

Betrouwbaarheid en validiteit van de LIRIK

Eindrapport LIRIK valideringsonderzoek

© 2015 Nederlands Jeugdinstituut Niets uit deze uitgave mag worden vermenigvuldigd en/of openbaar gemaakt door middel van druk, fotokopie, microfilm of op enige andere wijze zonder voorafgaande schriftelijke toestemming.

Dit rapport werd (mede) mogelijk gemaakt door Stichting Kinderpostzegels.

Redactie

Cora Bartelink

Leontien de Kwaadsteniet

Ingrid ten Berge

Cilia Witteman

Willemijn van Gastel

Nederlands Jeugdinstituut

Catharijnesingel 47

Postbus 19221

3501 DE Utrecht

Telefoon 030 - 230 63 44

Website www.nji.nl

E-mail info@nji.nl

Nederlands
Jeugd
instituut



Radboud Universiteit



juzt Bij opvoeden
en opgroeien

spirit!
voor hulp aan jeugd en opvoeders



Gelderland-Zuid

Woord vooraf

Inschatten of er sprake is van onveiligheid of risico's daarop is lastig. Ouders en jeugdigen moeten er op kunnen rekenen dat deze voor hen belangrijke beslissingen zo zorgvuldig mogelijk genomen worden. In 2007 is daarvoor het Licht Instrument Risicotaxatie Kindveiligheid (LIRIK) ontwikkeld. Dit instrument zou professionals moeten helpen om systematisch de veiligheid en risico's voor jeugdigen in gezinnen in te schatten. Steeds meer professionals zijn in de loop van de tijd met de LIRIK gaan werken.

Het Nederlands Jeugdinstituut wil eraan bijdragen dat professionals in de zorg voor jeugd werken met betrouwbare en valide instrumenten en met effectieve methoden. Daarbij hoort ook onderzoek doen naar de effecten daarvan. Het Nederlands Jeugdinstituut en de Radboud Universiteit Nijmegen hebben in samenwerking met Juzt, Spirit en GGD Gelderland-Zuid het afgelopen jaar onderzocht hoe de betrouwbaarheid en validiteit van de LIRIK zijn. Voorliggend rapport beschrijft de opzet en resultaten van dit onderzoek. Wij hopen dat het de instellingen voor jeugdgezondheidszorg, jeugdhulp en jeugdbescherming zal stimuleren om verder te werken aan verbeteringen in de oordeels- en besluitvorming over kindermishandeling.

Dit onderzoek was niet mogelijk geweest zonder de medewerking van medewerkers bij GGD Gelderland-Zuid, Juzt en Spirit. Op deze plek willen wij in de eerste plaats alle deelnemers aan het onderzoek hartelijk danken voor hun tijd en inzet. Daarnaast gaat onze dank uit naar de bestuurders, die het belang van dit onderzoek onderschreven en de uitvoering daarvan hebben ondersteund.

Enkele mensen willen wij in het bijzonder bedanken voor hun inzet gedurende het hele onderzoek. Mascha Kamphuis, Ignace Vermaes, Carlo Schuengel en Karin Kleine hebben als leden van de klankbordgroep meegedacht bij de opzet, uitvoering en rapportage van het onderzoek. Hun opbouwende kritiek hebben wij als zeer behulpzaam ervaren. Zij onderschrijven de conclusies die we in dit rapport hebben getrokken.

Floor Minkhorst en Raoul Hakkenberg van Gaasbeek hebben meegewerkt aan de uitvoering van de twee deelonderzoeken. Wij bedanken ook Pieter Groenestijn voor de technische ondersteuning bij het maken van de digitale vragenlijst van deelonderzoek 1 en Pierre Souren voor zijn adviezen over de statistische analyses. Het onderzoek is mogelijk gemaakt met financiering van Stichting Kinderpostzegels.

Mede namens het hele projectteam,

Cora Bartelink

Leontien de Kwaadsteniet

Ingrid ten Berge

Mariëtte Blik

Mariëtte van Brandenburg

Willemijn van Gastel

Mariëlle Hornstra

Carolien Konijn

Cilia Witteman

Utrecht, juli 2015

Inhoudsopgave

Woord vooraf	2
Inhoudsopgave	5
Samenvatting	7
1 Inleiding	9
1.1 Inleiding	9
1.1.1 Licht Instrument Risicotaxatie Kindveiligheid (LIRIK)	9
1.1.2 Aanleiding	11
1.2 Probleemstelling en relevantie	12
1.2.1 Probleemstelling	12
1.2.2 Belang van het onderzoek	14
1.3 Doel- en vraagstelling en hypothesen	15
1.3.1 Doelstelling	15
1.3.2 Vraagstelling	15
1.3.3 Hypothesen	15
1.4 Onderzoeksopzet	16
1.5 Leeswijzer	17
2 Deelonderzoek 1. Betrouwbaarheid	18
2.1 Onderzoeksvragen	18
2.2 Methode	18
2.2.1 Deelnemers	18
2.2.2 Procedure	20
2.2.3 Materialen	20
2.2.4 Gevalsbeschrijvingen	22
2.2.5 Analyses	22
2.3 Resultaten	23
2.3.1 Interbeoordelaarsovereenstemming over conclusies over veiligheid en risico's gemaakt met en zonder de LIRIK	23
2.3.2 Is de interbeoordelaarsovereenstemming over conclusies over veiligheid en risico's hoger wanneer de LIRIK wordt gebruikt dan wanneer de LIRIK niet wordt gebruikt?	26
2.3.3 Interbeoordelaarsovereenstemming voor de aanwezigheid van specifieke factoren die op mogelijke kindermishandeling of verhoogd of verlaagd risico kunnen duiden.	27
2.3.4 Evaluaties van deelnemers van de bruikbaarheid van de LIRIK	35
2.4 Discussie	36
2.4.1 Beperkingen	37
3 Deelonderzoek 2. Predictieve validiteit	39
3.1 Inleiding	39
3.2 Methode	40
3.2.1 Onderzoeksgroep	40

3.2.2	Procedure	41
3.2.3	Analyses.....	43
3.3	Resultaten voor de predictieve validiteit	44
3.3.1	Samenhang tussen veiligheids- en risico-inschattingen en uitkomsten –rangcorrelaties.....	46
3.3.2	Samenhang tussen veiligheids- en risico-inschattingen en uitkomsten – ordening naar niveau van onveiligheid en risico	47
3.3.3	Samenhang tussen de LIRIK-items en uitkomsten op T2.....	51
3.4	Resultaten voor de interne consistentie.....	55
3.5	Discussie	59
3.5.1	Beperkingen	62
4	Conclusies, discussie en aanbevelingen	63
4.1	Inleiding.....	63
4.2	Resultaten samengevat	63
4.2.1	Betrouwbaarheid: Interbeoordelaarsovereenstemming	63
4.2.2	Predictieve validiteit	64
4.3	Verklaringen voor bevindingen.....	64
4.4	Beperkingen van het onderzoek	65
4.5	Aanbevelingen	66
4.5.1	Verbeteringen in de praktijk	67
4.5.2	Verbetering van de LIRIK	68
4.5.3	Verder onderzoek en theorievorming.....	69
Referenties		70
Bijlagen		74
Bijlage 1.	Licht Instrument Risicotaxatie Kindveiligheid (LIRIK).....	75
Bijlage 2.	Ongestructureerd professioneel oordeel.....	79
Bijlage 3.	Checklist dossieronderzoek	80
Bijlage 4.	Interpretatie van overeenstemming en samenhang.....	87
Bijlage 5.	Uitkomsten LIRIK.....	88
Bijlage 6.	Gegevens over T1 en aanvullende gegevens over T2	95
Bijlage 7.	Oorspronkelijke rangcorrelaties (daadwerkelijke en ideale rangcorrelaties).....	97
Bijlage 8.	Zachte uitkomsten.....	98
Bijlage 9.	Klankbordgroep.....	101

Samenvatting

Het NJi heeft de LIRIK ontwikkeld om hulpverleners te ondersteunen bij een professionele beoordeling van de veiligheid van jeugdigen. Gezien het wijdverbreide en toenemende gebruik van de LIRIK en de implicaties voor jeugdige en ouders van mogelijk verkeerde oordelen, is het cruciaal dat betrouwbaarheid en validiteit van de LIRIK en de meerwaarde ten opzichte van het professioneel oordeel zonder instrument vastgesteld wordt.

Deelonderzoek 1: Betrouwbaarheid

Betrouwbaarheid van oordelen houdt in dat verschillende professionals op basis van dezelfde informatie tot dezelfde oordelen komen.

In dit deelonderzoek hebben professionals (ambulante begeleiders en behandelcoördinatoren) voor 6 van in totaal 12 gevalsbeschrijvingen veiligheidsoordelen en risico-inschattingen gegeven, waarbij zij de LIRIK gebruikten (36 deelnemers) of de LIRIK niet gebruikten (43 deelnemers). We vroegen ook naar gebruikerservaringen met de LIRIK.

Resultaten:

- De overeenstemming tussen professionals bleek onvoldoende voor de veiligheidsoordelen en risico-inschattingen gemaakt met de LIRIK. De overeenstemming over risico's in de nabije toekomst, rekening houdend met voorziene veranderingen, bleek het laagst. Overeenstemming voor de LIRIK-items voorafgaand aan deze conclusies was in de meeste gevallen ook onvoldoende, slechts voor enkele items goed.
- De overeenstemming over de veiligheidsoordelen en risico-inschattingen bleek niet beter wanneer de LIRIK was gebruikt dan wanneer de LIRIK niet was gebruikt.
- Deelnemers waren vrij positief over de ondersteuning van het professionele oordeel door de LIRIK.

Deelonderzoek 2: Predictieve validiteit

Predictieve validiteit is de mate waarin een meetinstrument in staat is een gebeuren te voorspellen. In dit onderzoek is onderzocht of de oordelen die professionals zich vormen over de huidige veiligheid en toekomstige risico's voorspellen of er een AMK-melding, raadsmelding, ondertoezichtstelling, uithuisplaatsing, interventie gericht op veiligheid of crisis geweest is binnen een half jaar na de beoordeling. Daarvoor zijn 278 beoordelingen met de LIRIK en 92 beoordeling zonder de LIRIK verzameld en is na een half jaar een dossieronderzoek uitgevoerd om na te gaan of één of meerdere van de gekozen uitkomstmaten zich hadden voorgedaan in deze zaken.

Gebleken is:

- De predictieve validiteit bleek over het geheel genomen laag voor de oordelen gemaakt met de LIRIK over de huidige veiligheid en de risico's voor de toekomst.
- De predictieve validiteit voor deze oordelen bleek niet beter wanneer de LIRIK was gebruikt dan wanneer de LIRIK niet was gebruikt. Bij sommige uitkomsten leek de LIRIK het beter te doen, bij andere het ongestructureerd professioneel oordeel.

