

De Sociale Plaatjes Test

Onderzoek naar de betrouwbaarheid en validiteit



Faculteit der Maatschappij- en Gedragwetenschappen
Universiteit van Amsterdam
Roeterstraat 15
1018 WB Amsterdam

Auteurs: Claire Verwey (0119822)
Jet Vleeming (0107085)
Begeleiding: Brenda Jansen
Datum: Februari 2005
Vakgroep: Klinische Ontwikkelingspsychologie

VOORWOORD	3
SAMENVATTING	4
1. INLEIDING	5
2. PILOTSTUDY	9
2.1 METHODE	9
<i>Proefpersonen</i>	9
<i>Materialen</i>	9
<i>Procedure</i>	10
2.2 RESULTATEN.....	11
<i>Proefpersonen</i>	11
<i>Agressievragenlijst</i>	11
<i>Afname ervaringen</i>	11
<i>Bepalen van het scoringsstelsel</i>	12
<i>Ontwerp van de definitieve versie van de test</i>	14
3. ONDERZOEK	14
3.1 METHODE	14
<i>Proefpersonen</i>	14
<i>Materialen</i>	15
<i>Procedure</i>	15
3.2. RESULTATEN.....	15
<i>Proefpersonen</i>	15
<i>Betrouwbaarheid agressievragenlijst</i>	15
<i>Betrouwbaarheid van de Sociale Plaatjes Test</i>	16
<i>Factoranalyse</i>	17
<i>Verschillen tussen agressieve en niet-agressieve kinderen</i>	17
<i>Lerarenoordeel</i>	18
<i>Relatie tussen intentie en score op de categorieën van de Sociale Plaatjes Test</i>	19
<i>Leeftijdverschillen in reacties op de Sociale Plaatjes Test</i>	22
4. DISCUSSIE.....	25
<i>Conclusie</i>	25
<i>Verklaring van onverwachte resultaten</i>	26
<i>Tekortkomingen van het onderzoek</i>	26
<i>Suggesties voor toekomstig onderzoek</i>	26
<i>Aansluiting bij het SIV-model</i>	27
LITERATUUR	28

Voorwoord

Met veel plezier en toewijding hebben wij het afgelopen half jaar gewerkt aan dit werkstuk. We zijn begonnen met het verzamelen van theoretische achtergrond informatie over het ontwerpen van een instrument dat sociaal gedrag meet. Nadat ons werkstukconcept was goedgekeurd konden we echt aan de slag met het ontwikkelen van het instrument. Dit hadden we niet kunnen doen zonder de hulp van Daphne Hoolwerf die op korte termijn alle 34 plaatjes voor ons heeft getekend. Daarna wilden we de kwaliteit van het instrument onderzoeken waarvoor we de medewerking van Basisschool De Golfbreker uit Koog aan de Zaan hebben gekregen. Hiervoor willen we ze graag bedanken. Na de verwerking van deze resultaten konden we aan het echte onderzoek beginnen. We willen basisschool de Zonnewijzer uit Hoorn bedanken voor hun gastvrijheid en hun bereidheid om mee te werken aan ons onderzoek. Ook basisschool het Dok willen wij bedanken voor hun medewerking. Natuurlijk willen wij ook alle kinderen bedanken die mee hebben gedaan aan ons onderzoek, zij hebben ons diverse onvergetelijke momenten bezorgd met hun open, originele en humoristische antwoorden. Tevens willen we Marieke Ooijevaar bedanken die ons heeft geholpen om de interbeoordelaarsbetrouwbaarheid te kunnen berekenen. In het bijzonder willen we Brenda Jansen bedanken die ons gedurende het hele proces met raad en advies terzijde heeft gestaan en gezorgd heeft voor inspiratie als wij die kwijt waren.

Jet Vleeming
Claire Verwey

Samenvatting

Volgens het Sociale Informatie Verwerkingsmodel (SIV) verloopt het proces van sociale informatieverwerking in vijf opeenvolgende stappen. In dit onderzoek werd geprobeerd een betrouwbaar en valide plaatjesinstrument te ontwikkelen dat de stappen twee en vier van het SIV-model meet en onderscheid kan maken tussen agressieve en niet-agressieve kinderen. Hiervoor werden 140 kinderen uit groep 2 tot en met 8 van een reguliere basisschool getest en 30 kinderen van een speciale basisschool. De kinderen maakten de Sociale Plaatjes Test (SPT) en de leerkrachten vulden voor elk kind een agressievragenlijst in. De SPT bleek een betrouwbaar instrument te zijn maar voor de hypothesen die validiteit betroffen werd geen steun gevonden. In de steekproef van dit onderzoek zaten nauwelijks agressieve kinderen. Verwacht wordt dat in een onderzoek met agressieve kinderen de validiteit wel bewezen kan worden zodat er een instrument ontstaat dat bruikbaar is in de praktijk.

1. Inleiding

Informatie verwerkingstheorieën beschrijven de mentale processen die verantwoordelijk zijn voor de respons op een stimulus. In sociale informatie verwerkingstheorieën wordt verondersteld dat sociaal gedrag kan worden voorspeld door de wijze waarop mensen informatie verwerken. Een theorie die veel in handboeken wordt beschreven is het Sociale Informatie Verwerkingsmodel (SIV) van Dodge (bijvoorbeeld in Kievit, Tak & Bosch, 2002 en Sameroff, Lewis & Miller, 2000). Het model beschrijft de cognitieve taken van perceptie en probleemoplossingvaardigheden en de emotionele taken om deze informatie te integreren met iemands eigen doelen, motivatie en arousal regulatie (Dodge, 1993). Er ontstaan individuele verschillen tussen kinderen in hun reacties op sociale situaties omdat ze verschillen in onder andere biologische capaciteiten en ervaringen in het verleden (Arsenio en Lemerise, 2004). Het SIV-model wordt ook gebruikt om te verklaren waarom sommige kinderen wel en andere niet agressief reageren in dezelfde of vergelijkbare situaties.

Het proces van sociale informatieverwerking verloopt volgens Dodge in een bepaalde sequentie van mentale stappen (Dodge, 1993). Bij kinderen met gedragsproblemen zijn er verschillende tekorten bij het uitvoeren van de stappen.

De eerste stap is het decoderen van relevante signalen uit de omgeving via sensorische input, selectieve aandacht voor sociale stimuli en het opslaan van die informatie in het korte termijn geheugen. De omgeving is zo complex en overweldigend dat een kind leert om zijn aandacht selectief te richten op bepaalde belangrijke aspecten van de omgeving. Het komt bijvoorbeeld meer van pas om te letten op de gezichtsuitdrukkingen van iemand tijdens een gesprek dan te kijken naar zijn kleding. Bij het encoderen richt de aandacht van agressieve kinderen zich eerder op agressieve signalen.

De tweede stap is het interpreteren van de gedecodeerde informatie waarbij een betekenis wordt gegeven aan de inkomende informatie. Het kind vormt een mentale representatie van de stimulus, welke wordt opgeslagen in het geheugen. Bij het interpreteren wordt een signaal door een agressief kind eerder als agressief geïnterpreteerd. Als een agressief kind bijvoorbeeld langs een groepje jongens loopt die lachen, kan dit door het kind geïnterpreteerd worden als het uitlachen van hemzelf, terwijl de jongens net zo goed kunnen lachen om een grap die ze net hebben gehoord.

De mentale representatie die in de tweede stap is gevormd roept in de derde stap één of meerdere gedrags en affectieve responsen op. Deze responsen worden in het geheugen gezocht en/of er worden nieuwe responsen gegenereerd. Deze responsen kunnen bewust of onbewust zijn zoals verbale en motorische reacties of autonome arousal en het ervaren van affect. Bij het zoeken naar een respons kennen agressieve kinderen meer agressieve en minder prosociale responsen dan hun leeftijdsgenoten. Tevens denken zij eerder aan agressieve responsen.

Kinderen willen of kunnen niet alle gevonden responsen uitvoeren. In de vierde stap wordt dan ook een respons gekozen. De responsen worden beoordeeld op hun moraliteit, aanvaardbaarheid en op de voorziene consequenties. Voor het kiezen van een respons overwegen agressieve kinderen minder consequenties van het eigen gedrag, zij evalueren een agressieve respons positiever dan een prosociale respons.

De vijfde stap is het uitvoeren van de gekozen respons. Veel gedrag heeft oefening en herhaling nodig om goed uitgevoerd te worden, hierdoor beïnvloeden de vaardigheden van een kind ook het uitvoeren van het gedrag. Voor het uitvoeren van de respons hebben agressieve kinderen maar een beperkt sociaal gedragsrepertoire en zij zijn niet zo goed in het uitvoeren van een prosociale respons (Nieuwenhuijzen, Wijnroks & Vermeer, 2000; Dodge, 1993).

Het model zoals hierboven beschreven biedt theoretisch inzicht in het sociale verwerkingsproces, maar om hier in de praktijk gebruik van te kunnen maken zal het model meetbaar gemaakt moeten worden. Met een instrument kan men er wellicht achter komen welke stappen verschillend verlopen waardoor bijvoorbeeld het ene kind agressief op een bepaalde situatie reageert en het andere kind niet. Met deze informatie kan er wellicht een gerichtere interventie plaatsvinden. Wanneer een kind bijvoorbeeld verkeerd interpreteert kan de interventie gericht worden op de juiste interpretatie van sociale situaties en niet op het gedrag. Dodge, Bates, Pettit & Valente (1995) gebruikten in hun onderzoek korte videofragmenten waarin twee kinderen een dagelijkse sociale situatie naspeelden. De proefpersonen kregen de opdracht zich in te leven in één van de twee kinderen in het videofragment. De fragmenten eindigden in een negatieve uitkomst voor het kind met wie ze zich moesten identificeren. De intentie van het andere kind was agressief, ambigu of goedaardig. Om stap één van het Dodge model te meten moest het kind direct na het zien van de negatieve uitkomst het fragment samenvatten. Deze samenvatting werd dan beoordeeld op de relevantie van de aspecten. Voor het meten van stap drie moesten de proefpersonen vertellen hoe zij zouden reageren in dezelfde situatie. Daarna werden de fragmenten verder afgespeeld waarin drie mogelijke reacties te zien waren, namelijk een agressieve reactie, een passieve reactie en een competente assertieve reactie. De proefpersonen moesten deze reacties evalueren op een schaal van 1 tot 4. Hiermee werd stap vier gemeten. Voor het meten van stap twee liet men de kinderen stripverhaaltjes zien met dezelfde soort situaties als de videofragmenten. Ook bij de strip moest het kind zich identificeren met één van de twee kinderen. De proefpersoon moest daarna aangeven hoe en waarom het andere kind zich op deze manier gedroeg.

De videofragmenten en stripverhalen van Dodge et. al. (1995) hebben een aantal nadelen. Ten eerste is er een videorecorder nodig die wellicht niet in iedere diagnostische setting aanwezig is. Deze instrumenten zijn door de schrijvers zelf gemaakt en wel getest met een pilot studie maar niet gestandaardiseerd. Verder worden niet alleen ambigue situaties gebruikt maar tevens duidelijke situaties. Uit onderzoek blijkt dat alleen bij gebruik van ambigue situaties verschillen worden gevonden in responsen van kinderen (Graybill & Heuvelman, 1993). Verder meet men met een videofragment wel eerst het encoderen van de kinderen, maar de kinderen worden niet verbeterd als er relevante dingen worden genegeerd. Hierdoor hebben kinderen niet dezelfde situatie voor ogen waardoor niet goed kan worden gemeten hoe kinderen stap drie en vier doorlopen op basis van dezelfde informatie.

Videofragmenten en stripverhalen zijn wel een goede vorm voor een instrument voor kinderen omdat deze kinderen meer aanspreken dan een vragenlijst en in mindere mate een beroep doen op het taalvermogen van een kind. Om met deze informatie een standaard instrument te maken dat makkelijk te hanteren is maar kinderen wel aanspreekt, kunnen er stripverhaaltjes worden gebruikt die slechts één plaatje bevatten. Een verouderd instrument dat hierbij als uitgangspunt kan dienen is de Rosenzweig Picture Frustration Test (RPF).

