

# Zeven instrumenten voor het meten van behandelangst. Betrouwbaarheid en convergente validiteit

Angst voor een tandheelkundige behandeling begint soms al in de kindertijd. Verschillende instrumenten kunnen bij kinderen deze vorm van angst meten. In dit onderzoek zijn de betrouwbaarheid en de convergente validiteit van 7 meetinstrumenten beschreven. Er bleek geen 'gouden standaard' te bestaan waarmee de validiteit van een instrument ondubbelzinnig is vast te stellen. Bovendien bleek dat zelfrapportageschalen relatief vaker zijn onderzocht dan gedragsobservatie-instrumenten, terwijl deze laatste meetmethode in de praktijk vaker wordt gebruikt. De meeste instrumenten zijn voldoende betrouwbaar, maar correleren onderling soms onvoldoende. Dit laatste heeft mogelijk te maken met het multidimensionele karakter van het begrip 'angst' en, in sommige gevallen, met de soort meetmethode. In vrijwel geen enkel onderzoek werden vooraf normen bepaald waarmee statistische indices ondubbelzinnig kunnen worden geïnterpreteerd. In het onderhavig onderzoek zijn deze normen wel vastgesteld.

Reinders JJ, Sorghabi W. Zeven instrumenten voor het meten van behandelangst. *Betrouwbaarheid en convergente validiteit*. *Ned Tijdschr Tandheelkd* 2009; 116: 551-557

## Inleiding

Angst voor tandheelkundige behandelingen bij kinderen vormt een serieus probleem dat meer consequenties heeft dan men doorgaans zou kunnen vermoeden. Vermijding van mondzorg is slechts 1 van de talrijke problemen (Aartman et al, 1996). Er zijn meerdere 'paden' waarlangs een dergelijke angst zich kan ontwikkelen tot, in het meest ernstige geval, een fobie of een bijzondere vorm van angststoornis (American Psychiatric Association, 2000). Vooral kinderen zijn vatbaar voor het ontwikkelen van dergelijke angsten. Uit het onderzoek van Ten Berge et al (2002) is gebleken dat 6% van de Nederlandse kinderen tussen 4 en 11 jaar oud een grote mate van angst voor tandheelkundige behandeling vertoont, terwijl 8% zich in de 'gevaarzone' bevindt.

Uit onderzoek blijkt dat angst voor tandheelkundige behandeling het resultaat is van een complexe interactie van verschillende factoren (Ten Berge, 2001; Veerkamp en Majstorovic, 2006). Daarnaast lijkt angstig gedrag van kinderen veel meer gecorreleerd aan directe conditionering dan aan indirect leren of aan psychologische factoren (Locker et al, 1999). Deze directe conditionering betreft dan de associatie die ontstaat tussen de behandeling en de negatieve consequenties van die behandeling. Het verschil tussen milde en extreme vormen van behandelangst manifesteert zich niet altijd even duidelijk. Een betrouwbaar en valide meetinstrument heeft een grote diagnostische waarde en kan bepalend zijn voor de vervolgbehandeling (Klingberg, 1995).

Er bestaat geen recente literatuur waarin instrumenten

## Wat weten we?

Van de Nederlandse kinderen tussen 4 en 11 jaar oud heeft 14% in meer of mindere mate last van serieuze tandartsangst. Er bestaan echter weinig recente onderzoeksartikelen waarin een vergelijking plaatsvindt van instrumenten die tandheelkundige behandelangst bij kinderen meten. Bovendien bestaan er geen overzichtsartikelen die verschillende meetmethoden vergelijken. Normen waarop kwaliteitsuitspraken rond betrouwbaarheid en validiteit zijn gebaseerd, zijn vrijwel nooit transparant.

## Wat is nieuw?

Een overzichtsartikel waarin een vergelijking plaatsvindt van instrumenten van verschillende meetmethoden die tandheelkundige behandelangst bij kinderen meten is nieuw. Daarnaast zijn expliciete normen gehanteerd voor verschillende statistische maten rond betrouwbaarheid en validiteit. Deze expliciete normen maken een vergelijking tussen onderzoeken en statistische indices beter mogelijk en leiden bovendien tot meer transparantie. Voorts is duidelijk geworden dat er geen 'gouden standaard' bestaat voor instrumenten die behandelangst bij kinderen meten. Zonder vergelijking met een oorspronkelijk instrument en een goede consensus over het construct 'tandheelkundige fobie', is het dus lastig de (convergente) validiteit te bepalen.

## Praktijktoepassing

Zowel de meettijd als het exacte meetdoel van een instrument is een belangrijke afweging voor gebruik in praktijk-situaties. De kwaliteit zou echter van een even groot belang moeten zijn. De meeste instrumenten blijken voldoende betrouwbaar.

van verschillende meetmethoden die angst voor tandheelkundige behandeling bij kinderen meten, worden vergeleken. Wel bestaat er literatuur waarin meerdere instrumenten van 1 methode worden vergeleken (Aartman et al, 1998). Daarnaast is niet transparant op welke normen kwalificaties rond betrouwbaarheid en validiteit zijn gebaseerd (Van de Plassche et al, 2003).

