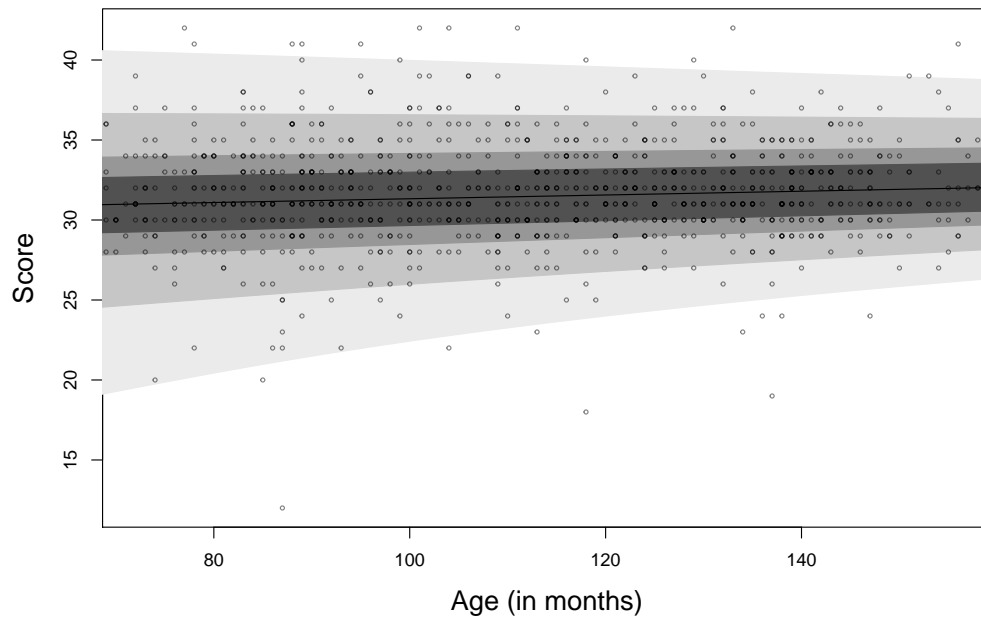
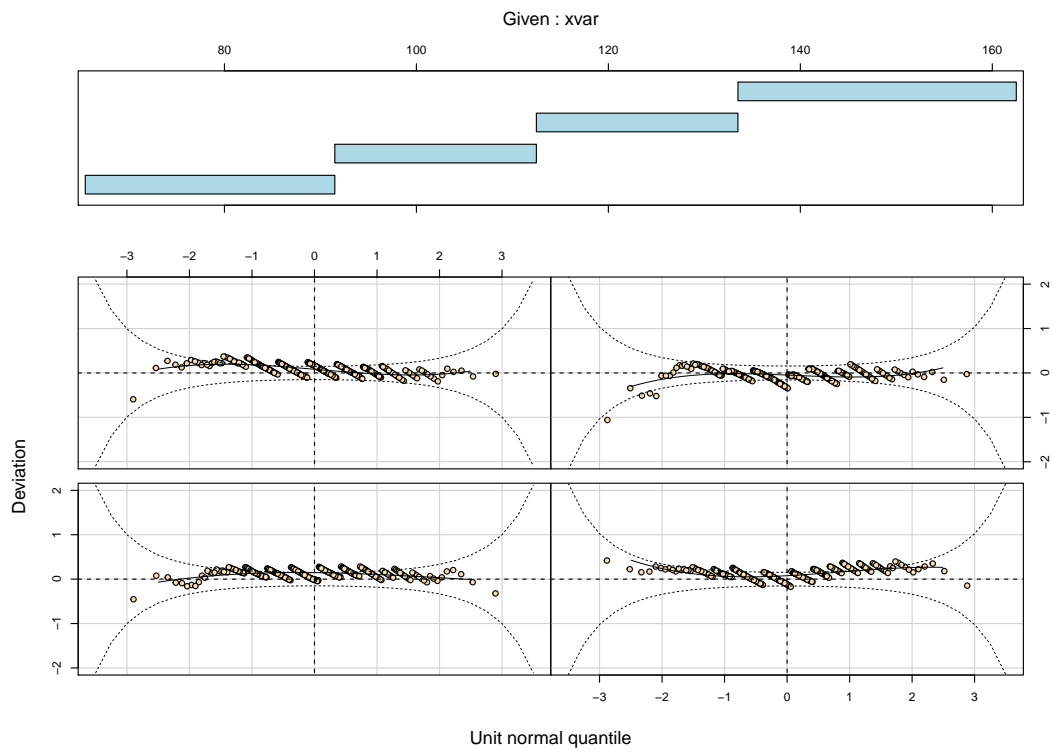


Blok 7 deel 2 minus Blok 7 deel 1

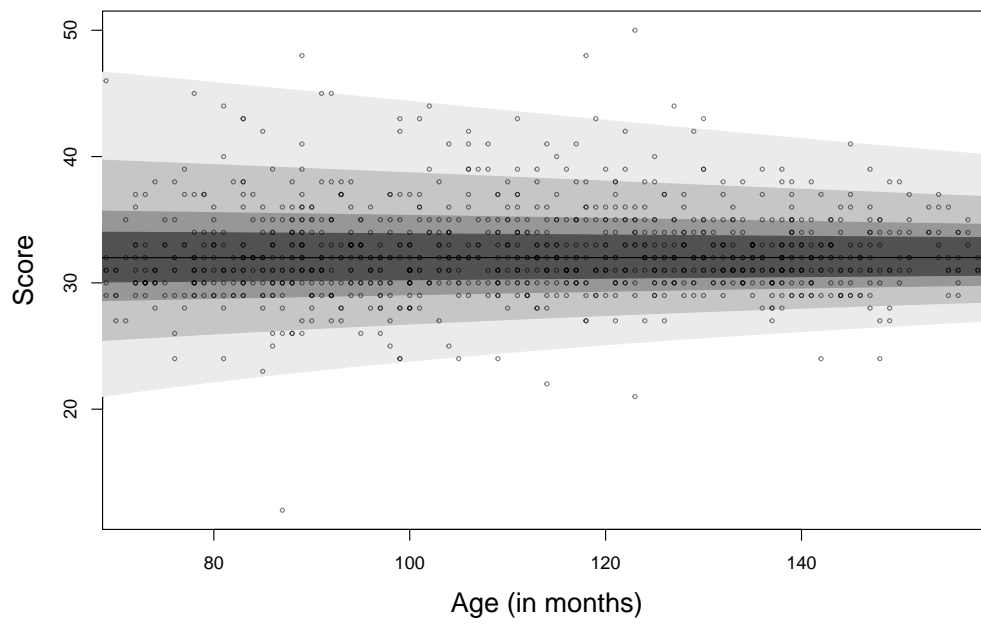
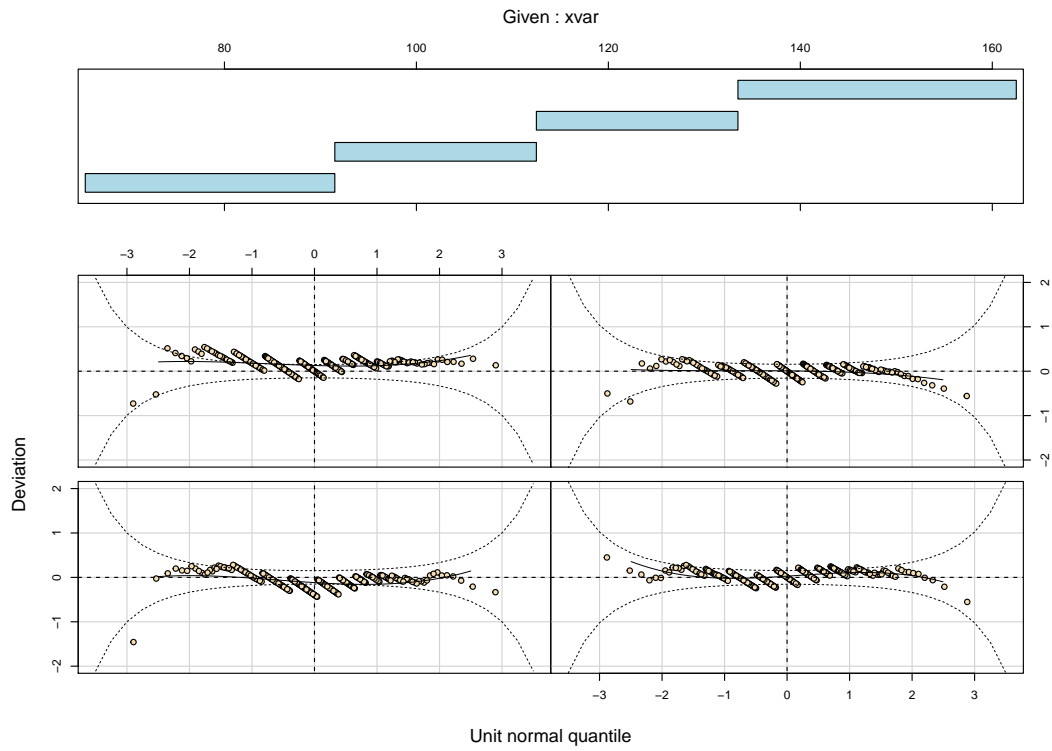


CONTRASTPARAMETERS FOUTEN

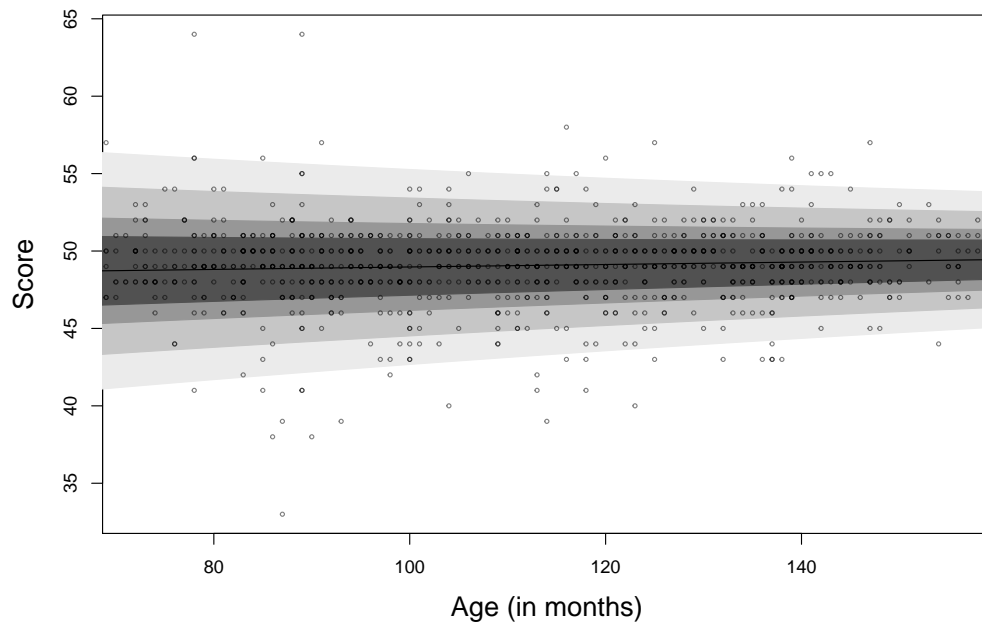
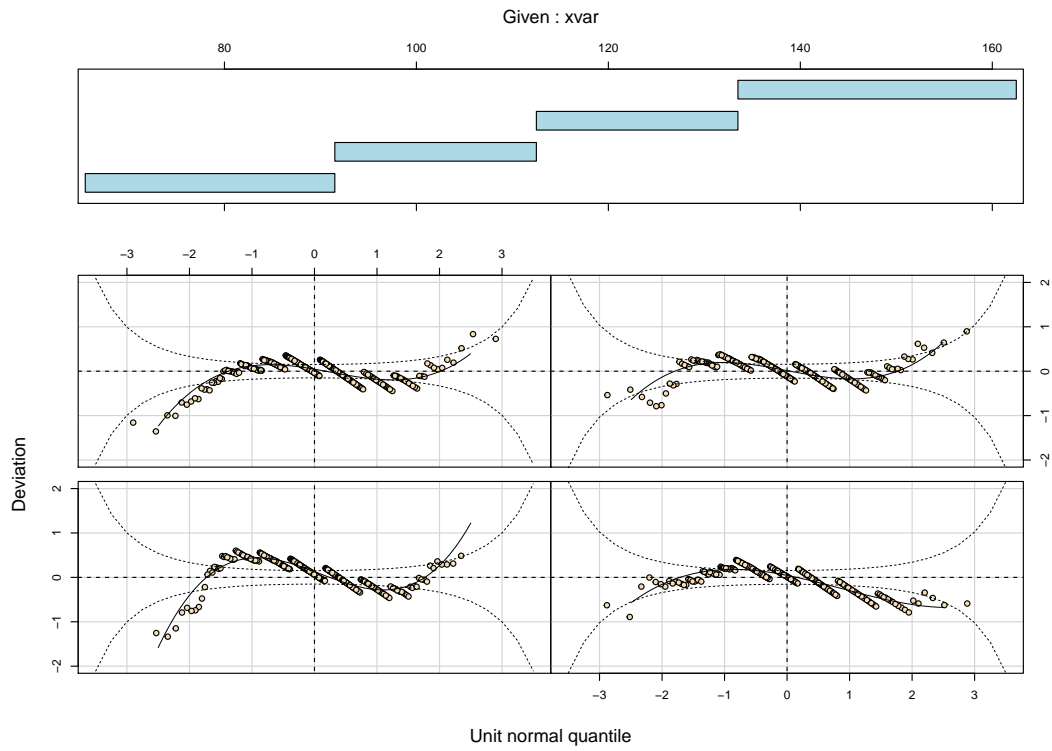
Blok 3 minus Blok 2



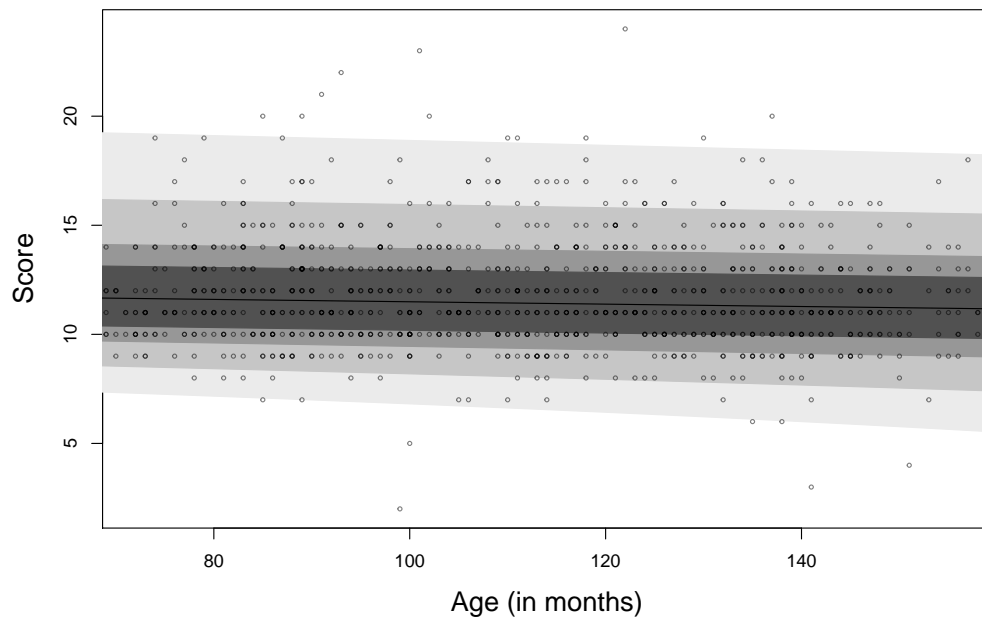
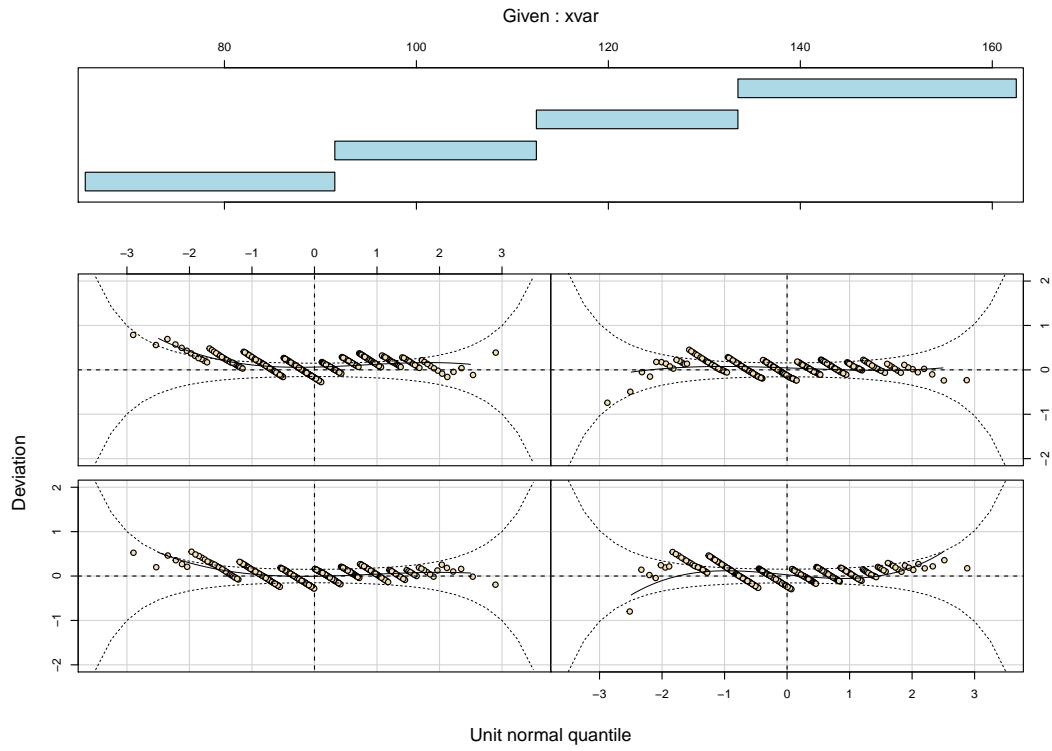
Blok 4 minus Blok 2