De RPF is een instrument waarin 24 plaatjes met een alledaagse sociale situatie worden getoond. Dit instrument is bedoeld om de reacties van kinderen in frustrerende situaties te meten. Als theoretische achtergrond voor het ontwerpen en scoren van het instrument gebruikt Rosenzweig een frustratie-agressie theorie (Rosenzweig, Fleming & Rosenzweig, 1946). Sommige plaatjes laten ego-blocking situaties zien, waarin obstakels, persoonlijke of niet-persoonlijke teleurstellingen en andere situaties die direct de persoon frustreren te zien zijn. Een aantal andere situaties worden beschreven als superego-blocking, wat betekent dat de persoon beschuldigd wordt van iets door iemand anders. Oorspronkelijk werd de RPF voor volwassenen ontworpen, maar vrijwel direct nadat dit instrument klaar was, werd duidelijk dat de methode beter geschikt zou zijn voor afname bij kinderen. De wijze van afname lijkt op een spel voor kinderen. Hierdoor verwachtten Rosenzweig et al. dat de kinderen minder sociaal wenselijke antwoorden zouden geven.

Op de plaatjes van het instrument staan minstens twee personen. Eén van de personen zegt iets en het kind moet aangeven wat hij zou antwoorden als hij de andere persoon was in deze situatie. Het kind identificeert zich bewust of onbewust met de figuur met een lege tekstballon. Het instrument is geschikt voor kinderen van 4 tot en met 13 jaar. Bij oudere kinderen is het de bedoeling dat ze zelf hun antwoorden in de tekstballon schrijven en kan het instrument individueel of groepsgewijs afgenomen worden. Bij kinderen onder de acht jaar wordt het instrument individueel afgenomen en schrijft de proefleider het antwoord van het kind op.

De antwoorden worden gescoord in elf verschillende categorieën van type responsen. De scores op deze elf categorieën worden gecombineerd om vervolgens zes categorieën van agressie te vormen: drie richtingen van agressie (naar buiten gerichte, naar binnen gerichte en het ontwijken van agressie) en drie typen van agressie (incident staat centraal, eigen persoon staat centraal en de oplossing staat centraal). Er werden ongeveer 600 afnamen gescoord en geanalyseerd op basis waarvan de normen werden gemaakt. Bij de normering zijn vijf leeftijdsgroepen onderscheidden (4-5, 6-7, 8-9, 10-11 en 12-13 jaar).

In 1978 deed Rosenzweig een onderzoek naar de betrouwbaarheid van de RPF. Hij onderzocht de twee leeftijdsgroepen 10-11 jaar en 12-13 jaar. Uit dit onderzoek bleek dat de betrouwbaarheid van de RPF significant was voor de meeste categorieën. Hierbij moet wel vermeld worden dat de betrouwbaarheden matig waren.

De validiteit van de RPF bleek in verschillende onderzoeken onder de maat te zijn (Graybill, Peterson & Williams, 1989; Graybill & Heuvelman, 1993). Graybill en Heuvelman meenden dat de theoretische achtergrond van de RPF verantwoordelijk was voor de slechte validiteit van het instrument. Zij deden onderzoek om te kijken of de RPF verbeterd kon worden door het SIV-model als theoretische achtergrond te gebruiken. In zowel de frustratie-agressie theorie van Rosenzweig als de theorie van Dodge wordt aangenomen dat er individuele verschillen zijn in de reacties op frustratie. Het SIV-model voegt hier echter aan toe dat er een mediërende variabele is, namelijk het idee dat kinderen hebben over de intentie van de persoon die de frustratie veroorzaakt (stap twee). Daarbij voorspelt Dodge dat er meer verschillen tussen agressieve en niet-agressieve kinderen zullen worden gevonden wanneer de intentie van de frustrerende persoon ambigu is.

Uit het onderzoek van Graybill en Heuvelman (1993) bleek inderdaad dat bij de items van de RPF waar duidelijk was of de situatie agressief of juist niet-agressief was alle kinderen de situatie hetzelfde interpreteerden en dat wanneer de situatie ambigu was er onderscheid te

maken viel tussen agressieve en niet-agressieve kinderen. De ambigue situaties werden door agressieve kinderen eerder geïnterpreteerd als agressief waardoor zij ook eerder een agressieve respons gaven dan niet-agressieve kinderen.

In vorige onderzoeken bleek ook dat de scores op de RPF niet overeenkwamen met het daadwerkelijke gedrag van de kinderen. Dit kan veroorzaakt zijn doordat slechts 8 van de 24 plaatjes ambigu zijn zodat het instrument weinig onderscheidend vermogen heeft. Uit de resultaten van Graybill en Heuvelman (1993) bleek dat het daadwerkelijke gedrag van de kinderen beter overeenkwam met hun antwoorden wanneer alleen de scores van de ambigue items werd gebruikt. Een nieuw instrument zal daarom enkel ambigue situaties moeten bevatten omdat een instrument pas waardevol is als de cognities van het kind, die in de antwoorden naar voren komen, overeenkomen met zijn/haar daadwerkelijke gedrag.

Volgens de COTAN beoordeling van 1992 door Evers, Van Vliet-Mulder en Ter Laak waren de betrouwbaarheid en validiteit van de RPF onvoldoende waardoor het instrument uit de COTAN werd geschrapt. Daarnaast zijn de plaatjes van de RPF erg verouderd, hierdoor sluiten ze niet goed meer aan bij de leefwereld van kinderen. Toch wordt het instrument nog door psychologen gebruikt in individueel onderzoek. Een reden hiervoor is dat er weinig andere instrumenten zijn die de cognitie achter het sociale gedrag van kinderen bestuderen. De Sociale Interpretatie-Test (Vijftigschild, Berger en Van Spaendonck, 1969), perspectiefneem taken zoals die van Selman (1979) en Theory of Mind taken zoals die van Ponsioen (2001) meten weliswaar sociale cognitie, maar meten niet specifiek de cognitie in frustrerende situaties. Vragenlijsten die gericht zijn op agressief gedrag refereren vaak naar gedrag in algemene situaties, bijvoorbeeld de Gedragsvragenlijst voor Kinderen (CBCL; F.C. Verhulst, J. van der Ende, H.M. Koot, 2003). Omdat het idee van de plaatjes een goede basis voor een instrument is, zal het doel van dit onderzoek zijn te onderzoeken of de RPF, met als theoretische achtergrond het SIV-model, een beter bruikbaar instrument zal worden dat discrimineert tussen agressieve en niet-agressieve kinderen. Daarbij zal een vereenvoudigde scoringsmethode worden gebruikt en moderne plaatjes worden ontwikkeld welke uitsluitend ambigue situaties voorstellen. Door de plaatjes eenvoudig te houden zullen er bij het encoderen (stap één) vrijwel geen verschillen zijn tussen kinderen. Voor het meten van stap twee zal bij ieder plaatje een vraag toegevoegd worden over de intentie van de frustrerende persoon. Met de reactie op de situaties wordt stap vier gemeten. Daarnaast is het doel van dit onderzoek om te weten te komen of er verschillen zijn in reageren op ambigue sociale situaties tussen kinderen van verschillende leeftijden. Uit eerder onderzoek blijkt namelijk dat jongere kinderen meer ruzie maken, bedreigen, vechten of dingen stuk maken en oudere kinderen vaker liegen (Stranger, Achenbach & Verhulst, 1997).

Met betrekking tot het bepalen van de betrouwbaarheid zijn er twee hypothesen opgesteld. Ten eerste wordt voorspeld dat het instrument een hoge interne consistentie zal hebben. Ten tweede wordt verwacht dat de antwoorden van de respondenten objectief gescoord kunnen worden door ontwikkelingspsychologen. Ook voor het bepalen van de validiteit van het instrument zijn er twee hypothesen opgesteld. Het instrument zal verschillen meten tussen agressieve en niet-agressieve kinderen en de scores op het instrument zullen samenhangen met de beoordeling van agressiviteit door de leraar. Wanneer aan de eisen van betrouwbaarheid en validiteit wordt voldaan, kan onderzocht worden of de score op de intentievraag overeenkomt met de score op het instrument. Verder kan nog worden gekeken hoe kinderen van verschillende leeftijden verschillen in het geven van antwoorden.

In dit onderzoek zal de Rosenzweig Picture Frustration Test worden aangepast en vernieuwd. Hiervoor zullen de door Graybill en Heuvelman (1993) aangewezen ambigue situaties

gebruikt worden. Daarnaast zullen de plaatjes moeten worden aangepast aan deze tijd en zullen er nieuwe plaatjes worden bedacht met ambigue situaties. Ook zal een vraag naar de intentie van de frustrerende persoon worden toegevoegd. Eerst zal er een pilotstudy plaatsvinden om te kijken of er plaatjes aangepast of verwijderd moeten worden. De pilotstudy moet ook informatie geven over de mogelijkheid om het instrument schriftelijk af te nemen bij kinderen van acht jaar en ouder. Daarnaast zal met behulp van de resultaten van de pilotstudy een nieuw scoringssysteem worden ontwikkeld. Vervolgens zal het verbeterde instrument afgenomen worden in de groepen 2 tot en met 8 van een reguliere basisschool en bij een kleine groep kinderen van een basisschool voor speciaal onderwijs.

2. Pilotstudy

2.1 Methode

Proefpersonen

In totaal deden er 62 kinderen van basisschool de Golfbreker in Koog aan de Zaan mee aan de pilotstudy. Deze kinderen zaten in groep 2 en groep 5. Een van de onderzoekers kent de directeur van deze basisschool en zo is het contact gelegd en werd toestemming verkregen om op deze school het onderzoek uit te voeren. De ouders kregen een brief met informatie over het onderzoek, zij konden deze brief op school inleveren wanneer ze geen toestemming wilden verlenen voor deelname van hun kind. Voor de pilotstudy werden twee versies van het instrument gebruikt, de kinderen werden at random aan een versie toegewezen.

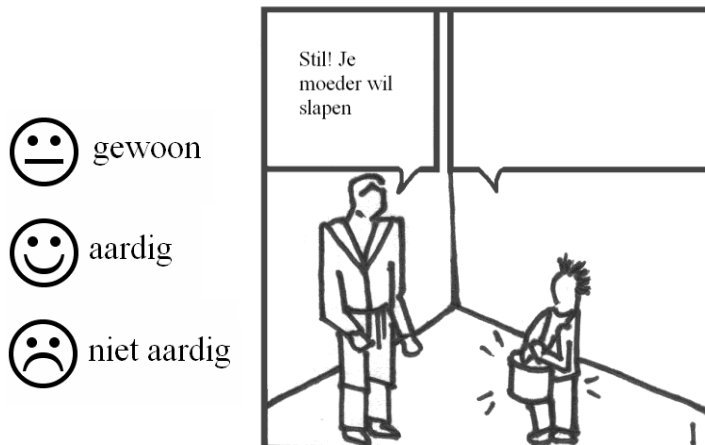
Materialen

Voor het meten van de reacties van kinderen in frustrerende situaties werd een plaatjesinstrument gebruikt. In totaal waren er 34 verschillende items die werden verdeeld over twee versies van de vragenlijst. Elke versie bestond uit 20 items. Elk item stelde een plaatje van een ambigue situatie voor. Een deel van de plaatjes is gebaseerd op het al bestaande instrument van Rosenzweig maar deze plaatjes zijn opnieuw getekend en aangepast aan deze tijd. Het andere deel van de plaatjes is ontwikkeld op basis van eigen ervaring in dagelijkse situaties. Op de plaatjes stonden minstens twee personen, waarbij één van de personen telkens iets zei. Het kind moest dan aangeven wat hij of zij zou antwoorden in de situatie. Op de helft van de plaatjes werd door een volwassene iets gezegd waarop het kind moest reageren terwijl op de andere helft van de plaatjes door een kind iets gezegd wordt. Op één van de plaatjes ziet men bijvoorbeeld een vader met een jongetje dat met een trommel speelt waarbij de vader zegt: ‘Stil, je moeder wil slapen’.

Naast elk plaatje stonden drie ‘smiley’s’: een lachend, een neutraal en een boos gezichtje. Met deze ‘smiley’s’ werd gevraagd naar de intentie van de pratende persoon. De kinderen konden antwoord op deze vraag geven door één van de drie gezichtjes naast het plaatje aan te kruisen. Deze vraag werd gescoord op een drie-puntsschaal namelijk, aardig, gewoon en niet aardig, waarbij een lachend gezichtje stond voor aardig, een neutraal gezichtje voor gewoon en een droevig gezichtje voor niet aardig. Voor elk kind werd dan het aantal keer aardig, gewoon en niet aardig opgeteld. Elk aantal kon variëren tussen de 0 en de 20 punten.