In totaal zijn 4 meetmethoden te onderscheiden: (zelf) rapportageschalen, gedragsobservatie, projectieve technieken en fysiologische metingen. Dit artikel richt zich specifiek op de jonge patiëntencategorie en vergelijkt 7 meetinstrumenten, inclusief enkele gemodificeerde versies, van 3 verschillende meetmethoden. De keuze voor deze 7 instrumenten is gemaakt na een initiële zoektocht in PubMed. Daaruit bleek dat deze 7 instrumenten recent het meest zijn onderzocht. Daarnaast is gekozen voor de beschrijving

Subjects:	humans
Language:	English, Dutch
Age:	all child: 0-18 year
Publication date:	from 1980
Key words:	reliability, convergent, congruent, validity, measures, dental, anxiety, fear, children

Tabel 1. Het zoekprofiel PubMed.

van een beperkt aantal instrumenten omwille van de overzichtelijkheid. De fysiologische meetmethode wordt in dit onderzoek buiten beschouwing gelaten omdat een dergelijk meting de tandheelkundige behandeling stoort, de interpretatie van bijvoorbeeld hartslagmetingen ook gerelateerd kan zijn aan de meethandeling zelf en er zeer weinig onderzoek is gedaan naar de betrouwbaarheid en validiteit van dergelijke metingen (Aartman et al, 1996). Diverse indices van betrouwbaarheid, afhankelijk van meetniveau en meetmethode, worden in dit artikel naast elkaar gelegd. Met betrouwbaarheid wordt bedoeld de mate van overeenstemming van een meting.

Voorts wordt in dit artikel aandacht besteed aan de validiteit van de diagnostische testen van angst voor tandheelkundige behandeling bij kinderen. Deze validiteit betreft de mate waarin de test meet wat hij zou moeten meten. In het onderzoek daarnaar werd een specifieke vorm van validiteit onderzocht: convergente validiteit. Convergente validiteit geeft de mate van samenhang aan tussen de gebruikte test en een andere soortgelijke oorspronkelijke test. Deze oorspronkelijke test zou men een 'gouden standaard' kunnen noemen.

Het belang van betrouwbaarheid en validiteit van diagnostische instrumenten is voor een tandarts algemeen-practicus zeer groot. Een juiste diagnose maakt een verwijzing mogelijk die relevant is voor de aard en de ernst van het probleem. De zorg is immers niet alleen voor nu, de zorg is ook voor later: zowel kinderen als angsten kunnen groeien. Een sterkere angst kan tot een sterkere vermijding leiden en van vermijding tot ongrijpbaarheid, van ongrijpbaarheid tot gebrek aan zorg en uiteindelijk tot verslechtering van de mondgezondheid.

## Materiaal en methode

Voor het selecteren van de literatuur werden het systematisch literatuuronderzoek en het literatuuronderzoek volgens de sneeuwbal methode gehanteerd. Het systematisch literatuuronderzoek werd uitgevoerd met behulp van PubMed waarbij de volgende afzonderlijke zoektermen werden gehanteerd: reliability, convergent, congruent, validity, measures, dental, anxiety, fear, children (tab. 1). De inhoudelijke relevantie van de gevonden literatuur werd bepaald aan de hand van de titel en de samenvatting.

Omwille van de uitputtendheid werd tevens gebruikgemaakt van de sneeuwbal methode. Additionele literatuur werd verzameld via oorspronkelijke bronnen en referentielijsten. Voorts werd deze werkwijze toegepast op alle nieuwe artikelen en bronnen. Hierbij werd de inclusie van nieuwe

Index	'Voldoende'	'Goed'	Bron
Cronbach's alpha (interne consistentie)	$\alpha > 0,70$	$\alpha > 0,80$	Nunnally, 1978; Nunnally en Bernstein, 1994
Correlatiecoëfficiënt (test-hertest)	$r > 0,80$	$r > 0,90$	Newton en Buck, 2000
Cohen's kappa (interbeoordelaarsbetrouwbaarheid)	$\kappa > 0,60$	$\kappa > 0,80$	Landia en Koch, 1977
Correlatiecoëfficiënt (interbeoordelaarsbetrouwbaarheid)	$r > 0,75$	$r > 0,90$	Streiner en Norman, 1995 Bonett, 2002
Percentage overeenstemming (interbeoordelaarsbetrouwbaarheid)	PO > 0,80	PO > 0,95	Rojahn en Schreuder, 1991 Krishnan et al, 2002
Correlatiecoëfficiënt (convergente validiteit)	$r > 0,60$	$r > 0,80$	Newton en Buck, 2000

Tabel 2. Normen per statistische index.

literatuur bepaald door dezelfde normen te hanteren die ook voor de eerdere systematische zoekstrategie werden gebruikt (tab. 1).

Bij het verzamelen van informatie over de psychometrische kwaliteit zoals betrouwbaarheid en validiteit, werden verschillende indices verzameld. Daarnaast werd ook een onderscheid gemaakt tussen de aard van de meetmethode, de testduur en het scorebereik per meetinstrument.