Blok 7 minus Blok 2

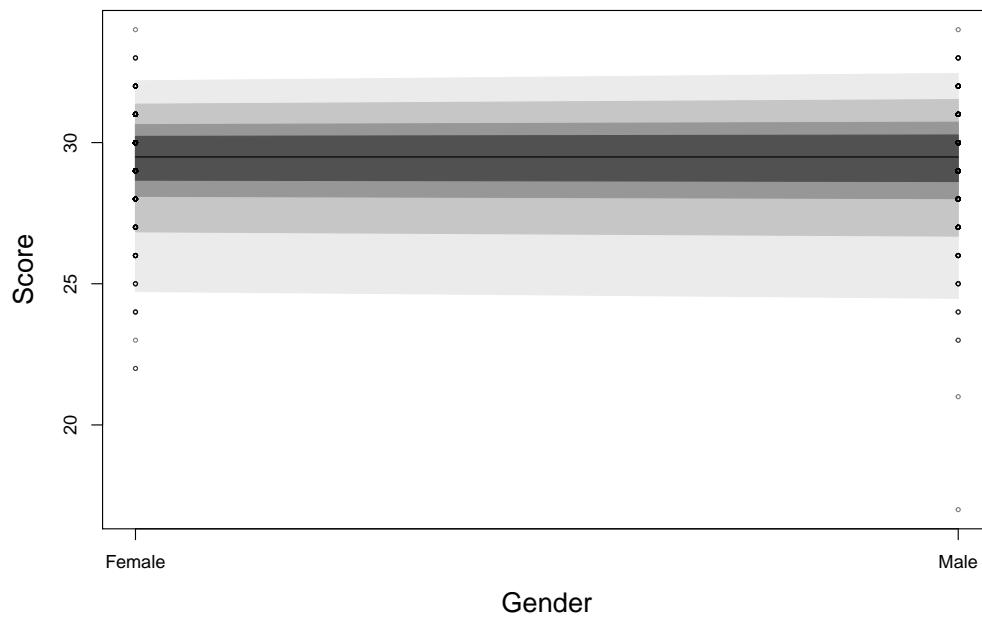
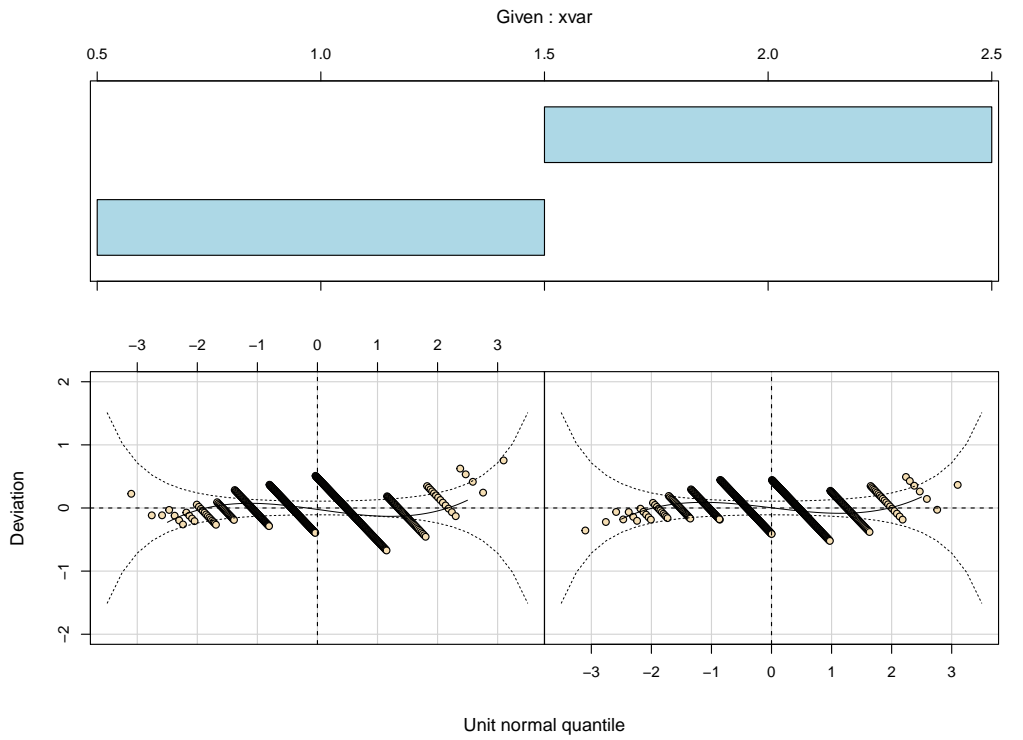


Blok 4 incompatible trials minus compatible trials



Blok 7 deel 2 minus Blok 7 deel 1

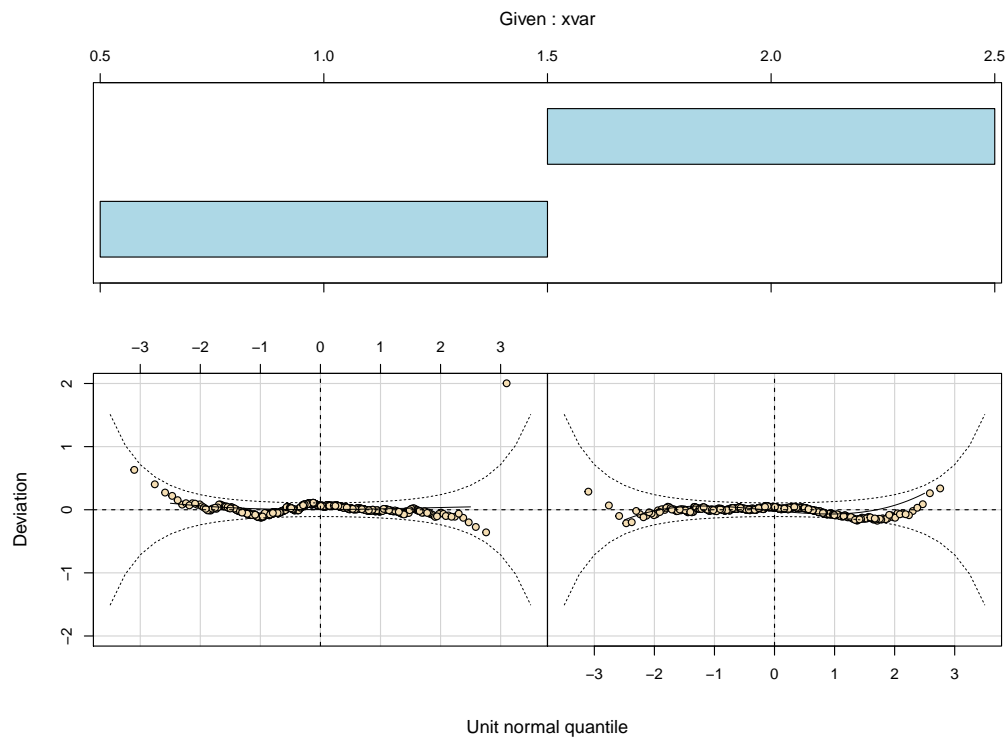
De scores op deze variabele, en daarmee de normen, zijn afhankelijk van sekse, en niet van leeftijd. De linker wormplot is voor de meisjes, de rechter voor de jongens.



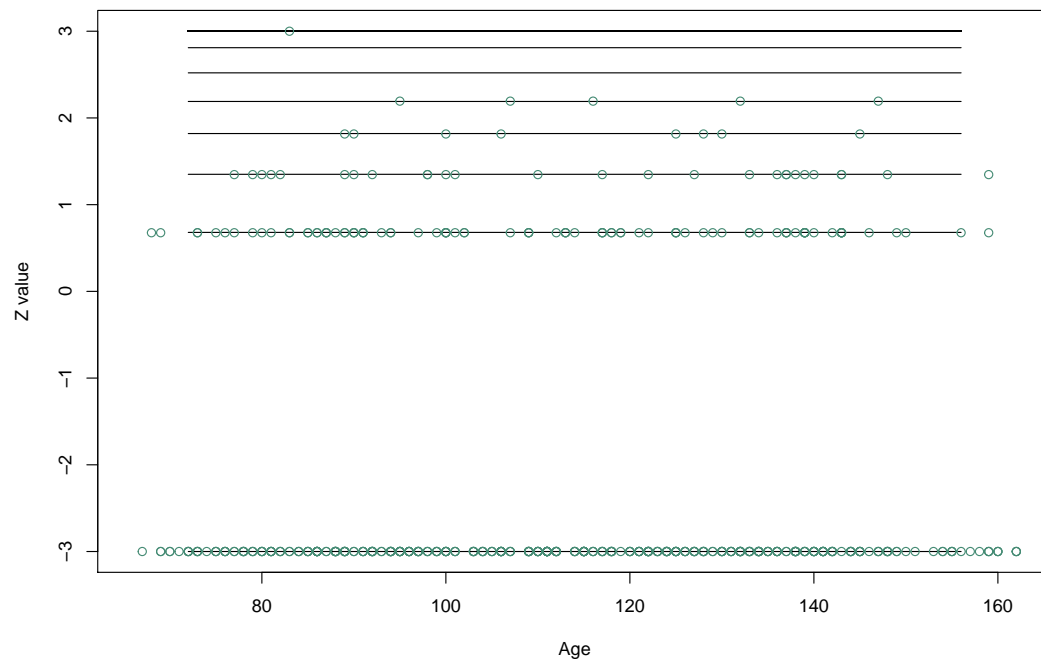
BLOKOVERKOEPELENDE AFWIJKENDE RESPONSEN

Aantal premature responsen

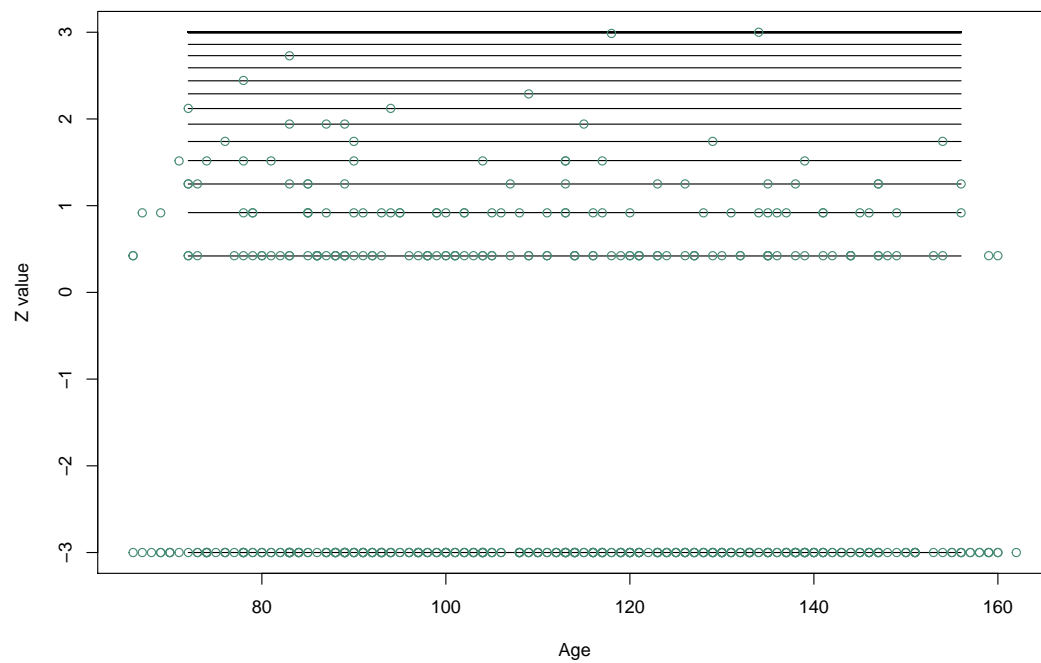
De scores op deze variabele, en daarmee de normen, zijn afhankelijk van sekse, en niet van leeftijd. De linker wormplot is voor de meisjes, de rechter voor de jongens.



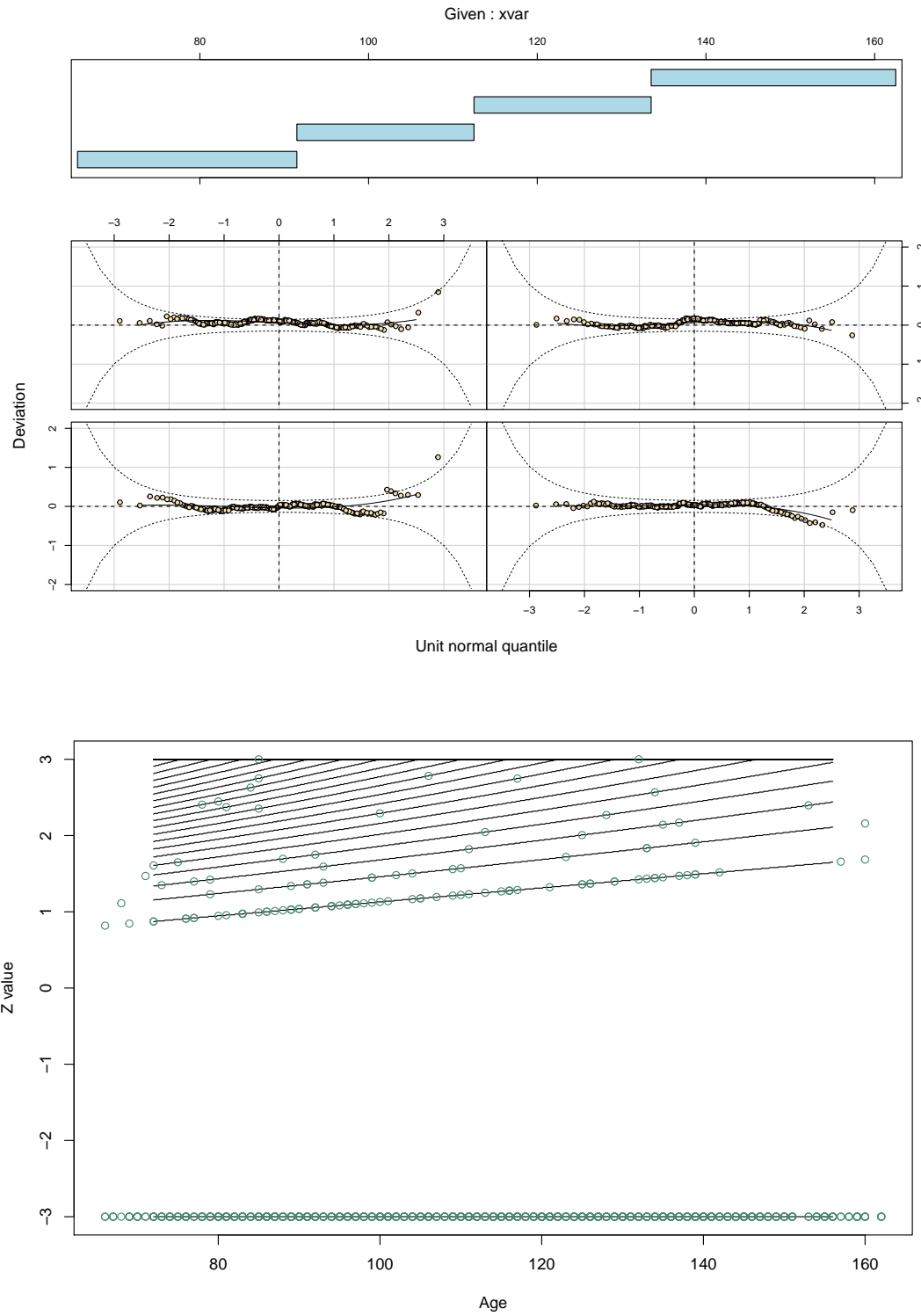
z-curveplot van Aantal premature responsen voor de meisjes



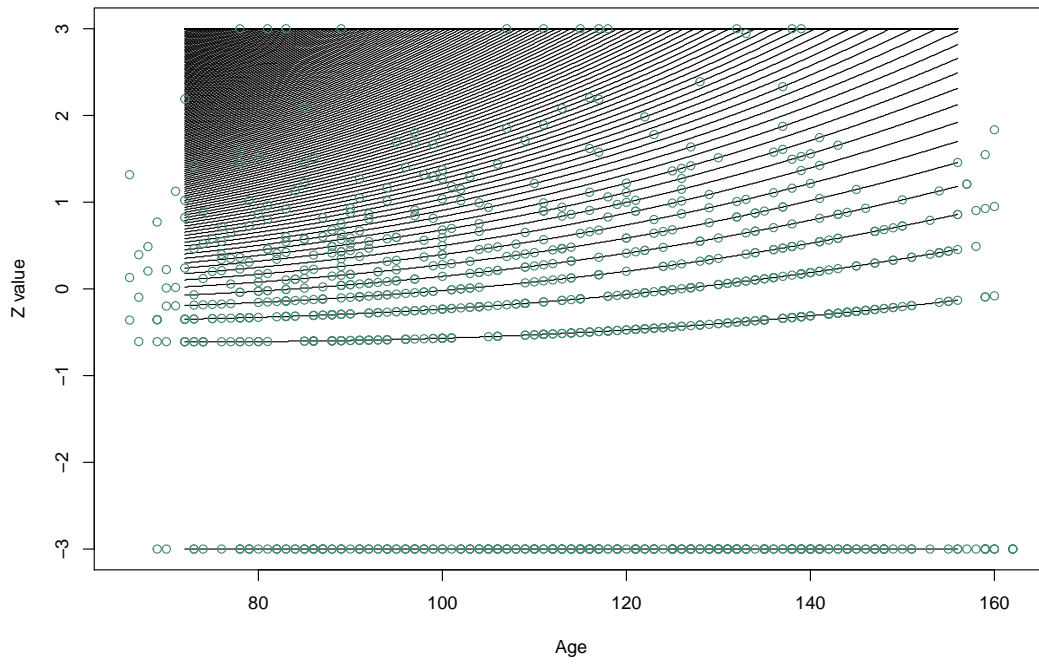
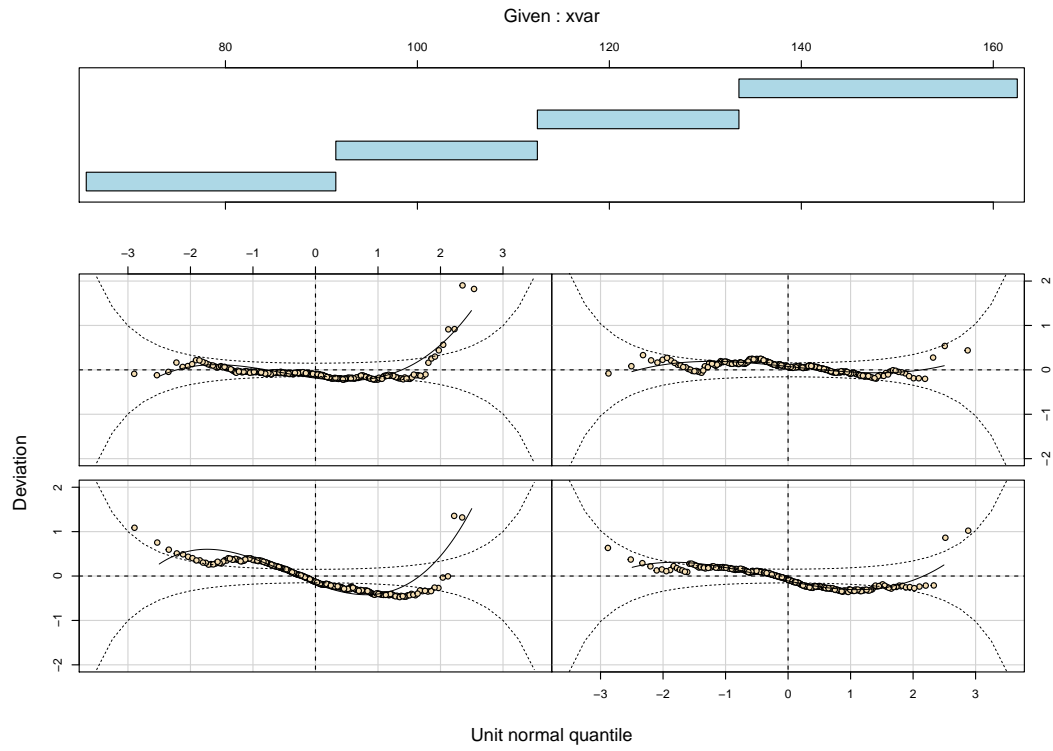
z-curveplot van Aantal premature responsen voor de jongens