De reacties op de plaatjes werden gescoord in 21 verschillende categorieën van type responsen. Deze categorieën varieerden van niet-agressieve, bijvoorbeeld excuses aanbieden,

tot erg agressieve antwoorden, bijvoorbeeld agressief ontkennen. De antwoorden van de kinderen werden beoordeeld op een drie-puntsschaal, waarbij de antwoordmogelijkheden waren: past geheel niet in deze categorie (1), past een beetje in deze categorie (2) en past helemaal in deze categorie (3). Wanneer alle antwoorden waren beoordeeld, werden de punten per categorie opgeteld. Hierdoor ontstond er een totaalscore per categorie, waarbij minimaal 20 punten en maximaal 60 punten gescoord konden worden.



Tenslotte werd er nog een vragenlijst over agressie (zie bijlage 1) voor elk kind door de leerkracht ingevuld (Waschbusch, Wiloughby & Pelham 1998, $\alpha = .92-.95$) Deze oorspronkelijk Engelse vragenlijst bestond uit zes items waarbij drie items over proactieve agressie gingen (bijvoorbeeld: gebruikt fysiek geweld om te domineren) en drie items over reactieve agressie (bijvoorbeeld: geeft anderen de schuld bij ruzies). Proactieve agressie is agressie die gebruikt wordt om anderen te domineren en om een bepaald doel te bereiken. Reactieve agressie komt voort uit het gevoel bedreigd te worden en uit het verkeerd beoordelen van de bedoelingen van anderen (Dodge, Lochman, Harnish, Bates & Petit, 1997). De antwoorden op de vragenlijst werden gemeten op een vijf-puntsschaal, waarbij de antwoordmogelijkheden lagen tussen nooit en bijna altijd. Er kon een minimum worden gescoord van 6 en een maximum van 36 punten. Hierbij stond 6 punten voor geheel niet agressief en 36 punten voor erg agressief.

Procedure

Binnen een tijdsbestek van drie dagen werd het onderzoek op de school afgenomen. Voor groep 5 werd het instrument klassikaal afgenomen. Onder toezicht van de docent werd het instrument onder schooltijd afgenomen. Aan de hand van een vast protocol legden de proefleiders aan de kinderen de procedure uit. Er werd ze verteld dat er geen foute antwoorden waren en dat het geen toets was. De instructie (zie bijlage 2) werd gezamenlijk doorgelezen en er werden twee voorbeelden behandeld. Deze voorbeelden dienden om de werking van het invullen van het instrument duidelijk te maken. Dit waren geen voorbeelden van ambigue situaties maar een duidelijk aardige en een duidelijk onaardige situatie ('Ik vind je trui mooi' en 'wat ben jij stom'). Bij het eerste voorbeeld werd een aantal voorbeelden van antwoorden gegeven. De kinderen mochten dan ook zelf zeggen of opschrijven wat ze zouden invullen. Bij het tweede voorbeeld mochten de kinderen meteen zelf het antwoord geven. Bij beide voorbeelden werd duidelijk aangegeven dat eerst de zin van de pratende persoon

gelezen moest worden, daarna bedacht moest worden welke smiley gekozen werd en tenslotte een reactie gegeven moest worden. Tijdens het invullen van het instrument kregen de kinderen de mogelijkheid tot vragen stellen door middel van het opsteken van hun vinger. Als ze klaar waren met het invullen van het instrument moesten ze in stilte iets voor zichzelf gaan doen. Tijdens het invullen van het instrument door de kinderen vulde de leraar het agressievragenlijstje in. Wanneer iedereen klaar was werden de instrumenten opgehaald en werd de kinderen gevraagd of ze het instrument moeilijk, makkelijk, leuk of saai vonden. Tenslotte werden de kinderen bedankt voor hun medewerking aan het onderzoek. Bij de kinderen van groep 2 werd het instrument individueel afgenomen. Zij werden één voor één uit de klas gehaald en meegenomen naar een rustige ruimte. De procedure verliep hetzelfde als bij de klassikale afname alleen werden de items voorgelezen en werden de antwoorden van de kinderen opgeschreven door de proefleider.

2.2 Resultaten

Proefpersonen

Alle 62 kinderen vulden het instrument volledig in. Er waren 36 meisjes en 24 jongens en twee kinderen vergaten hun sekse in te vullen. De kinderen waren tussen de vijf en tien jaar ($M=6.98$; $SD=1.57$) oud. Er waren 15 kinderen van vijf jaar, 16 kinderen van zes jaar, 18 kinderen van acht jaar, 12 kinderen van negen jaar en 1 kind van tien jaar. De kinderen waren afkomstig uit verschillende groepen 2 en één groep 5. Van elke groep kregen 15 kinderen versie 1 voorgelegd en 16 kinderen versie 2. Er werden geen significante verschillen gevonden voor sekse en leeftijd tussen de kinderen die versie 1 hadden ingevuld en de kinderen die versie twee hadden ingevuld.

Agressievragenlijst

De betrouwbaarheid van de agressievragenlijst die de leerkrachten hebben ingevuld ligt vrij hoog, Cronbach's alpha had een waarde van 0.80. De gemiddelde score op de vragenlijst was 11.85 ($SD=4.07$).

Afname ervaringen

De schriftelijke afname in groep 5 verliep goed. De kinderen snaptten de procedure en bleven rustig doorwerken. De vragen die werden gesteld gingen voornamelijk over een woord dat ze niet begrepen of een woord dat ze zelf wilden schrijven maar niet wisten hoe ze dat moesten spellen. Het enige wat veel kinderen niet snaptten was de manier waarop ze hun sekse moesten aangeven. Na de afname gaven de kinderen aan dat ze het leuk vonden om te doen. Sommigen vonden het heel makkelijk, anderen vonden het soms wel moeilijk om een antwoord te bedenken. Het eerste kind was na ongeveer 15 minuten klaar en het laatste kind na iets meer dan een half uur. De individuele afname bij de kinderen van groep 2 ging ook goed. Sommige plaatjes waren wel te moeilijk, maar ze snaptten de procedure van het instrument wel. Een afname duurde ongeveer 20 minuten. Wanneer de afname van het instrument langer zou duren, zou de concentratie verslappen. Ook deze kinderen vonden het leuk om te doen en vonden het wisselend moeilijk en makkelijk. De leraren vonden het instrument er leuk uitzien en zeiden achteraf ook dat de kinderen enthousiast waren. Het invullen van de agressievragenlijst vonden ze vaak lastig, maar de tijd die het kostte viel ze wel mee.

Bepalen van het scoringsysteem

Tijdens het scoren werd al duidelijk dat sommige categorieën nauwelijks werden gebruikt. Eerst werd daarom gekeken naar de gemiddelden en standaardafwijkingen van de categorieën. In Tabel 1 zijn deze gemiddelden en Cronbach's alpha te zien. De categorieën met een hoge betrouwbaarheid en een gemiddelde dat niet te dicht bij de minimale score lag, werden geselecteerd voor de factoranalyse. Aangezien een aantal categorieën belangrijk waren om het instrument goed te kunnen scoren werden deze ondanks een lage betrouwbaarheid toch geselecteerd voor de factoranalyse. De geselecteerde categorieën staan dikgedrukt in Tabel 1.

Tabel 1. Gemiddelde, standaardafwijkingen en Cronbach's alpha van de categorieën.

<i>Categorie</i>	<i>Versie 1</i>			<i>Versie 2</i>		
	Gem	SD	Alpha	Gem	SD	Alpha
Berusting	29.63	5.72	0.67	28.84	5.35	0.68
Geduld	20.10	0.40	-0.03	20.47	0.71	0.00
Schuld bekennen	20.40	0.67	-0.20	20.09	0.39	-0.06
Excuses aanbieden	21.37	2.51	0.65	21.63	1.74	0.39
Oplossing geven	23.57	4.05	0.73	23.59	4.23	0.75
Belofte voor de toekomst	20.37	0.72	-0.20	20.19	0.40	0.19
Probleem kleiner maken	21.40	1.79	0.45	22.16	2.37	0.50
Probleem niet gezien	24.57	2.18	-0.20	23.88	2.83	0.53
Probleem niet gesnapt	22.63	3.00	0.50	22.81	2.98	0.47
Probleem ontkennen	20.77	1.22	0.25	22.38	2.59	0.59
Ander de schuld geven	20.30	0.70	-0.07	20.09	0.39	-0.03
Toegeven met smoes	21.27	1.55	0.31	21.28	1.53	0.28
Ander laten oplossen	20.33	1.47	0.76	20.91	1.91	0.63
Agressieve ontkenning	20.73	1.05	0.48	20.41	1.21	0.77
Boze reactie	31.33	4.76	0.69	29.91	4.67	0.64
Schelden	20.57	1.70	0.78	20.22	0.61	0.26
Agressieve reactie	20.70	1.42	0.64	20.56	0.84	0.20
Reden vragen	21.43	1.65	0.09	22.38	3.65	0.74
Sarren of pesten	22.50	2.22	0.39	23.25	3.10	0.62
Zeuren	27.80	4.66	0.67	25.66	4.06	0.65
Brutaal antwoorden	22.57	3.13	0.65	22.13	2.08	0.32

Noot. De score heeft een minimum van 20 en een maximum van 60.

Om na te gaan of de categorieën in betekenisvolle factoren waren onder te verdelen werd een Principale Componenten Analyse op de somscores van de schalen uitgevoerd. Op deze manier werd gekeken of er categorieën waren die bij elkaar hoorden zodat een scoringsysteem met minder categorieën kon worden ontwikkeld. Voor versie 1 werden vier componenten geëxtraheerd met een eigenwaarde groter dan 1. Deze componenten verklaarden samen 70.8% van de totale gemeenschappelijke variantie (Tabel 2). Voor versie 2 werden vijf componenten geëxtraheerd met een eigenwaarde groter dan 1. Deze componenten verklaarden samen 68.9% van de totale gemeenschappelijke variantie (Tabel 3).

Tabel 2. Geroteerde factor matrix met eigenwaarden en communaliteiten van versie 1

<i>Categorie</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>Communa- liteiten</i>
Berusting	-.59	.44	.38	.23	.73
Excuses aanbieden	-.07	.11	-.90	-.08	.84
Oplossing geven	-.21	-.16	.11	.85	.81
Probleem kleiner maken	.51	.38	-.01	.46	.62
Probleem niet gezien	.26	.32	.27	.76	.83
Probleem ontkennen	.16	.75	.21	-.19	.67
Ander laten oplossen	-.02	-.07	-.91	-.17	.87
Boze reactie	.50	-.47	.03	-.62	.85
Schelden	.85	-.06	.08	-.07	.74
Agressieve reactie	.86	.07	.09	.04	.76
Reden vragen	.19	.66	.14	.16	.52
Sarren of pesten	.19	-.61	.32	-.02	.52
Zeuren	.43	.11	.36	-.33	.44
Brutaal antwoorden	.28	-.71	.26	-.27	.73
Eigenwaarden	2.80	2.60	2.26	2.24	
% Verklaarde variantie	20.01	18.64	16.15	15.97	

Tabel 3. Geroteerde factor matrix met eigenwaarden en communaliteiten van versie 2

<i>Categorie</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Factor 3</i>	<i>Factor 4</i>	<i>Factor 5</i>	<i>Communa- liteiten</i>
Berusting	-.61	.28	-.39	-.33	.12	.72
Excuses aanbieden	-.60	.02	-.11	.52	-.13	.66
Oplossing geven	.06	-.38	-.70	-.11	.01	.65
Probleem kleiner maken	.09	.16	.04	-.01	.88	.82
Probleem niet gezien	-.35	.68	-.05	-.13	.13	.62
Probleem ontkennen	-.02	.68	.09	.41	-.30	.73
Ander laten oplossen	-.08	-.55	-.10	-.09	-.16	.35
Boze reactie	.61	.01	.71	-.06	-.03	.87
Schelden	-.06	.55	.03	.01	-.52	.58
Agressieve reactie	.49	.49	-.05	.36	.14	.64
Reden vragen	-.03	.13	.07	.90	.04	.83
Sarren of pesten	.78	-.26	-.11	-.11	.12	.71
Zeuren	-.03	-.57	.66	-.06	.13	.79
Brutaal antwoorden	.80	.17	.06	-.02	.03	.67
Eigenwaarden	2.72	2.47	1.65	1.54	1.27	
% Verklaarde variantie	19.41	17.64	11.76	10.97	9.10	

Uit de geroteerde factormatrices van versie 1 en 2 blijkt dat op de eerste factor vooral de agressieve of onaardige categorieën laden. Berusting en excuses aanbieden laden negatief op deze factor en boze reactie, schelden, agressieve reactie, sarren of pesten, zeuren en brutaal antwoorden laden juist hoog positief op deze factor. Deze factor verklaart 20% van de variantie in versie 1 en 19% van de variantie in versie 2.