- De meeste afzonderlijke LIRIK-items die voorafgaan aan de conclusies bleken niet of nauwelijks met de uitkomsten samen te hangen, maar bij een aantal items was er een zwakke tot matige samenhang te zien.
- Tussen de LIRIK-items en de daaropvolgende oordelen over de huidige veiligheid en risico's bleek vaak meer samenhang te zijn. Dit betekent dat de oordelen logisch volgen op de systematische analyse van zorgen en risicofactoren. Dit gold echter niet voor de beschermende factoren.

Conclusies en implicaties:

- De betrouwbaarheid en predictieve validiteit voor oordelen gemaakt met de LIRIK bleken onvoldoende en niet beter dan voor ongestructureerde oordelen. Kanttekening daarbij moet wel zijn dat medewerkers die situaties ongestructureerd beoordeelden, meestal de LIRIK kenden en getraind waren in het gebruik ervan.
- De betrouwbaarheid en validiteit van andere risicotaxatie-instrumenten zijn ook beperkt en niet of nauwelijks beter dan de resultaten met de LIRIK. Bij gebruik van risicotaxatie-instrumenten hoort een grote mate van onzekerheid ten aanzien van de juistheid van oordelen.
- Professionals moeten zich ervan bewust zijn dat de LIRIK geen instrument is dat leidt tot een objectief oordeel over de veiligheid en risico's. Wanneer professionals zich hier niet bewust van zijn, kan het gebruik van instrumenten leiden tot overmatig vertrouwen in hun conclusies.
- Mogelijk kan de LIRIK hen wel helpen om met een brede blik een systematische analyse te maken en op basis daarvan tot een professioneel oordeel te komen.
- Mogelijk valt er winst te behalen als professionals veel meer in gezamenlijkheid de veiligheid en risico's beoordelen: samen met ouders en jeugdigen en samen met een kritisch meedenkende collega of gedragswetenschapper.

1 Inleiding

Ingrid ten Berge, Cilia Witteman, Cora Bartelink & Leontien de Kwaadsteniet

1.1 Inleiding

1.1.1 *Licht Instrument Risicotaxatie Kindveiligheid (LIRIK)*

Om kinderen en ouders goed hulp te kunnen verlenen, is het noodzakelijk een goede inschatting te maken van de huidige veiligheidssituatie van het kind (is het kind op dit moment thuis veilig?), factoren die daar in positieve of negatieve zin aan bijdragen en de risico's voor het kind in de toekomst.

Het Nederlands Jeugdinstituut heeft in 2007 met en voor de Bureaus Jeugdzorg (inclusief AMK's) het Licht Instrument Risicotaxatie Kindveiligheid (LIRIK; Ten Berge, Eijgenraam & Bartelink, 2014) ontwikkeld om hulpverleners te ondersteunen bij een professionele beoordeling van kindveiligheid. De inhoud van de LIRIK is gebaseerd op een uitgebreide literatuurstudie naar risico- en protectieve factoren voor kindermishandeling, kenmerken van mishandelende ouders en kinderen en kindsignalen. Daarnaast is gebruik gemaakt van de internationale onderzoeksliteratuur over besluitvorming in situaties van (vermoedelijke) kindermishandeling en de expertise van medewerkers van het Advies- en Meldpunt Kindermishandeling (AMK)¹.

De LIRIK is een checklist gericht op gestructureerde oordeelsvorming. Dat wil zeggen dat het bedoeld is om de hulpverlener te helpen om op grond van de beschikbare informatie systematisch de voorliggende casus te beoordelen. Het professionele oordeel van de hulpverlener staat daarbij centraal. De LIRIK is bedoeld om de hulpverlener te ondersteunen bij het expliciteren en onderbouwen van zijn oordeel. De LIRIK biedt op gestructureerde wijze relevante aandachtspunten om tot een expliciet en weloverwogen oordeel ten aanzien van de veiligheid te komen. De LIRIK is echter geen actuarieel taxatie-instrument, dat wil zeggen het geeft geen formules om tot een objectief oordeel te komen. De uiteindelijke beoordeling maakt de hulpverlener zelf door de zorgen, problemen en sterke kanten tegen elkaar af te wegen. De meerwaarde van de LIRIK ten opzichte van het eigen professionele oordeel ligt in de focus op (on)veiligheid, het expliciteren van de professionele intuïtie, het voorkomen van een tunnelvisie of andere valkuilen, het stilstaan bij feiten en het aanreiken van kennis over uit onderzoek bekende relevante aandachtspunten.

De LIRIK vraagt op gestructureerde wijze naar relevante signalen, risicofactoren en beschermende factoren, zodat werkers tot een expliciet en onderbouwd oordeel over aard en ernst van de problematiek komen dat de basis kan vormen voor verschillende beslissingen, bijvoorbeeld om een AMK-melding of Raadsmelding te doen of om hulp in te zetten. De checklist is ontwikkeld op basis van bestaande literatuur over factoren die samenhangen met (on)veiligheid en risico's op kindermishandeling. Voor het nemen van een besluit is naast de beoordeling van de problematiek ook een analyse nodig van de gewenste situatie en mogelijke oplossingsrichtingen. De LIRIK geeft daar geen richtlijnen voor.

De LIRIK bestaat uit twee delen om de beoordeling van de veiligheid van de jeugdige te ondersteunen:

- Deel 1: huidige veiligheidssituatie
- Deel 2: risicotaxatie

¹ Het Advies- en Meldpunt Kindermishandeling bestaat sinds 1 januari 2015 niet meer, maar is nu Veilig Thuis. Omdat het onderzoek grotendeels vóór 1 januari 2015 is uitgevoerd, spreken we nog van het AMK.

Beide delen van de LIRIK bestaan uit een aantal vetgedrukte kernvragen, aandachtspunten voor het beantwoorden van deze kernvragen, en een algemene conclusie. In principe worden altijd beide delen ingevuld.

Deel 1 ondersteunt de professional bij het beoordelen van de *huidige situatie* van de jeugdige. Het gaat om de vraag ‘Wat is er op dit moment aan de hand?’. Is de jeugdige op dit moment veilig of zijn er signalen die wijzen op kindermishandeling? Kernvragen in deel 1 zijn:

- Zijn er vermoedens of aanwijzingen dat de jeugdige op dit moment in een levensbedreigende situatie verkeert of fysiek in gevaar is?
- Zijn er aanwijzingen voor bedreigend handelen of nalaten van de ouder(s)?
- Zijn er kindsignalen die kunnen wijzen op kindermishandeling?
- Zijn er risicofactoren voor kindermishandeling (bij ouders, jeugdige, gezin en/of omgeving)?
- Zijn er beschermende factoren (bij ouders, jeugdige, gezin en/of omgeving)?

Deel 1 leidt tot een van de volgende conclusies over de huidige veiligheidssituatie:

- Er lijkt op dit moment geen sprake van kindermishandeling
- Er is mogelijk sprake van kindermishandeling: er lijkt sprake van bedreigend handelen of nalaten van de ouder(s); er zijn kindsignalen die daarop wijzen; er zijn risicofactoren bij de ouder(s); er zijn risicofactoren bij jeugdige, gezin en/of omgeving; ouder geeft geen informatie of inzicht in eigen handelen.
- Er is sprake van feitelijk aangetoonde kindermishandeling, op dit moment of in het verleden (gestopt)
- Er is sprake van een levensbedreigende situatie of direct fysiek gevaar
- Onvoldoende informatie om een oordeel te vormen

Deel 2 is gericht op risicotaxatie. Risicotaxatie is het inschatten van de risico's voor de jeugdige in de *toekomst*. Daarbij staat de vraag centraal ‘Wat kan er gaan gebeuren?’. Risicotaxatie is gebaseerd op informatie uit het heden: op grond van kenmerken van de huidige situatie probeer je toekomstige schade of bedreiging voor de jeugdige zo goed mogelijk in te schatten. Een goede risicotaxatie is dus alleen mogelijk als de huidige situatie met behulp van de kernvragen uit deel 1 goed in beeld is gebracht.

Kernvragen in deel 2 zijn:

- Wat kan er gebeuren? Aard en kans van diverse mogelijke risico's voor de jeugdige, i.e. levensbedreigende situatie / direct fysiek gevaar; ontstaan van kindermishandeling; voortduren/herhaling van kindermishandeling
- Ernst van de verwachte gevolgen voor de jeugdige
- Zijn er beschermende factoren die de risico's kunnen verminderen?

Deel 2 leidt tot een conclusie over aard en ernst (op een schaal van 1 tot 5) van de risico's voor de jeugdige op dit moment en bij voorziene veranderingen in de nabije toekomst.

In Bijlage 1 is te LIRIK in zijn geheel te vinden.

De LIRIK wordt inmiddels gebruikt door alle Bureaus Jeugdzorg en diverse instellingen voor hulp en steun aan jeugdigen en ouders (schoolmaatschappelijk werk, geïndiceerde jeugdzorg).

Met subsidie van Stichting Kinderpostzegels is een pilotstudie naar de bruikbaarheid van de LIRIK in LVB instellingen gedaan.² De LIRIK wordt voorts door verschillende instellingen ingezet bij stap 4 van de meldcode huiselijk geweld en kindermishandeling (het wegen van risico's, aard en ernst van het geweld). Deze meldcode is sinds medio 2013 verplicht voor alle professionals die werken met ouders en/of jeugdigen. Het NJi signaleert als gevolg hiervan een stijgende vraag naar de LIRIK vanuit voorliggende voorzieningen als CJG, jeugdgezondheidszorg en kinderopvang. Onderzoek naar de LIRIK is dus urgent.

1.1.2 Aanleiding

Veiligheid van kinderen is een belangrijk en maatschappelijk relevant thema. In Nederland wonen circa 3,5 miljoen jeugdigen tussen 0 en 18 jaar (bron: CBS). Uit Nederlands prevalentie-onderzoek (Alink et al., 2011) is bekend dat tenminste 3,4% van deze jeugdigen, dus 119.000 kinderen, slachtoffer is van kindermishandeling. Dit zijn dan de slachtoffers die door hulpverleners worden gezien. Het daadwerkelijke aantal mishandelde jeugdigen ligt vrijwel zeker veel hoger. In hetzelfde onderzoek geeft 9,9% van de jeugdigen tussen 12 en 17 jaar in zelfrapportage aan zich slachtoffer te voelen. Recent onderzoek van de Universiteit Tilburg (Lahlah, Van der Knaap & Bogarts, 2013) onder bijna 500 jongeren laat zien dat een op de vijf Nederlandse jongens en een op de drie Marokkaans-Nederlandse jongens ooit te maken heeft gehad met grof huiselijk geweld. Bijna de helft van de Marokkaans-Nederlandse jongens is ooit getuige geweest van fysiek geweld tussen de ouders; bij de Nederlandse jongens is dat bijna een op de vijf.