Om de interpretatie en uniforme vergelijking van verschillende indices rond betrouwbaarheid en validiteit eenduidig te maken, waren vooraf statistische normen voor de kwalificaties 'voldoende' en 'goed' vastgesteld. Deze zogenaamde 'kwaliteitsnormen' zijn ontleend aan (methodologische) onderzoeken naar betrouwbaarheid en validiteit.

Als maat voor de betrouwbaarheid van een meetinstrument bestaan verschillende indices. Voor dit onderzoek waren 3 relevante indices te onderscheiden: de interne consistentie, de test-hertestcorrelatie en de interbeoordelaarsbetrouwbaarheid. Bij een meetinstrument dat gebruik maakt van diverse items is het van belang de interne samenhang of consistentie vast te stellen. Een samenvattende maat voor de interne consistentie is Cronbach's alfa. Een andere vorm van betrouwbaarheid is de test-hertestcorrelatie (intrabeoordelaarsbetrouwbaarheid) en de interbeoordelaarsbetrouwbaarheid, uitgedrukt in een correlatie, Cohen's kappa of in een percentageovereenstemming (PO). De convergente validiteit werd berekend door 2 testen met elkaar te correleren (tab. 2). Om een transparante interpretatie van de resultaten mogelijk te maken, zijn voor elke statistische index vooraf normen voor 'voldoende' en 'goed' vastgelegd. Bovendien moeten deze normen leiden tot minder interpretatieverschillen tussen onderzoeken en de vergelijkbaarheid van verschillende statistische indices vergroten.

Meetmethode	Instrument	Aantal items	Scorebereik	Betrouwbaarheid		Convergente validiteit	Bron	n	
				Cronbach's interne consistentie	Test-hertest				
ZR	CFSS-DS	15	15-75	$\alpha = 0,90$ (G) $\alpha = 0,85$ (G)	r = 0,97 (G) r = 0,74 (O)	r = 0,87 met CDFP (G) r = -0,22 met FBRS (O) r = -0,68 met FBRS (V) r = -0,26 met FBRS (O) r = 0,78 met MCDAS (V) r = 0,80 met MCDASf (V) r = 0,60 met SFP (O)	Ten Berge et al, 1998	718	
							Arapostathis et al, 2008	260	
							Klingberg, 1994	52	
							Arapostathis et al, 2008	260	
							Klingberg 1995	146	
							Klorman et al, 1979	181	
							Klorman et al, 1978	101	
							Klorman et al, 1978	101	
							Wong et al, 1998	314	
							Howard en Freeman, 2007	206	
ZR	MCDAS	8	8-40	$\alpha = 0,84$ (G) $\alpha = 0,86$ (G)		r = 0,60 met SFP (O) r = 0,73 met DAS (V) r = 0,78 met CFSS-DS (V)	Buchanan, 2005	464	
							Wong et al, 1998	314	
							Christophorou et al, 2000	286	
							Buchanan, 2005	464	
ZR	MCDASf	8	8-40	$\alpha = 0,82$ (G)	r = 0,80 (O)	r = 0,80 met CFSS-DS (V)	Wong et al, 1998	314	
							Howard en Freeman, 2007	287	
							Howard en Freeman, 2007	287	
ZR	SFP	4	4-28	$\alpha = 0,80$ (V)	r = 0,80 (O)	r = 0,60 met MCDAS (O) r = 0,60 met CFSS-DS (O)	Howard en Freeman, 2007	206	
							Buchanan, 2005	464	
GO	VABRS	2	0-10			r = 0,78 (V) r = 0,98 (G)	Buchanan, 2005	100	
							Buchanan, 2005	241	
GO	FBRS	4	1-16	*	*	r = 0,82 (V)	Buchanan, 2005	241	
							meer onderzoek nodig	26	
GO	FBRS	4	1-16	*	*	r = 0,82 (V)	Venham et al, 1979	26	
							Venham et al, 1979	26	
							Venham et al, 1979	26	
							Del Gaudio en Nevid, 1991	440	
							$\kappa = 0,80$ (V)	Kantovitz et al, 2007	20
							$\kappa = 0,98$ (G)	Nakai et al, 2000	361
							r = -0,22 met CFSS-DS (O)	Klorman et al, 1979	181
							r = -0,68 met CFSS-DS (V)	Klorman et al, 1978	101
							r = -0,26 met CFSS-DS (O)	Klorman et al, 1978	101
							r = 0,44 met DAS (O)	Neverlien en	163
PT	CDFP	30	nvt			PO = 88,9% (V)	Backer Johnsen, 1991	23	
							r = -0,54 met CFSS-SF (O)	Folayan en Otuyemi, 2002	23
							Klingberg et al, 1995	146	
							Klingberg et al, 1995	146	

Type meetmethode: ZR = (zelf) rapportageschalen, GO = gedragsobservatie en PT = projectieve technieken

Instrumenten: CFSS-DS = Children's Fear Survey Schedule, Dental Subscale; MCDAS = Modified Child Dental Anxiety Scale; MCDASf = Modified Child Dental Anxiety Scale, Faces version; SFP = Smiley Faces Program; VABRS = Venham Anxiety and Behavior Rating Scales; FBRS = Frankl Behavior Rating Scale; CDFP = Children's Dental Fear Picture test

Kwaliteit: onvoldoende (O), voldoende (V) of goed (G)