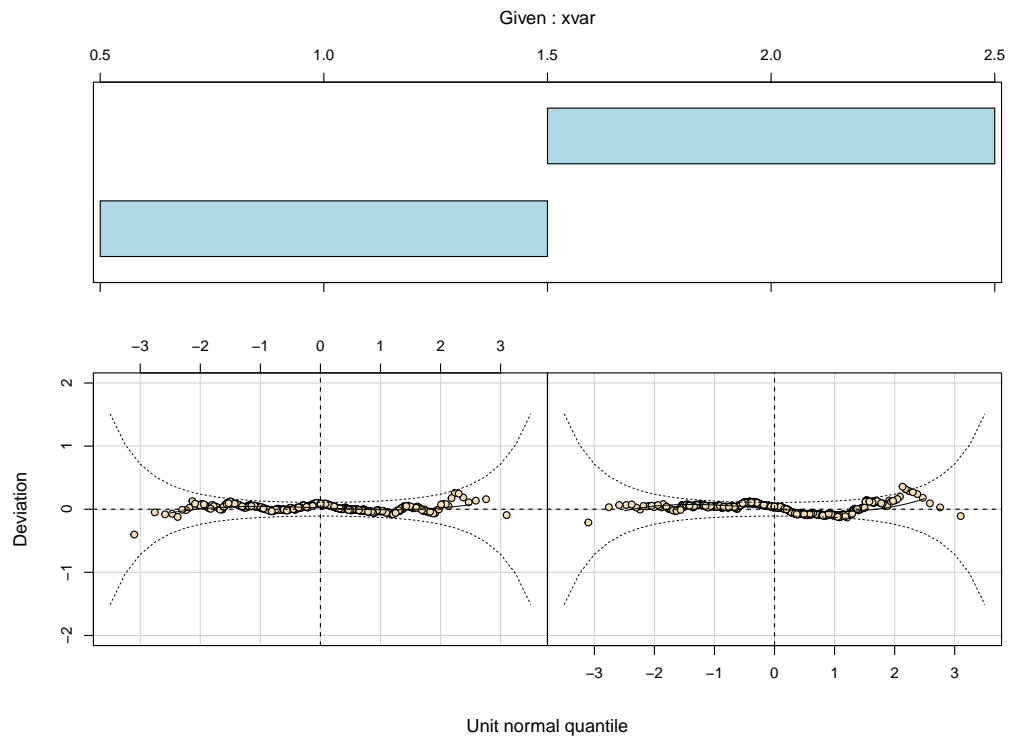
Aantal irrelevante knopdrukken



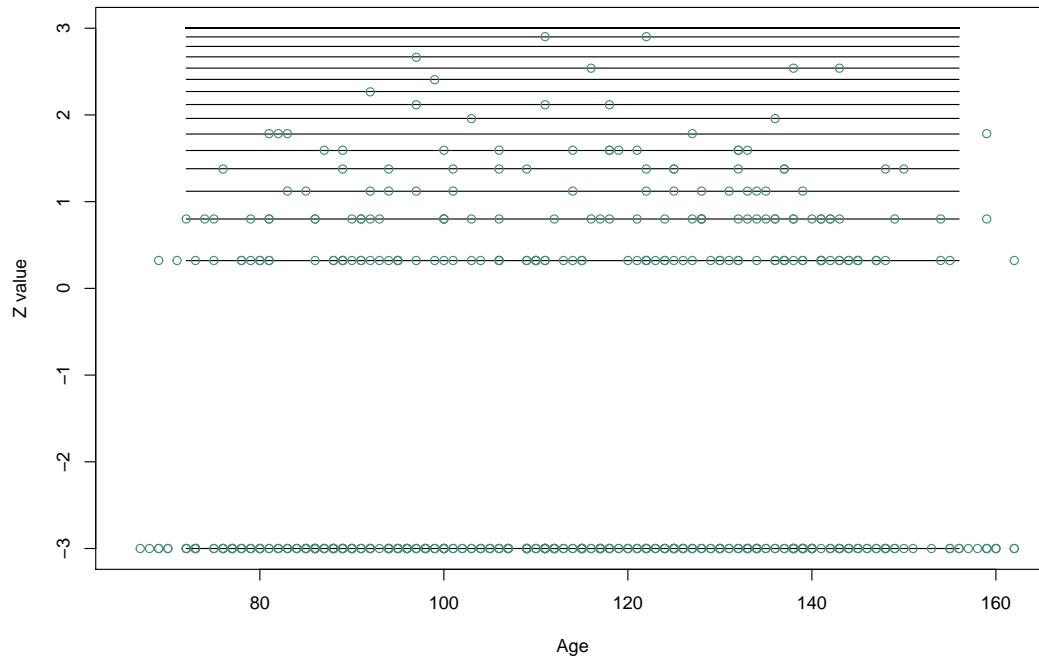
Aantal extra responsen



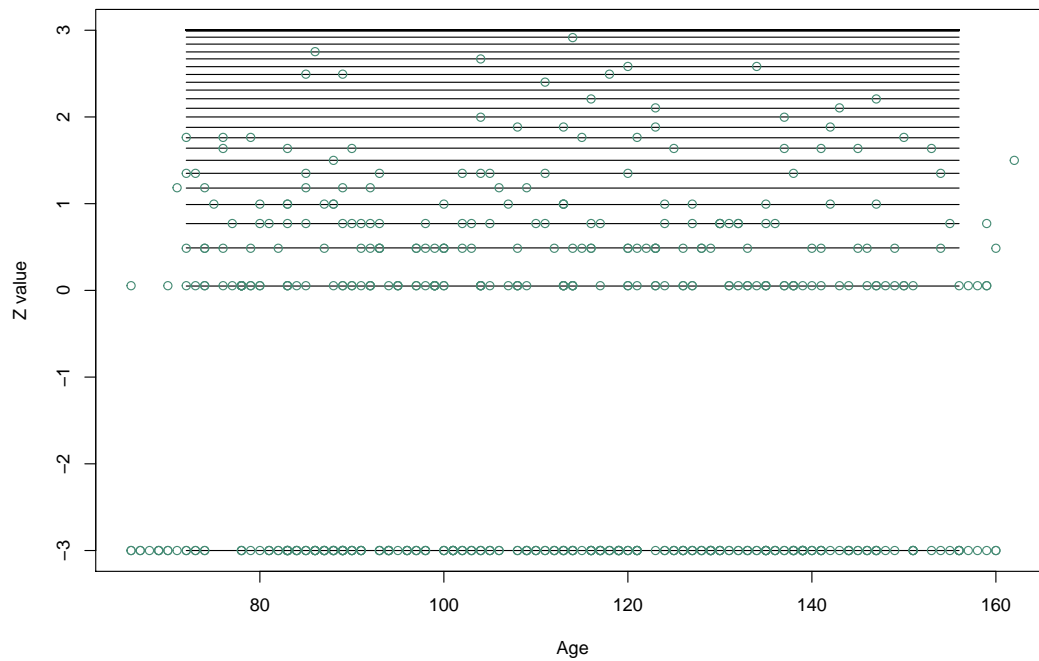
Aantal responscompetities



z-curveplot van Aantal responscompetities voor de meisjes



z-curveplot van Aantal responscompetities voor de jongens



1.3 Representativiteitsgegevens per leeftijdsgroep

In tabel 3 is de verdeling van migratieachtergrond weergegeven over drie leeftijdsgroepen: 66 t/m 95 maanden (5 t/m 7 jaar) (n=316), 96 t/m 119 maanden (8 t/m 9 jaar) (n=293) en 120 t/m 162 maanden (10 jaar t/m 13 jaar) (n=423). Hierin is te zien dat de percentages in deze leeftijdsgroepen variëren tussen de 23.9 en 30.1. Deze percentages liggen dichtbij de 25.5, het percentage dat representatief is voor Nederlandse kinderen tussen de 6 en 12 jaar. Toetsing van de samenhang tussen etniciteit en de leeftijdsgroepen laat zien dat er geen samenhang is tussen de variabelen ($X^2(2) = 3.59$, $p = 0.17$).

Tabel 3: Kruistabel leeftijdsgroepen en migratieachtergrond van de COTAPP normgroep.

Leeftijd		Heeft het kind een migratie achtergrond?		
		nee	ja	totaal
5-7 jaar	Aantal	236	80	316
	Percentage	74.7%	25.3%	100.0%
8-9 jaar	Aantal	204	88	292
	Percentage	69.9%	30.1%	100.0%
10-13 jaar	Aantal	321	101	422
	Percentage	76.1%	23.9%	100.0%
Totaal	Aantal	761	269	1030
	Percentage	73.9%	26.1%	100.0%

Noot: Van 2 kinderen ontbreekt informatie over de migratieachtergrond.

In tabellen 4 en 5 is de verdeling van opleidingsniveau van ouders weergegeven over de eerder genoemde drie leeftijdsgroepen. Er blijkt geen significante samenhang tussen leeftijdsgroep en opleidingsniveau van moeder ($X(2) = 5.15$, $p=.53$). Er is een kleine samenhang tussen leeftijdsgroep en opleidingsniveau van vader ($X(2) = 12.9$, $p=.05$), waarbij de vaders van de oudste groep kinderen iets minder vaak hoog opgeleid zijn, maar vaker middelbaar. Dit lijkt echter het gevolg van de -enkel voor subgroep analyses gehanteerde- leeftijdsgroepen. Middels regressie-analyses blijkt opleidingsniveau van vaders niet lineair of non-lineair (gecentreerde leeftijd in het kwadraat) samen te hangen met leeftijd van de kinderen: lineair $Beta=-0.01$, $p=.65$; non-lineair $Beta=-.04$, $p=.40$. Samenvattend geven deze resultaten aan dat er geen samenhang is tussen leeftijd van het kind en opleidingsniveau van moeders. Er is sprake van een kleine samenhang tussen leeftijd van het kind en opleidingsniveau van vaders wanneer er middels leeftijdsgroepen gerekend wordt, maar dit effect wordt niet teruggevonden wanneer leeftijd als continue variabele geanalyseerd wordt. Het opleidingsniveau van ouders is daarmee grotendeels gelijkmatig verdeeld over de leeftijdsrange van de normgroep.

Middels variantie analyse met regio als onafhankelijke variabele en geschatte intelligentie als covariaat en de genormeerde COTAPP parameters als afhankelijke variabelen, is nagegaan of er een samenhang is tussen regio en COTAPP scores die niet verklaard kan worden door eventuele regionale verschillen in intelligentieniveau, leeftijd en/of geslacht van de onderzochte kinderen. Voor 4/21 Basis rapportage parameters wordt een klein effect van regio gevonden, zie

Tabel 4: Kruistabel leeftijdsgroepen en opleidingsniveau moeder van de COTAPP normgroep.

		opleidingsniveau moeder				totaal	
		laag	middelbaar	hoog	onbekend		
Leeftijdsgroepen	5-7 jaar	Aantal	61	102	139	14	316
		Percentage	19.3%	32.3%	44.0%	4.4%	100.0%
	8-9 jaar	Aantal	56	90	140	7	293
		Percentage	19.1%	30.7%	47.8%	2.4%	100.0%
	10-13 jaar	Aantal	82	142	177	22	423
		Percentage	19.4%	33.6%	41.8%	5.2%	100.0%
Totaal	Aantal	199	334	456	43	1032	
	Percentage	19.3%	32.4%	44.2%	4.2%	100.0%	
Populatiepercentage		Vrouwen	16.5%	40.5%	41.3%	1.7%	100.0%

Tabel 5: Kruistabel leeftijdsgroepen en opleidingsniveau vader van de COTAPP normgroep.

		opleidingsniveau vader				totaal	
		laag	middelbaar	hoog	onbekend		
Leeftijds-groepen	5-7 jaar	Aantal	33	99	173	11	316
		Percentage	10.4%	31.3%	54.7%	3.5%	100.0%
	8-9 jaar	Aantal	48	86	154	5	293
		Percentage	16.4%	29.4%	52.6%	1.7%	100.0%
	10-13 jaar	Aantal	46	164	194	19	423
		Percentage	10.9%	38.8%	45.9%	4.5%	100.0%
Totaal	Aantal	127	349	521	35	1032	
	Percentage	12.3%	33.8%	50.5%	3.4%	100.0%	
Populatiepercentage		Mannen	18.5%	41.9%	37.9%	1.7%	100.0%

Tabel 6. Post-hoc paarsgewijze vergelijkingen tussen kinderen uit de regio Zuid versus elk van de overige regio's, wijst uit dat bij 1 van deze 4 parameters (Wachttijdafweging) de groep kinderen uit regio Zuid een betere score behaalt dan kinderen uit de regio West ($p=.05$). Ook blijkt de groep gemiddeld een iets snellere reactietijd te hebben dan kinderen uit regio Oost ($p=.05$). Gezien het grote aantal toetsen dat is uitgevoerd, lijken dit weinig betekenisvolle verschillen. Het effect van de ondervertegenwoordiging van kinderen uit regio Zuid op de representativiteit van de normen lijkt hiermee minimaal.

In tabellen 7 en 8 is de verdeling van regio en van stedelijkheid weergegeven over drie leeftijdsgroepen. Er is een scheve verdeling van regio over de leeftijdsgroepen ($X(2) = 27.5$, $p<.001$), maar niet van stedelijkheid ($X(2) = 2.04$, $p=.36$). 5-7 jarigen zijn relatief vaker afkomstig uit regio Noord en Oost en relatief minder vaak uit regio West. 8-9 jarigen zijn relatief minder vaak afkomstig uit regio Oost en relatief vaak uit regio West. De relatie tussen regio en COTAPP scores blijkt klein en niet systematisch wanneer rekening gehouden wordt met leeftijd, geslacht

en IQ (Tabel 6). Op basis hiervan -en het afwezige verband tussen leeftijd en stedelijkheid- lijkt het niet aannemelijk dat de ondervertegenwoordiging van 5-7 jarigen en oververtegenwoordiging van 8-9 jarigen vanuit regio West van invloed is op de representativiteit van de normen.

Tabel 6: Samenhang regio met COTAPP scores Basis rapportage gecorrigeerd voor leeftijd, geslacht en IQ.

	F; p	Zuid vs Noord ΔM; p	Zuid vs West ΔM; p	Zuid vs Oost ΔM; p
Verwerkingssnelheid				
Reactiesnelheid	1.39; .25	.31; .10	.34; .06	.37; .05
Variabiliteit in reactiesnelheid	1.02; .38	-.10; .59	.04; .81	.05; .77
Reactiesnelheid t.g.v. cognitieve belasting	3.17; .02	.32; .11	.04; .84	.17; .39
Versnellingscapaciteit in reactiesnelheid	0.59; .62	.17; .41	.16; .42	.23; .24
Variabiliteit in beslistijd	1.21; .30	.08; .69	.23; .22	.17; .39
Aandacht				
Volgehouden aandacht	2.45; .06	-.05; .80	.19; .30	.15; .45
Arousalregulatie	0.25; .86	.08; .68	.12; .50	.13; .48
Onoplettendheid	0.60; .62	-.03; .88	.06; .76	.11; .59
Executieve controle				
Slordigheid	0.95; .42	-.01; .97	.14; .50	.03; .89
Interferentiegevoeligheid	3.22; .02	-.09; .61	-.01; .97	.18; .29
Wachttijdafweging	2.89; .04	.21; .42	.49; .04	.45; .07
(Werk)geheugen				
Reactiesnelheid eenvoudig	0.09; .97	.07; .72	.09; .62	.08; .66
Fouten eenvoudig	1.21; .30	-.40; .23	-.13; .67	-.12; .71
Reactiesnelheid complex	0.79; .50	.27; .17	.26; .16	.30; .13
Fouten complex	0.87; .46	-.16; .55	-.03; .92	-.19; .47
Reactiesnelheid zeer complex	0.64; .59	-.06; .78	.08; .70	.02; .92
Fouten zeer complex	2.44; .06	-.24; .28	.04; .83	-.04; .86
Leersnelheid				
Leersnelheid	1.72; .16	.19; .58	.50; .13	.46; .18
Gedrag tijdens de taak				
Benodigde ondersteuning	6.20; <.01	-.24; .43	.36; .20	.13; .67
Verbaliseren	2.78; .04	-.49; .21	.03; .94	-.03; .94
Beweeglijkheid	2.20; .09	-.24; .28	.01; .98	.03; .90

Tabel 7: Kruistabel leeftijdsgroepen en regio van de COTAPP normgroep.

			Regio				Totaal
			Noord	West	Oost	Zuid	
Leeftijdsgroepen	5-7 jaar	Aantal	74	112	113	17	316
		Percentage	23.4%	35.4%	35.8%	5.4%	100.0%
	8-9 jaar	Aantal	48	164	69	12	293
		Percentage	16.4%	56.0%	23.5%	4.1%	100.0%
	10-13 jaar	Aantal	73	205	127	18	423
		Percentage	17.3%	48.5%	30.0%	4.3%	100.0%
Totaal	Aantal	195	481	309	47	1032	
	Percentage	18.9%	46.6%	29.9%	4.6%	100.0%	
Populatiepercentage			10.1%	47.5%	21.1%	21.2%	100.0%

Tabel 8: Kruistabel leeftijdsgroepen en stedelijkheid van de COTAPP normgroep.

			Stedelijkheid			Totaal
			(Zeer) sterk	matig	weinig tot niet	
Leeftijdsgroepen	5-7 jaar	Aantal	140	52	94	286
		Percentage	49.0%	18.2%	32.9%	100.0%
	8-9 jaar	Aantal	138	70	73	281
		Percentage	49.1%	24.9%	26.0%	100.0%
	10-13 jaar	Aantal	161	72	131	364
		Percentage	44.2%	19.8%	36.0%	100.0%
Totaal	Aantal	439	194	298	931	
	Percentage	47.2%	20.8%	32.0%	100.0%	
Populatiepercentage			48.0%	17.3%	34.7%	100.0%

1.4 Gegevens over verschillen tussen subgroepen

In deze sectie worden de gegevens over de verschillen tussen relevante subgroepen op basis van geslacht, migratieachtergrond en intelligentie (IQ) weergegeven. Op theoretische gronden verwachtten we een verschil in scoreverdeling op veel van de uitkomstparameters op basis van IQ (zie bijvoorbeeld Polderman et al., 2006; Rommelse et al., 2008), een mogelijk verschil op basis van migratieachtergrond (Neisser et al., 1996) en geen verschil op basis van geslacht. Deze laatste verwachting was er omdat voor tests met substantiële verschillen in scoreverdelingen tussen de geslachten, er aparte normen berekend zijn per geslacht.

Voor de dichotome variabelen geslacht (jongen; meisje) en migratieachtergrond (geen migratieachtergrond; wel migratieachtergrond) geven we per testonderdeel de gemiddelde score, SD en steekproefomvang per subgroep, en de resultaten van de bijbehorende *t*-procedure weer. De effectgrootte wordt uitgedrukt middels Cohen's *d* (indicatie van interpretatie: .2 = klein; .5 = medium; .8 = groot effect (Cohen, 1998)). Voor de variabele Totaal IQ groep (TIQ) (ondergemiddeld TIQ < 85, gemiddeld TIQ 85-115, bovengemiddeld TIQ > 115) geven we per testonderdeel de gemiddelde score per subgroep weer (de overige samenvattingsmaten laten we hier achterwege i.v.m. de overzichtelijkheid) en de resultaten van een univariate variantie-analyse (ANOVA). De effectgrootte wordt uitgedrukt middels ω^2 (indicatie van interpretatie: .01 = klein; .06 = medium; .14 = groot effect (Cohen, 1998)). Per gebruikte achtergrondvariabele toetsen we met een overall = 0.05, en corrigeren we voor multiple hypothese toetsing met Bonferroni correctie; we vergelijken de *p*-waarde met .05/21 (Basis rapportage) en .05/46 (Specialistische rapportage). Het verwachte aantal tests per achtergrondvariabele met een significant verschil is $0.05 \cdot 21$ (= 1.05) voor de Basisrapportage en $0.05 \cdot 46$ (= 2.3) voor de Specialistische rapportage.

1.4.1 Subgroepen naar migratieachtergrond

In Tabel 9 zijn per testonderdeel en per migratieachtergrond de gemiddelde score, SD en steekproefomvang, en de uitkomst van de *t*-procedure weergegeven. Er zijn 3 (van de 21) uitkomstparameters in de Basisrapportage waarop een significant verschil in gemiddelden tussen kinderen met en zonder een migratieachtergrond is gevonden. Dit is iets hoger dan verwacht of basis van kansniveau (1). Op de Specialistische rapportage is dit het geval voor 4 uitkomstparameters (van de 46), wederom iets boven kansniveau (2 a 3). De bijbehorende effectgroottes zijn echter klein tot (laag) medium (met een maximale waarde van Cohen's *d* van .35). Op deze testonderdelen presteren kinderen met een migratieachtergrond gemiddeld iets minder goed dan kinderen zonder een migratieachtergrond. Dit is in lijn met de verwachting gebaseerd op onderzoeken naar verschillen in cognitieve en didactische prestaties van kinderen met en zonder een migratieachtergrond, echter zijn er maar verschillen gevonden op enkele parameters en zijn de verschillen veelal klein. Daarnaast valt hieruit geen causaliteit af te leiden: een migratieachtergrond is met veel factoren geassocieerd (lagere SES, kwalitatief minder goed onderwijs, wonen in een achterstandswijk), welke gezamenlijk kunnen leiden tot minder gunstige omgevingsstimulatie van de cognitieve ontwikkeling. Om stigmatisering te voorkomen, zijn deze subgroepsanalyses bewust niet in de handleiding opgenomen.