Instellingen die werken met ouders en jeugdigen, onder meer in de jeugdzorg, werken primair samen met ouders vanuit vertrouwen in eigen kracht en regie van de cliënt. Toch zijn er helaas situaties waarin sprake is van opvoedingsnood of –crisis, waarbij de veiligheid van de jeugdige (ernstig) bedreigd wordt of dreigt te worden. In die situaties is het gewenst dat ouders ondersteund worden in het herstellen van de veiligheid. Het is daarbij cruciaal dat alleen daadwerkelijk onveilige situaties worden aangepakt. Oversignalering kan het vertrouwen van de ouders in de zorg ernstig schaden, ondersignalering is gevaarlijk voor de ontwikkeling van de jeugdige en kan in de toekomst problemen voor de participatie in de samenleving opleveren. Daarom hebben instellingen behoefte aan instrumenten die de beoordeling van de veiligheid van jeugdigen en de dialoog hierover tussen alle betrokkenen kunnen ondersteunen.

Uit onderzoek is bekend dat hulpverleners onderling aanzienlijk kunnen verschillen in hun beoordeling van veiligheid van kinderen (Regehr, Bogo, Shlonsky, & LeBlanc, 2010; Rossi, Schuerman, & Budde, 1996; Ten Berge, Bartelink, & De Kwaadsteniet, 2011). Dat is deels te wijten aan het feit dat er geen concreet beleid en geen concrete, op empirisch onderzoek gestoelde richtlijnen zijn (Osma & Benbenishty, 2004), waardoor werkers grotendeels op hun eigen ervaring en onderscheidingsvermogen moeten varen (Arad-Davidzon &

² Doel van deze studie was om door middel van een gebruikersevaluatie na te gaan of de LIRIK inhoudelijk voldoende relevant en dekkend is voor de specifieke doelgroep ouders en kinderen met LVB problematiek en wat randvoorwaarden zijn voor implementatie (Ten Berge & Meuwissen, 2013).

Benbenishty, 2008). Als gevolg daarvan is de betrouwbaarheid van de oordelen van sociaal werkers laag (Lindsey, 1992).

Om verkeerde beslissingen te voorkomen wordt het besluitvormingsproces vaak expliciet gemaakt en gestructureerd (zie o.m. De Bruyn, Ruijssenaars, Pameijer & Van Aarle, 2003; Munro, 2008; Pameijer & Van Beukering, 2004; Shlonsky & Wagner, 2005). Door gestructureerd te werken zijn professionals beter in staat om weloverwogen een beslissing te nemen over de verschillende mogelijkheden die zij hebben om in te grijpen.

Met behulp van een instrument voor gestructureerde oordeelsvorming kunnen hulpverleners naar verwachting de veiligheid en de veiligheidsrisico's van jeugdigen zorgvuldiger en beter beoordelen. Dit komt ten goede aan de kwaliteit van zorg voor deze doelgroep (Regehr et al., 2010).

1.2 Probleemstelling en relevantie

1.2.1 Probleemstelling

Gezien het wijdverbreide en toenemende gebruik van de LIRIK en de implicaties voor jeugdige, ouders en samenleving van mogelijk verkeerde oordelen (ten onrechte aangewezen als onveilig of hoog risico op kindermishandeling (vals positief) of ten onrechte niet herkend als onveilig of hoog risico (vals negatief)), is het cruciaal dat de kwaliteit van de LIRIK vastgesteld wordt. De bedoeling is immers dat de LIRIK bijdraagt aan een betere beoordeling van de veiligheid van jeugdigen, dat wil zeggen dat de LIRIK tot meer betrouwbare en valide oordelen moet leiden dan wanneer dit instrument niet wordt gebruikt. Dat was vóór dit onderzoek nog niet aangetoond. In feite gebruiken veel professionals dus een instrument waarvan niet is vastgesteld dat het beter meet wat het beoogt te meten. De Nederlandse jeugdzorg heeft daarnaast weinig alternatieven voor de LIRIK. Er zijn enkele checklists voor het beoordelen van veiligheid ('Veilig Thuis?'-kaart, Delta veiligheidslijst), of voor risicotaxatie (CARE-NL, CFRA), maar ook deze zijn niet of slechts beperkt onderzocht op betrouwbaarheid en validiteit (Bartelink & Kooijman, 2013; Ten Berge, 2008).

Onderzoek naar CARE-NL en CFRA

De Nederlandse versie van de Child Abuse Risk Evaluation-Nederland (CARE-NL; De Ruiter & De Jong, 2006) wordt in Nederland op beperkte schaal gebruikt voor risicotaxaties van kindermishandeling. Hiermee worden de 18 belangrijkste risicofactoren voor kindermishandeling geïnventariseerd. Zoals uit onderstaande blijkt, is dit instrument niet grondig onderzocht en zijn de eerste resultaten nog niet overtuigend, dus kan (nog) niet geconcludeerd worden dat het instrument betrouwbaar en valide is. Bovendien is de CARE-NL niet, zoals de LIRIK, ook bedoeld voor veiligheidstaxatie. Er zijn geen gebruikerservaringen van de CARE-NL bekend.

In een retrospectief onderzoek zijn de interbeoordelaarsbetrouwbaarheid en de predictieve validiteit van de CARE-NL onderzocht (De Ruiter, Hildebrand, & Van der Hoorn, 2012). Zestig afgesloten dossiers van AMK's werden door drie speciaal getrainde codeurs (studenten) beoordeeld met de CARE-NL. ICC's voor een beoordelaar varieerden van -.08 tot 1 voor de verschillende factoren (de meeste >.40) en waren .59 voor risico's ingeschat voor de verzorger en .56 voor de risico's ingeschat voor het gezin. Deze ICC's waren voor

consistentie, niet voor absolute overeenstemming die dus lager zal uitvallen. De speciaal getrainde codeurs waren geen AMK-medewerkers. Omdat het om afgesloten dossiers gaat en de codeurs speciaal getraind maar geen AMK-medewerkers waren, geeft dit onderzoek geen goed beeld van de overeenstemming die daadwerkelijk in de praktijk behaald had kunnen worden.

De beoordelingen zijn vervolgens gerelateerd aan wat er sinds aanmelding met de beschreven cliënten gebeurd is: her-aanmelding bij een AMK, uithuisplaatsing, een OTS-maatregel. De auteurs schrijven dat het oordeel met behulp van de CARE-NL een redelijke predictieve validiteit heeft voor OTS (AUC =.73) en een goed voorspellend vermogen voor uithuisplaatsing (AUC =.77). Als kanttekening valt daarbij te maken dat de wijze waarop oordelen en uitkomsten aan elkaar gerelateerd zijn niet duidelijk is. Het is bijvoorbeeld niet duidelijk of individuele beoordelingen gemaakt met de CARE-NL zijn gebruikt voor het berekenen van de AUC, of de gemiddelde oordelen van de drie getrainde codeurs. Ook is onduidelijk hoe de oordelen zijn gedichotomiseerd om de sensitiviteit en de specificiteit te berekenen. Daarnaast is een beperking van dit retrospectieve onderzoek aan de hand van dossiers dat de risico-inschattingen zijn gemaakt op basis van dezelfde dossiers als waar de uitkomstmaten zijn uitgehaald, de AMK-dossiers geven wellicht al een selectie van informatie.

De CARE-NL is ook onderzocht door AMK teams te vergelijken met derdejaarsstudenten psychologie bij de beoordeling van twee cases (De Ruiter, De Jong, & Reus, 2013). De beoordelingen werden inhoudelijk gecodeerd en vergeleken, en de frequenties waarmee de risicofactoren werden benoemd werden berekend en vergeleken. De auteurs vonden dat er door professionals niet meer relevante factoren genoemd werden dan door studenten, en in het algemeen werd aan de belangrijkste categorie, de ouderfactoren, zeer weinig aandacht besteed. De auteurs concluderen dat de CARE-NL kan helpen bij het systematisch inventariseren en beschouwen van relevante risicofactoren, hoewel hun bevindingen lieten zien dat mensen die gebruik maken van het instrument belangrijke risicofactoren onderbelichten.

Ook de Nederlandse versie van de California Family Risk Assessment wordt maar beperkt gebruikt. Het gaat om een instrument dat systematische beoordelingen van risico's op onveiligheid in de toekomst ondersteunt door niet-specialisten bij bijvoorbeeld Centra voor Jeugd en Gezin en Maatschappelijk werk, terwijl de LIRIK en de CARE-NL bedoeld zijn voor taxaties na aanmelding bij de jeugdzorg (BJZ, AMK). Gebruikerservaringen ervan zijn niet bekend.

Recent is ook onderzoek gedaan naar de Nederlandse vertaling van de California Family Risk Assessment (CFRA; Van der Elst, Sondeijker, Vogel, Jansen, & Hermanns, 2012). Na een training van twee dagdelen hebben gezinscoaches en interventieverpleegkundigen 500 maal de CFRA ingevuld voor gezinnen die ze begeleidden. Verwaarlozing en mishandeling na zes maanden werden gebruikt als uitkomstmaten om de voorspellende validiteit van de CFRA te bepalen. De predictieve validiteit bleek laag te zijn: van de gezinnen die op de CFRA als hoog tot zeer hoog risico waren beoordeeld, was na zes maanden 11% gemeld bij een AMK, en waren er voor 7% indicaties voor mishandeling in het KIDOS (elektronisch kinddossier). Voor de laag risicogezinnen is niet gerapporteerd hoeveel AMK-meldingen en indicaties voor mishandeling in het KIDOS er waren gevonden.

Onderzoek naar de LIRIK

De LIRIK is in een gebruikersevaluatie bij bureaus jeugdzorg en AMK's onderzocht op inhoudelijke relevantie, volledigheid en bruikbaarheid (Ten Berge & Van Rossum, 2009). De gebruikerservaringen waren positief. Naar aanleiding van deze evaluatie zijn enkele verbeteringen doorgevoerd om het gebruiksgemak nog te vergroten. Dit resulteerde in de LIRIK-JZ 2009. Naar aanleiding van nieuwe feedback is in 2014 opnieuw een herziene versie van de LIRIK uitgebracht.

De indrukvaliditeit van de LIRIK is onderzocht in verschillende gebruikersevaluaties (Ten Berge & Van Rossum, 2009; Faber, 2012). Hieruit blijkt dat de jeugdzorgwerkers de LIRIK inhoudelijk als relevant en volledig ervaren. Gebruikers geven ook aan dat de LIRIK hen helpt om systematisch tot een oordeel te komen over de huidige veiligheidssituatie en de veiligheidsrisico's en dit te expliciteren en te onderbouwen. De inhoud van de LIRIK lijkt daarmee representatief voor datgene wat we willen meten, namelijk huidige veiligheid van het kind en risico's voor de toekomst.

De kwaliteit van de LIRIK in termen van betrouwbaarheid en validiteit was nog niet onderzocht. Dat betekent dat niet eerder vastgesteld was of de LIRIK in dezelfde situaties daadwerkelijk tot dezelfde inschattingen leidt en of met de LIRIK de veiligheid beoordeeld wordt zoals die in feite is. Daarom hebben Juzt, Spirit en de GGD Gelderland-Zuid samenwerking gezocht met het Nederlands Jeugdinstituut en de Radboud Universiteit voor het huidige onderzoek. Juzt en Spirit werken sinds 2012 met de LIRIK, GGD Gelderland-Zuid voerde sinds 2014 de LIRIK in.