$\alpha$  = Cronbach's alfa,  $\kappa$  = Cohen's kappa, r = correlatiecoëfficiënt, PO = percentage overeenstemming

Statistische normen: interne consistentie [ $\alpha > 0,70 = V$ ,  $\alpha > 0,80 = G$ ]; test-hertest [ $r > 0,80 = V$ ,  $r > 0,90 = G$ ];

Interbeoordelaarbetrouwbaarheid [ $\kappa > 0,60 = V$ ,  $\kappa > 0,80 = G$  of  $r > 0,75 = V$ ,  $r > 0,90 = G$  of  $PO > 0,80 = V$  of  $PO > 0,95 = G$ ];

Convergente validiteit [ $r > 0,60 = V$ ,  $r > 0,80 = G$ ]

\* = wisselt door modificaties

Tabel 3. Overzicht meetinstrumenten.

## Resultaten

Uit dit literatuuronderzoek bleek dat het merendeel van de gevonden onderzoekspublicaties betrekking had op zelfrapportageschalen. De 7 onderzochte meetinstrumenten zijn te verdelen in 3 van de 4 typen methoden voor angstmeting: (zelf)rapportageschalen, gedragsobservatie en projectieve technieken (tab. 3). Hoewel de gedragsobservatie volgens Klingberg (1995) op dit gebied de meest gebruikte methode is, is de zelfrapportage als methode vaker onderzocht op betrouwbaarheid en validiteit (tab. 3).

### *Children's Fear Survey Schedule, Dental Subscale*

De Children's Fear Survey Schedule, Dental Subscale (CFSS-DS) is een zelfrapportageschaal (Cuthbert en Melamed, 1982). Deze telt 15 items die gerelateerd zijn aan de mate waarin een tandheelkundige behandeling invasief is. Op elke vraag kan op een vijfpuntenschaal worden gescoord (1 = helemaal niet bang; 5 = heel erg bang). Zodoende varieert de eindscore van 15 tot 75. Scores vanaf 38 en hoger zijn geassocieerd met klinische angst voor tandheelkundige behandeling (Klingman et al, 1984). Meestal wordt de CFSS-DS gebruikt in onderzoeken die de prevalentie van angst voor tandheelkundige behandeling onderzoeken naast mogelijke voorspellers en andere samenhangende factoren van behandelangst.

De betrouwbaarheid is doorgaans goed. De Nederlandse versie van het instrument heeft een Cronbach's alfa van 0,90 (Ten Berge et al, 1998). Bij de Griekse versie is een Cronbach's alfa van 0,85 gevonden (Arapostathis et al, 2008). Verder is er een test-hertestcorrelatie gevonden van  $r = 0,97$  ( $p < 0,001$ ) (Klingberg, 1994a).

Er is een grote correlatie gevonden tussen de CFSS-DS en de Children's Dental Fear Picture Test (CDFP). De CDFP verklaarde 76% van de variantie in scores van de CFSS-DS ( $r = 0,87$ ,  $p < 0,001$ ) (Klingberg et al, 1995).

### *Modified Child Dental Anxiety Scale*

De Modified Child Dental Anxiety Scale (MCDAS) is een op Corah's Dental Anxiety Scale gebaseerde zelfrapportageschaal voor kinderen (Corah, 1969). De schaal bevat 8 vragen over specifieke situaties van een tandheelkundige behandeling, zoals lokale anesthesie, extractie en het polijsten van het gebit. Een vijfpunten Likert-schaal wordt gebruikt om de mate van angst te bepalen. De eindscore ligt tussen 8 (= weinig of geen behandelangst) en 40 (= extreme behandelangst).

De interne consistentie is goed (Cronbach's alfa = 0,84) (Wong et al, 1998). Bij de Griekse versie van het instrument is een Cronbach's alfa van 0,86 gevonden (Christophorou et al, 2000). De correlatieve samenhang tussen de MCDAS, de Smiley Faces Program (SFP), de Dental Anxiety Scale (DAS voor volwassen patiënten) en de CFSS-DS is onvoldoende tot voldoende:  $r =$  respectievelijk 0,60, 0,73 en 0,78 (Wong et al, 1998; Buchanan, 2005).

### *Modified Child Dental Anxiety Scale, Faces version*

Een alternatieve vorm voor Modified Child Dental Anxiety Scale is de 'Faces version' (MCDASf). Hierbij is een analoge

gezichtenschaal toegevoegd boven de oorspronkelijke numerieke vorm. Op deze wijze wordt invullen van de zelfrapportageschaal gemakkelijker en dus ook mogelijk voor jonge kinderen.

De betrouwbaarheid is voldoende tot goed. De interne consistentie uitgedrukt in Cronbach's alfa is 0,82, terwijl de test-hertestcorrelatie 0,80 is. De MCDASf correleert voldoende met de CFSS-DS ( $r = 0,80$ ;  $p < 0,001$ ) (Howard en Freeman, 2007).