Tabel 9: Subgroepen naar migratieachtergrond

	Mean		SD		Valid N		t-test			
	wel	niet	wel	niet	wel	niet	<i>t</i> -value	df	p-value	Cohen's <i>d</i>
BASISRAPPOROTAGE										
Verwerkingssnelheid										
Reactiesnelheid	0.05	-0.01	1.04	0.98	269	761	0.83	448.22	0.41	0.06
Variabiliteit in reactiesnelheid	0.23	-0.08	0.99	0.99	269	761	4.41	471.17	0.00*	0.31
Reactiesnelheid ten gevolge van cognitieve belasting	-0.07	0.10	1.10	0.95	269	761	-2.24	417.96	0.03	0.17
Versnellingscapaciteit in reactiesnelheid	0.02	0.03	1.03	0.97	269	761	-0.13	449.61	0.89	0.01
Variabiliteit in beslistijd	0.25	-0.09	0.99	0.99	269	761	4.90	470.85	0.00*	0.35
Aandacht										
Volgehouden aandacht	0.16	0.01	1.04	0.96	269	761	2.19	437.46	0.03	0.16
Arousalregulatie	0.16	0.07	1.03	0.95	269	761	1.25	441.19	0.21	0.09
Onoplettendheid	0.05	-0.12	1.04	1.00	269	761	2.28	455.15	0.02	0.16

Executieve controle										
Slordigheid	-1.32	-1.68	1.24	1.09	269	761	4.20	422.91	0.00*	0.32
Interferentiegevoeligheid	0.00	0.01	1.05	0.99	269	761	-0.14	446.83	0.89	0.01
Wachttijdafweging	0.13	0.03	1.32	1.22	269	761	1.14	439.76	0.26	0.08
(Werk)geheugen										
Reactiesnelheid eenvoudig	0.07	0.00	1.00	1.01	269	761	0.96	472.56	0.34	0.07
Fouten eenvoudig	-0.95	-0.93	1.73	1.70	269	761	-0.17	463.72	0.87	0.01
Reactiesnelheid complex	-0.04	0.02	1.03	0.98	247	713	-0.82	410.96	0.41	0.06
Fouten complex	-0.48	-0.59	1.49	1.46	247	713	0.97	419.16	0.33	0.07
Reactiesnelheid zeer complex	0.00	0.01	1.09	0.96	187	564	-0.09	286.74	0.93	0.01
Fouten zeer complex	-0.09	-0.36	1.08	1.13	187	564	2.94	331.96	0.00	0.24
Leersnelheid										
Leersnelheid	-1.18	-1.58	1.74	1.67	269	761	3.29	455.30	0.00	0.24
Gedrag tijdens de taak										
Benodigde ondersteuning	-1.92	-2.30	1.85	1.57	268	761	3.02	410.56	0.00	0.23
Verbaliseren	-1.06	-1.31	1.89	1.87	268	760	1.89	463.69	0.06	0.14
Beweeglijkheid	-0.01	-0.25	1.04	1.06	268	761	3.25	476.16	0.00	0.23

SPECIALISTISCHE RAPPORTAGE

Mediane reactiesnelheid

Blok 1	0.08	0.00	1.14	0.97	269	761	0.97	413.30	0.33	0.07
Blok 2	0.03	0.03	1.06	0.97	269	761	-0.00	437.41	1.00	0.00
Blok 3	0.00	0.02	1.03	0.98	269	761	-0.30	453.05	0.77	0.02
Blok 4 compatibele trials	0.03	0.00	1.03	1.00	269	761	0.34	457.25	0.74	0.02
Blok 4 incompatibele trials	-0.01	0.02	1.04	0.98	269	761	-0.44	448.26	0.66	0.03
Blok 5 1-back	0.07	0.00	1.00	1.01	269	761	0.96	472.56	0.34	0.07
Blok 5 2-back	-0.04	0.02	1.03	0.98	247	713	-0.82	410.96	0.41	0.06
Blok 5 2-back PI	0.00	0.01	1.09	0.96	187	564	-0.09	286.74	0.93	0.01

Blok 6	-0.05	0.02	1.06	0.97	269	761	-0.96	437.70	0.34	0.07
Blok 7 deel 1	0.10	-0.04	1.03	0.99	269	761	1.94	452.89	0.05	0.14
Blok 7 deel 2	0.14	-0.05	1.00	0.98	269	761	2.68	462.94	0.01	0.19
Intra-individuele coefficient van variatie										
Blok 1	0.21	-0.07	1.06	0.96	269	761	3.84	435.68	0.00*	0.28
Blok 2	0.13	-0.04	1.03	0.98	269	761	2.42	452.99	0.02	0.17
Blok 3	0.11	-0.04	0.98	1.00	269	761	2.20	479.47	0.03	0.15
Blok 4 compatibele trials	0.08	-0.03	0.96	1.01	269	761	1.63	492.03	0.10	0.11
Blok 4 incompatibele trials	0.07	-0.03	0.95	1.01	269	761	1.50	499.51	0.14	0.10
Blok 5 1-back	-0.00	0.00	1.06	0.97	269	761	-0.10	437.76	0.92	0.01
Blok 5 2-back	-0.01	0.00	0.96	1.01	247	713	-0.15	445.58	0.88	0.01
Blok 5 2-back PI	0.10	-0.03	1.04	0.98	187	564	1.56	304.85	0.12	0.13
Blok 6	0.25	-0.09	0.99	0.99	269	761	4.90	470.85	0.00*	0.35
Blok 7 deel 1	0.16	-0.05	1.01	0.99	269	761	2.95	460.91	0.00	0.21
Blok 7 deel 2	0.25	-0.09	0.98	0.99	269	761	4.90	475.68	0.00*	0.35
Aantal fouten										
Blok 2	-0.57	-0.46	1.51	1.38	269	761	-1.13	435.99	0.26	0.08
Blok 3	-0.23	-0.26	1.13	1.17	269	761	0.49	485.22	0.62	0.03
Blok 4 compatibele trials	-0.97	-0.73	1.67	1.61	269	761	-2.05	455.40	0.04	0.15
Blok 4 incompatibele trials	-0.37	-0.32	1.31	1.24	269	761	-0.56	448.02	0.58	0.04
Blok 5 1-back	-0.95	-0.93	1.73	1.70	269	761	-0.17	463.72	0.87	0.01
Blok 5 2-back	-0.48	-0.59	1.49	1.46	247	713	0.97	419.16	0.33	0.07
Blok 5 2-back PI	-0.09	-0.36	1.08	1.13	187	564	2.94	331.96	0.00	0.24
Blok 7 deel 1	-0.90	-0.86	1.66	1.70	269	761	-0.35	478.63	0.73	0.02
Blok 7 deel 2	-1.49	-1.42	1.80	1.84	269	761	-0.58	479.50	0.57	0.04
Contrastparameters reactiesnelheid										
Blok 2 minus Blok 1	-0.07	0.10	1.10	0.95	269	761	-2.24	417.96	0.03	0.17
Blok 3 minus Blok 2	0.02	0.03	1.03	0.97	269	761	-0.13	449.61	0.89	0.01

	Blok 4 minus Blok 2	0.00	0.01	1.05	0.99	269	761	-0.14	446.83	0.89	0.01
	Blok 7 versus Blok 2	0.16	0.01	1.04	0.96	269	761	2.19	437.46	0.03	0.16
Blok 4	incompatibele trials minus compatibele trials	-0.07	0.03	1.07	0.97	269	761	-1.26	433.97	0.21	0.09
	Blok 7 deel 2 minus Blok 7 deel 1	0.16	0.07	1.03	0.95	269	761	1.25	441.19	0.21	0.09
	Contrastparameters fouten										
	Blok 3 minus Blok 2	0.10	0.06	0.96	1.00	269	761	0.65	489.07	0.52	0.04
	Blok 4 minus Blok 2	-0.01	0.04	0.99	1.00	269	761	-0.67	472.88	0.50	0.05
	Blok 7 minus Blok 2	-0.01	0.01	1.01	0.93	269	761	-0.22	439.07	0.82	0.02
Blok 4	incompatibele trials minus compatibele trials	0.09	0.05	1.00	0.97	269	761	0.62	461.29	0.53	0.04
	Blok 7 deel 2 minus Blok 7 deel 1	-0.02	-0.01	0.98	1.00	269	761	-0.22	477.18	0.83	0.02
	Blokoeverkoepelende afwijkende responsen										
	Aantal premature responsen	-1.60	-1.92	1.88	1.77	269	761	2.47	446.27	0.01	0.18
	Aantal irrelevante knopdrukken	-2.19	-2.60	1.76	1.27	269	761	3.53	371.48	0.00*	0.29
	Aantal extra responsen	-0.37	-0.71	1.54	1.56	269	761	3.16	475.31	0.00	0.22
	Aantal responscompetities	-1.26	-1.42	1.95	1.87	269	761	1.17	454.47	0.24	0.08

*p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

1.4.2 Subgroepen naar geslacht

In Tabel 10 zijn per testonderdeel en per geslacht de gemiddelde score, SD en steekproefomvang, en de uitkomst van de t -procedure weergegeven. Er zijn 4 (van de 21) uitkomstparameters in de Basisrapportage waarop een significant verschil is gevonden in gemiddelden tussen jongens en meisjes. Dit is iets hoger dan verwacht of basis van kansniveau (1). In de Specialistische rapportage is dit het geval voor 3 uitkomstparameters, wat net boven kansniveau ligt (2 a 3). De bijbehorende effectgroottes zijn echter klein (met een maximale waarde van Cohen's d van .29). Ter controle hebben we het normeringsmodel van het testonderdeel met de hoogste Cohen's d , (Werk)geheugen reactiesnelheid complex (Mediane reactiesnelheid Blok 5 2-back), nader bekeken. De scoreverdeling voor dit onderdeel is gemodelleerd met de BCPE verdeling, die afhankelijk is van 4 parameters (μ, σ, ν, τ). Om te bepalen in hoeverre geslacht en/of leeftijd gebruikt zouden moeten worden als predictor in het model (en dus de normen leeftijd- en/of geslachtsafhankelijk zouden moeten zijn), zijn er vijf modellen gefit, met de aangegeven effecten voor de 4 parameters van de BCPE verdeling. Deze modellen waren:

1. Model 0: Geen predictoren voor μ, σ, ν, τ
2. Model 1: Leeftijd als predictor voor μ, σ, ν, τ
3. Model 2: Geslacht als predictor voor μ, σ, ν, τ
4. Model 3: Leeftijd en geslacht (de hoofdeffecten) als predictoren voor μ, σ, ν, τ
5. Model 4: Leeftijd en geslacht, en het interactie-effect hiervan, als predictoren voor μ, σ, ν, τ .

Het best passende model van deze vijf is geselecteerd op basis van de (laagste) BIC en AIC waarden, met visuele controle op modelfit middels de wormplot en de z -score curves. In geval er een verschil was tussen het model aangegeven door de BIC en de AIC selecteerden we het model op basis van de visuele weergaven van de modelfit. Voor het testonderdeel (Werk)geheugen reactiesnelheid complex (Mediane reactiesnelheid Blok 5 2-back) had Model 1 (leeftijd als predictor) de laagste BIC en AIC waarden, en de visuele representaties lieten geen afwijkingen zien. Daarom is voor dit model gekozen. Nu er voor dit onderdeel een klein geslachtseffect in de gemiddelden lijkt te zijn, hebben we ter controle aanvullend een schatting gedaan van Model 5: Leeftijd als predictor voor μ, σ, ν, τ , en voor alleen de parameter μ geslacht als predictor en de interactie van leeftijd en geslacht. De BIC van dit model was weinig lager dan van het gekozen Model 1. Als dit model behoord had tot de reeks van te fitten modellen, dan was dit model dus naar voren gekomen. Inspectie van de centielcurves liet zien dat de verschillen in geschatte centielcurves tussen de jongens en meisjes weinig verschillen. Dit was ook te verwachten omdat anders wel Model 3 (leeftijd en geslacht als

predictor) door de BIC geïndiceerd zou zijn - immers, dan zou de modelfit van Model 3 ten opzichte van Model 1 zoveel beter zijn, dat de toename in aantal te schatten parameters 'gerechtvaardigd' zou zijn.

Tabel 10: Subgroepen naar geslacht

	Mean		SD		Valid N		t-test			
	jongen	meisje	jongen	meisje	jongen	meisje	t-value	df	p-value	Cohen's d
BASISRAPPORTAGE										
Verwerkingssnelheid										
Reactiesnelheid	-0.06	0.06	1.06	0.93	516	516	-2.05	1013.16	0.04	0.13
Variabiliteit in reactiesnelheid	0.10	-0.11	1.01	0.97	516	516	3.37	1028.73	0.00*	0.21
Reactiesnelheid tgv cognitieve belasting	0.04	0.06	1.02	0.98	516	516	-0.42	1028.53	0.68	0.03
Versnellingscapaciteit reactiesnelheid	0.03	0.02	0.99	0.99	516	516	0.15	1030.00	0.88	0.01
Variabiliteit in beslistijd	0.08	-0.08	1.04	0.95	516	516	2.69	1021.44	0.01	0.17
Aandacht										
Volgehouden aandacht	0.14	-0.05	0.98	0.97	516	516	3.12	1029.88	0.00	0.19
Arousalregulatie	0.17	0.02	0.97	0.97	516	516	2.39	1029.97	0.02	0.15
Onoplettendheid	-0.06	-0.09	1.00	1.01	516	516	0.40	1029.88	0.69	0.02
Executieve controle										
Slordigheid	-1.43	-1.73	1.16	1.11	516	516	4.29	1027.48	0.00*	0.27
Interferentiegevoeligheid	0.05	-0.04	1.05	0.95	516	516	1.51	1020.50	0.13	0.09
Wachttijdafweging	0.16	-0.04	1.26	1.22	516	516	2.55	1029.30	0.01	0.16
(Werk)geheugen										
Reactiesnelheid eenvoudig	-0.03	0.07	1.06	0.95	516	516	-1.68	1019.22	0.09	0.10
Fouten eenvoudig	-0.86	-1.00	1.74	1.68	516	516	1.38	1028.84	0.17	0.09
Reactiesnelheid complex	-0.15	0.14	1.00	0.97	471	471	-4.47	954.01	0.00*	0.29
Fouten complex	-0.61	-0.52	1.52	1.41	471	471	-0.95	945.81	0.34	0.06
Reactiesnelheid zeer complex	-0.09	0.10	0.99	0.99	372	372	-2.75	749.51	0.01	0.20
Fouten zeer complex	-0.30	-0.29	1.21	1.03	372	372	-0.15	726.83	0.88	0.01

Leersnelheid											
Leersnelheid	-1.31	-1.64	1.72	1.66	516	516	3.15	1028.87	0.00	0.20	
Gedrag tijdens de taak											
Benodigde ondersteuning	-2.10	-2.30	1.74	1.55	515	515	1.94	1015.03	0.05	0.12	
Verbaliseren	-1.03	-1.46	1.84	1.89	514	514	3.72	1027.42	0.00*	0.23	
Beweeglijkheid	-0.18	-0.19	1.04	1.08	515	515	0.24	1027.64	0.81	0.02	

SPECIALISTISCHE RAPPORTAGE

Mediane reactietijd

Blok 1	-0.09	0.13	1.05	0.98	516	516	-3.51	1025.44	0.00*	0.22
Blok 2	-0.06	0.12	1.04	0.94	516	516	-3.00	1019.02	0.00	0.19
Blok 3	-0.09	0.12	1.04	0.93	516	516	-3.45	1017.52	0.00*	0.21
Blok 4 compatibele trials	-0.05	0.06	1.06	0.95	516	516	-1.81	1017.76	0.07	0.11
Blok 4 incompatibele trials	-0.04	0.07	1.07	0.91	516	516	-1.84	1004.33	0.07	0.11
Blok 5 1-back	-0.03	0.07	1.06	0.95	516	516	-1.68	1019.22	0.09	0.10
Blok 5 2-back	-0.15	0.14	1.00	0.97	471	471	-4.47	954.01	0.00*	0.29
Blok 5 2-back PI	-0.09	0.10	0.99	0.99	372	372	-2.75	749.51	0.01	0.20
Blok 6	-0.03	0.03	1.04	0.95	516	516	-0.87	1022.86	0.38	0.05
Blok 7 deel 1	0.00	-0.01	1.07	0.93	516	516	0.12	1009.06	0.90	0.01
Blok 7 deel 2	0.04	-0.03	1.03	0.96	516	516	1.04	1025.00	0.30	0.06

Intra-individuele coëfficiënt van variatie

Blok 1	0.07	-0.07	0.98	1.01	516	516	2.28	1028.46	0.02	0.14
Blok 2	0.10	-0.09	1.03	0.96	516	516	3.11	1024.77	0.00	0.19
Blok 3	0.08	-0.09	1.02	0.98	516	516	2.74	1028.18	0.01	0.17
Blok 4 compatibele trials	0.05	-0.05	0.99	1.01	516	516	1.55	1029.64	0.12	0.10
Blok 4 incompatibele trials	0.05	-0.06	1.03	0.97	516	516	1.75	1026.75	0.08	0.11
Blok 5 1-back	0.08	-0.08	0.98	1.01	516	516	2.61	1029.49	0.01	0.16
Blok 5 2-back	0.04	-0.04	0.99	1.01	471	471	1.13	958.75	0.26	0.07

Blok 5 2-back PI	0.04	-0.04	1.00	1.00	372	372	1.09	749.61	0.28	0.08
Blok 6	0.08	-0.08	1.04	0.95	516	516	2.69	1021.44	0.01	0.17
Blok 7 deel 1	0.06	-0.06	1.02	0.97	516	516	2.04	1026.75	0.04	0.13
Blok 7 deel 2	0.04	-0.05	0.99	1.00	516	516	1.49	1029.76	0.14	0.09
Aantal fouten										
Blok 2	-0.34	-0.63	1.39	1.42	516	516	3.38	1029.36	0.00*	0.21
Blok 3	-0.21	-0.29	1.10	1.22	516	516	1.06	1019.48	0.29	0.07
Blok 4 compatibele trials	-0.65	-0.92	1.64	1.61	516	516	2.64	1029.67	0.01	0.16
Blok 4 incompatibele trials	-0.32	-0.34	1.28	1.24	516	516	0.23	1028.87	0.82	0.01
Blok 5 1-back	-0.86	-1.00	1.74	1.68	516	516	1.38	1028.84	0.17	0.09
Blok 5 2-back	-0.61	-0.52	1.52	1.41	471	471	-0.95	945.81	0.34	0.06
Blok 5 2-back PI	-0.30	-0.29	1.21	1.03	372	372	-0.15	726.83	0.88	0.01
Blok 7 deel 1	-0.81	-0.94	1.67	1.70	516	516	1.26	1029.61	0.21	0.08
Blok 7 deel 2	-1.30	-1.58	1.85	1.80	516	516	2.47	1029.22	0.01	0.15
Contrastparameters reactiesnelheid										
Blok 2 minus Blok 1	0.04	0.06	1.02	0.98	516	516	-0.42	1028.53	0.68	0.03
Blok 3 minus Blok 2	0.03	0.02	0.99	0.99	516	516	0.15	1030.00	0.88	0.01
Blok 4 minus Blok 2	0.05	-0.04	1.05	0.95	516	516	1.51	1020.50	0.13	0.09
Blok 7 minus Blok 2	0.14	-0.05	0.98	0.97	516	516	3.12	1029.88	0.00	0.19
Blok 4 incompatibele trials minus compatibele trials	-0.01	0.01	1.01	0.98	516	516	-0.23	1029.03	0.82	0.01
Blok 7 deel 2 minus Blok 7 deel 1	0.17	0.02	0.97	0.97	516	516	2.39	1029.97	0.02	0.15
Contrastparameters fouten										
Blok 3 minus Blok 2	0.14	0.00	1.05	0.92	516	516	2.19	1013.15	0.03	0.14
Blok 4 minus Blok 2	0.07	-0.01	1.05	0.94	516	516	1.28	1016.96	0.20	0.08
Blok 7 minus Blok 2	-0.07	0.08	1.00	0.89	516	516	-2.58	1016.34	0.01	0.16
Blok 4 incompatibele trials minus compatibele trials	0.10	0.03	1.03	0.92	516	516	1.19	1017.30	0.23	0.07
Blok 7 deel 2 minus Blok 7 deel 1	0.00	-0.03	0.99	0.99	516	516	0.52	1030.00	0.60	0.03
Blokoeverkoepelende afwijkende responsen										

Aantal premature responsen	-1.68	-1.99	1.85	1.74	516	516	2.85	1026.45	0.00	0.18
Aantal irrelevante knopdrukken	-2.36	-2.62	1.58	1.24	516	516	2.99	975.33	0.00	0.19
Aantal extra responsen	-0.46	-0.78	1.53	1.57	516	516	3.30	1029.20	0.00	0.21
Aantal responscompetities	-1.21	-1.56	1.88	1.89	516	516	2.94	1029.96	0.00	0.18

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

1.4.3 Subgroepen naar Totaal IQ

In Tabel 11 zijn per uitkomstparameter de gemiddelde score per Totaal IQ groep (TIQ_i85, TIQ 85-115, TIQ_i115), en de uitkomst van de ANOVA weergegeven. In lijn met de verwachting zijn er 14 (van de 21) uitkomstparameters in de Basisrapportage waarop een significant verschil is gevonden in gemiddelden tussen de Totaal IQ groepen. Dit is duidelijk hoger dan het verwachte aantal van ongeveer 1 uitkomstparameter. Ook in de Specialistische rapportage is het aantal (23 parameters) duidelijk hoger dan verwacht op kansniveau. De bijbehorende effectgroottes variëren van klein tot (laag) groot (met een range in ω^2 van .01 tot .11). Voor de uitkomstparameters met significante verschillen in gemiddeldes tussen de drie IQ-groepen behalen kinderen met een benedengemiddeld IQ over de hele linie en voor alle cognitieve parameters (verwerkingssnelheid, aandacht, executieve controle, (werk)geheugen, leersnelheid en gedrag) een minder goede prestatie op de COTAPP dan kinderen met een gemiddelde intelligentie. Kinderen met een bovengemiddelde intelligentie behalen een betere prestatie op de COTAPP dan kinderen met een gemiddelde intelligentie. Deze verschillen zijn voor zowel de Basis als Specialistische rapportage duidelijk zichtbaar op de uitkomstparameters die zijn gebaseerd op reactietijden en in iets mindere mate op uitkomstparameters die zijn gebaseerd op fouten. Echter, door het self-paced karakter van de COTAPP is de spreiding in het aantal fouten relatief klein en is ook de kans op verschillen bij de foutenparameters kleiner. We zouden kunnen concluderen dat kinderen met een hoger IQ sneller basale en complexe informatie kunnen waarnemen, verwerken en onthouden dan kinderen met een lager IQ met behoud van nauwkeurigheid (zie ook Rommelse, Luman, Kievit, 2020). Deze subgroepverschillen zijn geheel in lijn met de verwachtingen, waarbij de cognitieve domeinen bemeten met de COTAPP (informatieverwerking, aandacht en executieve functies) relateren aan (en in enige mate ten grondslag liggen aan) de kennis en vaardigheden met betrekking tot de cognitieve domeinen bemeten met intelligentietests.