Met behulp van deze informatie en met de ervaring van het scoren zijn de categorieën ingedeeld in vijf categorieën. De eerste is “aardige reactie”, dit zijn positieve reacties waaruit blijkt dat de kinderen de situatie accepteren zoals hij is. Hieronder vallen de voormalige categorieën: Berusting, Oplossing geven en Excuses aan bieden. De tweede is “Probleem niet gezien”, dit zijn verschillende reacties die niet onder een negatieve categorie vallen maar ook niet positief zijn. Hieronder vallen de voormalige categorieën: Probleem niet gezien, Probleem klein maken, Reden vragen en Probleem ontkennen. De derde categorie is “Zeuren”, dit zijn antwoorden waaruit blijkt dat het kind zeurt. Hieronder vallen de voormalige categorieën: Zeuren en Ander laten oplossen. De vierde categorie is “Onaardige reactie”, dit zijn antwoorden die onaardig zijn tegenover de andere persoon. Hieronder vallen de voormalige categorieën: Boze reactie, Sarren of pesten en Brutaal. De vijfde en laatste categorie is “Agressieve reactie”, dit zijn antwoorden waaruit blijkt dat een kind agressief reageert tegenover de andere persoon. Hieronder vallen de voormalige categorieën: Agressieve reactie en Schelden. Voor een uitgebreide instructie over het scoren van het instrument wordt naar bijlage 3 verwezen.

Ontwerp van de definitieve versie van de test

Om te beslissen welke plaatjes geschikt waren om in het uiteindelijke instrument op te nemen werden de plaatjes op verschillende punten beoordeeld. Ten eerste werd gekeken of de plaatjes wel begrepen werden door de meeste kinderen. Wanneer veel kinderen bepaalde plaatjes niet begrepen, werden die plaatjes eruit gehaald. Daarnaast werd gekeken naar het onderscheidend vermogen van de plaatjes. Er werd gekeken op welke categorieën per plaatje werd gescoord. Als er bijvoorbeeld alleen maar aardige reacties werden gegeven en nauwelijks onaardige reacties werd het plaatje niet geselecteerd. Tenslotte werd gekeken naar de verdeling van de intentievraag per plaatje. Wanneer vrijwel elk kind de persoon bestempelde als niet aardig en niemand als aardig werd het plaatje ook niet geselecteerd. Uiteindelijk bleven er 20 plaatjes over. Deze 20 plaatjes werden in een nieuwe volgorde geplaatst, waarbij weer werd gelet op de volgorde volwassen/kind en de verhouding jongens/meisjes. Bij de instructie werd alleen het aangeven van de sekse aangepast omdat bleek dat de kinderen de oude manier vaak lastig vonden.

3. Onderzoek

3.1 Methode

Proefpersonen

In totaal deden er 140 kinderen uit de groepen 2 tot en met 8 van basisschool de Zonnewijzer in Hoorn mee aan het onderzoek. Een van de onderzoekers kent de directeur van deze basisschool en zo is het contact gelegd en werd toestemming verkregen om op deze school het onderzoek uit te voeren. De ouders kregen een brief met informatie over het onderzoek (bijlage 4), zij konden deze brief op school inleveren wanneer ze geen toestemming wilden verlenen voor deelname van hun kind. Daarnaast deden nog 30 kinderen met een leerniveau van groep 5 van basisschool het Dok in Zaandam mee aan het onderzoek. Het Dok is een school voor speciaal onderwijs. De kinderen komen op de school als het ze op de een of andere manier niet gelukt is om mee te komen in het reguliere onderwijs. De school verdeelt de kinderen in drie stromen: praktijk, theorie en een gemengde stroom. Daarnaast wordt bij de vorming van klassen gelet op leeftijd en sociaal-emotionele ontwikkeling. Ook hier waren

persoonlijke contacten met een leerkracht en na overleg met de directeur werd er toestemming gegeven om op deze school onderzoek te doen.

Materialen

Voor het meten van de reacties van kinderen in frustrerende situaties is aan de hand van de resultaten van de pilotstudy het plaatjesinstrument van 20 plaatjes ontwikkeld (bijlage 5). De antwoorden werden gescoord in vijf verschillende categorieën van type responsen. Deze categorieën waren: Aardige reactie, Probleem niet gezien, Zeuren, Onaardige reactie en Agressieve reactie. De antwoorden van de kinderen werden weer beoordeeld op een 3-puntsschaal, waarbij de antwoordmogelijkheden waren: past geheel niet in deze categorie (1), past een beetje in deze categorie (2) en past helemaal in deze categorie (3). Wanneer alle antwoorden waren beoordeeld werden de punten per categorie opgeteld. Hierdoor ontstond er een totaalscore per categorie, waarbij minimaal 20 punten en maximaal 60 punten gescoord konden worden (zie bijlage 6 voor het scoringsformulier). Verder werd bij elk plaatje wederom gevraagd naar de intentie van de pratende persoon en werd de agressievragenlijst voor elk kind door de leerkracht ingevuld.

Procedure

Binnen een tijdsbestek van vier dagen werd het onderzoek op de school afgenomen. Bij groep 5 tot en met groep 8 en bij de 30 kinderen van het speciaal onderwijs werd het instrument klassikaal afgenomen. Bij groep 2 tot en met 4 werd het instrument individueel afgenomen. De verdere procedure verliep zoals beschreven in de pilotstudy.

3.2. Resultaten

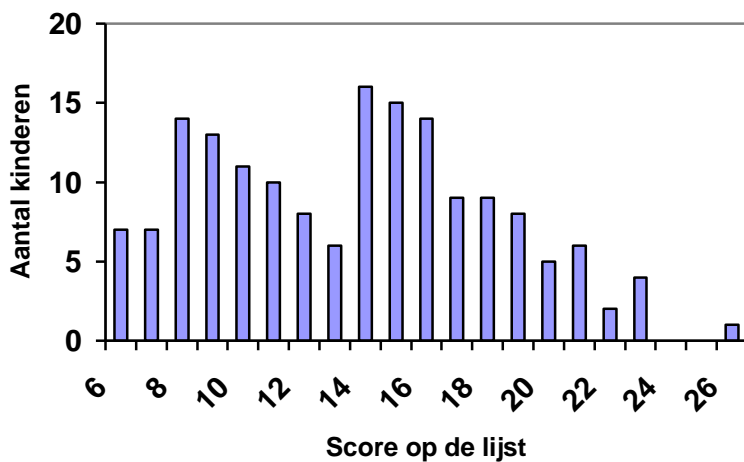
Proefpersonen

Van de in totaal 170 kinderen vulden vijf kinderen de tekstballonnen van meer dan drie plaatjes niet in, zij werden daarom uitgesloten van verdere analyses. Van de overige 165 kinderen waren er 76 meisjes (46.1%) en 89 jongens (53.9%). De leeftijd van de kinderen lag tussen de 5 en de 12 jaar (gem = 9.3; sd = 2.1). De kinderen waren afkomstig uit groep 2 tot en met groep 8.

Betrouwbaarheid agressievragenlijst

De betrouwbaarheid van de agressievragenlijst kwam uit op 0.87, dit is goed te noemen. Ook de betrouwbaarheid van het proactieve deel (0.84) en het reactieve deel (0.76) zijn goed te noemen. De gemiddelde score op de vragenlijst was 13.59 (SD = 4.59). In Figuur 1 is te zien dat de kinderen die meededen aan het onderzoek door de leraren over het algemeen niet als erg agressief beoordeeld werden. De maximale score die behaald zou kunnen worden is 36 en slechts één kind heeft boven de 24 gescoord.

Figuur 1. Frequentieverdeling van de scores op de agressievragenlijst



Betrouwbaarheid van de Sociale Plaatjes Test

In Tabel 4 is de betrouwbaarheid te zien van de verschillende categorieën. Cronbach's alpha is goed te noemen voor Aardige reactie, Zeuren en Onaardige reactie. Voor Probleem niet gezien en Agressieve reactie is Cronbach's alpha minder goed. De splithalve betrouwbaarheid is alleen voor Agressieve reactie minder goed.

Tabel 4. Cronbach's alpha en splithalve betrouwbaarheid van de categorieën

Categorieën	betrouwbaarheid	Splithalve betrouwbaarheid
Aardige reactie	0.75	0.76
Probleem niet gezien	0.62	0.61
Zeuren	0.75	0.77
Onaardige reactie	0.78	0.77
Agressieve reactie	0.56	0.73

In Tabel 5 is de interbeoordelaarsbetrouwbaarheid te zien. Cohen's kappa is uitgerekend tussen de onderzoekers van dit onderzoek en tussen de onderzoekers en een laatstejaars psychologie studente. In Tabel 5 is te zien dat de onderzoekers een goede overeenstemming bereiken bij de categorieën Aardig, Onaardig en agressief maar niet zo goed bij Zeuren en Probleem niet gezien. De totale interbeoordelaarsbetrouwbaarheid is wel weer goed te noemen. In Tabel 5 is verder te zien dat de onderzoekers met de studente een goede overeenstemming bereiken. Er zijn op de schalen wel een aantal verschillen maar de totale interbeoordelaarsbetrouwbaarheid is voor ieder koppel vrijwel gelijk. Deze totale betrouwbaarheid is goed te noemen.

Tabel 5. Cohen's kappa tussen de onderzoeker en met de psychologiestudente

Categorieën	Tussen de onderzoekers	Onderzoeker 1 & Studente	Onderzoeker 2 & Studente
Aardige reactie	0.86	0.82	0.83
Probleem niet gezien	0.49	0.70	0.68
Zeuren	0.56	0.64	0.72
Onaardige reactie	0.70	0.80	0.70
Agressieve reactie	0.95	0.98	0.97
Totaal	0.78	0.79	0.78

Factoranalyse

Om na te gaan of de categorieën in betekenisvolle factoren waren onder te verdelen werd een Principale Componenten Analyse uitgevoerd. Er werden twee principale componenten geëxtraheerd met een eigenwaarde groter dan 1. Deze componenten verklaarden samen 66.7% van de totale gemeenschappelijke variantie. In Tabel 6 is te zien dat op de eerste factor Onaardige reactie en Agressieve reactie hoog positief laden en Aardige reactie hoog negatief. Dit zal een factor zijn die maken heeft met hoe onaardig iemand reageert. Op de tweede factor laden Probleem niet gezien en Zeuren hoog positief en Aardige reactie hoog negatief. Dit zal een factor zijn die te maken heeft met het inzicht in de situatie. Bij de categorie Zeuren wordt vaak gezegd ‘Waarom dan niet?’. Dit antwoord duidt op enig onbegrip van de situatie evenals bij de categorie Probleem niet gezien.

Tabel 6. Geroteerde factor matrix met eigenwaarden en communaliteiten

<i>Categorieën</i>	<i>Factor 1</i>	<i>Factor 2</i>	<i>Communaliteiten</i>
Aardige reactie	-.75	-.62	.95
Probleem niet gezien	-.15	.65	.44
Zeuren	-.03	.69	.47
Onaardige reactie	.93	-.11	.88
Agressieve reactie	.66	-.40	.60
Eigenwaarde	1.89	1.45	
Verklaarde variantie	37.74	28.96	

Verschillen tussen agressieve en niet-agressieve kinderen

Om de verschillen tussen agressieve en niet-agressieve kinderen te meten is de groep kinderen van het speciale onderwijs op twee manieren gekoppeld aan een controlegroep. Ten eerste is de groep van het speciaal onderwijs gekoppeld aan groep 5, deze komen namelijk overeen wat betreft het niveau van leren. Ten tweede is de groep gematcht met een groep kinderen met ongeveer dezelfde geboortedatum en hetzelfde geslacht. Het aantal jongens, meisjes, de gemiddelde leeftijd en standaarddeviaties en de gemiddelde score op de agressievragenlijst van deze groepen staan weergegeven in Tabel 7.