### *Smiley Faces Program*

De Smiley Faces Program (SFP) is een digitale zelfrapportageschaal met 4 items waarbij afbeeldingen van gezichtsuitdrukkingen worden gebruikt. De items hebben betrekking op: 1. het krijgen van een tandheelkundige behandeling op de volgende dag, 2. zitten in de wachtkamer, 3. boren in een gebitselement en 4. het krijgen van lokale anesthesie door middel van een injectie. Op een computerscherm kan het kind 1 van de 7 gezichtsuitdrukkingen (4 = 'neutraal') kiezen die het beste bij zijn eigen emotie past.

De interne consistentie blijkt voldoende (Cronbach's alfa = 0,80), maar de test-hertestcorrelatie toont onvoldoende samenhang ( $r = 0,80$ ;  $p < 0,001$ ) (Buchanan, 2005). In hetzelfde onderzoek is tevens gecorreleerd met MCDAS ( $r = 0,60$ ;  $p < 0,01$ ) en de CFSS-DS ( $r = 0,60$ ;  $p < 0,01$ ). Deze laatste resultaten duiden op onvoldoende samenhang.

### *Venham Anxiety and Behavior Rating Scales*

De Venham Anxiety and Behavior Rating Scales (VABRS) is een gedragsobservatietest die bestaat uit 2 items waarmee behandelangst en oncoöperatief gedrag van het kind kunnen worden gemeten. Elk schaal omvat 6 scoremogelijkheden (0 = geen angst of erg coöperatief gedrag; 5 = grote angst of erg oncoöperatief gedrag). Hogere scores indiceren een grotere mate van behandelangst of weinig coöperatie.

Voldoende tot grote correlaties zijn gevonden tussen de gedragsbeoordelingen van getrainde observatoren ( $r = 0,78$  en  $r = 0,98$ ) (Venham et al, 1979). Door wisselende resultaten en afwezigheid van correlaties met andere meetinstrumenten, zal de validiteit van VABRS nog nader moeten worden onderzocht (Aartman et al, 1996).

### *Frankl Behavior Rating Scale*

Met de Frankl Behavior Rating Scale (FBRS) kunnen mond-zorgverleners het gedrag van een kind in verschillende situaties categoriseren. Met behulp van een vierpunteschaal is het gedrag te typeren als 'zeer positief', 'positief', 'negatief' en 'zeer negatief'. De oorspronkelijke gedragsobservatielijst bestaat uit 4 items met een scorebereik van 1 tot 16 (Frankl et al, 1962). Er bestaan echter meerdere gemodificeerde versies van de FBRS (Aartman et al, 1996).

De interbeoordelaarsbetrouwbaarheid is in ten minste 3 verschillende onderzoeken aan de orde geweest en is voldoende tot goed. Del Gaudio en Nevid (1991) vonden een grote betrouwbaarheid van  $r = 0,82$  ( $p < 0,001$ ). In 2 andere onderzoeken zijn kappacoëfficiënten gevonden van 0,80 en 0,98 (Nakai et al, 2000; Kantovitz et al, 2007).

De correlatie met andere instrumenten zoals de DAS, de CFSS-DS en de verkorte versie hiervan (CFSS-SF) is gering (Klorman et al, 1978; Klorman et al, 1979; Folyan en Otuymi, 2002; Neverlien en Backer Johnsen, 1991).

### Children's Dental Fear Picture test

De Children's Dental Fear Picture test (CDFP) is gebaseerd op een projectieve meetmethode en heeft 3 verschillende subtesten: 1. de Dental Setting Pictures (CDFP-DS), 2. de Pointing Pictures (CDFP-PP) en 3. de Sentence Completion Task (CDFP-SC). Het totale aantal items is 30 (tab. 3).

De interbeoordelaarsbetrouwbaarheid van de subtesten CDFP-DS, -PP en -SC was respectievelijk 81,5%, 85,2% en 85,2%. De interbeoordelaarsbetrouwbaarheid voor het totale instrument was 88,9%. Dit duidt op een voldoende betrouwbaarheid. Daarnaast verklaarde de CDFP 76% van de variantie in de scores van CFSS-DS ( $r = 0,87$ ;  $p < 0,001$ ) (Klingberg en Hwang, 1994; Klingberg et al, 1995).

### Discussie

In dit onderzoek is bij 7 meetinstrumenten die angst voor tandheelkundige behandeling bij kinderen meten de betrouwbaarheid en de convergente validiteit beschreven. Gegeven de vooraf vastgestelde normen, is de betrouwbaarheid van deze 7 meetinstrumenten doorgaans voldoende.

Een kanttekening bij het onderhavige onderzoek is de bruikbaarheid van de gehanteerde normen. De keuze van deze normen, zoals de indeling van Landis en Koch (1977) voor Cohen's kappa, kan ter discussie worden gesteld. Echter, daar staat tegenover dat de verschillende onderzoeken die dezelfde statistische maten hanteren ook op uniforme wijze met elkaar kunnen worden vergeleken. Bovendien zijn er opvallend veel onderzoeken waarin niet expliciet de grens tussen 'onvoldoende', 'voldoende' en 'goed' wordt aangegeven. Zo kan het gebeuren dat de ene onderzoeker een correlatie als 'klein' betitelt, terwijl een andere onderzoeker dezelfde correlatie 'aanzienlijk', 'goed' of 'groot' zal noemen.