1.2.2 Belang van het onderzoek

Maatschappelijke relevantie

Kindermishandeling vormt een ernstige bedreiging voor de ontwikkeling en het welzijn van jeugdigen. Jeugdigen die opgroeien in een onveilige situatie kunnen daar langdurig last van hebben. Tijdig goede hulp aan jeugdigen en hun ouders is daarom van groot belang. Een betrouwbaar en valide instrument kan bijdragen aan het onderkennen van kindermishandeling en mogelijke risico's daarop. Beslissingen die genomen moeten worden in situaties waarin (mogelijk) kindermishandeling speelt zijn ingrijpend voor gezinnen. Daarom is het zaak dat een instrument in staat is om die jeugdigen aan te wijzen die ook daadwerkelijk te maken hebben met kindermishandeling en daarbij zo min mogelijk gezinnen onterecht van kindermishandeling te beschuldigen, maar ook zo min mogelijk gezinnen onterecht over het hoofd te zien. Dit onderzoek biedt inzicht in de bijdrage die de LIRIK levert aan het onderkennen van (risico's op) kindermishandeling in gezinnen.

Wetenschappelijke relevantie

In breder perspectief draagt dit onderzoek bij aan de kennis over de waarde van risicotaxatie instrumenten. In Nederland zijn instrumenten voor veiligheids- en risicotaxatie nog maar zeer beperkt onderzocht. Internationaal is er wel enig onderzoek naar risicotaxatie instrumenten gedaan. De toegevoegde waarde van een risicotaxatie instrument op het ongestructureerd professioneel oordeel is echter nog nauwelijks onderzocht. Het is van belang dat Nederlandse professionals beschikken over een betrouwbaar en valide instrument dat ook van waarde is boven het ongestructureerde professionele oordeel.

1.3 Doel- en vraagstelling en hypothesen

1.3.1 Doelstelling

Doel van dit onderzoek was vast te stellen of hulpverleners met behulp van de LIRIK komen tot betrouwbare en valide beoordelingen van de veiligheid van jeugdigen nu en in de nabije toekomst waardoor passende interventies ingezet kunnen worden om onveiligheid op te heffen c.q. de risico's te beperken.

1.3.2 Vraagstelling

Hoofdvraag van het onderzoek was of de LIRIK bijdraagt aan een betere beoordeling van de veiligheid van jeugdigen dan wanneer geen risicotaxatie-instrument wordt gebruikt (een ongewapend klinisch oordeel).

Deze hoofdvraag wordt beantwoord aan de hand van de volgende twee onderzoeksvragen:

- Worden met behulp van de LIRIK betrouwbare en valide oordelen gevormd over de veiligheid van kinderen?
- Is het oordeel met behulp van de LIRIK meer betrouwbaar en valide dan het ongestructureerde professionele oordeel van de hulpverlener?

1.3.3 Hypothesen

De verwachting was dat hulpverleners die gebruik maken van de LIRIK meer betrouwbare en valide oordelen geven dan als zij hun ongestructureerde professionele oordeel geven. Het is echter niet plausibel om te verwachten dat met een dergelijk instrument 100% betrouwbaarheid en een grote voorspellende waarde bereikt kan worden. De LIRIK is geen genormeerde en gestandaardiseerde vragenlijst met scherp afgebakende criteria en normscores, maar een checklist voor gestructureerde oordeelsvorming waarbij de weging van de afzonderlijke aandachtspunten afhankelijk blijft van de gebruiker.

In eerder onderzoek naar gestructureerde besluitvorming over vermoedens van kindermishandeling bleek dat de interbeoordelaarsbetrouwbaarheid varieerde tussen 60% en 90 % voor ja/nee-oordelen (Ten Berge, Bartelink, & De Kwaadsteniet, 2011). Ook internationaal onderzoek naar consensus-based en actuariële risicotaxatie-instrumenten laat zien dat de interbeoordelaarsbetrouwbaarheid ten aanzien van inschatting van kans op (herhaling van) kindermishandeling wisselend is en mede afhankelijk van de mate van feitelijkheid van de opgenomen items (Baird, Wagner, Healy, & Johnson, 1999; Knoke & Trocmé, 2005). Geen enkel instrument bereikt 100% betrouwbaarheid als niet objectief waarneembare feiten worden geregistreerd (bv. het kind heeft een blauwe plek), maar afgeleide oordelen (in dat voorbeeld: het kind wordt geslagen). Concrete items als leeftijd van het kind en aantal eerdere meldingen worden consistenters gescoord dan items die een zeker mate van subjectieve evaluatie en/of klinische observatie vereisen, zoals de vraag of sprake is van excessieve discipline of problematische ouder-kind interactie.

Uit internationaal onderzoek naar de voorspellende waarde van risicotaxatie-instrumenten blijkt dat die over het algemeen niet hoog is (Knoke & Trocmé, 2005). Herhaling van kindermishandeling wordt voor ongeveer een derde voorspeld door factoren die in de instrumenten opgenomen zijn. Vals positieven variëren tussen 14% en 29% en vals negatieven tussen 14% en 66%. Actuariële instrumenten, dat wil zeggen instrumenten waarbij gebruik gemaakt wordt van statistische modellen om geïnformeerde beslissingen over onzekere toekomstige situaties te nemen, leiden tot betere risico-inschattingen dan ongestructureerde oordelen van individuele hulpverleners of consensusoordelen (Baird & Wagner, 2000; Knoke & Trocmé, 2005). Echter,

onderzoek naar de Nederlandse versie van de CFRA laat zien dat de voorspellende waarde van dit actuariële instrument alsnog zeer laag is en het daarmee slechter deed dan de Amerikaanse versie (zie paragraaf 1.2). Ook met behulp van instrumenten blijft het moeilijk om veiligheidsrisico's voor kinderen betrouwbaar en valide in te schatten. Wel verwachtten we dat hulpverleners met de LIRIK veiligheidsrisico's voor kinderen beter zouden inschatten dan zonder de LIRIK, dat wil zeggen meer betrouwbaar en valide, omdat zij hierdoor gestuurd worden in het verzamelen van relevante informatie en daardoor mogelijk minder gevoelig zijn voor vertekeningen zoals tunnelvisie.

1.4 Onderzoeksopzet

Praktijkonderzoek in een setting waarin veel medewerkers zijn getraind, het instrument is geïmplementeerd en medewerkers beperkte tijd hebben voor deelname, brengt beperkingen met zich mee. Een strikt experimenteel design (RCT) is niet mogelijk. Derhalve is hier gekozen voor een opzet bestaande uit twee deelonderzoeken gericht op de verschillende onderzoeksvragen (zie paragraaf 1.3):

1. Vignetonderzoek naar de uniformiteit in veiligheids- en risicotaxatie, i.e. interbeoordelaarsovereenstemming.
2. Dossieronderzoek naar de predictieve validiteit van de LIRIK.

In deze paragraaf wordt de opzet van de deelonderzoeken kort toegelicht. De deelonderzoeken worden in de volgende hoofdstukken gedetailleerd beschreven.

Deelonderzoek 1: Vignetonderzoek naar de interbeoordelaarsovereenstemming

De vraag die in deelonderzoek 1 centraal stond, was of hulpverleners met de LIRIK tot betrouwbare oordelen over de huidige veiligheid en risico's komen. Specifiek is hier gekeken of professionals het meer met elkaar eens zijn over de huidige veiligheid en risico's voor jeugdigen wanneer zij de LIRIK gebruiken bij hun oordeelsvorming dan wanneer zij dat niet doen.

Daarvoor zijn twaalf vignetten, gebaseerd op echte casus, opgesteld. Deelnemers aan het onderzoek kregen op twee momenten drie vignetten voorgelegd (in totaal zes vignetten) en beoordeelden daarvoor de huidige veiligheid en de risico's voor de nabije toekomst in een digitale vragenlijst. Een deel van de deelnemers deed dit met de LIRIK en een deel zonder de LIRIK. Deze vragenlijst bevatte:

- Items m.b.t. achtergrondinformatie van respondenten (leeftijd, opleidingsniveau, werksetting, werkervaring van de respondent).
- Items over kennis van en ervaring met de LIRIK en ervaren nut.
- Voor de groep die de LIRIK gebruikte: de gedigitaliseerde versie van de LIRIK.
- Voor de controlegroep: vragen naar veiligheid en risico's.

De mate van overeenstemming tussen professionals is berekend (Krippendorff's alpha).

In de oorspronkelijke opzet was het de bedoeling niet alleen naar de interbeoordelaarsovereenstemming te kijken, maar ook naar de test-hertestbetrouwbaarheid. Respondenten zouden op twee meetmomenten dezelfde vignetten beoordelen om daarmee te onderzoeken of oordelen over de tijd meer consistent zijn wanneer de LIRIK wordt gebruikt dan wanneer de LIRIK niet wordt gebruikt. Omdat het, ondanks extra inspanningen, niet lukte voldoende deelnemers te werven voor dit deelonderzoek, is besloten de opzet voor de test-hertest betrouwbaarheid te laten vervallen en deelnemers op twee meetmomenten niet dezelfde maar

verschillende vignetten voor te leggen. Dat resulteerde in genoeg gegevens om in ieder geval een uitspraak te kunnen doen over de interbeoordelaarsovereenstemming.

Deelonderzoek 2: Dossieronderzoek naar de predictieve validiteit

De vraag die in deelonderzoek 2 centraal stond, was of we met de LIRIK meten wat we beogen te meten, dat wil zeggen: voorspelt de LIRIK correct welke gezinnen een hoog risico en welke een laag risico op kindermishandeling hebben en hebben inschattingen gemaakt met de LIRIK een betere voorspellende waarde dan zonder instrument?

Deze studie onderzoekt de validiteit van de LIRIK in de praktijk aan de hand van dossiers van Spirit, Juzt en GGD Gelderland-Zuid. Het onderzoek kende twee meetmomenten:

- Op T1 vond de beoordeling van veiligheid en veiligheidsrisico's bij 370 gezinnen door individuele hulpverleners plaats (of in tweetallen, als het een beginner met werkbegeleider betrof). De beoordeling vond plaats in 2 condities: met en zonder de LIRIK.
- Op T2 (na 6 maanden) is teruggekeken in het dossier van de jeugdigen die op T1 zijn beoordeeld om de vraag te beantwoorden of de initiële beoordelingen van veiligheid en risico's samen hingen met uitkomstmaten van onveiligheid. Deze dossieranalyse is uitgevoerd door een onafhankelijk onderzoeker.