Tabel 7. Aantal meisjes, jongens, gemiddelde leeftijd en SD van de groepen

Groep	Totaal aantal	Aantal meisjes	Aantal jongens	Gem. leeftijd	SD	Gem. Agressie vragenlijst
1.Speciaal onderwijs	30	12	18	10.63	0.49	14.17
2.Groep 5	28	15	13	8.36	0.49	11.41
3.Gematchte groep	30	13	17	10.53	0.57	13.50

In Tabel 7 is te zien dat de groepen over het algemeen laag scoren op de agressievragenlijst. De groep speciaal onderwijs scoort wel hoger dan de andere twee groepen maar dit verschil is met de gematchte groep niet significant. Met groep 5 is dit verschil wel significant ($t(55) = -2.29$, $p < 0,05$), maar het verschil is klein.

De gemiddelde score en de standaarddeviatie per categorie van den SPT van de groep speciaal onderwijs en de controle groep (gematcht) staan weergegeven in Tabel 8. In Tabel 8 is te zien dat de controle groep, zoals verwacht, gemiddeld hoger scoort op Aardige reactie en, tegen de verwachting in, ook hoger scoort op Agressieve reactie. De controle groep scoort hoger op Probleem niet gezien en Zeuren. Zoals verwacht scoort de groep van het speciaal onderwijs hoger op Onaardige Reactie.

Tabel 8. Gemiddelde score en SD per categorie van het speciaal onderwijs en de gematchte groep van het regulier onderwijs.

<i>School type</i>	<i>Aardige reactie</i>		<i>Probleem niet gezien</i>		<i>Zeuren</i>		<i>Onaardige reactie</i>		<i>Agressieve reactie</i>	
	Gem	SD	Gem	SD	Gem	SD	Gem	SD	Gem	SD
Speciaal	36,84	5,93	24,57	2,89	29,34	6,15	28,87	5,06	20,34	0,80
Gematcht	37,48	6,85	27,20	3,73	30,00	4,41	28,48	5,32	20,43	1,00

Met behulp van een t-toets kon de groep van de speciale school vergeleken worden met de controle groep. De verschillen tussen deze groepen waren alleen significant voor de categorie Probleem niet gezien ($t(56) = 4.053, p < 0,05$). Dit verschil is tegengesteld aan het verwachte verschil.

De gemiddelde score en standaarddeviatie per categorie van de groep speciaal onderwijs en de controle groep (groep 5) staan weergegeven in Tabel 9. In Tabel 9 is te zien dat de controle groep, zoals verwacht, hoger scoort op Aardige reactie. Verder scoort de controle groep hoger op Probleem niet gezien. Zoals verwacht scoort de groep van het speciaal onderwijs hoger op Zeuren, Onaardige reactie en Agressieve reactie.

Tabel 9. Gemiddelde score en SD per categorie van het speciaal onderwijs en groep 5 van het regulier onderwijs.

<i>School type</i>	<i>Aardige reactie</i>		<i>Probleem niet gezien</i>		<i>Zeuren</i>		<i>Onaardige reactie</i>		<i>Agressieve reactie</i>	
	Gem	SD	Gem	SD	Gem	SD	Gem	SD	Gem	SD
Speciaal	36,84	5,93	24,57	2,89	29,34	6,15	28,87	5,05	20,34	0,80
Groep 5	38,19	6,60	25,19	3,62	26,54	4,74	28,35	3,97	20,19	0,49

Met behulp van een t-toets konden de groepen met elkaar worden vergeleken. De verschillen bleken niet significant te zijn.

Lerarenoordeel

Om te controleren of de score op de agressievragenlijst samenhangt met de score op de SPT, zijn de correlaties berekend tussen deze scores. De correlaties zijn te zien in Tabel 10. Zoals verwacht correleert de categorie Aardige reactie, wanneer de correlaties van de gehele test worden berekend, negatief met de agressievragenlijst. Onaardige reactie correleert positief met de agressievragenlijst. Tegen de verwachting in correleert Agressieve reactie negatief met de agressievragenlijst. Alle correlaties waren echter niet significant. Ook wanneer de correlaties apart berekend werden voor de volwassenen en kind plaatjes werden er, zoals te zien in Tabel 10, geen significante correlaties gevonden. In Tabel 11 zij de correlaties te zien tussen de scores op de intentievraag en de score op de agressievragenlijst. Er is een positieve correlatie tussen de agressievragenlijst en Aardig en Niet aardig, deze correlaties zijn echter niet significant. Er is een negatieve correlatie tussen de agressievragenlijst en Gewoon, deze correlatie is wel significant maar zeer laag.

Tabel 10. Pearson correlatie van de agressievragenlijst met de score op de categorieën van de SPT

Sociale plaatjes test	Aardige reactie	Probleem niet gezien	Zeuren	Onaardige reactie	Agressieve reactie
Gehele test	-.011	-.081	-.037	.091	-.033
Volwassen plaatjes	.032	-.040	-.076	.072	-.042
Kind plaatjes	-.060	-.108	.020	.082	-.019

Tabel 11. Pearson correlatie van de agressievragenlijst met de score op de intentievraag

	Aardig	Gewoon	Niet Aardig
Agressievragenlijst	.042	-.155	.129

Relatie tussen intentie en score op de categorieën van de Sociale Plaatjes Test

Voor elk plaatje werd een chikwadraat toets uitgevoerd voor elke categorie. Zo ontstonden er per categorie 20 frequentietabellen met de intentie als rijen en de categorieën van de SPT in kolommen. Deze tabellen werden bij elkaar opgeteld zodat uiteindelijk vijf tabellen overbleven, voor elke categorie één.

Om erachter te komen of de kinderen hoog scoorden in de categorie Aardige reactie wanneer zij aangaven de persoon aardig te vinden, werd een chikwadraat toets uitgevoerd op de frequenties van Tabel 12. Het resultaat van deze toets bleek significant ($X^2(4) = 701.61$, $p < 0.05$). Dit betekent dat de geobserveerde frequenties significant afwijken van de frequenties die je op basis van onafhankelijkheid zou verwachten. Over het algemeen geldt dat er een verband is tussen de intentie aardig en de categorie aardig. In Tabel 12 is te zien dat wanneer de persoon in het plaatje aardig gevonden wordt, de kinderen meer dan verwacht op basis van onafhankelijkheid een aardige reactie geven. Daarnaast is te zien dat wanneer kinderen aangeven de persoon niet aardig te vinden, zij minder vaak een aardig antwoord gaven dan verwacht op basis van onafhankelijkheid. Als zij de intentie van de persoon als “gewoon” beoordelen, geven zij minder vaak dan op basis van onafhankelijkheid wordt verwacht een aardige reactie.

Tabel 12. Kruistabel van de intentie met de categorie Aardige reactie.

		Aardig			Totaal
		Niet van toepassing	Beetje van toepassing	Wel van toepassing	
Intentie	<i>Aardig</i>	Aantal	94	18	248
		Verw. Aantal	198,2	16,4	145,4
	<i>Gewoon</i>	Aantal	348	61	805
		Verw. Aantal	717,8	59,5	526,7
	<i>Niet aardig</i>	Aantal	1230	67	240
		Verw. Aantal	846,0	70,1	620,9
Totaal			1762	146	1293
					3201

De kruistabel van intentie met de categorie Probleem niet gezien, is te zien in Tabel 13. Het resultaat van de chikwadraat toets bleek significant ($X^2(4) = 82.93, p < 0.05$). Over het algemeen geldt dat de categorie Probleem niet gezien vaker niet van toepassing was dan op basis van onafhankelijkheid verwacht kan worden. Wanneer de reactie van het kind “beetje van toepassing” is in de categorie Probleem niet gezien, vindt het kind minder vaak, dan verwacht op basis van onafhankelijkheid, “gewoon”, maar vaker “niet aardig”. Wanneer kinderen het probleem niet zien, vinden ze een persoon minder vaak, dan verwacht op basis van onafhankelijkheid, “aardig” en vaker “gewoon”.

Tabel 13. Kruistabel van de intentie met de categorie Probleem niet gezien.

		<i>Probleem niet gezien</i>				
		Niet van toepassing	Beetje van toepassing	Wel van toepassing	Totaal	
Intentie	<i>Aardig</i>	Aantal	314	22	24	360
		Verw. Aantal	271,5	21,3	67,15	
	<i>Gewoon</i>	Aantal	1254	79	248	1581
		Verw. Aantal	1192,5	93,6	294,9	
	<i>Niet aardig</i>	Aantal	1234	119	201	1554
		Verw. Aantal	1172,1	92,0	289,9	
Totaal			2802	220	473	3495

De kruistabel van intentie met de categorie Zeuren is te zien in Tabel 14. Het resultaat was significant ($X^2(4) = 151.91, p < 0.05$). Wanneer kinderen aangaven de persoon niet aardig te vinden, zeurden zij meer dan op basis van onafhankelijkheid verwacht kan worden. Daarnaast zeurden kinderen minder dan verwacht wanneer zij aangaven de persoon aardig of gewoon te vinden.

Tabel 14. Kruistabel van de intentie met de categorie Zeuren.

		<i>Zeuren</i>				
		Niet van toepassing	Beetje van toepassing	Wel van toepassing	Totaal	
Intentie	<i>Aardig</i>	Aantal	313	21	13	347
		Verw. Aantal	256,4	55,3	35,4	
	<i>Gewoon</i>	Aantal	1095	142	139	1376
		Verw. Aantal	1016,7	219,1	140,3	
	<i>Niet aardig</i>	Aantal	991	354	179	1524
		Verw. Aantal	1126,0	242,7	155,4	
Totaal			2399	517	331	3247

Om erachter te komen of de kinderen een onaardig antwoord gaven wanneer zij aangaven de persoon niet aardig te vinden werd een chikwadraat toets uitgevoerd op de kruistabel van de intentie met de categorie Onaardige reactie, welke te zien is in Tabel 15. Het resultaat van deze toets bleek significant ($X^2(4) = 657.22, p < 0.05$). Over het algemeen geldt dat er een verband is tussen de intentie niet aardig en de categorie onaardige reactie. Dit betekent dat wanneer de persoon in het plaatje niet aardig gevonden wordt, de kinderen meer dan verwacht op basis van onafhankelijkheid een onaardige reactie gaven. Dit betekent tevens dat wanneer de kinderen aangaven de persoon aardig of gewoon te vinden, zij minder vaak een onaardige reactie gaven dan verwacht op basis van onafhankelijkheid.

Tabel 15. Kruistabel van de intentie met de categorie Onaardige reactie.

		<i>Onaardige reactie</i>				
		Niet van toepassing	Beetje van toepassing	Wel van toepassing	Totaal	
Intentie	<i>Aardig</i>	Aantal	297	32	18	347
		Verw. Aantal	200,3	79,9	38,7	
	<i>Gewoon</i>	Aantal	1117	181	78	1376
		Verw. Aantal	794,2	316,7	153,5	
	<i>Niet aardig</i>	Aantal	635	604	300	1539
		Verw. Aantal	888,3	354,2	171,7	
Totaal			2049	817	396	3550

Tenslotte werd nog een chikwadraat toets uitgevoerd op kruistabel van Tabel 16. Het resultaat bleek significant ($X^2(4) = 23.12, p < 0.05$). Over het algemeen geldt dat er weinig agressieve reacties gegeven werden maar als deze categorie een beetje van toepassing was, werd de intentie vaker als niet aardig beoordeeld en minder vaak als gewoon, dan verwacht op basis van onafhankelijkheid.

Tabel 16. Kruistabel van de intentie met de categorie Agressieve reactie.

		<i>Agressieve reactie</i>				
		Niet van toepassing	Beetje van toepassing	Wel van toepassing	Totaal	
Intentie	<i>Aardig</i>	Aantal	354	2	0	356
		Verw. Aantal	349,9	5,2	0,9	
	<i>Gewoon</i>	Aantal	1367	9	1	1377
		Verw. Aantal	1353,4	20,2	3,4	
	<i>Niet aardig</i>	Aantal	1495	37	7	1539
		Verw. Aantal	1512,7	22,6	3,8	
Totaal			3216	48	8	3272

Leeftijdsverschillen in reacties op de Sociale Plaatjes Test

Om te toetsen of er verschillen zijn in reacties in de ontwikkeling van tussen kinderen van verschillende leeftijden werden de kinderen in vier groepen ingedeeld.