De term 'validiteit' wordt vaak als een 'parapluterm' gebruikt, maar behelst in feite een scala aan psychometrische nuances. In dit onderzoek wordt validiteit gedefinieerd, terwijl in sommige validiteitsonderzoeken het begrip validiteit onduidelijk blijft.

Vanwege het ontbreken van een unanieme definiëring van validiteit is het lastig sterke validiteitsuitspraken te doen. Wel is het interessant dat sommige zelfrapportageschalen slecht met elkaar correleren. Dit is mogelijk toe te schrijven aan het multidimensionele karakter van het begrip 'angst'. Er is tevens een kanttekening te plaatsen ten aanzien van correlaties tussen instrumenten van verschillende meetmethoden. De vooraf vastgestelde statistische normen zijn geschikt voor vergelijkingen tussen meetinstrumenten van dezelfde meetmethode, maar correlaties tussen verschillende meetmethoden vallen vrijwel altijd kleiner uit (Aartman et al, 1998). Wellicht is het nodig minder strenge statistische normen te hanteren wanneer instrumenten van verschillende meetmethoden met elkaar worden vergeleken. Een kritische vraag is in welke mate verschillende

statistische maten voor betrouwbaarheid, zoals percentage overeenstemming en Cohen's kappa, zijn te vergelijken. Niettemin zijn in dit onderzoek gelijke kwaliteitsnormen gehanteerd en is de strengheid van de statistische normen per maat aangepast om een gelijkwaardige vergelijking te bevorderen.

Een beperking van dit onderzoek is het ontbreken van informatie over de klinische toepasbaarheid van de instrumenten. Dit komt vooral doordat hier in de literatuur weinig aandacht aan wordt besteed. Verder is in dit onderzoek geen rekening gehouden met selectiecriteria als de leeftijd van kinderen. In sommige onderzoeken hebben kinderen (of hun ouders) een zelfrapportageschaal ingevuld direct vóór de behandeling, in andere onderzoeken hebben ze dat thuis gedaan. Het is niet ondenkbaar dat het moment en de situatie van een meting van invloed zijn geweest op de mate van angst.

### Conclusie

Hoewel de 7 onderzochte instrumenten hetzelfde beogen te meten, bestaat er soms weinig of geen samenhang tussen de testen. Op basis van de resultaten van dit onderzoek kan niet zonder meer worden vastgesteld welke van de beschreven instrumenten het meest valide of bruikbaar is. Uit dit onderzoek blijkt dat er geen duidelijke 'gouden standaard' bestaat en dat waarschijnlijk niet elk instrument exact dezelfde dimensies van het multidimensionele begrip 'angst' kan meten. Met 'gouden standaard' wordt bedoeld dat niet 1 maar meerdere testen als referentiepunt worden gebruikt bij het vaststellen van de validiteit. De uitkomsten van verschillende meetinstrumenten kunnen elkaar overlappen.

Wel kan, op basis van dit onderzoek, een aantal praktische aanbevelingen worden gedaan. De betrouwbaarheid van de diverse instrumenten is doorgaans voldoende, hetgeen betekent dat er geen overweging hoeft te zijn het ene instrument boven het andere te verkiezen. Tijdsinvestering is wel een praktische overweging. Daarvoor kan het totale aantal items per meetinstrument als uitgangspunt worden genomen. In dat geval zou de SFP het meest bruikbare instrument zijn. Daarentegen is de meetmethode ook zeer bepalend voor de tijdsinvestering. Projectieve methoden zijn arbeidsintensiever dan zelfrapportageschalen. Gedragsobservaties zijn doorgaans het minst tijdrovend indien deze slechts door een enkele persoon worden uitgevoerd. Vanzelfsprekend kosten videoregistratie en het bepalen van een eindoordeel door meerdere beoordelaars meer tijd. Tot slot wordt de keuze bepaald door de gewenste specificiteit van het antwoord.

Wanneer men vooral is geïnteresseerd is in hoeverre het kind angstig is voor invasieve behandelingen, zoals 'boren' en 'injecties', ten opzichte van minder invasieve behandelingen, kiest men de CFSS-DS. Wil men nagaan welke situaties angst oproepen, dan kiest men voor instrumenten als de MCDAS of de voor kinderen meer 'gebruiksvriendelijke' MCDASf. De SFP is geschikt voor het meten van angst voor allerlei situaties of voor het meten van diverse soorten invasieve behandelingen. Het gebruik van de SFP vereist echter de beschikbaarheid van een computer. Wanneer de

mate van oncoöperatief gedrag centraal staat, is de VABRS geschikt. Evenwel kunnen bij jonge kinderen (4 tot 6 jaar), naast de behandelangst, ook andere factoren oncoöperatief gedrag veroorzaken, zoals van de moeder gescheiden zijn of een nieuwe, onbekende situatie (Aartman et al, 1996). De FBRS lijkt een zeer handzaam instrument voor gedragsobservatie en wordt om die reden veel in de praktijk gebruikt. Het wordt onder andere veel gebruikt als selectie-instrument bij onderzoeken met sedatie. De CDFP kan worden gekozen wanneer de directe beleving van het kind onderwerp van onderzoek is. Het kind kan meer zijn eigen unieke verhaal kwijt. Deze projectieve test is echter zeer arbeidsintensief en is lastig te interpreteren.