Tabel 11: Subgroepen naar Totaal IQ

	Mean			ANOVA				
	onder- gemiddeld	gemiddeld	boven- gemiddeld	F-value	df1	df2	p-value	ω^2
BASISRAPPORTAGE								
Verwerkingssnelheid								
Reactiesnelheid	0.71	-0.04	-0.50	62.67	2	979	0.00*	0.11
Variabiliteit in reactiesnelheid	0.53	-0.02	-0.37	32.98	2	979	0.00*	0.06
Reactiesnelheid tgv cognitieve belasting	0.25	0.09	-0.29	11.94	2	979	0.00*	0.02

Versnellingscapaciteit in reactiesnelheid	0.15	0.01	-0.05	1.58	2	979	0.21	0.00
Variabiliteit in beslistijd	0.36	0.00	-0.38	21.06	2	979	0.00*	0.04
Aandacht								
Volgehouden aandacht	0.27	0.01	-0.02	4.44	2	979	0.01	0.01
Arousalregulatie	0.35	0.10	-0.12	8.51	2	979	0.00*	0.02
Onoplettendheid	0.33	-0.10	-0.36	18.29	2	979	0.00*	0.03
Executieve controle								
Slordigheid	-1.18	-1.60	-1.81	11.91	2	979	0.00*	0.02
Interferentiegevoeligheid	0.08	-0.00	-0.05	0.64	2	979	0.53	-0.00
Wachttijdafweging	0.38	0.05	-0.24	9.38	2	979	0.00*	0.02
(Werk)geheugen								
Reactiesnelheid eenvoudig	0.45	-0.02	-0.33	24.20	2	979	0.00*	0.05
Fouten eenvoudig	-0.36	-0.98	-1.09	9.18	2	979	0.00*	0.02
Reactiesnelheid complex	0.01	0.02	-0.23	3.89	2	913	0.02	0.01
Fouten complex	-0.26	-0.55	-0.87	5.86	2	913	0.00	0.01
Reactiesnelheid zeer complex	-0.20	0.04	-0.12	2.86	2	711	0.06	0.01
Fouten zeer complex	0.44	-0.25	-0.85	35.69	2	711	0.00*	0.09
Leersnelheid								
Leersnelheid	-0.77	-1.52	-1.94	19.32	2	979	0.00*	0.04
Gedrag tijdens de taak								
Benodigde ondersteuning	-1.43	-2.30	-2.59	22.33	2	979	0.00*	0.04
Verbaliseren	-0.78	-1.32	-1.34	5.10	2	979	0.01	0.01
Beweeglijkheid	0.15	-0.19	-0.45	12.44	2	979	0.00*	0.02
SPECIALISTISCHE RAPPORTAGE								
Mediane reactietijd								
Blok 1	0.61	-0.07	-0.18	30.67	2	979	0.00*	0.06
Blok 2	0.60	0.01	-0.42	42.62	2	979	0.00*	0.08

Blok 3	0.69	-0.00	-0.51	60.17	2	979	0.00*	0.11
Blok 4 compatibele trials	0.60	-0.03	-0.42	42.33	2	979	0.00*	0.08
Blok 4 incompatibele trials	0.57	-0.02	-0.47	43.97	2	979	0.00*	0.08
Blok 5 1-back	0.45	-0.02	-0.33	24.20	2	979	0.00*	0.05
Blok 5 2-back	0.01	0.02	-0.23	3.89	2	913	0.02	0.01
Blok 5 2-back PI	-0.20	0.04	-0.12	2.86	2	711	0.06	0.01
Blok 6	-0.05	-0.01	0.08	0.76	2	979	0.47	-0.00
Blok 7 deel 1	0.58	-0.04	-0.43	41.18	2	979	0.00*	0.08
Blok 7 deel 2	0.69	-0.03	-0.41	53.08	2	979	0.00*	0.10
Intra-individuele coefficient van variatie								
Blok 1	0.48	-0.04	-0.27	24.00	2	979	0.00*	0.04
Blok 2	0.27	-0.00	-0.16	7.34	2	979	0.00*	0.01
Blok 3	0.32	-0.02	-0.23	11.95	2	979	0.00*	0.02
Blok 4 compatibele trials	0.16	0.00	-0.11	2.91	2	979	0.05	0.00
Blok 4 incompatibele trials	0.26	-0.00	-0.20	8.04	2	979	0.00*	0.01
Blok 5 1-back	0.32	-0.00	-0.31	14.80	2	979	0.00*	0.03
Blok 5 2-back	0.11	-0.02	-0.13	1.91	2	913	0.15	0.00
Blok 5 2-back PI	0.04	-0.03	-0.05	0.22	2	711	0.80	-0.00
Blok 6	0.36	0.00	-0.38	21.06	2	979	0.00*	0.04
Blok 7 deel 1	0.42	-0.00	-0.33	21.76	2	979	0.00*	0.04
Blok 7 deel 2	0.45	-0.02	-0.33	23.86	2	979	0.00*	0.04
Aantal fouten								
Blok 2	-0.70	-0.44	-0.38	2.42	2	979	0.09	0.00
Blok 3	-0.69	-0.20	0.01	15.36	2	979	0.00*	0.03
Blok 4 compatibele trials	-1.07	-0.75	-0.57	3.80	2	979	0.02	0.01
Blok 4 incompatibele trials	-0.48	-0.31	-0.18	2.10	2	979	0.12	0.00
Blok 5 1-back	-0.36	-0.98	-1.09	9.18	2	979	0.00*	0.02
Blok 5 2-back	-0.26	-0.55	-0.87	5.86	2	913	0.00	0.01

Blok 5 2-back PI	0.44	-0.25	-0.85	35.69	2	711	0.00*	0.09
Blok 7 deel 1	-0.94	-0.82	-1.08	1.64	2	979	0.19	0.00
Blok 7 deel 2	-1.55	-1.41	-1.43	0.35	2	979	0.70	-0.00
Contrastparameters reactiesnelheid								
Blok 2 minus Blok 1	0.25	0.09	-0.29	11.94	2	979	0.00*	0.02
Blok 3 minus Blok 2	0.15	0.01	-0.05	1.58	2	979	0.21	0.00
Blok 4 minus Blok 2	0.08	-0.00	-0.05	0.64	2	979	0.53	-0.00
Blok 7 minus Blok 2	0.27	0.01	-0.02	4.44	2	979	0.01	0.01
Blok 4 incompatibele trials minus compatibele trials	0.20	-0.02	-0.20	5.99	2	979	0.00	0.01
Blok 7 deel 2 minus Blok 7 deel 1	0.35	0.10	-0.12	8.51	2	979	0.00*	0.02
Contrastparameters fouten								
Blok 3 minus Blok 2	-0.19	0.07	0.28	8.38	2	979	0.00*	0.01
Blok 4 minus Blok 2	-0.06	0.02	0.14	1.52	2	979	0.22	0.00
Blok 7 minus Blok 2	0.06	0.01	-0.11	1.34	2	979	0.26	0.00
Blok 4 incompatibele trials minus compatibele trials	0.08	0.06	0.10	0.16	2	979	0.85	-0.00
Blok 7 deel 2 minus Blok 7 deel 1	0.01	-0.04	0.07	0.84	2	979	0.43	-0.00
Blokoeverkoepelende afwijkende responsen								
Aantal premature responsen	-1.47	-1.84	-1.99	3.36	2	979	0.04	0.00
Aantal irrelevante knopdrukken	-2.19	-2.49	-2.69	4.65	2	979	0.01	0.01
Aantal extra responsen	-0.11	-0.65	-0.94	11.20	2	979	0.00*	0.02
Aantal responscompetities	-1.78	-1.36	-1.07	5.27	2	979	0.01	0.01

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

1.5 Informatie over de meetnauwkeurigheid

In deze sectie wordt informatie gegeven over de nauwkeurigheid van de meting. De standaardmeetfout van test X (SEM_X) kan, uitgaande van de klassieke testtheorie, berekend worden met de volgende formule:

$$SEM_X = \sqrt{\sigma_X^2(1 - \rho_{xx'})}$$

met σ_X^2 de variantie en $\rho_{xx'}$ de betrouwbaarheid van test X (zie bijv. Mellenbergh, 2011; p. 121).

Om de SEM van genormeerde scores voor een COTAPP testonderdeel te berekenen, kan voor $\rho_{xx'}$ de geschatte betrouwbaarheid voor die uitkomstparameter gebruikt worden. Deze zijn weergegeven in Tabel 20. Voor σ_X^2 kan de waarde 1 gebruikt worden, omdat de genormeerde scores uitgedrukt zijn in z -scores, en deze dus een variantie van 1 hebben.

Een betrouwbaarheidsinterval voor de ware score ('true score') in klassiek testtheoretische zin kan op twee manieren worden berekend, namelijk op basis van het individuele model en het groepsmodel (zoals bijv. beschreven in Mellenbergh, 2011, sectie 5.3). Voor de benodigde formules verwijzen we naar Mellenbergh (2011) voor het individuele model (sectie 5.3.1) en voor het groepsmodel (sectie 5.3.2).

1.6 Gegevens over de scoreverdelingen

In deze sectie worden de gegevens over de scoreverdelingen weergegeven voor zowel de ruwe scores, als de genormeerde scores. Als samenvattingsmaten worden weergegeven de gemiddelde (mean), de standaarddeviatie (sd), scheefheid (skewness), kurtosis (kurtosis), en de beschikbare steekproefomvang voor het testonderdeel (valid n).

Voor de vier parameters gebaseerd op factor-analyse (Reactiesnelheid, Variabiliteit in reactiesnelheid, Slordigheid en Wachtijdafweging) worden de samenvattingsmaten alleen gegeven voor de genormeerde testcores (uitgedrukt in genormaliseerde *IQ*-scores); de 'ruwe' scores zijn niet interessant, omdat deze de som zijn van de genormeerde (i.e., genormaliseerde *z*-) scores van de variabelen behorende bij deze geaggregeerde parameters.

De genormeerde testcores zijn *z*-scores die beperkt zijn tot het -3 tot +3 interval (zie Tabel 7.7 en AII.2 van de handleiding). Bij testonderdelen waarbij de ruwe score zeer scheef verdeeld is, is de gemiddelde *z*-score niet 0. Zoals beschreven in de handleiding (paragraaf 7.4.3) is hiervoor gecorrigeerd bij de omzetting van de *z*-scores in percentielscores in de Basis en Specialistische rapportage.

Tabel 12: Samenvattingsmaten van de geobserveerde testcores voor de Basis en specialistische rapportage

	mean	sd	skewness	kurtosis	valid n
BASISRAPPORTAGE					
Verwerkingssnelheid					
Reactiesnelheid	-	-	-	-	-
Variabiliteit in reactiesnelheid	-	-	-	-	-
Reactiesnelheid tgv cognitieve belasting	266.34	198.27	-6.20	75.18	1032
Versnellingscapaciteit in reactiesnelheid	5.87	100.66	0.82	7.53	1032
Variabiliteit in beslistijd	0.31	0.13	1.03	1.46	1032
Aandacht					
Volgehouden aandacht	106.78	185.27	3.13	22.43	1032
Arousalregulatie	122.15	202.17	2.68	28.30	1032
Onoplettendheid	25.53	6.96	0.40	0.22	1032
Executieve controle					
Slordigheid	-	-	-	-	-
Interferentiegevoeligheid	86.93	140.13	7.90	156.16	1032
Wachttijdafweging	-	-	-	-	-
(Werk)geheugen					
Reactiesnelheid eenvoudig	1137.71	369.04	1.95	7.61	1032
Fouten eenvoudig	1.51	2.08	2.53	8.39	1032
Reactiesnelheid complex	1429.87	476.70	1.42	3.55	961
Fouten complex	2.12	2.13	1.67	3.25	961
Reactiesnelheid zeer complex	1501.12	504.61	1.41	2.86	752
Fouten zeer complex	3.78	2.11	0.39	0.25	752
Leersnelheid					
Leersnelheid	57.42	7.53	1.70	3.22	1032
Gedrag tijdens de taak					
Benodigde ondersteuning	0.51	1.80	6.20	47.60	1031
Verbaliseren	1.64	2.56	2.00	3.96	1030
Beweeglijkheid	5.72	3.18	0.38	-0.35	1031
SPECIALISTISCHE RAPPORTAGE					
Mediane reactiesnelheid					
Blok 1	375.86	227.38	7.62	86.26	1032
Blok 2	642.20	186.78	2.33	11.94	1032
Blok 3	648.07	197.15	3.35	27.31	1032
Blok 4 compatibele trials	692.40	219.53	7.00	116.65	1032
Blok 4 incompatibele trials	765.84	261.40	7.19	122.87	1032
Blok 5 1-back	1137.71	369.04	1.95	7.61	1032
Blok 5 2-back	1429.87	476.70	1.42	3.55	961

Blok 5 2-back PI	1501.12504.61	1.41	2.86	752
Blok 6	1585.10676.75	1.12	2.30	1032
Blok 7 deel 1	748.98277.74	3.78	34.89	1032
Blok 7 deel 2	871.13371.84	2.81	14.23	1032
Intra-individuele coëfficiënt van variatie				
Blok 1	0.12 0.07	2.93	15.21	1032
Blok 2	0.15 0.06	1.68	5.36	1032
Blok 3	0.16 0.05	1.26	2.66	1032
Blok 4 compatibele trials	0.19 0.07	0.70	0.44	1032
Blok 4 incompatibele trials	0.15 0.07	1.07	1.33	1032
Blok 5 1-back	0.20 0.07	1.14	2.65	1032
Blok 5 2-back	0.23 0.08	0.95	1.67	961
Blok 5 2-back PI	0.25 0.09	0.64	0.44	752
Blok 6	0.31 0.13	1.03	1.46	1032
Blok 7 deel 1	0.17 0.08	1.44	2.96	1032
Blok 7 deel 2	0.16 0.09	1.81	6.49	1032
Aantal fouten				
Blok 2	2.37 2.44	2.08	6.05	1032
Blok 3	4.06 2.91	1.18	2.17	1032
Blok 4 compatibele trials	1.48 1.62	1.58	2.89	1032
Blok 4 incompatibele trials	3.26 2.69	1.32	2.11	1032
Blok 5 1-back	1.51 2.08	2.53	8.39	1032
Blok 5 2-back	2.12 2.13	1.67	3.25	961
Blok 5 2-back PI	3.78 2.11	0.39	0.25	752
Blok 7 deel 1	1.44 1.87	3.01	17.50	1032
Blok 7 deel 2	0.76 1.30	3.69	22.44	1032
Contrastparameters reactiesnelheid				
Blok 2 minus Blok 1	266.34198.27	-6.20	75.18	1032
Blok 3 minus Blok 2	5.87100.66	0.82	7.53	1032
Blok 4 minus Blok 2	86.93140.13	7.90	156.16	1032
Blok 7 minus Blok 2	106.78185.27	3.13	22.43	1032
Blok 4 incompatibele trials minus compatibele trials	73.44 88.67	1.80	10.77	1032
Blok 7 deel 2 minus Blok 7 deel 1	122.15202.17	2.68	28.30	1032
Contrastparameters fouten				
Blok 3 minus Blok 2	1.70 3.11	-0.28	3.30	1032
Blok 4 minus Blok 2	2.37 3.50	0.68	3.49	1032
Blok 7 minus Blok 2	-0.93 2.57	-0.44	5.60	1032
Blok 4 incompatibele trials minus compatibele trials	1.78 2.44	0.80	2.00	1032
Blok 7 deel 2 minus Blok 7 deel 1	-0.68 1.53	-1.48	7.01	1032
Blokoeverkoepelende afwijkende responsen				
Aantal premature responsen	0.60 1.88	14.39	309.37	1032
Aantal irrelevante knopdrukken	0.29 1.44	9.41	115.42	1032

Aantal extra responsen	5.60	27.19	21.38	551.25	1032
Aantal responscompetities	1.32	2.63	3.48	14.84	1032

Tabel 13: Samenvattingsmaten van de genormeerde testcores voor de Basis en specialistische rapportage

	mean	sd	skewness	kurtosis	valid n
BASISRAPPORTAGE					
Verwerkingssnelheid					
Reactiesnelheid	-0.00	1.00	0.05	-0.07	1032
Variabiliteit in reactiesnelheid	-0.00	1.00	0.31	-0.12	1032
Reactiesnelheid tgv cognitieve belasting	0.05	1.00	0.02	0.13	1032
Versnellingscapaciteit in reactiesnelheid	0.02	0.99	0.07	0.02	1032
Variabiliteit in beslistijd	0.00	1.00	0.01	-0.10	1032
Aandacht					
Volgehouden aandacht	0.05	0.98	0.09	0.03	1032
Arousalregulatie	0.09	0.97	0.18	-0.05	1032
Onoplettendheid	-0.08	1.01	0.03	-0.04	1032
Executieve controle					
Slordigheid	-1.58	1.14	0.83	0.63	1032
Interferentiegevoeligheid	0.01	1.00	-0.02	0.05	1032
Wachttijdafweging	0.06	1.24	-0.82	0.68	1032
(Werk)geheugen					
Reactiesnelheid eenvoudig	0.02	1.01	0.00	0.01	1032
Fouten eenvoudig	-0.93	1.71	-0.09	-1.35	1032
Reactiesnelheid complex	0.00	1.00	0.02	-0.07	961
Fouten complex	-0.56	1.47	-0.43	-0.51	961
Reactiesnelheid zeer complex	0.01	1.00	0.05	-0.08	752
Fouten zeer complex	-0.30	1.12	-0.42	0.19	752
Leersnelheid					
Leersnelheid	-1.48	1.70	0.31	-1.72	1032
Gedrag tijdens de taak					
Benodigde ondersteuning	-2.20	1.65	1.65	0.95	1031
Verbaliseren	-1.25	1.88	0.26	-1.67	1030
Beweeglijkheid	-0.18	1.06	-0.12	0.54	1031
SPECIALISTISCHE RAPPORTAGE					
Mediane reactiesnelheid					
Blok 1	0.02	1.02	0.09	0.05	1032
Blok 2	0.03	1.00	0.00	0.00	1032
Blok 3	0.02	1.00	0.01	-0.10	1032
Blok 4 compatibele trials	0.02	1.00	0.01	-0.10	1032
Blok 4 incompatibele trials	0.01	1.00	-0.03	0.04	1032
Blok 5 1-back	0.02	1.01	0.00	0.01	1032
Blok 5 2-back	0.00	1.00	0.02	-0.07	961