Tabel 17. Indeling van leeftijdsgroepen met gemiddelden en standaardafwijkingen

<i>Groepen</i>	<i>Aantal</i>	<i>Leeftijd</i>	<i>Gemiddeld</i>	<i>SD</i>
Groep 1	30	5-6	6.1	0.10
Groep 2	44	7-8	8.2	0.09
Groep 3	42	9-10	10.0	0.09
Groep 4	49	11-12	11.6	0.08

In Tabel 17 worden de aantallen, gemiddelde leeftijd en standaarddeviatie per groep vermeld. Eerst werd gekeken of er verschillen tussen de leeftijdsgroepen waren in de antwoorden op de intentievraag. In Tabel 18 is te zien dat de jongste groep kinderen de personen op de plaatjes over het algemeen wat vaker 'aardig' vonden dan de oudere kinderen. Dit was echter geen significant verschil. Ook is te zien dat de 5 en 6 jarigen minder vaak de personen 'gewoon' vonden dan de oudere kinderen. Dit verschil bleek significant te zijn, $F(3,161) = 6.05, p < 0.05$. De 5 en 6 jarigen vonden de personen vaker 'niet aardig' dan de oudere groepen kinderen. Ook dit verschil bleek significant te zijn, $F(3, 161) = 3.57, p < 0.05$.

Tabel 18. Gemiddelden en standaardafwijkingen van de intentie

<i>Leeftijd</i>	<i>Aardig</i>		<i>Gewoon</i>		<i>Niet aardig</i>	
	<i>Gem</i>	<i>SD</i>	<i>Gem</i>	<i>SD</i>	<i>Gem</i>	<i>SD</i>
5-6	1.42	0.24	3.87	0.45	4.70	0.43
7-8	0.89	0.20	5.96	0.37	3.15	0.36
9-10	0.67	0.20	6.14	0.38	3.14	0.36
11-12	0.80	0.19	5.80	0.35	3.12	0.34

Vervolgens werd gekeken of er ook verschillen tussen de leeftijdsgroepen waren in reacties op de plaatjes. In Tabel 19 is te zien dat er bij de categorieën Aardig, Probleem niet gezien en Agressieve reactie slechts kleine verschillen zijn tussen de groepen. Deze verschillen bleken ook niet significant te zijn. Bij de categorie Zeuren is te zien dat de 9 en 10 jarigen meer zeuren dan de kinderen in de andere leeftijdsgroepen. Dit verschil bleek significant te zijn, $F(3,161) = 5.33, p < 0.05$. Bij een post hoc Tukey HSD bleek dat het verschil significant was tussen groep 1 en 3 en tussen 2 en 3. In Tabel 19 is tenslotte te zien dat bij de categorie Onaardige reactie de jongste kinderen het hoogste scoren en de 9 en 10 jarigen het laagst. Ook dit verschil bleek significant te zijn, $F(3,161) = 3.84, p < 0.05$. Een post hoc Tukey gaf aan dat dit verschil zat tussen groep 1 en 3. Voor normtabellen van de leeftijdsgroepen wordt naar bijlage 7 verwezen.

Tabel 19. Gemiddelden en standaardafwijkingen van de somscores op de categorieën

<i>Leeftijd</i>	<i>Aardige reactie</i>		<i>Probleem niet gezien</i>		<i>Zeuren</i>		<i>Onaardige reactie</i>		<i>Agressieve reactie</i>	
	<i>Gem</i>	<i>SD</i>	<i>Gem</i>	<i>SD</i>	<i>Gem</i>	<i>SD</i>	<i>Gem</i>	<i>SD</i>	<i>Gem</i>	<i>SD</i>
5-6	35.80	1.37	25.40	0.78	25.80	0.93	32.13	1.03	20.80	0.18
7-8	36.73	1.13	24.62	0.64	25.90	0.77	30.30	0.85	20.37	0.15
9-10	37.87	1.16	25.87	0.66	29.74	0.79	27.68	0.87	20.18	0.16
11-12	37.28	1.07	25.77	0.61	27.82	0.73	29.95	0.80	20.37	0.14

Leeftijdsverschillen en verschillen in reactie tegen volwassenen en kinderen

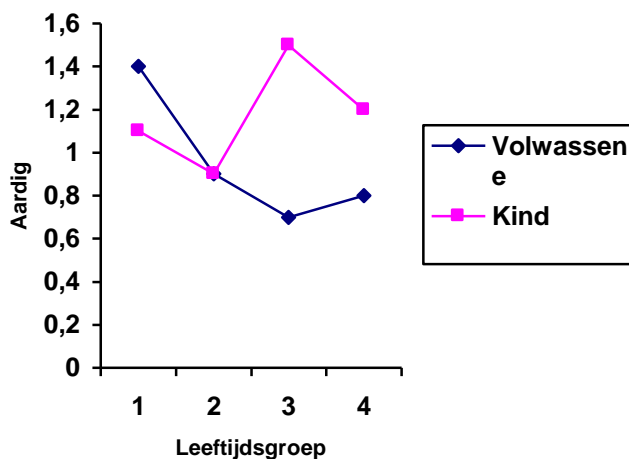
De plaatjes bestonden voor de helft uit plaatjes waarop een volwassene iets zei en voor de andere helft uit plaatjes waarop een kind iets zei. Er werd een Herhaalde metingen analyse uitgevoerd om te onderzoeken of er verschillen waren in antwoorden op de intentievraag wanneer de persoon op het plaatje een kind of een volwassene was en of er interactieverschillen waren tussen leeftijd en de persoon op het plaatje. In Tabel 20 is te zien dat kinderen over het algemeen de volwassene minder vaak aardig vinden dan een kind, dit verschil is significant ($F(1,161) = 4.33, p < 0.05$). Ook is te zien dat kinderen de volwassene vaker gewoon vinden dan kinderen, dit verschil is ook significant ($F(1,161) = 189.52, p < 0.05$). Tenslotte vinden kinderen volwassen minder vaak onaardig dan kinderen, ook dit verschil is significant ($F(1,161) = 153.29, p < 0.05$).

Tabel 20. Gemiddelden en standaardafwijkingen van de intentie

Persoon in het plaatje	Aardig		Gewoon		Niet aardig	
	Gem	SD	Gem	SD	Gem	SD
Volwassene	0.95	0.10	5.44	0.20	3.53	0.19
Kind	1.20	0.10	2.70	0.14	5.99	0.15

Bij het aardig vinden van de persoon werd er een interactie-effect gevonden tussen de leeftijd en de persoon in het plaatje. Dit was een significant effect, $F(3,161) = 4.13, p < 0.05$). In Figuur 2 is te zien dat de jongste twee groepen volwassenen en kinderen ongeveer even vaak aardig vinden en de oudste twee groepen kinderen vaker aardig vinden dan volwassenen. Het is opvallend dat kinderen de intentie niet vaak als aardig beoordelen. De plaatjes lijken inderdaad voldoende ambig of als frustrerend te worden geïnterpreteerd.

Figuur 2. Gemiddelden van de leeftijdsgroepen bij kinderen en volwassenen



Er werd tevens een herhaalde metingen analyse uitgevoerd om te onderzoeken of er verschillen waren in reacties wanneer de persoon in het plaatje een volwassene of een kind was en of er interactieverschillen waren tussen leeftijd en de persoon in het plaatje. In Tabel 21 is te zien dat kinderen over het algemeen meer aardige reacties geven tegen volwassenen dan tegen kinderen. Dit verschil is significant, ($F(1,161) = 95.34, p < 0.05$). Bij Probleem niet gezien werden geen significante verschillen gevonden. Verder is in Tabel 21 te zien dat kinderen meer zeuren tegen volwassenen dan tegen kinderen, dit bleek een significant verschil te zijn ($F(1,161) = 8.53, p < 0.05$). Kinderen gaven significant meer onaardige reacties tegen

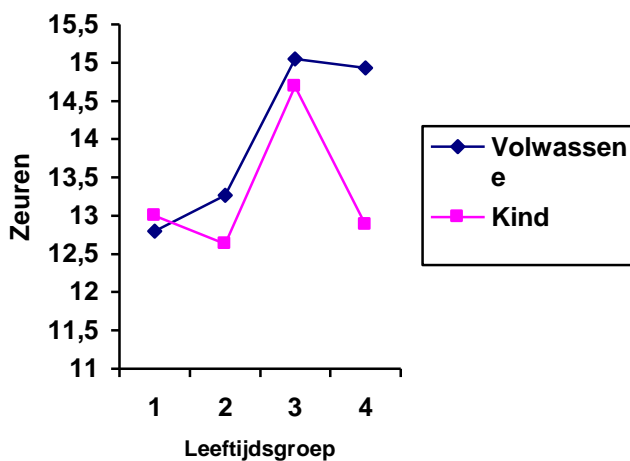
kinderen ($F(1, 161) = 108.00, p < 0.05$), en waren ook significant agressiever tegen kinderen dan volwassenen ($F(1, 161) = 6.13, p < 0.05$).

Tabel 21. Gemiddelden en standaardafwijkingen van de somscores op de categorieën

Persoon in het plaatje	Aardige reactie		Probleem niet gezien		Zeuren		Onaardige reactie		Agressieve reactie	
	Gem	SD	Gem	SD	Gem	SD	Gem	SD	Gem	SD
Volwassene	20.30	0.39	12.84	0.23	14.01	0.25	13.52	0.27	10.14	0.04
Kind	16.62	0.30	12.58	0.17	13.30	0.22	16.49	0.26	10.29	0.06

Bij de categorie Zeuren werd tevens een significant interactie-effect gevonden tussen de leeftijd en de persoon in het plaatje ($F(3, 161) = 4.04, p < 0.05$). In Figuur 3 is te zien dat de eerste drie groepen ongeveer evenveel zeuren tegen kinderen en volwassenen maar dat de oudste groep meer zeurt tegen volwassenen dan tegen kinderen.

Figuur 3. Gemiddelden van de leeftijdsgroepen bij kinderen en volwassenen



4. Discussie

Conclusie

De resultaten van het beschreven onderzoek geven gedeeltelijk steun aan de hypothese dat de Sociale Plaatjes Test (SPT) een betrouwbaar instrument is. Voor drie van de vijf categorieën (Aardige reactie, Zeuren en Onaardige reactie) werd een goede betrouwbaarheid gevonden en voor de twee andere categorieën (Agressieve reactie en Probleem niet gezien) werd een redelijke betrouwbaarheid gevonden. Er werd steun gevonden voor de hypothese over de interbeoordelaarsbetrouwbaarheid. Uit de resultaten bleek dat de antwoorden van de respondenten objectief gescoord kunnen worden door ontwikkelingspsychologen met de vernieuwde scoringsmethode en een korte instructie. De hypothese dat het instrument verschillen zou meten tussen agressieve en niet-agressieve kinderen werd niet ondersteund. Ook de hypothese welke de samenhang tussen de beoordeling van agressiviteit door de leraar en de score op de SPT beschrijft, werd niet bevestigd. Ondanks dat de hypothesen met betrekking tot de validiteit niet werden bevestigd, werd toch gekeken naar de overeenkomst tussen de intentievraag en de score op het instrument. Deze hypothese werd door de resultaten gesteund: wanneer kinderen de persoon aardig vonden, gaven zij vaak een aardig antwoord en als kinderen de persoon onaardig vonden gaven zij juist vaak een onaardig antwoord.

Kinderen van verschillende leeftijden bleken verschillend te antwoorden op de SPT. Ten eerste werd gevonden dat de jongste kinderen personen vaker niet aardig en minder vaak gewoon vonden dan oudere kinderen. Ten tweede werd gevonden dat de jongste kinderen vaker onaardige reacties gaven dan oudere kinderen en dat kinderen van negen en tien jaar het meest zeurden. Deze verschillen tussen de leeftijdsgroepen waren echter maar kleine verschillen. De gevonden leeftijdsverschillen zouden het gevolg kunnen zijn van de manier van afname, de jongere kinderen werden immers individueel getest en de oudere kinderen klassikaal. Dit is onderzocht door naar de resultaten van de achtjarige kinderen apart te kijken. De achtjarige kinderen kunnen zowel in groep 4 als 5 zitten, het gevolg hiervan is dat een deel individueel (groep 4) is getest en een deel klassikaal (groep 5). Dit biedt de mogelijkheid om te toetsen of de leeftijdsverschillen veroorzaakt worden door de manier van afname. De scores van de achtjarigen die individueel getest waren bleken niet significant te verschillen van de scores van de achtjarigen die klassikaal waren getest. De verschillen worden blijkbaar veroorzaakt door leeftijd en niet door de manier van afname.