Vervolgonderzoek is nodig naar de validiteit van meetinstrumenten voor het meten van angst voor tandheelkundige behandeling (bij kinderen). De convergente validiteit van de Venham Anxiety and Behavior Rating Scale (VABRS) is namelijk nauwelijks onderzocht. Bovendien zijn zelfrapportageschalen vaker onderzocht dan observatie-instrumenten. Observatie-instrumenten worden in de praktijk echter het meest toegepast. Voorts is het van belang een 'gouden standaard' voor angstmeting te ontwikkelen zodat kan worden nagegaan hoe valide de bestaande instrumenten werkelijk zijn. Deze 'gouden standaard' zou bijvoorbeeld een multidimensionale meting van de behandelangst bij kinderen mogelijk moeten maken. In vervolgonderzoek zou men moeten nagaan welke test een 'gouden standaard' zou kunnen zijn of waaraan een test zou moeten voldoen om een dergelijk referentiepunt voor andere testen te kunnen vormen.

#### Literatuur

- \* Aartman IHA, Everdingen T van, Hoogstraten J, Schuurs AHA. Appraisal of behavioral measurement techniques for assessing dental anxiety and fear in children: a review. *J Psychopathol Behav Assess* 1996; 18:153-171.
- \* Aartman IH, Everdingen T van, Hoogstraten J, Schuurs A. Self-report measurements of dental anxiety and fear in children: a critical assessment. *ASDC J Dent Child* 1998; 65: 252-258, 229-230.
- \* American Psychiatric Association. Diagnostic and statistical manual of mental disorders. Washington: American Psychiatric Association, 2000.
- \* Arapostathis KN, Coolidge T, Emmanouil D, Kotsanos N. Reliability and validity of the Greek version of the Children's Fear Survey Schedule-Dental Subscale. *Int J Paediatr Dent* 2008; 18: 374-379.
- \* Berge ten M. Dental fear in children. Prevalence, aetiology and risk factors. Amsterdam: Universiteit van Amsterdam, 2001. Academisch proefschrift.
- \* Berge ten M, Hoogstraten J, Veerkamp JSJ, Prins PJM. The Dental Subscale of the Children's Fear Survey Schedule: a factor analytic study in the Netherlands. *Community Dent Oral Epidemiol* 1998; 26: 340-343.
- \* Berge ten M, Veerkamp JS, Hoogstraten J, Prins PJM. Childhood dental fear in the Netherlands: prevalence and normative data. *Community Dent Oral Epidemiol* 2002; 30: 101-107.
- \* Bonett, DG. Sample size requirements for estimating intraclass correlations with desired precision. *Statistics in Medicine* 2002; 21: 1331-1335.
- \* Buchanan H. Development of a computerised dental anxiety scale for children: validation and reliability. *Br Dent J* 2005; 199:359-362.
- \* Christophorous S, Lee GTR, Humphris G. The reliability and validity of the Modified Child Dental Anxiety Scale: a study of Greek Cypriot school children. *Eur J Paediatr Dent* 2000; 1: 75-81.
- \* Corah NL. Development of a dental anxiety scale. *J Dent Res* 1969; 48: 596.
- \* Cuthbert MI, Melamed BG. A screening device: children at risk for dental fears and management problems. *ASDC J Dent Child* 1982; 49: 432-436.
- \* Del Gaudio DJ, Nevid JS. Training dentally anxious children to cope. *ASDC J Dent Child* 1991; 58: 31-37.
- \* Folan MO, Otuyemi OD. Reliability and validity of a short form of the dental subscale of the child fear survey schedule used in a Nigerian children population. *Niger J Med* 2002; 11: 161-163.
- \* Frankl SN, Shiere FR, Fogels HR. Should the parent remain with the child in the dental operatory? *ASDC J Dent Child* 1962; 29: 150-163.
- \* Howard KE, Freeman R. Reliability and validity of a faces version of the Modified Child Dental Anxiety Scale. *Int J Paediatr Dent* 2007; 17: 281-288.
- \* Kantovitz KR, Puppini-Rontani RM, Gaviao MB. Sedative effect of oral diazepam and chloral hydrate in the dental treatment of children. *J Indian Soc Pedod Prev Dent* 2007; 25: 69-75.
- \* Klingberg G. Dental fear and behavior management problems in children. A study of measurement, prevalence, concomitant factors, and clinical effects. *Swed Dent J* 1995; 103 (Suppl.): 1-78.
- \* Klingberg G. Reliability and validity of the Swedish version of the Dental Subscale of the Children's Fear Survey Schedule, CFSS-DS. *Acta Odontol Scand* 1994a; 52: 255-256.
- \* Klingberg G, Hwang CP. Children's Dental Fear Picture test (CDFP): a projective test for the assessment of child dental fear. *ASDC J Dent Child* 1994b; 61: 89-96.
- \* Klingberg G, Löfqvist LV, Hwang CP. Validity of the Children's Dental Fear Picture test (CDFP). *Eur J Oral Sci* 1995; 103: 55-60.
- \* Klingman A, Melamed BG, Cuthbert MI, Hermecz DA. Effects of participant modeling on information acquisition and skill utilization. *J Consult Clin Psychol* 1984; 52: 414-422.
- \* Klorman R, Michael R, Hilpert PL, Sveen OA. A further assessment of predictors of the child's behavior in dental treatment. *J Dent Res* 1979; 58: 2338-2343.
- \* Klorman R, Ratnem J, Arata CLG, King JB, Sveen OA. Predicting the child's uncooperativeness in dental treatment from maternal trait, state and dental anxiety. *ASDC J Dent Child* 1978; 45: 62-67.
- \* Krishnan LV, O'Kane KS, Gill-Body KM. Reliability of a modified version of the Dynamic Gait Index-a pilot study. *Neurol Rep* 2002; 26: 8-14.
- \* Landis JR, Koch GG. The measurement of observer agreement for categorical data. *Biometrics* 1977; 33: 159-174.
- \* Locker D, Liddell A, Dempster L, Shapiro A. Age of onset of dental anxiety. *J Dent Res* 1999; 78: 790-796.
- \* Nakai Y, Hirakawa T, Milgrom P, et al. The Children's Fear Survey Schedule-Dental Subscale in Japan. *Community Dent Oral Epidemiol* 2005; 33: 196-204.
- \* Nakai Y, Milgrom P, Mancl L, Coldwell SE, Domoto PK, Ramsay DE. Effectiveness of local anesthesia in pediatric dental practice. *J Am Dent Assoc* 2000; 131: 1699-1705.
- \* Neverlien PO, Backer Johnsen TO. Optimism-pessimism dimension and dental anxiety in children aged 10-12 years. *Community Dent Oral Epidemiol* 1991; 19: 342-346.