De dossieranalyse op T2 vond plaats aan de hand van een checklist. Met deze checklist zijn achtergrondkenmerken nagegaan over het gezin (aantal kinderen, leeftijd, geslacht), risico- en veiligheidsinschattingen al of niet met de LIRIK gemaakt en informatie over de actuele veiligheidssituatie (dossier is afgerond zonder interventies t.a.v. veiligheid, gezin is gemeld bij het AMK, bij de Raad voor de Kinderbescherming, kind is uit huis geplaatst) opgenomen. Geanalyseerd is in hoeverre de inschattingen van risico en veiligheid op T1 de situatie op T2 konden voorspellen.

1.5 Leeswijzer

De volgende hoofdstukken beschrijven de specifieke vraagstelling, opzet en resultaten van de twee deelonderzoeken: het onderzoek naar de betrouwbaarheid in hoofdstuk 2 en het onderzoek naar de validiteit in hoofdstuk 3. In hoofdstuk 4 wordt op basis van alle onderzoeksresultaten het antwoord op de hoofdvraag besproken. Ook doen we daar aanbevelingen voor verbetering en voor vervolgonderzoek.

Ter bevordering van de leesbaarheid worden in deze publicatie de termen ouder/opvoeder en kind/jeugdige door elkaar gebruikt. Met ouder/opvoeder wordt de primaire opvoeder van het kind bedoeld, dat wil zeggen diegene die de opvoeding en verzorging van het kind grotendeels op zich neemt. Kind/jeugdige verwijst naar kinderen van alle leeftijden. Waar in de tekst in een algemene context gesproken wordt over 'hij' of de mannelijke vorm, kan ook 'zij' of de vrouwelijke vorm worden gelezen.

De termen betrouwbaarheid en overeenstemming worden afwisselend gebruikt. Waar de term betrouwbaarheid wordt gebruikt bedoelen we de mate van overeenstemming in oordelen tussen hulpverleners.

2 Deelonderzoek 1. Betrouwbaarheid

Leontien de Kwaadsteniet, Floor Minkhorst & Cora Bartelink

2.1 Onderzoeksvragen

We hebben de interbeoordelaarsovereenstemming onderzocht van veiligheidsoordelen en risico-inschattingen gemaakt met de LIRIK en deze vergeleken met veiligheidsoordelen en risico-inschattingen waarbij geen instrument werd gebruikt (ongestructureerd), om te zien of de LIRIK in de praktijk tot (meer) betrouwbare oordelen kan leiden. Daarnaast hebben we gekeken naar de overeenstemming over de aan de algemene conclusies voorafgaande specifieke items. Ook hebben we gevraagd hoe professionals denken over de bruikbaarheid van de LIRIK om veiligheid en risico's in te schatten.

Samengevat waren dit onze onderzoeksvragen in deelonderzoek 1:

1. Hoe is de interbeoordelaarsovereenstemming voor conclusies over kindveiligheid en risico's wanneer hulpverleners in de jeugdzorg de LIRIK gebruiken?
2. Is de interbeoordelaarsovereenstemming voor conclusies over kindveiligheid en risico's hoger wanneer hulpverleners in de jeugdzorg de LIRIK gebruiken dan wanneer zij op ongestructureerde wijze tot deze conclusies komen (geen risicotaxatie-instrument gebruiken)?
3. Hoe is de interbeoordelaarsovereenstemming voor de beoordeling van de aanwezigheid van specifieke signalen in de ouder-kind-interactie, kindsignalen, risicofactoren en beschermende factoren met de LIRIK?
4. Vinden hulpverleners in de jeugdzorg de LIRIK geschikt om hen te ondersteunen bij hun professionele oordeelsvorming over de huidige veiligheid van kinderen en risico's in de nabije toekomst?

2.2 Methode

2.2.1 Deelnemers

Het doel was om zowel bij Juzt als Spirit 50 behandelcoördinatoren en 50 ambulante hulpverleners te werven, om deel te nemen in werktijd. Daarnaast zijn nog 8 deelnemers geworven bij de Koraalgroep. Medewerkers werden door hun leidinggevende gevraagd om deel te nemen en ontvingen een brief met informatie. In totaal ontvingen we emailadressen van 106 medewerkers, 79 medewerkers namen ook werkelijk deel.

De aangeschreven deelnemers werden ingedeeld in twee condities en gevraagd om op twee momenten hun veiligheidsoordelen te geven voor vignetten. In één conditie gebruikten deelnemers hierbij de LIRIK, in de andere niet. Bij Juzt gebruikten alle medewerkers de LIRIK al, zij werden willekeurig verdeeld over de twee condities. Bij Spirit gebruikten nog niet alle werknemers de LIRIK al. Deelnemers van wie bekend was dat zij de LIRIK nog niet hadden gebruikt werden in de niet-LIRIK-conditie ingedeeld, de overige deelnemers werden willekeurig ingedeeld. De deelnemers van de Koraalgroep werkten allen nog niet met de LIRIK, zij werden in de niet-LIRIK-conditie ingedeeld. Aangeschreven deelnemers deden uiteindelijk (na herhaalde herinneringen) niet allemaal mee, of vulden de vragenlijsten niet op beide momenten (geheel) in (zie procedure).

In Tabel 1 staan kenmerken van de deelnemers in beide groepen (geslacht, leeftijd, opleiding, werkervaring, ervaring met de LIRIK). Verschillen tussen de groepen bleken niet statistisch significant (zie Tabel 1). Niet alle deelnemers hebben de vragen over persoonlijke gegevens ingevuld, omdat zij de vragenlijsten niet op beide momenten hebben ingevuld, of niet compleet.

De meeste deelnemers hadden een training (van een hele of halve dag) gevolgd voor het gebruik van de LIRIK: 59.3% (16 van de 27) in de LIRIK-conditie, 70.8% (17 van de 24) in de niet-LIRIK conditie (dit verschil is niet statistisch significant, $z=-0.86$, $p=.39$). Deelnemers hadden de LIRIK ook leren kennen door instructies van collega's, de handleiding te lezen, of de LIRIK gewoon te gebruiken. In de LIRIK-conditie gaf 55.5% van de deelnemers aan (15 van de 27) en in de niet-LIRIK-conditie 50% (12 van de 24) dat zij de LIRIK altijd of bijna altijd gebruikten voor een nieuwe cliënt; 18.5% (LIRIK) en 29.2% (niet-LIRIK) soms; 25.9% (LIRIK) en 20.8% (niet-LIRIK) nooit (geen statistisch significant verschil, $U=322.5$; $z=-0.03$; $p=.96$). Voor cliënten die niet nieuw zijn gebruikte in de LIRIK-conditie 11.1% de LIRIK altijd of bijna altijd (3 van de 27), versus 20.8% (5 van de 24) in de niet-LIRIK groep; 51.9% (LIRIK) en 58.3% (niet-LIRIK) soms; 37.0% (LIRIK) en 20.8% (niet-LIRIK) nooit ($U=256.0$; $z=-1.43$; $p=.17$). In de LIRIK-conditie gaf 14.8% (4 van de 27) en in de niet-LIRIK-conditie 20.8% (5 van de 24) aan dat zij de LIRIK altijd of bijna altijd met een collega invulden; 33.3% in zowel de LIRIK als niet-LIRIK-conditie soms; 51.9% en 45.8% nooit ($U=303.5$; $z=-0.42$; $p=.68$).

We hadden verwacht dat in de LIRIK-conditie de deelnemers vaker een training zouden hebben gehad, vaker en sinds langer de LIRIK gebruikten, maar dit bleek juist niet het geval te zijn. We hebben aanvullende analyses gedaan voor de deelnemers die een training hadden gevolgd (zie resultaten).

Tabel 1. Beschrijvende kenmerken van de deelnemers in de LIRIK-conditie en de niet-LIRIK-conditie

	LIRIK (<i>n</i> =36)		Geen LIRIK (<i>n</i> =43)		Toets voor verschillen
Vrouw	85.3%	(29/34)	84.2%	(32/38)	$z=0.13$, $p=.90$
Leeftijd	$M=40.7$	$SD=9.7$	$M=38.8$	$SD=10.0$	$t(70)=0.77$, $p=.44$
Opleiding					
MBO	3.1%	1	5.4%	2	
HBO	46.9%	15	48.6%	18	
Universitair	50.0%	16	45.9%	17	$\chi^2=0.28$ $df=2$, $p=.87$
Ervaring					
In jeugdzorg	$M=14.2$	$SD=8.7$	$M=12.2$	$SD=7.6$	$t(70)=1.04$, $p=.30$
In huidige functie	$M=6.9$	$SD=3.6$	$M=7.6$	$SD=4.4$	$t(70)=-0.79$, $p=.43$
Gebruik LIRIK sinds					
≥ 1 jaar	48.1%	13	45.8%	11	
6-12 maanden	22.2%	6	29.2%	7	
1-6 maanden	7.4%	2	12.5%	3	
≤ 1 maand	22.2%	6	12.5%	3	$U=315.5$, $z=-0.17$, $p=.86$

Noot. Frequenties tellen niet op tot het totaal aantal deelnemers, doordat deelnemers die alleen de tweede vragenlijst invulden geen persoonlijke gegevens hadden ingevuld en deelnemers die alleen de eerste vragenlijst invulden geen gegevens over hun ervaring met de LIRIK hadden ingevuld; daarnaast hadden sommige deelnemers deze vragen niet beantwoord.

2.2.2 Procedure

Deelnemers ontvingen op twee momenten een link naar een online vragenlijst, T1 en T2. Op beide momenten bevatten de vragenlijsten drie verschillende casusbeschrijvingen met bijbehorende vragen over kindveiligheid en risico's. In de vragenlijsten van T1 werden daaraan voorafgaand vragen gesteld over leeftijd, geslacht, opleiding, werkervaring en werksetting van deelnemers. In de vragenlijsten van T2 werden na de casusbeschrijvingen vragen gesteld over de kennis van en het gebruik van de LIRIK door deelnemers, naar hun mening over de LIRIK en over wat zij belangrijk vinden voor een goed risico-taxatie-instrument. Tussen het toesturen van T1 en T2 zat meestal drie maanden, behalve voor de laatste deelnemers, voor wie er tenminste twee weken tijd tussen zat. We verstuurdten voor ieder meetmoment twee keer een herinnering. Op meetmoment T1 verzonden we 49 links naar LIRIK-vragenlijsten (naar deelnemers in de LIRIK-conditie). Daarvan werden er 31 volledig ingevuld, 3 deelnemers hadden vragen voor één casusbeschrijving beantwoord. We verzonden 57 links naar niet-LIRIK-vragenlijsten (naar deelnemers in de niet-LIRIK-conditie). Daarvan werden er 35 volledig ingevuld, 3 deelnemers hadden de vragen voor één (2) of twee (1) gevalsbeschrijvingen ingevuld. Op meetmoment T2 verzonden we 48 links naar LIRIK-vragenlijsten en 53 links naar niet-LIRIK-vragenlijsten. Vijf deelnemers hadden aangegeven dat zij niet (meer) konden deelnemen, omdat hun contract was beëindigd, vanwege zwangerschapsverlof, ziekte, of tijdsgebrek. In de LIRIK-conditie vulden 27 deelnemers de vragenlijsten volledig in, 5 niet volledig. In de niet-LIRIK-conditie vulden 26 deelnemers de vragenlijsten volledig in, 1 deelnemer twee gevalsbeschrijvingen. Er waren 7 deelnemers die alleen aan T2 deelnamen en voor wie persoonlijke gegevens daardoor ontbreken (2 in de LIRIK-groep, 5 in de niet-LIRIK-groep). Op T1 waren er twee deelnemers die vanwege technische problemen de lijst niet konden invullen. Eén deelnemer liet weten de lijst niet te kunnen invullen omdat hij daarvoor te weinig informatie had. Op T2 ontvingen we geen opmerkingen over problemen. Op T2 namen minder mensen deel dan op T1. Een mogelijke verklaring hiervoor is een hoge werkdruk, plus het feit dat dit onderzoek plaatsvond tijdens de stelselwijziging van de jeugdzorg (in 2014), waardoor de instellingen te maken hadden met grote reorganisaties en bezuinigingen, en werknemers met veranderingen en onzekerheid.