Tevens werd er gekeken naar verschillen tussen responsen op plaatjes met volwassenen en die op de plaatjes met kinderen. Op de vraag naar de intentie van de persoon (vind je deze persoon aardig, gewoon of niet aardig) kozen kinderen bij de plaatjes met volwassenen over het algemeen vaker gewoon en minder vaak aardig en niet aardig dan bij plaatjes waarop twee kinderen stonden afgebeeld. Naast dit hoofdeffect werd een interactie-effect gevonden. De jongste twee groepen kiezen even vaak aardig bij plaatjes met volwassenen en kinderen, terwijl de oudste twee groepen kinderen vaker aardig vinden dan volwassenen. Er werden tevens hoofdeffecten gevonden voor de reacties die in de tekstballon werden geschreven. Kinderen geven vaker een aardige reactie, maar zeuren ook meer tegen volwassenen dan tegen kinderen. Ze geven minder vaak een onaardige of agressieve reactie tegen een volwassene. Naast dit hoofdeffect werd een interactie-effect gevonden. Wanneer de oudste kinderen zeurden was dit vaker tegen een volwassene dan tegen een kind, terwijl de andere leeftijdsgroepen hierin geen onderscheid maakten.

Verklaring van onverwachte resultaten

De onverwacht lage betrouwbaarheid voor de categorie Agressieve reactie wordt waarschijnlijk veroorzaakt doordat er over het algemeen weinig tot niet gescoord werd in deze categorie. De lage betrouwbaarheid van de categorie Probleem niet gezien werd wellicht veroorzaakt doordat dit een soort restcategorie was, waarin antwoorden van kinderen werden geplaatst die niet in een van de andere categorieën thuis hoorden. Dit kan tevens de reden zijn van de lage interbeoordelaarsbetrouwbaarheid tussen de onderzoekers bij deze categorie. Door een miscommunicatie tussen de onderzoekers bij het maken van de afspraken voor het scoren is de betrouwbaarheid van de categorie zeuren lager uitgevallen dan verwacht. Er was een afspraak gemaakt over het veel voorkomende antwoord 'dat is niet eerlijk', maar één van de onderzoekers heeft deze bij het scoren over het hoofd gezien.

Uit dit onderzoek zijn geen aanwijzingen voor validiteit van de SPT gebleken. Dat er geen verschillen zijn gevonden tussen de kinderen van de reguliere en de speciale basisschool kan liggen aan het feit dat de kinderen op de speciale basisschool minder agressief waren dan verwacht. Dit was ook te zien bij de resultaten van de agressievragenlijst. De score van de speciale basisschool verschilde nauwelijks van de score van de reguliere basisschool. Deze score duidde niet op een hoge mate van agressie. Dat er weinig agressieve kinderen meededen aan het onderzoek kan ook gevolgen hebben gehad voor de correlatie tussen de agressievragenlijst en de scores op de categorieën. Omdat de categorie Agressieve reactie een bodemeffect en een kleine standaardafwijking heeft, is het een logisch gevolg dat er geen correlatie gevonden werd met de agressievragenlijst. De agressievragenlijst heeft een bodemeffect: er is alleen maar laag tot gemiddeld en niet hoog op gescoord.

Tekortkomingen van het onderzoek

Op de speciale school bleken weinig agressieve kinderen te zijn. De verwachting dat op een speciale school meer agressieve kinderen zitten dan op een reguliere school bleek niet te kloppen. De groep was daardoor eigenlijk niet geschikt voor het uitvoeren van het valideringsonderzoek. De reguliere basisschool die meedeed aan het onderzoek was een school in een middenstandswijk waardoor het lastig is te generaliseren naar een grotere groep kinderen op het reguliere basisonderwijs. Daarnaast kwamen alle proefpersonen van dezelfde school en zijn telkens alle kinderen uit een klas getest. Hierdoor zijn de proefpersonen niet onafhankelijk van elkaar.

Suggesties voor toekomstig onderzoek

Ook al voldoet het instrument nog niet aan de eisen voor een betrouwbaar en valide instrument, dit onderzoek is wel een goed begin voor het ontwikkelen van een bruikbaar instrument. Het instrument sprak de kinderen aan, ze vulden de antwoorden met plezier in en ze snapten de werking van het instrument goed. De test was ook goed te begrijpen voor kinderen met een leerachterstand. Het is gelukt om een eenvoudiger scoringssysteem te maken welke objectieve resultaten oplevert. In de toekomst zou het beter zijn om met dit instrument een onderzoek op te zetten waarbij de kinderen van tevoren zijn gediagnosticeerd als zijnde agressief. Dit kan bijvoorbeeld door op scholen eerst de Child Behavior Checklist af te nemen en dan het onderzoek voort te zetten met de kinderen die hoog op de agressieschaal scoren. Daarnaast kan ook gezocht worden naar scholen waar de resultaten van de CBCL al beschikbaar zijn. Het zou tevens beter zijn om de controlegroep samen te stellen met kinderen van verschillende basisscholen zodat de resultaten beter te generaliseren zijn.

Aansluiting bij het SIV-model

Het is gebleken dat de ambigue situaties verschillende reacties opleverden en dat tevens de intentie van de persoon op het plaatje verschillend werd beoordeeld. De situaties werden door verschillende kinderen op diverse manieren geïnterpreteerd. Hieruit blijkt dat kinderen in stap twee van het SIV-model dezelfde informatie anders interpreteren. Omdat in dit onderzoek geen echt agressieve kinderen participeerden, kan nog niet worden gezegd of de voorspelling van Dodge (1993) uitkomt. Hij voorspelde dat verschillen tussen agressieve en niet-agressieve kinderen eerder naar voren zullen komen in ambigue situaties.

De resultaten lieten zien dat de reacties van de kinderen vrij vaak overeenkwamen met hun oordeel over de intentie van de persoon in het plaatje. Als zij de persoon aardig vonden, reageerden ze ook vaak aardig en als ze de persoon niet aardig vonden, reageerden ze juist vaker onaardig. Deze intentievraag kan in de klinische praktijk nuttig zijn. Wanneer een kind tegen personen die hij wel aardig vindt toch steeds onaardig is, kan dit een teken zijn dat dit kind vooral moeite heeft met stap vier van het SIV-model. Het kind kiest dan steeds een onaardige reactie terwijl het toch de intentie van de persoon positief geïnterpreteerd heeft. De meeste kinderen beoordeelden niet alle personen op de plaatjes hetzelfde. Ze kozen een aantal keer voor aardig, een aantal keer voor gewoon en een aantal keer voor niet aardig. Het komt echter ook voor dat een kind iedereen onaardig vindt. Dit kan een signaal zijn dat een kind moeite heeft met stap twee van het SIV-model. Wanneer één van bovenstaande gevallen voorkomt kan de therapie van een kind hierop worden aangepast.

In de praktijk zijn er weinig instrumenten beschikbaar om sociale informatieverwerking te meten, waardoor het moeilijker is om geschikte interventies uit te voeren. De instrumenten die er zijn, zijn vaak onaantrekkelijk voor kinderen en zeker niet geschikt voor jonge kinderen. Bestaande instrumenten meten vaak enkel de algemene sociale informatieverwerking en niet die in frustrerende situaties. Juist frustrerende situaties kunnen problematisch zijn voor kinderen waardoor ze agressief gedrag kunnen gaan vertonen. Met de SPT wordt getracht een instrument te maken dat dit probleemgedrag kan voorspellen. Het is daarom belangrijk om onderzoek naar de SPT voort te zetten met de eerder beschreven suggesties zodat er een instrument beschikbaar komt dat wel geschikt en aantrekkelijk voor kinderen is. Om het SIV-model volledig meetbaar te maken, zullen ook de overige stappen in dit model geoperationaliseerd moeten worden in vervolgonderzoek.

Literatuur

- Arsenio, W.F. & Lemerise, E.A. (2004). Aggression and moral development: intergrating social information processing and moral domain models. *Child Development*, 75, 987-1002.
- Dodge, K.A.(1993). Social-cognitive mechanisms in the development of conduct disorder and depression. *Annual Review of Psychology*, 44, 559-584.
- Dodge, K.A., Bates, J.E., Pettit, G.S. & Valente, E. (1995). Social information-processing patterns partially mediate the effect of early physical abuse on later conduct problems. *Journal of Abnormal Psychology*, 104, 632-643.
- Dodge, K.A., Lochman, J.E., Harnish, J.D., Bates, J.E. & Petit, G.S. (1997). Reactive and proactive aggression in school children en psychiatrically impaired chronically assaultive youth. *Journal of Abnormal Psychology*, 106, 37-51.
- Evers, A., Van Vliet-Mulder, J.C. & Ter Laak, J. (1992). *Documentatie van Tests en Testresearch in Nederland*. Assen/Maastricht: Van Gorcum.
- Graybill, D. & Heuvelman, L.R. (1993). Validity of the children's Picture-Frustration Study: A social-cognitive perspective. *A Journal of Personality Assessment*, 60, 379-389.
- Graybill, D., Peterson, S.P. & Williams, P.G. (1989). Variability of responses within the aggression categories on the Rosenzweig Picture-Frustration Study, Children's form. *Journal of Personality Assessment*, 53, 472-477.
- Kievit, Tak & Bosch (2002) *Handboek psychodiagnostiek voor de hulpverlening aan kinderen*. Utrecht, De Tijdstroom Uitgeverij.
- Nieuwenhuijzen, M., Wijnroks, L. & Vermeer, A. (2000). Sociale informatieverwerking en emotionele en gedragsproblemen bij licht verstandelijk gehandicapten kinderen en jongeren.
- Ponsioen, A.J.G.B. (2001). *Cognitieve vaardigheden van lichtverstandelijk gehandicapten en jongeren*. Amsterdam: EPOS.
- Rosenzweig, S. (1978). An investigation of the reliability of the Rosenzweig Picture-Frustration (P-F) Study, Children's form. *Journal of Personality Assessment*, 42, 483-488.
- Rosenzweig, S., Fleming, E.E. & Rosenzweig, L. (1948). The Children's form of the Rosenzweig Picture-Frustration Study. *Journal of Psychology*, 26, 141-191.
- Sameroff, A.J., Lewis, M., & Miller, S.M.(2000). *Handbook of Developmental Psychopathology*. New York, Kluwer Academic.
- Selman, R.L. (1979). *Assessing interpersonal understanding: an interview and scoring manual in five parts constructed by the Harvard-Judge Baken social reasoning project*. Harvard-Judge Baken social reasoning project, Cambridge, MA.

Stranger, C., Achenbach, T.M. & Verhulst, F.C. (1997). Accelerated longitudinal comparisons of aggressive versus delinquent syndromes. *Development and Psychopathology*, 9, 43-58.

Vijftigschild, W., Berger, H.J. Chr. & Van Spaendonck, J.A.S. (1969). Sociale Interpretatie Test.

Waschbusch, D. A., Willoughby, M. T. & Pelham, W. E. (1998). Criterion Validity and the Utility of Reactive and Proactive Aggression: Comparisons to Attention Deficit Hyperactivity Disorder, Oppositional Defiant Disorder, Conduct Disorder, and Other Measures of Functioning. *Journal of Clinical Child Psychology*, 27, 396-405

Bijlage 1

Agressievragenlijst

Voornaam:

Geboortedatum:

	nooit	zelden	soms	vaak	bijna altijd
1 Gebruikt fysiek geweld om te domineren	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
2 Zet anderen tegen klasgenootjes op	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
3 Bedreigt en pest andere kinderen	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
4 Als het kind geplaagd wordt, pikt hij het niet	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
5 Geeft anderen de schuld bij ruzies	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>
6 Reageert overdreven boos bij ongelukjes	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>	<input type="radio"/>

Bijlage 2

Klasseninstructie:

Hallo allemaal, ik ben.... en dit is... Wij komen van de universiteit. We zijn bezig met het maken van een taakje voor kinderen. We weten nog niet zeker of het een goed taakje is en daarom hebben we jullie hulp nodig. We willen kijken of jullie het snappen en te weten komen wat kinderen antwoorden. Dat taakje ligt voor jullie. Voorop staat een uitleg die gaan we nu even samen doorlezen:

Je zult zo meteen een aantal plaatjes zien waarop mensen staan die verschillende dingen zeggen en doen. Bekijk de plaatjes één voor één. Er is altijd één persoon die wat zegt. Lees wat deze persoon zegt. Naast de pratende persoon staan telkens drie gezichtjes waarvan je er één moet aankruisen. Kruis ☺ aan als je de persoon aardig vindt, kruis ☹ aan als je de persoon gewoon vindt, dus niet aardig maar ook niet onaardig en kruis ☹ aan als je de persoon niet aardig vindt. Daarna schrijf je in de lege ruimte van het plaatje op wat je denkt dat de jongen of het meisje zal antwoorden. Schrijf het antwoord op dat het eerst in je op komt, er zijn geen goede of foute antwoorden. Hieronder staan eerst twee voorbeelden die we samen zullen doen.