Blok 5 2-back PI	0.011.00	0.05	-0.08	752
Blok 6	-0.001.00	0.00	-0.09	1032
Blok 7 deel 1	-0.001.00	-0.04	-0.09	1032
Blok 7 deel 2	0.000.99	0.04	-0.15	1032
Intra-individuele coëfficiënt van variatie				
Blok 1	0.001.00	-0.00	-0.10	1032
Blok 2	0.001.00	0.02	-0.15	1032
Blok 3	-0.001.00	0.00	-0.07	1032
Blok 4 compatibele trials	-0.001.00	0.01	-0.15	1032
Blok 4 incompatibele trials	-0.001.00	-0.01	-0.15	1032
Blok 5 1-back	-0.001.00	-0.01	-0.01	1032
Blok 5 2-back	-0.001.00	0.00	-0.09	961
Blok 5 2-back PI	0.001.00	0.01	-0.04	752
Blok 6	0.001.00	0.01	-0.10	1032
Blok 7 deel 1	0.001.00	0.04	-0.13	1032
Blok 7 deel 2	-0.011.00	0.02	-0.09	1032
Aantal fouten				
Blok 2	-0.481.41	-0.44	-0.17	1032
Blok 3	-0.251.16	-0.33	0.42	1032
Blok 4 compatibele trials	-0.791.63	-0.27	-1.17	1032
Blok 4 incompatibele trials	-0.331.26	-0.44	0.35	1032
Blok 5 1-back	-0.931.71	-0.09	-1.35	1032
Blok 5 2-back	-0.561.47	-0.43	-0.51	961
Blok 5 2-back PI	-0.301.12	-0.42	0.19	752
Blok 7 deel 1	-0.871.69	-0.18	-1.33	1032
Blok 7 deel 2	-1.441.83	0.44	-1.51	1032
Contrastparameters reactiesnelheid				
Blok 2 minus Blok 1	0.051.00	0.02	0.13	1032
Blok 3 minus Blok 2	0.020.99	0.07	0.02	1032
Blok 4 minus Blok 2	0.011.00	-0.02	0.05	1032
Blok 7 minus Blok 2	0.050.98	0.09	0.03	1032
Blok 4 incompatibele trials minus compatibele trials	0.001.00	-0.01	-0.00	1032
Blok 7 deel 2 minus Blok 7 deel 1	0.090.97	0.18	-0.05	1032
Contrastparameters fouten				
Blok 3 minus Blok 2	0.070.99	0.03	-0.01	1032
Blok 4 minus Blok 2	0.031.00	0.04	0.15	1032
Blok 7 minus Blok 2	0.000.95	-0.39	1.83	1032
Blok 4 incompatibele trials minus compatibele trials	0.060.98	0.19	-0.14	1032
Blok 7 deel 2 minus Blok 7 deel 1	-0.010.99	0.02	0.39	1032
Blokoeverkoepelende afwijkende responsen				
Aantal premature responsen	-1.841.80	0.96	-0.91	1032
Aantal irrelevante knopdrukken	-2.491.43	2.49	4.41	1032

Aantal extra responsen	-0.621.56	-0.45	-0.78	1032
Aantal responscompetities	-1.381.89	0.44	-1.53	1032

2 Betrouwbaarheid

2.1 Parallellie van testhelften

Bij neuropsychologische tests zoals de COTAPP wordt een sterk beroep gedaan op concentratie, reactiesnelheid en geheugen. Factoren als vermoeidheid, motivatie, stress etc beïnvloeden de beschikbare cognitieve capaciteit binnen een -voor dat individu- specifieke range van capaciteit. Dit is een wezenlijk verschil met andere prestatietests zoals intelligentietests waar aan/afwezige kennis en vaardigheden worden bemeten (Danner et al., 2011). Om deze reden wordt afgeraden test-hertest betrouwbaarheid te gebruiken in de beoordeling van taken als de COTAPP (Farnsworth et al., 2017). Interne betrouwbaarheidsmaten worden hiervoor beter geschikt geacht (Gregory 2011).

In Appendix I van de Handleiding (Stimulusprogrammering, blz 129) staat per blok beschreven welke stimuli getoond zijn in welke volgorde per blok. Hieruit is af te leiden dat de verdeling van de type stimuli per blok nagenoeg gelijk is voor beide testhelften op basis van even versus oneven trial nummering. Tevens is bewust gekozen voor een split-half op basis van even en oneven genummerde trial nummers, en niet op basis van de eerste versus de tweede helft van de trials, om rekening te houden met time-on-task effecten in snelheid en accuratesse die doorgaans optreden gedurende een taak, bijvoorbeeld door het optreden van vermoeidheid of juist door leereffecten.

De betrouwbaarheid van de uitkomstparameters is berekend op basis van de split-half betrouwbaarheid. De onderliggende aanname hierbij is dat de testdelen van de oneven en de even trials parallel zijn, oftewel gelijke ware scores en gelijke meetfoutvarianties hebben. Op theoretische gronden is dit een volstrekt redelijke aanname, omdat de trials inhoudelijk gezien replicaties van elkaar zijn, en eventuele leereffecten in de loop van de taak gelijkelijk verdeeld verwacht mogen worden over de oneven en even trials.

Als parallellie opgaat voor twee testhelften, dan heeft dat implicaties voor de te verwachte scores op deze test (zie bijv. Mellenbergh, 2011; sectie 5.2.3). De volgende implicaties zijn hierbij van belang.

1. de verwachte waarden van de gemiddelden van beide testhelften zijn gelijk;
2. de verwachte waarden van de varianties van beide testhelften zijn gelijk;
3. de correlaties tussen de beiden testhelften met een willekeurige derde test zijn gelijk.

Deze drie implicaties zijn hier van belang omdat deze gebruikte kunnen worden om de aannemelijkheid van de aanname van parallellie voor twee testhelften te onderzoeken (Mellenbergh, 2011; p. 125-126). Dat kan door in de steekproef na te gaan in hoeverre deze verwachte effecten opgaan.

Voor de uitkomstparameters van de COTAPP Basisrapportage en Specialistische rapportage zijn daarom de aannemelijkheid van de drie implicaties onderzocht. Als willekeurige derde test is de subschaal 'Conflict' van de Relatie Vragenlijst Leerlingen (RVL) genomen. Deze is gebaseerd op zelfrapportage van de kinderen. Dit is gedaan door

1. de gemiddelde scores van de testhelften te toetsen met een paired t-test;
2. de varianties van de testhelften te toetsen met een Levene's test voor gelijke varianties;
3. de correlaties van de testhelften met de subschaal 'Conflict' zijn getoetst met een z-test, na Fisher-z-transformatie.

Voor alle drie typen toetsen die toegepast zijn, is de nulhypothese dat de statistieken (i.e., gemiddelde, variantie, correlatie) gelijk zijn voor de twee testhelften. Voor elke toets wordt de steekproefstatistiek per testhelft gegeven, de toetsgrootte, de vrijheidsgraden (df) en de p -waarde.

Voor de Basisrapportage zijn in totaal 48 toetsen uitgevoerd (16x3), voor de Specialistische rapportage in totaal 138 toetsen (46x3). Dit betekent dat er sprake is van een hoge mate van multiple hypothesetoetsing. We gebruiken een overall $\alpha = .05$, en per statistiek een Bonferroni correctie, en bekijken hoeveel en welke tests een p -waarde onder deze grens hebben ($< .05/48 = 0.00104$ Basisrapportage; $< .05/138 = 0.000363$ Specialistische rapportage). Deze staan aangegeven met * in de tabellen, en in vervolg aangeduid als kritieke waarde. Als aan de aanname van de Bonferroni correctie voldaan is, en de nulhypothese opgaat, dan is de verwachte waarde van het aantal tests met een p -waarde kleiner dan de kritieke waarde gelijk aan 2.4 ($48 * .05$) in de Basisrapportage en gelijk aan 6.9 ($138 * .05$) in de Specialistische rapportage. Per teststatistiek is de verwachte waarde van het aantal tests met een p -waarde kleiner dan de kritieke waarde gelijk aan 0.8 ($16 * .05$) in de Basisrapportage en gelijk aan 2.3 ($46 * .05$) in de Specialistische rapportage.

Uit de vergelijking van gemiddelden (zie tabellen 14 en 17) blijkt dat de p -waarde kleiner is dan de kritieke waarde voor 3 tests in de Basisrapportage, en voor 12 tests in de Specialistische rapportage. Deze aantallen zijn groter dan de verwachte waarden van 0.8 en 2.3, respectievelijk.

Inspectie van de effect sizes van deze respectievelijk 3 en 12 tests leert dat de effectgrootten (Cohen's d) variëren tussen verwaarloosbaar (3), klein (5), medium (5) en groot (2). Variabelen waarvoor een groot ($d=.5$) verschil gevonden wordt tussen beide testhelften, betreft het aantal fouten in het zeer complexe werkgeheugenblok. Omdat deze variabele in beide rapportages gegenereerd wordt, betreft het feitelijk één en dezelfde variabele. Variabelen waarvoor een gemiddeld groot ($d=.3$) verschil gevonden wordt tussen beide testhelften, betreffen (m.u.v. 1) foutenparameters of contrastparameters tussen blokken. Dit zijn variabelen die sterk kunnen gaan fluctueren wanneer ze op een te klein aantal trials worden gebaseerd:

- Aangezien de COTAPP een self-paced taak betreft (het kind kan op eigen tempo werken en wordt niet opgejaagd), worden er gemiddeld weinig fouten gemaakt en zijn deze parameters scheef verdeeld. Bij opdeling in testhelften, is de kans groter dan bij normaal verdeelde variabelen (zoals reactietijden) dat er een toevallig verschil tussen testhelften ontstaat zonder dat dit reflecteert dat de ene testhelft echt moeilijker is dan de andere testhelft.

- Evenzo is de verwachting dat contrastparameters tussen blokken sneller onderhevig zijn aan toevallige verschillen tussen testhelften. Deze parameters zijn an sich al moeilijker betrouwbaar te krijgen vanwege de combinatie van meetonnauwkeurigheid van twee variabelen. Bij berekening op een te klein aantal trials (immers, de helft van het aantal wat we als noodzakelijk achten om dit goed te kunnen meten), kunnen deze scores sneller dan andere scores gaan fluctueren.

Er is met andere woorden een grotere kans dat de scores op juist deze specifieke set uitkomstparameters meer fluctueren en dat om die reden parallelle niet voor alle uitkomstparameters aangetoond kan worden. Het telkens wisselen van tekens van de effect sizes ondersteunt verder de hypothese dat er geen systematisch verschil is in moeilijkheid tussen de even en oneven trials.

De aannemelijkheid van parallelle van beide testhelften wordt verder ondersteund uit de vergelijking van varianties (zie tabellen 15 en 18), waaruit blijkt dat de p -waarde kleiner is dan de kritieke waarde voor 1 test in de Basisrapportage, en voor geen van de tests in de Specialistische rapportage. Deze aantallen zijn in lijn met de verwachte waarden. Ook uit de vergelijking van correlaties (zie tabellen 16 en 19) blijkt dat testhelften vergelijkbaar correleren met een externe variabele (geen tests met een p -waarde kleiner dan de kritieke waarden). Samenvattend kan geconcludeerd worden de parallelle van testhelften zeker aannemelijk lijkt op basis van de resultaten op deze toetsen.

Tabel 14: Basisrapportage: Vergelijking van gemiddelden van testhelften

	Mean deel 1	Mean deel 2	Vershil in means	t-value	df	p-value	Cohen's d
Verwerkingssnelheid							
Reactiesnelheid	4016.84	4003.66	13.19	1.52	959.00	0.13	0.01
Variabiliteit in reactiesnelheid	0.67	0.66	0.01	2.71	959.00	0.01	0.06
Reactiesnelheid tgv cognitieve belasting	260.05	260.42	-0.37	-0.09	959.00	0.93	-0.00
Versnellingscapaciteit in reactiesnelheid	8.97	2.25	6.72	1.81	959.00	0.07	0.06
Variabiliteit in beslistijd	0.31	0.31	-0.00	-0.77	959.00	0.44	-0.02
Aandacht							
Volgehouden aandacht	114.34	104.38	9.95	2.35	959.00	0.02	0.05
Arousal regulatie	108.45	135.73	-27.28	-3.31	959.00	0.00*	-0.12
Onoplettendheid	13.03	12.82	0.22	1.59	959.00	0.11	0.06
Executieve controle							
Slordigheid	2.62	2.55	0.07	0.75	959.00	0.46	0.01
Interferentiegevoeligheid	84.91	92.97	-8.06	-2.47	959.00	0.01	-0.05
(Werk)geheugen							
Reactiesnelheid eenvoudig	1189.57	1077.73	111.84	13.10	959.00	0.00*	0.30
Fouten eenvoudig	0.69	0.80	-0.11	-2.65	959.00	0.01	-0.09
Reactiesnelheid complex	1397.23	1449.14	-51.91	-4.47	860.00	0.00*	-0.11
Fouten complex	1.10	0.96	0.14	3.27	860.00	0.00	0.12
Reactiesnelheid zeer complex	1504.92	1513.05	-8.12	-0.62	705.00	0.53	-0.02
Fouten zeer complex	2.24	1.55	0.69	10.17	705.00	0.00*	0.50

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

Tabel 15: Basisrapportage: Vergelijking van varianties van testhelften

	SD deel 1	SD deel 2	F-value	df	p-value
Verwerkingsnelheid					
Reactiesnelheid	1240.83	1186.92	0.04	1918.00	0.84
Variabiliteit in reactiesnelheid	0.20	0.20	0.06	1918.00	0.81
Reactiesnelheid tgv cognitieve belasting	209.48	211.82	0.06	1918.00	0.81
Versnellingscapaciteit in reactiesnelheid	103.45	121.22	0.12	1918.00	0.73
Variabiliteit in beslistijd	0.14	0.13	0.82	1918.00	0.37
Aandacht					
Volgehouden aandacht	207.16	202.49	0.11	1918.00	0.74
Arousal regulatie	224.38	237.03	0.36	1918.00	0.55
Onoplettendheid	3.83	3.85	0.03	1918.00	0.87
Executieve controle					
Slordigheid	12.51	11.84	0.03	1918.00	0.86
Interferentiegevoeligheid	168.59	136.95	1.08	1918.00	0.30
(Werk)geheugen					
Reactiesnelheid eenvoudig	399.36	343.49	7.74	1918.00	0.01
Fouten eenvoudig	1.25	1.14	4.01	1918.00	0.05
Reactiesnelheid complex	464.02	503.63	1.25	1720.00	0.26
Fouten complex	1.21	1.21	0.04	1720.00	0.84
Reactiesnelheid zeer complex	506.62	545.53	1.71	1410.00	0.19
Fouten zeer complex	1.45	1.31	12.24	1410.00	0.00*

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

Tabel 16: Basisrapportage: Vergelijking van correlaties van testhelften met derde variabele Y (Subschaal 'Conflict')

	R(deel 1,Y)	R(deel 2,Y)	Z-value	p-value
Verwerkingssnelheid				
Reactiesnelheid	0.03	0.03	0.04	0.97
Variabiliteit in reactiesnelheid	0.16	0.14	0.31	0.76
Reactiesnelheid tgv cognitieve belasting	0.02	0.02	0.11	0.92
Versnellingscapaciteit in reactiesnelheid	0.02	0.02	0.10	0.92
Variabiliteit in beslistijd	0.12	0.08	0.89	0.37
Aandacht				
Volgehouden aandacht	0.10	0.09	0.12	0.90
Arousal regulatie	0.08	0.10	0.51	0.61
Onoplettendheid	0.05	0.05	0.07	0.94
Executieve controle				
Slordigheid	0.07	0.10	0.63	0.53
Interferentiegevoeligheid	0.02	0.03	0.26	0.79
(Werk)geheugen				
Reactiesnelheid eenvoudig	-0.05	-0.01	0.74	0.46
Fouten eenvoudig	-0.01	0.02	0.67	0.50
Reactiesnelheid complex	-0.05	-0.05	0.01	0.99
Fouten complex	0.07	0.05	0.37	0.71
Reactiesnelheid zeer complex	-0.01	0.00	0.20	0.85
Fouten zeer complex	0.09	0.08	0.24	0.81

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

Tabel 17: Specialistische rapportage: Vergelijking van gemiddelden van testhelften

	Mean deel 1	Mean deel 2	Vershil in Means	t-value	df	p-value	Cohen's d
Mediane reactiesnelheid							
Blok 1	374.85	369.92	4.93	1.49	959.00	0.14	0.02
Blok 2	634.90	630.34	4.56	2.10	959.00	0.04	0.03
Blok 3	643.87	632.59	11.28	4.49	959.00	0.00*	0.06
Blok 4 compatibel	671.32	698.30	-26.99	-6.34	959.00	0.00*	-0.12
Blok 4 incompatibel	768.29	748.32	19.97	5.38	959.00	0.00*	0.08
Blok 5 1-back	1189.57	1077.73	111.84	13.10	959.00	0.00*	0.30
Blok 5 2-back	1397.23	1449.14	-51.91	-4.47	860.00	0.00*	-0.11
Blok 5 2-back PI	1504.92	1513.05	-8.12	-0.62	705.00	0.53	-0.02
Blok 6	1595.64	1592.15	3.48	0.34	959.00	0.74	0.01
Blok 7 deel 1	749.24	734.73	14.51	3.74	959.00	0.00*	0.05
Blok 7 deel 2	857.69	870.45	-12.77	-1.89	959.00	0.06	-0.04
Intra-individuele coëfficiënt van variatie							
Blok 1	0.12	0.11	0.01	2.28	959.00	0.02	0.07
Blok 2	0.15	0.14	0.00	0.84	959.00	0.40	0.03
Blok 3	0.15	0.15	-0.00	-1.92	959.00	0.06	-0.06
Blok 4 compatibele trials	0.19	0.17	0.02	6.18	959.00	0.00*	0.23
Blok 4 incompatibele trials	0.15	0.15	-0.00	-1.07	959.00	0.29	-0.04
Blok 5 1-back	0.19	0.18	0.01	2.69	959.00	0.01	0.10
Blok 5 2-back	0.22	0.21	0.01	2.22	860.00	0.03	0.09
Blok 5 2-back PI	0.25	0.24	0.01	1.74	705.00	0.08	0.08
Blok 6	0.31	0.31	-0.00	-0.77	959.00	0.44	-0.02
Blok 7 deel 1	0.17	0.16	0.00	0.94	959.00	0.35	0.03
Blok 7 deel 2	0.16	0.16	0.00	0.19	959.00	0.85	0.01
Aantal fouten							
Blok 2	1.19	1.18	0.01	0.24	959.00	0.81	0.01
Blok 3	2.03	2.08	-0.05	-0.75	959.00	0.45	-0.03
Blok 4 compatibele trials	0.78	0.72	0.06	1.63	959.00	0.10	0.06
Blok 4 incompatibele trials	1.44	1.86	-0.42	-9.76	959.00	0.00*	-0.28
Blok 5 1-back	0.69	0.80	-0.11	-2.65	959.00	0.01	-0.09
Blok 5 2-back	1.10	0.96	0.14	3.27	860.00	0.00	0.12
Blok 5 2-back PI	2.24	1.55	0.69	10.17	705.00	0.00*	0.50
Blok 7 deel 1	0.74	0.71	0.02	0.68	959.00	0.50	0.02
Blok 7 deel 2	0.40	0.36	0.03	1.20	959.00	0.23	0.04
Contrastparameters snelheid							
Blok 2 minus Blok 1	260.05	260.42	-0.37	-0.09	959.00	0.93	-0.00
Blok 3 minus Blok 2	8.97	2.25	6.72	1.81	959.00	0.07	0.06
Blok 4 minus Blok 2	84.91	92.97	-8.06	-2.47	959.00	0.01	-0.05
Blok 7 minus Blok 2	114.34	104.38	9.95	2.35	959.00	0.02	0.05
Blok 4 incompatibele trials minus compatibele trials	96.98	50.02	46.96	9.38	959.00	0.00*	0.40
Blok 7 deel 2 minus Blok 7 deel 1	108.45	135.73	-27.28	-3.31	959.00	0.00	-0.12
Contrastparameters fouten							
Blok 3 minus Blok 2	0.85	0.91	-0.06	-0.77	959.00	0.44	-0.03
Blok 4 minus Blok 2	1.03	1.41	-0.38	-5.29	959.00	0.00*	-0.19
Blok 7 minus Blok 2	-0.45	-0.46	0.01	0.25	959.00	0.80	0.01
Blok 4 incompatibele trials minus compatibele trials	0.65	1.14	-0.48	-8.74	959.00	0.00*	-0.32
Blok 7 deel 2 minus Blok 7 deel 1	-0.34	-0.35	0.01	0.16	959.00	0.87	0.01
Blokoeverkoepelende afwijkende responsen							
Aantal premature responsen	0.29	0.32	-0.03	-1.22	959.00	0.22	-0.03
Aantal irrelevante knopdrukken	0.16	0.13	0.02	1.09	959.00	0.27	0.03
Aantal extra responsen	2.85	2.76	0.09	0.79	959.00	0.43	0.01
Aantal responscompetities	0.64	0.71	-0.06	-1.54	959.00	0.12	-0.04