In een pilot-onderzoek bleek dat het invullen van de LIRIK-vragenlijsten ongeveer een uur kostte (voor drie gevalsbeschrijvingen, dus per meetmoment), hoewel dit verschilde voor verschillende deelnemers. De niet-LIRIK-vragenlijsten werden ingevuld in minder dan een uur. Deelnemers kregen de instructie dat zij de lijsten konden invullen op een zelfgekozen plaats en moment en dat zij konden stoppen wanneer zij langer dan een uur bezig waren.

2.2.3 Materialen

De vragenlijsten werden gemaakt met het online survey programma Perseus Survey Solutions (Perseus Development Corporation, 2003). In de instructie vooraf werd uitgelegd dat de vragen alleen moesten worden ingevuld voor het betreffende kind, niet voor eventuele andere kinderen in het gezin, dat deelnemers tussendoor konden stoppen om op een later moment verder te gaan en er werd verzocht de gevalsbeschrijvingen niet met collega's te bespreken om onafhankelijkheid van de oordelen te waarborgen. Na het lezen van een gevalsbeschrijving beantwoordden deelnemers vragen, waarbij zij de gevalsbeschrijving in een apart scherm konden teruglezen. De vragen naar algemene conclusies moesten worden ingevuld voordat deelnemers naar de volgende pagina konden, de overige vragen waren niet verplicht. Bij de meeste

gesloten vragen konden deelnemers meerdere opties tegelijk aanvinken (zoals in de papieren versie van de LIRIK). Deelnemers konden terug naar eerdere pagina's tot de vragenlijst was verzonden. Gegevens van niet volledig ingevulde vragenlijsten werden bewaard op de server, zodat geen gegevens verloren gingen.

Vragenlijsten in de LIRIK-conditie.

Na iedere gevalsbeschrijving vulden deelnemers een LIRIK in. Kort voor de start van het onderzoek was de LIRIK herzien naar aanleiding van gebruikerservaringen. Verschillen tussen de eerdere en huidige versie bestonden uit tekstuele veranderingen en een volgordeverandering. De vragen van de LIRIK werden gekopieerd in de digitale vragenlijst (de nieuwe versie moest nog gedigitaliseerd worden).

De LIRIK bestaat uit twee delen. Het eerste deel leidt naar een conclusie over de huidige veiligheid van een kind, het tweede deel naar de inschatting van risico's nu en risico's bij voorziene veranderingen in de nabije toekomst. In het eerste deel, subonderdeel A, vraagt de LIRIK naar een eerste inschatting van de directe veiligheid. Subonderdeel B vraagt naar signalen in de omgang ouder(s)-jeugdige, subonderdeel C naar signalen bij de jeugdige, subonderdeel D vraagt naar risicofactoren en beschermende factoren bij de ouder(s), de jeugdige, gezin en omgeving. Aan het eind van ieder subonderdeel wordt naar een samenvattende conclusie gevraagd over of er geen, weinig of veel signalen of factoren aanwezig zijn, of dat dit onbekend is. Deel D eindigt met de hoofdconclusie over de huidige veiligheidssituatie. Deel 2 vraagt naar aanvullende risicofactoren bij mogelijke of aangetoonde kindermishandeling (A), mogelijke gevolgen (B) en beschermende factoren (C) en eindigt met de hoofdconclusies over risico's op dit moment en bij voorziene veranderingen in de nabije toekomst.

In de online vragenlijsten begonnen de verschillende onderdelen steeds op een nieuwe pagina, alleen onderdeel 1D (risicofactoren en beschermende factoren) bestond uit meerdere pagina's om de lijst overzichtelijk te houden. Na de risico-inschattingen werd op een nieuwe pagina gevraagd hoe ernstig deelnemers de gevalsbeschrijving vonden in vergelijking met gevallen in hun dagelijkse praktijk.

Er waren twaalf gevalsbeschrijvingen die werden verdeeld over zes versies van de vragenlijst met drie gevalsbeschrijvingen voor T1 en zes versies met drie gevalsbeschrijvingen voor T2, zodanig dat iedere deelnemer in totaal zes verschillende beschrijvingen beoordeelde en de verschillende beschrijvingen ongeveer even vaak beoordeeld zouden worden, in verschillende volgordes. Uiteindelijk bleek in de LIRIK-conditie iedere gevalsbeschrijving door 13 tot 17 verschillende deelnemers te zijn beoordeeld.

Vragenlijsten in de niet-LIRIK-conditie.

In de niet-LIRIK-conditie vroegen we deelnemers na het lezen van een gevalsbeschrijving om een conclusie over de huidige veiligheidssituatie en over risico's op dit moment en bij voorziene veranderingen in de nabije toekomst. Deze vragen waren precies hetzelfde geformuleerd als de hoofdconclusies in de LIRIK. Hierna werd ook, op een nieuwe pagina, gevraagd hoe ernstig deelnemers de gevalsbeschrijving vonden in vergelijking met gevallen in hun dagelijkse praktijk. Deelnemers ontvingen één van zes versies op T1 en T2, de combinaties van gevalsbeschrijvingen in de versies waren hetzelfde voor de LIRIK en de niet-LIRIK-conditie. In de niet-LIRIK-conditie bleek uiteindelijk iedere gevalsbeschrijving door 11 tot 19 verschillende deelnemers te zijn beoordeeld.

2.2.4 Gevalsbeschrijvingen.

De gevalsbeschrijvingen waren samenvattingen van 521 tot 732 woorden ($M=607.5$) van bestaande verslagen van cliënten van Spirit en Altra. De beschrijvingen werden gekozen uit een grotere set van 20, die eerder in scriptie-onderzoek naar veiligheidsoordelen waren gebruikt (Veenhuizen, 2013). De beschrijvingen waren aangepast zodat ze niet tot de cliënten te herleiden zouden zijn. De gevalsbeschrijvingen werden zo geselecteerd dat er uiteindelijk evenveel over jongens als meisjes zouden gaan, van verschillende leeftijden (5 tot 16), met verschillende familiesituaties (een/twee-oudergezinnen, (geen) broertjes/zusjes), met verschillende sociale en culturele achtergronden (hoge/lage SES, autochtoon/allochtoon) en met verschillende aard en ernst van problemen (wijzend op lichamelijke, emotionele, seksuele en pedagogische mishandeling/verwaarlozing).

Scores die de deelnemers gaven voor de ernst van de problemen in de gevalsbeschrijvingen in vergelijking met gevallen in hun dagelijkse praktijk wezen erop dat deelnemers de problemen in de gevalsbeschrijvingen gemiddeld niet ernstiger inschatten. Op een schaal van 1 (veel minder ernstig) tot 5 (veel ernstiger) varieerden de gemiddelde scores (gemiddeld over deelnemers, per casus) van 2.13 tot 3.50 ($M=2.95$, $SD=.46$) in de LIRIK-conditie en van 1.74 tot 3.47 ($M=2.83$, $SD=.66$) in de niet-LIRIK-conditie (het verschil tussen gemiddelden over beschrijvingen per conditie was niet statistisch significant $t(22)=.59$, $p=.56$). Wanneer je echter de frequenties van de gekozen opties voor de verschillende hoofdconclusies (huidige veiligheidssituatie en risico's op dit moment en bij voorziene veranderingen in de nabije toekomst) in dit onderzoek naar de betrouwbaarheid van de LIRIK vergelijkt met de frequenties van deze oordelen in het validiteitsonderzoek (deelonderzoek 2, zie Tabel 3 van hoofdstuk 3), zie je dat de conclusies in dit onderzoek wijzen op wat ernstigere problematiek in de gevalsbeschrijvingen.

2.2.5 Analyses

We hebben eerst de interbeoordelaarsovereenstemming berekend voor de hoofdconclusies over de huidige veiligheid van het kind en risico's op dit moment en bij voorziene veranderingen in de nabije toekomst, voor de LIRIK en de niet-LIRIK-groep. We hebben hierbij aanvullende analyses gedaan voor deelnemers die een training voor het gebruik van de LIRIK hadden gevolgd. Vervolgens hebben we gekeken naar de interbeoordelaarsovereenstemming over de aanwezigheid van signalen in de omgang tussen ouder(s) en kind, kindsignalen, risicofactoren en beschermende factoren in de LIRIK-conditie (in de niet-LIRIK-conditie werden hierover geen vragen gesteld).

We hebben Krippendorff's alfa (α) gebruikt als maat voor de interbeoordelaarsovereenstemming (Hayes & Krippendorff, 2007). Dit is een maat gebaseerd op gevonden niet-overeenstemming en verwachte niet-overeenstemming ($\alpha = 1 - (\text{geobserveerde niet-overeenstemming} / \text{verwachte niet-overeenstemming})$, Krippendorff, 2004). Overeenstemming wordt hiermee uitgedrukt op een schaal van 0 (overeenstemming niet meer dan verwachte toevalsovereenstemming, ofwel statistische onafhankelijkheid van casus en categorieën) tot 1 (perfecte overeenstemming), waarbij de beoordelaars als inwisselbaar (willekeurige selectie van beoordelaars) worden gezien. Deze maat kan gebruikt worden voor oordelen gemaakt op nominaal, ordinaal, interval- of ratio-meetniveau, voor verschillende aantallen beoordelaars en cases, ook als niet iedere beoordelaar iedere casus heeft beoordeeld. De maat geeft absolute overeenstemming weer (in welke mate verschillende beoordelaars tot dezelfde oordelen komen), niet slechts associatie of consistentie (of beoordelaars het eens zijn over welke gevallen hogere en lagere beoordelingen krijgen, zonder dat oordelen

hetzelfde hoeven te zijn)³. In vergelijking met kappamaten voor oordelen op nominaal en ordinaal meetniveau leidt Krippendorff's alfa tot meer algemene uitspraken over overeenstemming, omdat kappamaten uitgaan van hoe vaak beoordelaars in een onderzoek categorieën gebruiken, terwijl Krippendorff's alfa populatieprevalenties van categorieën schat. We hebben ook bootstrap-betrouwbaarheidsintervallen berekend voor de alfa's.