Op het eerste plaatje zie je een jongen en een meisje, de jongen zegt wat heb je een mooie trui aan. Dan geef je dus eerst aan wat je van die jongen vindt, vind je hem aardig? Dan kruis het lachende gezichtje aan, vind je hem gewoon? Dan kruis je het middelste gezichtje aan en vind je hem niet aardig dan kruis het verdrietige gezichtje aan. Daarna bedenk je wat je terug kan zeggen op : Wat heb je een mooie trui aan. Je zou bijvoorbeeld kunnen zeggen: dankjewel of die heb ik net nieuw of ik vind jouw trui ook mooi of ik vind hem lelijk. Schrijf gewoon het eerste wat in je opkomt op het kan nooit fout zijn.

Bij het tweede voorbeeld zie je meisje en een jongen, het meisje zegt: Wat ben jij stom! Nu geef je weer eerst aan wat je van het meisje vindt, het gaat dan dus altijd om de persoon die al wat zegt! En daarna bedenk je weer wat je terug kan zeggen en schrijf je dat in het lege vak.

Zijn er nog vragen?

Individuele Instructie:

Hallo ik ben... en hoe heet jij? Hoe oud ben je. In welke groep zit je? Mijn vriendin en ik zijn bezig met het maken van een taakje voor kinderen. We weten nog niet zeker of het een goed taakje is en daarom hebben we jou hulp nodig. We willen kijken of jij en je klasgenootjes het snappen en te weten komen wat kinderen antwoorden. Dat taakje ligt voor je. Voorop staat een uitleg die gaan we nu even samen doornemen:

Verder hetzelfde als de klassikale instructie.

Bijlage 3

Instructies

Elke categorie bestaat uit een aantal antwoordmogelijkheden, de categorieën staan hieronder beschreven.

1. **Aardige reactie**, dit zijn positieve reacties waaruit blijkt dat de kinderen de situatie accepteren zoals hij is.
 - Berusting
 - Oplossing geven
 - Excuses aanbieden

Voorbeelden van antwoorden zijn:

Oké, sorry, oké is goed, jammer, ik zal het nooit meer doen, nou ja dan maak ik een nieuwe, oké nu ga ik jou tikken, hahaha dit is leuk, we kunnen een paraplu meenemen.

2. **Probleem niet gezien**, dit zijn verschillende reacties die niet onder een negatieve categorie vallen maar ook niet positief zijn.
 - Probleem niet gezien
 - Probleem klein maken
 - Reden vragen
 - Probleem ontkennen

Voorbeelden van antwoorden zijn:

Waarom (niet)?, Dat heb ik niet gedaan. Verder antwoorden waaruit blijkt dat het probleem niet gesnapt wordt, zoals bij plaatje 16: dankjewel

3. **Zeuren**, dit zijn antwoorden waaruit blijkt dat het kind zeurt. Vaak beginnen deze met: 'maar.....'
 - Zeuren
 - Ander laten oplossen

Voorbeelden van antwoorden:

Maar ik wil hem zo graag, maar ik wil erop, maar wat moet ik dan doen?, ik wil nog spelen, ik wil het zo graag zien, maar ik wou zo graag daar naar toe, ik heb zo'n zin in een snoepje.

4. **Onaardige reactie**, dit zijn antwoorden waaruit blijkt dat het kind de situatie negatief interpreteert.
 - Boze reactie
 - Sarren/pesten
 - Brutaal

Voorbeelden van antwoorden:

Kun je dat niet op een aardige manier vragen?, ik wil eruit!, Nee, ik ga naar beneden, hou je mond, nee! Lekker puh!

5. **Agressieve reactie**, dit zijn antwoorden waaruit blijkt dat het kind de situatie negatief interpreteert en daarbij blijkt dat het kind agressie gebruikt, zowel fysiek als schelden.
 - Agressieve reactie
 - Schelden

Voorbeelden van antwoorden:

Rotwijf, ik vind je heel erg stom ik ram je in elkaar, stommerd, ik duw je eraf.

Elk antwoord kan onder meerdere categorieën vallen en kan beoordeeld worden met een 1, 2 of 3. Wanneer het antwoord helemaal niet van toepassing is op de categorie krijgt dit antwoord op deze categorie een 1. Wanneer het een beetje van toepassing is een 2 en wanneer het helemaal van toepassing is een 3.

Verder hebben we voor een aantal antwoorden afgesproken hoe we die zouden beoordelen:

Plaatje 5:

Dan wil ik maar 1 ding → 2 aardige reactie, 2 niet gezien

Dan wil ik alleen de fiets → 2 aardige reactie, 2 niet gezien

Plaatje 6:

Nee het is mijn step → onaardige reactie 3

Nee → onaardige reactie 2

Plaatje 8: **Mag ik ook** → zeuren 2, niet gezien 2

Plaatje 10:

Ik ben aan de beurt → onaardige reactie 2, niet gezien 2

Plaatje 11:

Ik maak hem heus niet vies → niet gezien 3, zeuren 2

Plaatje 12:

Nee, we gaan omstebeurt → onaardige reactie 2, zeuren 2

Plaatje 14:

Ik ga naar juffie/mammie enz. toe → 2 zeuren, onaardige reactie 2

Plaatje 16:

Stommerd → onaardige reactie 3, agressieve reactie 2

Plaatje 17:

Niet ik ben groot → zeuren 2, onaardige reactie 2, niet gezien 2

Plaatje 19:

Maar je had het beloofd, snik → zeuren 2, niet gezien 2

Dat is niet eerlijk → zeuren 2, onaardige reactie 2

Ja maar → zeuren 3, niet gezien 2

Oh nee he → zeuren 2, berusting 2

Plaatje 20:

Nou ik ga ook niet met jou spelen → onaardige reactie 2

Dat vind ik niet leuk → onaardige reactie 2

Bijlage 4

Brief voor de ouders

Amsterdam, 28 oktober, 2004

Aan de ouders(s)/begeleider(s) van kinderen van de Golfbreker,

In het kader van de ontwikkeling van een instrument dat reacties van kinderen in sociale situaties meet, wordt in de maand november op de school van uw kind een onderzoek verricht. Dit onderzoek wordt uitgevoerd door twee vierdejaars studenten van de vakgroep Ontwikkelingspsychologie van de Universiteit van Amsterdam.

Het onderzoek houdt in dat uw kind tijdens de les een aantal plaatjes te zien krijgt. Op de plaatjes staan minstens twee personen. Eén van de personen zegt iets en het kind moet aangeven wat hij of zij zou antwoorden in deze situatie. De kinderen zullen het leuk vinden om antwoord te geven omdat de plaatjes aansprekende dagelijkse situaties laten zien.

De directie, als mede de leerkrachten, zijn met het onderzoek akkoord gegaan. Uiteraard worden de antwoorden van de kinderen anoniem en uiterst zorgvuldig behandeld. Wanneer u niet wilt dat uw kind deelneemt aan het onderzoek, kunt u onderstaand strookje invullen en inleveren bij de leerkracht vóór Uw kind zal dan niet deelnemen aan het onderzoek.

Met vriendelijke groeten,

Claire Verwey

Jet Vleeming

De ouder(s) / begeleider(s) van
geven **GEEN** toestemming voor het onderzoek.

Datum:

Handtekening:

Bijlage 6

Scoringsformulier

Sociale Plaatjes Test

Naam _____
 Geslacht _____
 Groep _____
 Testleider _____

	Jaar	Maand	Dag
Testdatum			
Geboortedatum			
Leeftijd			

Categorieën	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1. Aardige Reactie																				
2. Probleem Niet Gezien																				
3. Zeuren																				
4. Onaardige Reactie																				
5. Agressieve Reactie																				

Intentie	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Aardig																				
Gewoon																				
Niet Aardig																				

Categorieën	Totaalscore	Decielscore
1. Aardige Reactie	_____	_____
2. Probleem Niet Gezien	_____	_____
3. Zeuren	_____	_____
4. Onaardige Reactie	_____	_____
5. Agressieve Reactie	_____	_____

Categorieën	Aantal keer	Z-score
Aardig	_____	_____
Gewoon	_____	_____
Niet Aardig	_____	_____

Bijlage 7

Normtabellen voor de totaalscores van de categorieën voor 4 leeftijdsgroepen

<i>Decielen</i>	<i>Scores Aardige reactie</i>			
	5-6	7-8	9-10	11-12
1 ^e deciel	20 - 22	20 - 25	20 - 30	20 - 30
2 ^e deciel	23 - 24	26 - 31	31 - 32	31 - 32
3 ^e deciel	25 - 28	32 - 33	33 - 34	33 - 34
4 ^e deciel	29 - 34	34 - 37	35	35 - 36
5 ^e deciel	-	-	36	37
6 ^e deciel	35 - 37	38 - 40	37 - 38	38
7 ^e deciel	38 - 41	41	39 - 41	39 - 40
8 ^e deciel	42 - 45	42 - 43	42 - 44	41 - 43
9 ^e deciel	46 - 50	44	45 - 46	44 - 45
10 ^e deciel	51 - 60	45 - 60	47 - 60	46 - 60

<i>Decielen</i>	<i>Scores Probleem niet gezien</i>			
	5-6	7-8	9-10	11-12
1 ^e deciel	20	20	20 - 21	20 - 22
2 ^e deciel	21	21	-	-
3 ^e deciel	22	22	22 - 23	23 - 24
4 ^e deciel	-	23	24	-
5 ^e deciel	23	24	-	25
6 ^e deciel	24 - 25	25	25 - 26	-
7 ^e deciel	26	26	27 - 28	26
8 ^e deciel	27 - 28	27 - 28	29	27 - 28
9 ^e deciel	29 - 31	29	30 - 31	29 - 31
10 ^e deciel	32 - 60	30 - 60	32 - 60	32 - 60

<i>Decielen</i>	<i>Scores Zeuren</i>			
	5-6	7-8	9-10	11-12
1 ^e deciel	-	-	20 - 22	20 - 23
2 ^e deciel	-	20 - 22	23 - 24	24
3 ^e deciel	20 - 22	-	-	25
4 ^e deciel	23	23 - 24	25 - 27	26
5 ^e deciel	24	-	-	27
6 ^e deciel	25	25 - 26	28 - 31	28
7 ^e deciel	26 - 27	27	32 - 33	29
8 ^e deciel	28 - 29	28 - 29	-	30 - 32
9 ^e deciel	30 - 37	30 - 31	34 - 37	33
10 ^e deciel	38 - 60	32 - 60	38 - 60	34 - 60

<i>Decielen</i>	<i>Scores Onaardige reactie</i>			
	5-6	7-8	9-10	11-12
1 ^o deciel	20 - 23	20 - 23	20 - 22	20 - 24
2 ^o deciel	24	24 - 26	23	25 - 26
3 ^o deciel	25 - 26	27	24 - 25	-
4 ^o deciel	27 - 30	28	26	27 - 28
5 ^o deciel	31	29	27	29
6 ^o deciel	32	30	28	30
7 ^o deciel	33 - 35	31 - 32	29	31 - 32
8 ^o deciel	36 - 40	33 - 35	30 - 31	-
9 ^o deciel	41 - 44	36 - 38	32 - 33	33 - 37
10 ^o deciel	45 - 60	39 - 60	34 - 60	38 - 60

<i>Decielen</i>	<i>Scores Agressieve reactie</i>			
	5-6	7-8	9-10	11-12
1 ^o deciel	-	-	-	-
2 ^o deciel	-	-	-	-
3 ^o deciel	-	-	-	-
4 ^o deciel	-	-	-	-
5 ^o deciel	-	-	-	-
6 ^o deciel	-	-	-	-
7 ^o deciel	20	-	-	-
8 ^o deciel	-	-	-	20 - 21
9 ^o deciel	21 - 23	20 - 21	20	22
10 ^o deciel	24 - 60	22 - 60	21 - 60	23 - 60