- \* *Newton JT, Buck DJ* Anxiety and pain measures in dentistry: a guide to their quality and application. *J Am Dent Assoc* 2000; 131: 1449-1457.
- \* *Nunnally J* Psychometric theory. New York/London: McGraw-Hill, 1978.
- \* *Nunnally J, Bernstein J* Psychometric theory. New York/London: McGraw-Hill, 1994.
- \* *Plassche EE van de, Jaspers JPC, Pelkwijk BJ ter, Linden van den Heuvel GFE van, Oort RP van* Angst voor tandheelkundige behandelingen. Een overzicht van meetinstrumenten. *Ned Tijdschr Tandheelkd* 2003; 110: 395-398.
- \* *Rojahn J, Schroeder SR* Behavioral assessment. In: Matson JL, Mulick JA (eds.). *Handbook of Mental Retardation*. New York: Pergamon Press. 1991.
- \* *Streiner DL, Norman GR* Health measurement scales: A practical guide to their development and use. New York: Oxford University Press. 1995.
- \* *Veerkamp JSJ, Majstorovic M* Angst voor tandheelkundige behandeling en naaldangst bij kinderen. Een verband? *Ned Tijdschr Tandheelkd* 2006; 113: 226-229.
- \* *Venham LL, Murray P, Gaulin-Kremer E* Child-rearing variables affecting the preschool child's responses to dental stress. *J Dent Res* 1979a; 58: 2042-2045.
- \* *Venham LL, Murray P, Gaulin-Kremer E* Personality factors affecting the preschool child's responses to dental stress. *J Dent Res* 1979b; 58: 2046-2051.
- \* *Wong HM, Humphris GM, Lee GTR* Preliminary validation and reliability of the Modified Child Dental Anxiety Scale. *Psychol Rep* 1998; 83:1179-1186.

## Summary

### Seven means of measuring fear of treatment: reliability and convergent validity

*Fear of dental treatment sometimes already begins during childhood: Different kinds of diagnostic instruments can measure children's dental anxiety. In this study the reliability and convergent validity of 7 instruments are compared. Results show that there is no 'gold standard' with which convergent validity can be established unambiguously. Furthermore, self report measures are investigated more often than observational measures even though the latter are more commonly used. Generally, the compared measures are satisfactory in respect to their reliability. Nevertheless, sufficient interest coherence is sometimes lacking. This could be explained by the multidimensional character of the construct 'fear' and, in some cases, the type of measuring-method. In almost none of the compared studies were norms defined prior to the study by means of which statistical indicators could be unambiguously interpreted. In the present study, such norms were established.*

## Bron

J.J. Reinders<sup>1,2</sup>, W. Sorghabi<sup>3</sup>

Uit <sup>1</sup>de sectie Mondgezondheidszorg en Klinische Epidemiologie van het Centrum voor Tandheelkunde en Mondzorgkunde, Universitair Medisch Centrum Groningen, <sup>2</sup>het lectoraat Transparante Zorgverlening van de Hanzehogeschool Groningen en uit <sup>3</sup>het Centrum voor Tandheelkunde en Mondzorgkunde van het Universitair Medisch Centrum Groningen

Datum van acceptatie: 12 februari 2009

Adres: J.J.Reinders, UMCG, postbus 196, 9700 AD Groningen

j.j.reinders@med.umcg.nl

## Dankwoord

De auteurs bedanken mevrouw F.H. Bloemers voor haar inhoudelijk kritische feedback en aanvullingen op het onderzoek.