Om een betekenisvolle uitkomst te krijgen voor een overeenstemmingsmaat zijn voldoende oordelen nodig. Dit aantal hangt af van het aantal beoordelaars, het aantal beoordeelde cases en het aantal categorieën. Hoewel we streefden naar 50 beoordelaars die ieder 6 cases zouden beoordelen zodat we per groep (LIRIK en niet-LIRIK) 300 oordelen zouden hebben, is het aantal van in totaal 181 oordelen in de LIRIK-groep en 183 in de niet-LIRIK-groep nog voldoende (Krippendorff, 2011)⁴. Hiernaast moeten de verschillende categorieën ook vaak genoeg gebruikt worden. Als beoordelaars bepaalde categorieën niet of bijna niet gebruiken, is de maat niet even informatief voor alle categorieën (zie resultaten voor de controle hiervan). Na de analyses met betrekking tot de overeenstemming beschrijven we de resultaten met betrekking tot de mening van deelnemers over de LIRIK en over wat belangrijk is voor een goed risicotaxatie-instrument.

2.3 Resultaten

2.3.1 *Interbeoordelaarovereenstemming over conclusies over veiligheid en risico's gemaakt met en zonder de LIRIK*

De LIRIK leidt tot drie hoofdconclusies over veiligheid en risico's. Deelnemers beoordeelden de huidige veiligheid van een kind met een keuze uit de volgende opties: het kind lijkt op dit moment thuis veilig, er is mogelijk sprake van kindermishandeling, er is sprake van aangetoonde kindermishandeling, er is sprake van een levensbedreigende situatie of direct fysiek gevaar, of er is onvoldoende informatie om hierover te oordelen. Vervolgens beoordeelden deelnemers de huidige risico's voor het kind, en de risico's in de nabije toekomst bij voorziene veranderingen, op een schaal van 1 (laag) tot 5 (zeer hoog). Tabel 2 laat de frequenties zien van de gekozen opties voor iedere hoofdconclusie en voor iedere casus en het totaal aantal deelnemers dat een casus beoordeelde. Bij de eerste conclusie konden deelnemers meerdere opties kiezen, waardoor het aantal gekozen opties niet overeenkomt met het totaal aantal deelnemers.

Bij het analyseren van de overeenstemming over de conclusie over de huidige veiligheid van het kind waren er twee problemen. Het eerste probleem was dat deelnemers hier meerdere opties konden kiezen, net als bij het gebruik van de LIRIK in de praktijk, hoewel dit niet expliciet was benoemd. Ten tweede vormen de eerste vier opties samen een ordinale meetschaal voor geen tot ernstige bedreiging van veiligheid, terwijl de vijfde optie (onvoldoende informatie) van deze ordening afwijkt. We transformeerden daarom de antwoorden van deelnemers naar een variabele met vier antwoordmogelijkheden of een ontbrekende score. Om te controleren

³ We gebruiken in deze tekst soms betrouwbaarheid in plaats van (interbeoordelaars-)overeenstemming. In dit onderzoek hebben we de overeenstemming in de strikte betekenis van absolute overeenstemming onderzocht, niet consistentie.

⁴ Deze aantallen zijn groot genoeg om met voldoende betrouwbaarheid ($p \leq .05$) te besluiten of Krippendorff's alfa tenminste .80 is (zie Krippendorff (2011): Tabel 1 op pagina 105). Deze ondergrens van .80 lijkt onrealistisch om te verwachten gezien eerdere bevindingen van matige overeenstemming in dit domein. Om betrouwbaar te besluiten of alfa van tenminste een nog lagere waarde is, zijn minder oordelen nodig. Hoe hoog overeenstemming tenminste zou moeten zijn hangt af van het belang van beslissingen die op basis van de oordelen wordt genomen. Men zou .80 nog te laag kunnen vinden, gezien de aard van de problemen en de mogelijk grote gevolgen die oordelen kunnen hebben (zie resultaten).

voor mogelijke vertekeningen van de resultaten hierdoor, voerden we de transformatie op drie verschillende manieren uit.

De vijfde optie, onvoldoende informatie, geeft niet aan hoe een deelnemer de veiligheid van een kind beoordeelt. Bij de analyse van overeenstemming kunnen categorieën die niet informatief zijn het best worden vervangen door informatieve categorieën, of buiten de analyse worden gehouden (Krippendorff, 2011). We keken eerst of deelnemers meer dan een optie hadden gekozen. Wanneer dit het geval was en een van twee gekozen opties was 5 (onvoldoende informatie), namen we alleen de andere optie mee in de analyses. In het geval dat twee of meer van de categorieën 1 tot en met 4 waren gekozen, namen we ofwel de hoogste score (getransformeerde variabele “veiligheid-hoog”), de laagste score (getransformeerde variabele “veiligheid-laag”), of we maakten er een ontbrekende score van (getransformeerde variabele “veiligheid-missing”). Bij alle drie de transformaties lieten we dus keuzen van de vijfde categorie buiten beschouwing, ofwel door ze te vervangen door een gelijktijdig gekozen andere optie, of door een ontbrekende score. We berekenden vervolgens drie interbeoordelaarsovereenstemmingscoëfficiënten voor de conclusie over de huidige veiligheid van het kind.

Tabel 2. Frequenties van gekozen opties voor de hoofdconclusies in de condities met en zonder LIRIK en het totaal aantal deelnemers dat een casus beoordeelde

LIRIK		Huidige veiligheid ^a					Risico-inschatting nu (1 (laag) - 5 (hoog))					Risico-inschatting in de nabije toekomst bij voorziene veranderingen (1 (laag) - 5 (hoog))					
Case	Conditie	1	2	3	4	5	1	2	3	4	5	1	2	3	4	5	Totaal
1	Met	6	7	4	1	2	0	2	4	6	3	0	8	3	3	1	15
	Zonder	3	3	6	1	7	1	2	5	9	1	1	4	4	6	3	18
2	Met	2	8	4	1	3	0	1	2	8	4	0	3	8	3	1	15
	Zonder	0	7	6	3	1	0	1	3	7	4	0	2	5	4	4	15
3	Met	1	12	3	0	3	0	2	4	9	2	0	3	9	3	2	17
	Zonder	3	5	6	1	0	0	3	5	6	1	0	4	6	4	1	15
4	Met	2	7	3	0	5	0	3	6	5	2	3	0	7	6	0	16
	Zonder	2	9	2	1	3	0	1	9	7	0	1	3	10	3	0	17
5	Met	14	2	0	0	2	6	7	1	1	0	7	6	2	0	0	15
	Zonder	15	3	1	0	2	6	11	2	0	0	8	8	3	0	0	19
6	Met	0	5	7	2	0	0	0	2	5	7	0	1	7	2	4	14
	Zonder	0	4	13	4	0	0	0	0	11	6	0	0	7	3	7	17
7	Met	10	3	0	0	2	3	4	7	1	0	1	8	4	1	1	15
	Zonder	8	3	2	0	6	2	7	7	1	0	5	2	8	2	0	17
8	Met	0	6	10	2	1	0	0	2	9	6	1	4	5	6	1	17

	Zonder	1	3	7	2	0	0	0	4	7	1	0	2	6	3	1	12
9	Met	0	9	6	0	1	0	0	2	11	2	0	4	9	1	1	15
	Zonder	0	6	6	0	1	0	0	3	6	4	0	2	5	5	1	13
10	Met	9	8	2	0	0	1	3	5	6	0	6	5	2	2	0	15
	Zonder	4	5	1	0	1	0	3	4	4	0	3	3	4	1	0	11
11	Met	2	9	2	0	1	0	4	8	2	0	2	5	3	4	0	14
	Zonder	5	9	1	0	0	0	3	4	8	0	1	5	4	5	0	15
12	Met	9	3	0	1	5	2	1	5	4	1	2	5	4	2	0	13
	Zonder	8	0	0	1	6	2	7	3	2	0	2	4	5	2	1	14

Noot. ^a Bij de conclusie over de huidige veiligheid kozen deelnemers één of meer van de volgende opties: 1 het kind lijkt op dit moment veilig thuis, 2 (er is mogelijke sprake van kindermishandeling), 3 (er is sprake van feitelijk aangetoonde kindermishandeling), 4 (er is sprake van een levensbedreigende situatie of direct fysiek gevaar) en 5 (onvoldoende informatie).

In Tabel 3 staan Krippendorff alfa's (α) met bootstrap-betrouwbaarheidsintervallen voor de veiligheidsconclusies en de inschattingen van huidige risico's en risico's in de nabije toekomst bij voorziene veranderingen (bij de risicotaxaties konden deelnemers maar een optie kiezen). Om te kunnen spreken van een goede overeenstemming zouden Krippendorff's alfa's tenminste .80 moeten zijn (Krippendorff, 2004). De Commissie Testaangelegenheden Nederland (COTAN) geeft ook aan dat een overeenstemmingcoëfficiënt tenminste .80 zou moeten zijn om te kunnen spreken van goede overeenstemming, tenminste .60 om te kunnen spreken van voldoende overeenstemming (Evers, Lucassen, Meijer, & Sijtsma, 2010) (zie bijv. ook Cichetti, 2001). Wat echter als goed of voldoende kan worden gezien hangt af van waarvoor de oordelen gebruikt worden. Gezien de mogelijk verstrekende gevolgen die conclusies over veiligheid en risico's kunnen hebben voor kinderen en families, zowel wanneer aanwezige veiligheidsrisico's niet worden erkend, als wanneer een veilige situatie als onveilig wordt gezien, zou overeenstemming zo hoog mogelijk moeten zijn, liefst hoger dan .80. In onderzoek naar overeenstemming voor zowel op consensus gebaseerde als actuariële risicotaxatie-instrumenten is echter nooit zo'n hoge overeenstemming gevonden. In dit onderzoek bleek de overeenstemming over veiligheid en risico's in beide condities (met en zonder LIRIK) steeds onder .50 en kan dus als laag tot matig worden beschouwd. Bootstrap-betrouwbaarheidsintervallen bevatten de waarde .80 niet (en .60 ook niet), wat aantoont dat met voldoende betrouwbaarheid kan worden geconcludeerd dat overeenstemming niet deze gewenste criteria bereikt. De beoordeling van risico's in de nabije toekomst bij voorziene veranderingen bleek het laagst.