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

Tabel 18: Specialistische rapportage: Vergelijking van varianties van testhelften met derde variabele Y (Subschaal 'Conflict')

	SD deel 1	SD deel 2	F-value	df	p-value
Mediane reactiesnelheid					
Blok 1	231.10	239.83	0.08	1918.00	0.77
Blok 2	172.18	166.66	0.01	1918.00	0.93
Blok 3	179.41	188.49	0.64	1918.00	0.43
Blok 4 compatibele trials	246.37	191.58	0.19	1918.00	0.66
Blok 4 incompatibele trials	281.54	249.57	0.15	1918.00	0.70
Blok 5 1-back	399.36	343.49	7.74	1918.00	0.01
Blok 5 2-back	464.02	503.63	1.25	1720.00	0.26
Blok 5 2-back PI	506.62	545.53	1.71	1410.00	0.19
Blok 6	662.49	706.65	1.94	1918.00	0.16
Blok 7 deel 1	286.48	266.00	0.08	1918.00	0.78
Blok 7 deel 2	347.06	373.49	0.11	1918.00	0.74
Intra-individuele coëfficiënt van variatie					
Blok 1	0.08	0.08	0.06	1918.00	0.80
Blok 2	0.07	0.06	0.25	1918.00	0.62
Blok 3	0.06	0.06	0.56	1918.00	0.45
Blok 4 compatibele trials	0.08	0.08	3.91	1918.00	0.05
Blok 4 incompatibele trials	0.07	0.08	1.42	1918.00	0.23
Blok 5 1-back	0.09	0.08	2.56	1918.00	0.11
Blok 5 2-back	0.09	0.09	0.02	1720.00	0.89
Blok 5 2-back PI	0.11	0.10	0.69	1410.00	0.41
Blok 6	0.14	0.13	0.82	1918.00	0.37
Blok 7 deel 1	0.09	0.08	0.25	1918.00	0.62
Blok 7 deel 2	0.10	0.10	0.15	1918.00	0.69
Aantal fouten					
Blok 2	1.38	1.36	0.24	1918.00	0.62
Blok 3	1.80	1.75	0.27	1918.00	0.61
Blok 4 compatibele trials	0.97	0.99	1.21	1918.00	0.27
Blok 4 incompatibele trials	1.45	1.55	8.05	1918.00	0.00
Blok 5 1-back	1.25	1.14	4.01	1918.00	0.05
Blok 5 2-back	1.21	1.21	0.04	1720.00	0.84
Blok 5 2-back PI	1.45	1.31	12.24	1410.00	0.00
Blok 7 deel 1	1.13	1.07	0.25	1918.00	0.62
Blok 7 deel 2	0.80	0.74	0.84	1918.00	0.36
Contrastparameters reactiesnelheid					
Blok 2 minus Blok 1	209.48	211.82	0.06	1918.00	0.81
Blok 3 minus Blok 2	103.45	121.22	0.12	1918.00	0.73
Blok 4 minus Blok 2	168.59	136.95	1.08	1918.00	0.30
Blok 7 minus Blok 2	207.16	202.49	0.11	1918.00	0.74
Blok 4 incompatibele trials minus compatibele trials	113.22	121.96	1.94	1918.00	0.16
Blok 7 deel 2 minus Blok 7 deel 1	224.38	237.03	0.36	1918.00	0.55
Contrastparameters fouten					
Blok 3 minus Blok 2	1.92	1.95	0.19	1918.00	0.66
Blok 4 minus Blok 2	1.95	2.10	5.64	1918.00	0.02
Blok 7 minus Blok 2	1.54	1.54	0.44	1918.00	0.51
Blok 4 incompatibele trials minus compatibele trials	1.49	1.51	0.21	1918.00	0.65
Blok 7 deel 2 minus Blok 7 deel 1	1.06	1.04	0.08	1918.00	0.78
Blokoeverkoepelende afwijkende responsen					
Aantal premature responsen	1.00	1.05	0.33	1918.00	0.56
Aantal irrelevante knopdrukken	0.86	0.74	0.43	1918.00	0.51
Aantal extra responsen	14.54	13.72	0.02	1918.00	0.90
Aantal responscompetities	1.45	1.49	0.81	1918.00	0.37

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

Tabel 19: Specialistische rapportage: Vergelijking van correlaties van testhelften met derde variabele Y (Subschaal 'Conflict')

	R(deel 1,Y)	R(deel 2,Y)	Z-value	p-value
Mediane reactiesnelheid				
Blok 1	-0.03	-0.04	0.22	0.83
Blok 2	-0.01	-0.02	0.20	0.84
Blok 3	-0.00	-0.01	0.06	0.95
Blok 4 compatibele trials	0.01	0.01	0.04	0.97
Blok 4 incompatibele trials	-0.00	-0.00	0.07	0.94
Blok 5 1-back	-0.05	-0.01	0.78	0.44
Blok 5 2-back	-0.05	-0.05	0.01	0.99
Blok 5 2-back PI	-0.01	0.00	0.24	0.81
Blok 6	-0.11	-0.11	0.00	1.00
Blok 7 deel 1	0.06	0.06	0.15	0.88
Blok 7 deel 2	0.10	0.11	0.04	0.97
Intra-individuele coëfficiënt van variatie				
Blok 1	0.08	0.01	1.65	0.10
Blok 2	0.10	0.09	0.12	0.91
Blok 3	0.08	0.08	0.15	0.88
Blok 4 compatibele trials	0.07	0.06	0.23	0.82
Blok 4 incompatibele trials	0.05	0.12	1.67	0.10
Blok 5 1-back	0.04	0.07	0.57	0.57
Blok 5 2-back	0.01	0.08	1.71	0.09
Blok 5 2-back PI	0.04	-0.02	1.29	0.20
Blok 6	0.12	0.08	0.94	0.35
Blok 7 deel 1	0.15	0.08	1.70	0.09
Blok 7 deel 2	0.10	0.13	0.69	0.49
Aantal fouten				
Blok 2	0.07	0.06	0.29	0.77
Blok 3	-0.02	0.05	1.54	0.12
Blok 4 compatibele trials	0.01	0.01	0.02	0.98
Blok 4 incompatibele trials	0.07	0.05	0.34	0.73
Blok 5 1-back	-0.01	0.02	0.70	0.48
Blok 5 2-back	0.07	0.05	0.41	0.68
Blok 5 2-back PI	0.09	0.08	0.30	0.77
Blok 7 deel 1	0.06	0.07	0.28	0.78
Blok 7 deel 2	0.04	-0.02	1.56	0.12
Contrastparameters reactiesnelheid				
Blok 2 minus Blok 1	0.02	0.02	0.11	0.91
Blok 3 minus Blok 2	0.02	0.02	0.11	0.91
Blok 4 minus Blok 2	0.02	0.03	0.26	0.79
Blok 7 minus Blok 2	0.10	0.09	0.13	0.90
Blok 4 incompatibele trials minus compatibele trials	-0.02	-0.02	0.12	0.90
Blok 7 deel 2 minus Blok 7 deel 1	0.08	0.10	0.53	0.60
Contrastparameters fouten				
Blok 3 minus Blok 2	-0.07	0.00	1.64	0.10
Blok 4 minus Blok 2	0.01	0.01	0.00	1.00
Blok 7 minus Blok 2	-0.02	-0.00	0.43	0.67
Blok 4 incompatibele trials minus compatibele trials	0.06	0.05	0.26	0.79
Blok 7 deel 2 minus Blok 7 deel 1	-0.03	-0.09	1.41	0.16
Blokoeverkoepelende afwijkende responsen				
Aantal premature responsen	0.10	0.05	1.20	0.23
Aantal irrelevante knopdrukken	0.07	0.04	0.59	0.56
Aantal extra responsen	0.06	0.10	0.81	0.42
Aantal responscompetities	0.04	0.08	0.87	0.39

* p-waarde kleiner dan de kritieke waarde (zie tekst voor detail)

2.2 Split-half betrouwbaarheid op basis van leeftijdsgroepen

De split-half betrouwbaarheden zijn voor de hele normgroep en per leeftijdsgroep afzonderlijk berekend middels intra-klasse correlaties voor absolute agreement (Koo en Li, 2006). De coëfficiënten per leeftijdsgroep zijn vergeleken met de coëfficiënten gebaseerd op de gehele normgroep middels Fishers r-z transformaties ($p < .05$). In de handleiding zijn echter Pearsons correlaties gebruikt i.p.v. ICCs. Bij nader inzien blijken Pearsons correlaties niet de beste methode om de mate van overeenkomst te berekenen wanneer het gepaarde waarnemingen binnen één persoon betreffen die onderling uitwisselbaar zijn / geen rangordening hebben. Middels de Pearsons correlatie wordt uitgegaan van een lineaire relatie tussen de twee metingen (bijv. om de aanname te toetsen dat een score op de ene testhelpt altijd hoger is dan de score op de andere testhelpt); er wordt echter niet getoetst of de metingen overeen komen (Bland en Altman, 1986). Gezien we tussen twee testhelpten geen lineaire relatie verwachten, is een maat waarbij zowel de correlatie als de overeenkomst wordt meegenomen, zoals de ICC, meer wenselijk. Daarnaast worden middels de ICC alle datapunten gepoold en gestandaardiseerd i.p.v. apart voor beide testafnames (of testhelpten) bepaald zoals middels Pearsons correlaties. Het eerste ligt meer voor de hand, aangezien het inwisselbare waarnemingen betreffen. Er zijn verschillende ICC modellen en voor de test-hertest en split-half betrouwbaarheid hebben we gebruik gemaakt van de 'two-way mixed' methode met absolute overeenkomst, waarbij het 'mixed' model het meest passend is wanneer het metingen betreffen waarbij geen subjectieve beoordelaars-effecten een rol spelen zoals het geval is bij een gecomputeriseerde taak als de COTAPP (Koo en Li, 2006). Het is overigens in geen geval zo dat de ICCs systematisch hoger uitvallen dan Pearson correlaties. Er zijn meerdere scenario's waarin de Pearson correlatie zeer hoog kan zijn (1.0 of een benadering ervan), maar de ICC (tav consistentie en/of absolute overeenkomst) aanzienlijk lager (Koo en Li, 2006).

De ICCs voor de hele normgroep vallen voor een aantal uitkomstparameters iets hoger uit dan de Pearson correlaties gerapporteerd in de Handleiding. Voor leeftijdsgroep specifieke ICCs worden er weinig betekenisvolle verschillen gevonden i.v.t. de ICCs voor de hele normgroep. De betrouwbaarheidscoëfficiënten van de gehele normgroep zijn daarmee niet optimistischer dan voor de afzonderlijke leeftijdsgroepen.

Op 2 uitkomstparameters uit de Basisrapportage komen verschillen naar voren die goed te duiden zijn vanuit de cognitieve ontwikkeling van kinderen, waardoor deze parameters betrouwbaarder zijn wanneer de COTAPP wordt afgenomen in een bepaalde leeftijdsfase:

- de betrouwbaarheid van de uitkomstmaat Versnellingscapaciteit Reactiesnelheid is hoger bij oudere kinderen (8 t/m 9 jaar $r=.78$, 10 t/m 13 jaar $r=.80$) dan bij jongere kinderen (5 t/m 7 jaar $r=.50$); waarschijnlijk omdat bij de jongste kinderen een 2-keuze reactietijdtaak een relatief grote cognitieve inspanning vraagt vergeleken met oudere kinderen. Er is hierdoor weinig ruimte om nog verder te versnellen, omdat dit een zekere mate van geautomatiseerdheid vraagt die deze jonge kinderen nog niet beheersen.
- de betrouwbaarheid van de uitkomstmaat Fouten Zeer Complex blijkt hoger bij oudere kinderen dan bij de jongere kinderen (5 t/m 7 jaar $r=-.13$, 8 t/m 9 jaar $r=.14$; versus 10 t/m 13 jaar $r=.34$). Dit is waarschijnlijk te verklaren doordat deze zeer complexe variant op kans-niveau

wordt gemaakt door een deel van de jonge kinderen. Deze variant vraagt gemiddeld genomen (aanzienlijk) meer van de werkgeheugencapaciteit dan het kind ter beschikking heeft op 5-7 jarige leeftijd. Om die reden wordt dit blok alleen afgenomen als de prestatie op het voorgaande blok met minimaal 85 procent nauwkeurigheid wordt gemaakt; de zeer complexe variant is wel informatief om plafond-effecten te voorkomen bij cognitief sterke jonge kinderen. In ieder geval zijn fouten op de zeer complexe variant daarmee doorgaans bij jongere kinderen minder indicatief voor de werkgeheugencapaciteit dan bij oudere kinderen. Steun voor deze aanname is dat een vergelijkbaar effect zichtbaar is voor aantal fouten op de eenvoudige variant, die bij de oudste leeftijdsgroep minder betrouwbaar is dan bij de jongere kinderen: deze variant vraagt dusdanig weinig van de werkgeheugencapaciteit van gemiddelde 10-13 jarigen, dat fouten weinig indicatief zijn voor het werkgeheugenproces en eerder toevalstreffers reflecteren / incidentele fouten (NB dit ligt uiteraard anders op individueel interpretatieniveau wanneer er een groter aantal fouten wordt gemaakt door een ouder kind).

Vanuit de Specialistische rapportage wordt dit beeld verder ondersteund: foutenparameters en contrastparameters gebaseerd op fouten bij 'eenvoudigere' blokken (Blok 2, 3, 7) zijn minder betrouwbaar in de oudste leeftijdsgroep dan de jongste leeftijdsgroepen, waarschijnlijk vanwege bovengenoemde reden. Bij jongere kinderen zijn deze overwegend betrouwbaarder vergeleken de hele normgroep. De Basisrapportage bevat echter geen parameters gebaseerd op fouten, m.u.v. het Werkgeheugenblok vanwege goede Validiteitsgegevens.

Samennemend kan daarom geconcludeerd worden dat de split-half betrouwbaarheidscoëfficiënten berekend voor de hele normgroep voor de Basisrapportage van toepassing zijn ongeacht de leeftijd van het onderzochte kind. Foutenparameters en contrastparameters gebaseerd op fouten laten betekenisvolle verschillen zien in betrouwbaarheid die zinvol zijn om mee te wegen wanneer de testgebruiker ook de gegevens uit de Specialistische rapportage mee wil (en kan, zie vereist opleidingsniveau) wegen.

Tabel 20: Split-half betrouwbaarheidsgegevens voor de uitkomstmaten van de Basisrapportage en Specialistische rapportage voor gehele groep en per leeftijdsgroep.

Uitkomstmaat	hele groep n=1032	5- 7 jaar n=316	8-9 jaar n=293	10-13 jaar n=423
BASISRAPPORTAGE				
Verwerkingssnelheid				
Reactiesnelheid	.99	.98	.99	.99
Variabiliteit in reactiesnelheid	.89	.88	.88	.87
Reactiesnelheid t.g.v. van cognitieve belasting	.89	.89	.90	.85
Versnellingscapaciteit in reactiesnelheid	.64	.50	.78	.80
Variabiliteit in beslistijd	.80	.78	.79	.76
Aandacht				
Volgehouden aandacht	.87	.85	.88	.86
Arousalregulatie	.60	.53	.65	.63
Onoplettendheid	.59	.61	.59	.55
Executieve controle				
Slordigheid	.99	.96	.95	.99
Interferentiegevoeligheid	.86	.88	.78	.81
Wachttijdafweging	.85	.83	.87	.85
(Werk)geheugen				
Reactiesnelheid eenvoudig	.84	.81	.75	.78
Fouten eenvoudig	.63	.66	.60	.40
Reactiesnelheid complex	.87	.92	.82	.79
Fouten complex	.61	.63	.60	.51
Reactiesnelheid zeer complex	.88	.88	.82	.86
Fouten zeer complex	.26	-.13	.14	.34
Leersnelheid				
Leersnelheid	n.v.t.			
Gedrag tijdens de taak				
Benodigde ondersteuning	.94*			
Verbaliseren	.73*			
Beweeglijkheid	.70*			
SPECIALISTISCHE RAPPORTAGE				
Mediane reactiesnelheid				
Blok 1	.95	.96	.96	.89
Blok 2	.96	.95	.94	.96
Blok 3	.96	.94	.95	.96
Blok 4 compatibele trials	.91	.89	.85	.89
Blok 4 incompatibele trials	.95	.94	.92	.92
Blok 5 1-back	.84	.81	.75	.78

Blok 5 2-back	.87	.92	.82	.79
Blok 5 2-back PI	.88	.88	.82	.86
Blok 6	.94	.95	.94	.94
Blok 7 deel 1	.95	.94	.95	.94
Blok 7 deel 2	.92	.88	.94	.92
Intra-individuele coëfficiënt van variatie				
Blok 1	.72	.77	.60	.72
Blok 2	.61	.66	.59	.52
Blok 3	.67	.69	.65	.60
Blok 4 compatibele trials	.49	.38	.55	.48
Blok 4 incompatibele trials	.51	.43	.60	.45
Blok 5 1-back	.47	.47	.43	.49
Blok 5 2-back	.40	.49	.30	.41
Blok 5 2-back PI	.38	.54	.46	.24
Blok 6	.80	.78	.79	.76
Blok 7 deel 1	.74	.73	.70	.71
Blok 7 deel 2	.70	.69	.66	.63
Aantal fouten				
Blok 2	.70	.78	.71	.57
Blok 3	.53	.61	.60	.39
Blok 4 compatibele trials	.54	.53	.52	.56
Blok 4 incompatibele trials	.73	.75	.73	.72
Blok 5 1-back	.63	.66	.60	.40
Blok 5 2-back	.61	.63	.60	.51
Blok 5 2-back PI	.26	-.13	.14	.34
Blok 7 deel 1	.62	.77	.62	.39
Blok 7 deel 2	.58	.79	.46	.22
Contrastparameters reactiesnelheid				
Blok 2 minus Blok 1	.89	.89	.90	.85
Blok 3 minus Blok 2	.64	.50	.78	.80
Blok 4 minus Blok 2	.86	.88	.78	.81
Blok 7 minus Blok 2	.87	.85	.88	.86
Blok 4 incompatibele trials minus compatibele trials	.22	.21	.27	-.02
Blok 7 deel 1 minus Blok 7 deel 2	.60	.53	.65	.63
Contrastparameters fouten				
Blok 3 minus Blok 2	.41	.57	.49	.13
Blok 4 minus Blok 2	.60	.68	.59	.53
Blok 7 minus Blok 2	.52	.67	.51	.29
Blok 4 incompatibele trials minus compatibele trials	.49	.45	.47	.53
Blok 7 deel 1 minus Blok 7 deel 2	.15	.31	.24	-.05
Blokoeverkoepelende afwijkende responsen				
Aantal premature responsen	.87	.79	.69	.93

Aantal irrelevante knopdrukken	.78	.79	.76	.73
Aantal extra responsen	.98	.95	.94	.99
Aantal responscompetities	.79	.70	.85	.81

Noot. N.v.t. = Niet van toepassing, kan alleen voor gehele taak berekend worden.

* interbeoordelaarsbetrouwbaarheid

2.3 Test-hertest betrouwbaarheid

De test-hertest betrouwbaarheden zijn herberekend middels intra-klasse correlaties voor absolute agreement (Koo en Li, 2006). Dit geeft beter de individuele test-hertest betrouwbaarheid weer dan de groeps-gebaseerde analyses middels correlaties (Pearsons r) zoals nu gepresenteerd in de handleiding (Tabel 8.3), zie ook Bland en Altman, 1986. In Tabel 15 staan de test-hertest gegevens weergegeven. De coëfficiënten zijn over het geheel genomen sterker dan de groepsgebaseerde Pearsons r (Tabel 8.3 handleiding).

Voor 10 van de 20 uitkomstparameters van de Basisrapportage (Leersnelheid is buiten beschouwing gelaten i.v.m. het korte test-hertest interval, vandaar 20 en niet 21), zijn de test-hertest coëfficiënten .70 of hoger. Voor 5 van de 10 overige parameters tussen de .60 en .70. De parameters met de laagste test-hertest betrouwbaarheid liggen rond de .40 en zijn alle in de te verwachte richting. Verder zijn de test-hertest resultaten over het geheel vergelijkbaar voor het subsample kinderen uit de normgroep en het subsample kinderen met ADHD. In ogenschouw genomen dat een COTAPP testuitslag altijd gebaseerd wordt op het geheel van scores in onderlinge samenhang, geven deze resultaten steun voor de generaliseerbaarheid van het cognitieve profiel bemeten met de COTAPP over de tijd heen. Ondanks dat de test-hertestbetrouwbaarheid niet de aangewezen methode is om de betrouwbaarheid vast te stellen bij neuropsychologische tests waar met name reactietijd en ook accuratesse gemeten worden (Farnsworth, 2017), dragen deze resultaten wel bij aan de ondersteuning van de betrouwbaarheid van de COTAPP.

Tabel 21: Test-hertest betrouwbaarheidsgegevens voor de uitkomstmaten van de Basisrapportage en Specialistische rapportage.

Uitkomstmaat	hele groep n=117	normgroep n=59	ADHD n=58
Verwerkingssnelheid			
Reactiesnelheid	.92	.92	.92
Variabiliteit in reactiesnelheid	.76	.76	.77
Reactiesnelheid ten gevolge van cognitieve belasting	.69	.79	.62
Versnellingscapaciteit in reactiesnelheid	.41	.37	.42
Variabiliteit in beslistijd	.70	.71	.67
Aandacht			
Volgehouden aandacht	.60	.36	.67
Arousalregulatie	.40	.62	.24
Onoplettendheid	.65	.66	.60
Executieve controle			
Slordigheid	.43	.52	.26
Interferentiegevoeligheid	.70	.57	.75
Wachttijdafweging	.80	.76	.83
(Werk)geheugen			
Reactiesnelheid eenvoudig	.79	.77	.80

Fouten eenvoudig	.53	.41	.61
Reactiesnelheid complex	.86	.85	.86
Fouten complex	.38	.37	.34*
Reactiesnelheid zeer complex	.75	.70	.78
Fouten zeer complex	.62	.75	.42**
Leersnelheid			
Leersnelheid***	nvt	nvt	nvt
Gedrag tijdens de taak			
Benodigde ondersteuning	.78	.57	.83
Verbaliseren	.90	.93	.86
Beweeglijkheid	.61	.67	.55

SPECIALISTISCHE RAPPORTAGE

Mediane reactiesnelheid

Blok 1	.77	.71	.80
Blok 2	.93	.94	.91
Blok 3	.89	.91	.86
Blok 4 compatibele trials	.89	.88	.89
Blok 4 incompatibele trials	.93	.94	.92
Blok 5 1-back	.79	.77	.80
Blok 5 2-back	.86	.85	.86
Blok 5 2-back PI	.75	.70	.78
Blok 6	.79	.79	.80
Blok 7 deel 1	.86	.83	.87
Blok 7 deel 2	.80	.77	.82

Intra-individuele coëfficiënt van variatie

Blok 1	.45	.50	.42
Blok 2	.49	.56	.38
Blok 3	.62	.73	.46
Blok 4 compatibele trials	.46	.56	.33
Blok 4 incompatibele trials	.62	.61	.62
Blok 5 1-back	.23	.08	.36
Blok 5 2-back	.54	.53	.54
Blok 5 2-back PI	.58	.46	.75
Blok 6	.70	.71	.67
Blok 7 deel 1	.72	.76	.66
Blok 7 deel 2	.62	.57	.65

Aantal fouten

Blok 2	.49	.43	.53
Blok 3	.56	.60	.54
Blok 4 compatibele trials	.42	.38	.46
Blok 4 incompatibele trials	.51	.57	.48

Blok 5 1-back	.53	.41	.61
Blok 5 2-back	.38	.37	.34*
Blok 5 2-back PI	.62	.75	.42**
Blok 7 deel 1	.50	.50	.51
Blok 7 deel 2	.46	.61	.18
Contrastparameters reactiesnelheid			
Blok 2 minus Blok 1	.69	.79	.62
Blok 3 minus Blok 2	.41	.37	.42
Blok 4 minus Blok 2	.70	.57	.75
Blok 7 minus Blok 2	.60	.36	.67
Blok 4 incompatibele trials minus compatibele trials	.48	.06	.63
Blok 7 deel 1 minus Blok 7 deel 2	.40	.62	.24
Contrastparameters fouten			
Blok 3 minus Blok 2	.28	.24	.33
Blok 4 minus Blok 2	.30	.23	.34
Blok 7 minus Blok 2	.19	.06	.31
Blok 4 incompatibele trials minus compatibele trials	.26	.30	.24
Blok 7 deel 1 minus Blok 7 deel 2	.13	.06	.19
Blokoeverkoepelende afwijkende responsen			
Aantal premature responsen	.27	.47	.04
Aantal irrelevante knopdrukken	.36	.24	.46
Aantal extra responsen	.41	.42	.37
Aantal responscompetities	.54	.49	.57

*Coëfficiënt is berekend op 45/58 kinderen met ADHD. Bij 13/58 kinderen is dit niveau niet afgenomen omdat de prestatie op voorgaande niveau te zwak was.

** Coëfficiënt is berekend op 26/58 kinderen met ADHD. Bij 32/58 kinderen is dit niveau niet afgenomen omdat de prestatie op voorgaande niveau's te zwak was. Dit heeft de ICC waarschijnlijk negatief beïnvloed.

*** Test-hertest voor Leersnelheid kan niet beoordeeld worden met een kort test-hertest interval (4-6 weken) omdat kinderen bij hertest de instructie nog kenden van de eerste afname.

2.4 Vergelijking huidige betrouwbaarheidsgegevens met die van de handleiding

Cohen's richtlijnen zoals beschreven in de handleiding zijn bij nader inzien ongeschikt voor het bepalen van de betrouwbaarheid van een test voor het doen van individuele uitspraken. In dit document hanteren we daarom enkel de richtlijn van de COTAN t.a.v. de betrouwbaarheidsbeoordeling van tests voor minder belangrijke beslissingen op individueel niveau (immers, er wordt er nooit enkel op basis van een neuropsychologische test een belangrijke beslissing gemaakt op individueel niveau).

We menen dat met inachtneming van onderstaande overwegingen, de betrouwbaarheid van de COTAPP Basisrapportage als minimaal voldoende beoordeeld kan worden:

1. Betrouwbaarheid Basisrapportage is leidend

Zoals onderbouwd in de handleiding, hanteren gebruikers voor het maken van individueel rapport in principe altijd alleen de Basisrapportage. Psychometrisch zwakkere parameters die klinisch wel relevant kunnen zijn op indicatie worden weergegeven in de Specialistische rapportage. Juist om het goed meewegen van deze psychometrische mindere aspecten bij interpretatie, wordt herhaaldelijk aangegeven dat gebruikers hiervoor ruime ervaring met neuropsychologische diagnostiek moeten hebben (i.e. 2 jarige post-doctorale opleiding tot GZ-psycholoog i.c.m. minimaal 1 jaar ervaring met neuropsychologische diagnostiek).

2. Interne betrouwbaarheid is leidend

Er is een wezenlijk belangrijke onderscheid tussen neuropsychologische tests (zoals de COTAPP) waarbij een sterk beroep wordt gedaan op de op dat moment beschikbare cognitieve capaciteit van het kind (afhankelijk van concentratie, reactiesnelheid en geheugen) en andersoortige psychodiagnostische instrumenten zoals vragenlijsten en prestatietests die aanwezige vaardigheden/kennis meten. Om deze reden wordt afgeraden test-hertest betrouwbaarheid als leidend te laten zijn in de beoordeling van neuropsychologische tests en worden interne betrouwbaarheidsmaten hiervoor beter geschikt geacht (Farnsworth, 2017; Gregory 2011).

3. Split-half betrouwbaarheid is bruikbaar gezien de sterke aannemelijkheid van parallelle van testhelften

In dit document is uitgebreid onderzoek gepresenteerd naar de parallelle van testhelften, waaruit redelijkerwijs de conclusie volgt dat de testhelften parallel zijn. Nu deze onderbouwing gegeven is, kunnen de split-half coëfficiënten meegewogen worden in de bepaling van de betrouwbaarheid van de COTAPP.

4. Op basis van de voorlopige beoordeling zijn we nagegaan welke mogelijkheden er waren de parameters in de Basisrapportage aan te passen t.b.v. verbetering van de betrouwbaarheid

Dit heeft tot 2 belangrijke aanpassingen geleid. De parameter 'Interferentiecontrole' is opnieuw berekend met een betere betrouwbaarheid tot resultaat. Oorspronkelijk was deze parameter gebaseerd op het contrast tussen blok 2 (30 trials) en de compatibele trials uit Blok 2 (20 trials). Bij herberekening middels het contrast tussen blok 2 en alle trials uit blok 4 (40 trials) bleek de nieuwe parameter aanzienlijk betrouwbaarder. De correlatie tussen de nieuwe en oorspronkelijke parameter was .95 voor de ruwe score en .92 voor de gestandaardiseerde score. Deze zeer hoge correlaties geven aan dat de aanpassing in operationalisatie niet heeft geleid tot een inhoudelijk verschil. Uit simulaties bleek dat de hogere betrouwbaarheid van de nieuw geoperationaliseerde parameter verklaard kan worden door een groter aantal trials op basis waarvan de nieuwe parameter berekend is. Een tweede aanpassing betreft het verwijderen van de parameter 'Responsinhibitie' uit de Basisrapportage vanwege een te lage betrouwbaarheid (de parameter wordt alleen nog in de Specialistische rapportage weergegeven). Deze twee aanpassingen hebben ons inziens geleid tot een verbetering van de betrouwbaarheid van de Basisrapportage als geheel.

5. Split-half betrouwbaarheidscoëfficiënten vallen gunstiger uit nu de juiste statistische methode wordt gehanteerd

Zoals beschreven onder punt 5.3 blijken de groepsgebaseerde Pearson correlaties zoals gepresenteerd in de handleiding niet de juiste manier om de betrouwbaarheid op individueel niveau te berekenen. Middels intra-class correlaties voor absolute agreement (Koo Li, 2006) zijn de split-half coëfficiënten (en test-hertest coëfficiënten)

berekend. Voor 14 van de 20 uitkomstparameters van de Basisrapportage (split-half coëfficiënt kan niet voor de parameter 'Leersnelheid' berekend worden, vandaar 20 ipv 21), zijn de split-half coëfficiënten minimaal .70, maar veelal boven de .80. Voor 5 van de 6 overige parameters liggen deze rond de .60. Voor 1 parameter (fouten zeer complex werkgeheugen) is de split-half betrouwbaarheid lager (.26), zeer waarschijnlijk samenhangend met het gebrek aan parallelie voor testhelften voor specifiek deze parameter waardoor de uitkomst van de split-half ICC waarschijnlijk niet goed geschat kan worden. Bevestiging van deze hypothese volgt vanuit een aanmerkelijk hogere test-hertest betrouwbaarheid voor deze parameter (.75). De 6 parameters met een iets lagere betrouwbaarheid laten verder zonder uitzondering goede resultaten voor validiteit zien en zijn om die reden in de Basisrapportage opgenomen. De keuze voor de uiteindelijke weergave van parameters in de Basisrapportage is gebaseerd op een combinatie van betrouwbaarheid en validiteit.

6. Klinisch gebruik van de test is relevant in de beoordeling van betrouwbaarheid (en uiteraard validiteit)

Het is ons inziens relevant in de beoordeling van de betrouwbaarheid mee te wegen dat er **nooit** op basis van individuele uitkomstparameters conclusies worden getrokken, maar altijd op het profiel van sterke en zwakke scores. Tevens wordt neuropsychologische diagnostiek ingezet binnen een proces van klinisch redeneren (analoog aan Bayesiaanse gevolgtrekking), waarbij het cognitieve profiel de a-priori hypothese van bijv een ontwikkelingsstoornis (on)aannemelijker maakt, maar **nooit** als enig instrument gebruikt wordt. Beoordeling van parameters in isolatie van elkaar lijkt ons daarmee niet passend.

3 Criteriumvaliditeit

3.1 Representativiteit steekproeven o.b.v. achtergrondgegevens

Informatie over schoolbeloop was beschikbaar voor +/-83 procent van de normgroep (n=861 ja/nee speciaal onderwijs; n=858 ja/nee vroege start groep 3; n=860 ja/nee klas blijven zitten). Deze groep verschilde niet voor wat betreft leeftijd, geslacht, of geschat IQ van de totale normgroep (Tabel 22). Ook de groepen met LVS gegevens op verschillende domeinen verschilden niet van de totale normgroep (Tabellen 23-27). Uitzondering vormt de relatief kleine groep met LVS gegevens van Studievaardigheden. Deze LVS toets wordt op z'n vroegst vanaf groep 5 afgenomen, meestal pas in groep 7 en 8. Dit verklaart de hogere leeftijd van deze steekproef t.o.v. de normgroep. Op overige kenmerken is ook deze subgroep vergelijkbaar met de normgroep.

Tabel 22: Karakteristieken steekproef met beschikbare gegevens over schoolbeloop.

	Normgroep	Subgroep schoolbeloop
Leeftijd	M=9.4, SD=2.0, min=5.5, max=13.5	M=9.3, SD=1.9, min=5.5, max=13.5
% meisjes	50.0	50.9
Geschat IQ	M=100, SD=16.2, min=45, max=145	M=102, SD=15.5, min=45, max=145

Tabel 23: Karakteristieken steekproef met beschikbare LVS Rekenen-Wiskunde gegevens.

	Normgroep	Subgroep LVS gegevens n=863
Leeftijd	M=9.4, SD=2.0, min=5.5, max=13.5	M=9.5, SD=1.8, min=6.0, max=13.5
% meisjes	50.0	51.6
Geschat IQ	M=100, SD=16.2, min=45, max=145	M=101, SD=15.8, min=52, max=145

Tabel 24: Karakteristieken steekproef met beschikbare LVS DMT gegevens.

	Normgroep	Subgroep LVS gegevens n=831
Leeftijd	M=9.4, SD=2.0, min=5.5, max=13.5	M=9.6, SD=1.8, min=6.0, max=13.5
% meisjes	50.0	51.6
Geschat IQ	M=100, SD=16.2, min=45, max=145	M=101, SD=15.9, min=52, max=145

Tabel 25: Karakteristieken steekproef met beschikbare LVS Spelling gegevens.

	Normgroep	Subgroep LVS gegevens n=832
Leeftijd	M=9.4, SD=2.0, min=5.5, max=13.5	M=9.5, SD=1.8, min=6.0, max=13.5
% meisjes	50.0	52.4
Geschat IQ	M=100, SD=16.2, min=45, max=145	M=102, SD=15.5, min=52, max=145

Tabel 26: Karakteristieken steekproef met beschikbare LVS Begrijpend lezen gegevens.

	Normgroep	Subgroep LVS gegevens n=741
Leeftijd	M=9.4, SD=2.0, min=5.5, max=13.5	M=10.0, SD=1.6, min=6.7, max=13.5
% meisjes	50.0	53.6
Geschat IQ	M=100, SD=16.2, min=45, max=145	M=102, SD=15.4, min=52, max=145

Tabel 27: Karakteristieken steekproef met beschikbare LVS Studievaardigheden gegevens.

	Normgroep	Subgroep LVS gegevens n=98
Leeftijd	M=9.4, SD=2.0, min=5.5, max=13.5	M=11.3, SD=0.9, min=9.3, max=13.5
% meisjes	50.0	48.5
Geschat IQ	M=100, SD=16.2, min=45, max=145	M=101, SD=15.2, min=52, max=136

4 Referenties

Bechger, T., Hemker, B. & Maris, G. (2009). *Over het gebruik van continue normering*. Arnhem: CITO.

Bland, J.M., and Altman, D.G. 1986. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*, 1, 307-310.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Lawrence Erlbaum Associates.

Danner, D., Hagemann, D., Schankin, A., Hager, M., Funke, J. (2011). Beyond IQ: A latent state-trait analysis of general intelligence, dynamic decision making, and implicit learning. *Intelligence*, 39,323–334.

Evers, A., Lucassen, W., Meijer, R. & Sijtsma, K. (2010).

COTAN beoordelingssysteem voor de kwaliteit van tests (geheel herziene versie). Amsterdam: NIP.

Farnsworth JL 2nd, Dargo L, Ragan BG, Kang M. (2017). Reliability of Computerized Neurocognitive Tests for Concussion Assessment: A Meta-Analysis. *J Athl Train*, Sep;52(9):826-833.

Gregory RJ (2011). Concepts of Reliability. In RJ Gregory, *Psychological testing: History, principles and applications* (6th ed.) (pp. 87-108). Pearson Education, Boston.

Grob, A., Hagmann-von Arx, P., Ruiter, S., Timmerman, M. & Visser, L. (2018).

Koo, T.K. Li, M.Y. (2006). *A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research*. *J Chiropr Med*;15(2):155-63.

Mellenbergh, G. (2011). *A conceptual introduction to psychometrics* Den Haag: Eleven International Publishing.

Neisser, U., Boodoo, G., Bouchard, Jr T.J., Boykin, AW, Brody, N., Ceci, S.J., Halpern, D.F., Loehlin, J.C., Perloff, R., Sternberg, R.J., Urbina, S. (1996). Intelligence: Knowns and unknowns. *American Psychologist* 51, 77-101.

Polderman, T.J., Gosso, M.F., Posthuma, D., Van Beijsterveldt, T.C., Heutink, P., Verhulst, F.C., Boomsma, D.I. (2006). A longitudinal twin study on IQ, executive functioning, and attention problems during childhood and early adolescence. *Acta Neurologica Belgica*, 106, 191-207.

Rigby, R. A., & Stasinopoulos, D. M. (2005). Generalized additive models for location, scale and shape. *Journal of the Royal Statistical Society Series C, Applied Statistics*, 54, 507–554.<http://dx.doi.org/10.1111/j.1467-9876.2005.00510.x>.

Rommelse, N.N., Altink, M.E., Oosterlaan, J., Buschgens, C.J., Buitelaar, J., Sergeant, J.A. (2008). Support for an independent familial segregation of executive and intelligence endophenotypes in ADHD families. *Psychological Medicine*, 38, 1595-1606.

Rommelse, N., Hartman, C., Brinkman, A., Slaats-Willemse, D., de Zeeuw, P. & Luman, M. *COTAPP. COGNITIEVE TEST APPLICATIE. Test voor aandacht, informatieverwerking en executieve functies. Handleiding*. Amsterdam: Boom uitgevers.

Rommelse, N., Luman, M., Kievit, R. (2020). Slow processing speed: a cross-disorder phenomenon with significant clinical value, and in need of further methodological scrutiny. *European Child Adolescent Psychiatry*. DOI 10.1007/s00787-020-01639-9.

Swaab, H., Bouma, A., Hendriksen, J., Konig, C. (red.) (2016). *Klinische Kinderneuropsychologie* (Editie 2). Boom: Amsterdam.

Timmerman, M. E., Voncken, L., & Albers, C. J. (2020). A tutorial on regression-based norming of psychological tests with GAMLSS. *Psychological Methods*. <https://osf.io/mdc9u>.