Krippendorff's alfa geeft alleen voldoende informatie over overeenstemming voor de verschillende antwoordmogelijkheden afzonderlijk, wanneer deze mogelijkheden ieder ook vaak genoeg gekozen worden. De opties die het minst vaak werden gekozen wezen niet eenduidig op veilige of juist onveilige situaties, dit waren de opties "direct fysiek gevaar" bij oordelen over huidige veiligheid, laag risico (1) voor oordelen over huidig risico en hoog risico (5) voor oordelen over risico in de nabije toekomst bij voorziene veranderingen. Deze opties werden niet voldoende vaak gekozen (7 tot 11 keer) om met voldoende betrouwbaarheid ($p=.05$) te zien of overeenstemming (Krippendorff's alfa) tenminste .5 (of .6 of .8) is, voor deze categorieën afzonderlijk. De antwoordcategorieën laag risico (1) voor huidig risico en hoog risico (5) bij voorziene

veranderingen werden wel vaak genoeg gekozen om met $p=.10$ te zien of overeenstemming (Krippendorff's alfa) tenminste .5 is (Krippendorff, 2011).

De lage gevonden overeenstemming is ook te zien in Tabel 2. Oordelen over de huidige veiligheid gemaakt met de LIRIK liepen uiteen over twee categorieën bij vier gevalsbeschrijvingen, over drie categorieën in zes gevallen, en bij twee gevalsbeschrijvingen zelfs over vier categorieën (categorie 5, onvoldoende informatie, is buiten beschouwing gelaten). Dit betekent dat verschillende deelnemers de veiligheidssituatie van hetzelfde kind als veilig tot (zeer) onveilig beoordeelden. De oordelen over risico's op dit moment, gemaakt met de LIRIK, liepen uiteen over drie categorieën (bij vier gevalsbeschrijvingen), vier categorieën (zeven gevalbeschrijvingen) tot zelfs vijf categorieën (een gevalsbeschrijving). De oordelen over risico's in de nabije toekomst bij voorziene veranderingen liepen uiteen over drie categorieën (bij een gevalsbeschrijving), vier categorieën (negen gevalsbeschrijvingen) en vijf categorieën (twee gevalsbeschrijvingen) (zie Tabel 2). Verschillende deelnemers schatten dus risico's als laag tot (zeer) hoog in voor hetzelfde kind.

Tabel 3. Krippendorff's alfa's voor interbeoordelaarsovereenstemming (met bootstrap 95%-betrouwbaarheidsintervallen) voor deelnemers die wel en niet de LIRIK gebruikten in het onderzoek.

	LIRIK (n=36)		geen LIRIK (n=43)	
Conclusie veiligheid (hoog) ^a	0.41	(.32, .49)	0.40	(.31, .49)
Conclusie veiligheid (laag) ^b	0.47	(.39, .55)	0.37	(.27, .45)
Conclusie veiligheid (missing) ^c	0.48	(.40, .56)	0.42	(.33, .50)
Risico-inschatting nu	0.39	(.33, .45)	0.46	(.40, .51)
Risico-inschatting in de nabije toekomst bij voorziene veranderingen	0.19	(.10, .26)	0.25	(.18, .32)

Noot. ^a Wanneer meer opties gekozen waren bij de conclusie over de huidige veiligheid, is hierbij gekozen voor de hoogste waarde (de optie "onvoldoende informatie" is buiten beschouwing gelaten).

^b Wanneer meer opties gekozen waren bij de conclusie over de huidige veiligheid, is hierbij gekozen voor de laagste waarde (de optie "onvoldoende informatie" is buiten beschouwing gelaten).

^c Wanneer meer opties gekozen waren bij de conclusie over de huidige veiligheid, is hiervoor een missing gemaakt (de optie "onvoldoende informatie" is buiten beschouwing gelaten).

2.3.2 Is de interbeoordelaarsovereenstemming over conclusies over veiligheid en risico's hoger wanneer de LIRIK wordt gebruikt dan wanneer de LIRIK niet wordt gebruikt?

In Tabel 3 is te zien dat het gebruik van de LIRIK niet duidelijk tot meer overeenstemming over conclusies over veiligheid en risico's leidt. De overeenstemming over de huidige veiligheid was wat hoger wanneer de LIRIK werd gebruikt, de mate waarin verschilt voor de drie transformatiemethoden. Overeenstemming over risico's was hoger wanneer de LIRIK niet werd gebruikt. De 95%-betrouwbaarheidsintervallen berekend voor de bootstrap-steekproeven (1000 steekproeven) tonen statistisch significante verschillen in overeenstemming tussen de twee groepen (de intervallen overlappen niet of nauwelijks, wat een statistisch significant verschil impliceert, zie bijv. Belia, Fidler, Williams & Cumming, 2005; Cumming & Finch, 2005). Deze verschillen zijn dus in tegengestelde richtingen voor de verschillende oordelen.

In de niet-LIRIK-groep werden dezelfde opties het minst vaak gekozen als in de LIRIK-groep. De optie "hoog risico" bij voorziene veranderingen werd nog voldoende vaak gekozen, 18 keer, om met voldoende betrouwbaarheid ($p=.05$) te zien of overeenstemming (Krippendorff's alfa) tenminste .5 is, voor deze categorie afzonderlijk. De antwoordcategorieën "laag risico" bij huidig risico en "direct fysiek gevaar" werden

wel vaak genoeg gekozen om met $p=.10$ te zien of overeenstemming (Krippendorff's alfa) tenminste .5 is, voor de categorieën afzonderlijk (Krippendorff, 2011).

Net als in de LIRIK-conditie, liepen conclusies voor hetzelfde kind in de niet-LIRIK-conditie nogal uiteen. Veiligheidsoordelen liepen uiteen over twee categorieën voor een gevalsbeschrijving, over drie categorieën voor de helft van de gevalsbeschrijvingen en over alle vier categorieën voor vijf gevalsbeschrijvingen. Conclusies over huidige risico's liepen uiteen over twee categorieën (een gevalsbeschrijving), drie categorieën (zes gevalsbeschrijvingen), vier categorieën (vier gevalsbeschrijvingen), tot zelfs vijf categorieën (een gevalsbeschrijving). Conclusies over risico's in de nabije toekomst bij voorziene veranderingen liepen uiteen over drie (twee gevalsbeschrijvingen), vier (acht gevalsbeschrijvingen) en ook vijf categorieën (twee gevalsbeschrijvingen) (zie Tabel 2).

In beide condities waren deelnemers al bekend met de LIRIK, hadden ze een training gevolgd voor het gebruik van de LIRIK en gebruikten ze de LIRIK al. Het volgen van een training zou professionals zich al meer bewust kunnen maken van relevante factoren om mee te nemen in hun oordeelsvorming ten aanzien van veiligheid van en risico's voor een kind, ook wanneer zij de LIRIK niet invullen. Daarom hebben we apart naar de overeenstemming gekeken bij deelnemers die een training hebben gevolgd. In Tabel 4 staan de overeenstemmingscoëfficiënten voor deelnemers die een training hadden gevolgd uit beide condities samen en voor deelnemers die een training hadden gevolgd en de LIRIK gebruikten in dit onderzoek. Deze blijken niet hoger dan de coëfficiënten in Tabel 3, voor beide condities (LIRIK en niet-LIRIK). We hebben geen aparte analyses gedaan voor de hoeveelheid ervaring die deelnemers hadden met de LIRIK, omdat we hier geen eenduidige maat voor hadden.

Tabel 4. Krippendorff's alfa's voor interbeoordelaarsovereenstemming (met bootstrap 95%-betrouwbaarheidsintervallen) voor deelnemers die een training hebben gevolgd en de LIRIK wel en niet gebruikten in het onderzoek.

	Training & LIRIK (n=16)		Training (n=33)	
Conclusie veiligheid (hoog) ^a	0.38	(.28, .47)	0.43	(.34, .51)
Conclusie veiligheid (laag) ^b	0.40	(.31, .49)	0.40	(.31, .48)
Conclusie veiligheid (missing) ^c	0.40	(.32, .49)	0.45	(.37, .52)
Risico-inschatting nu	0.37	(.30, .44)	0.44	(.39, .51)
Risico-inschatting in de nabije toekomst bij voorziene veranderingen	0.29	(.21, .37)	0.23	(.16, .31)

Noot. ^a Wanneer meer opties gekozen waren bij de conclusie over de huidige veiligheid, is hierbij gekozen voor de hoogste waarde (de optie "onvoldoende informatie" is buiten beschouwing gelaten).

^b Wanneer meer opties gekozen waren bij de conclusie over de huidige veiligheid, is hierbij gekozen voor de laagste waarde (de optie "onvoldoende informatie" is buiten beschouwing gelaten).

^c Wanneer meer opties gekozen waren bij de conclusie over de huidige veiligheid, is hiervoor een missing gemaakt (de optie "onvoldoende informatie" is buiten beschouwing gelaten).

2.3.3 Interbeoordelaarsovereenstemming voor de aanwezigheid van specifieke factoren die op mogelijke kindermishandeling of verhoogd of verlaagd risico kunnen duiden.

Met de LIRIK wordt de aanwezigheid van specifieke factoren nagegaan die volgens de (empirische) literatuur indicatief zijn voor onveiligheid en risico op kindermishandeling en verwaarlozing: signalen in de interactie tussen ouder(s) en kind, kindsignalen, risicofactoren en beschermende factoren bij de ouders, het kind, het

gezin en de omgeving. De LIRIK vraagt of factoren al dan niet aanwezig zijn (soms dat er zorgen zijn over iets, of iets adequaat is), of dat dit onbekend is. Hoewel volgens de instructie van de LIRIK alleen geconcludeerd kan worden dat een factor niet aanwezig is, wanneer op basis van de beschikbare informatie duidelijk is dat een factor in een gezinssituatie geen rol speelt, is het niet uit te sluiten dat professionals aangeven dat een factor niet aanwezig is gezien de beschikbare informatie, terwijl zij geloven dat met meer informatie een factor toch aanwezig kan blijken te zijn. Wij hebben daarom onderzocht in welke mate deelnemers het eens waren over de aanwezigheid van factoren, zonder verder te kijken of zij het ook eens waren over of een factor afwezig is of dat dit onbekend is gegeven beschikbare informatie. We hebben hiertoe overeenstemmingscoëfficiënten berekend voor twee categorieën: aanwezig (“ja”) versus niet aanwezig (“nee” of “onbekend”). In de praktijk kan het belangrijk zijn om te zien wat relevant maar onbekend is, zodat duidelijk is welke informatie nog nodig is. We hebben daarom ook de overeenstemming berekend over alle drie de categorieën, zoals in de LIRIK (“ja”, “nee”, “onbekend”, een nominale meetschaal). Bij de items over beschermende factoren was er geen optie “onbekend”.

Na ieder onderdeel binnen deel 1 (A, B, C en D) en in deel 2D werd om een samenvattende conclusie gevraagd over of een deelnemer vond dat er geen, enkele of veel factoren in een specifiek domein aanwezig waren, of dat dit onbekend was. We maakten van de gekozen opties “onbekend” (die maar weinig voorkwamen) ontbrekende scores, zodat de overige opties een ordinale schaal van drie categorieën vormden waarvoor we Krippendorff's alfa's berekenden. We hebben ook analyses gedaan waarbij we “onbekend” wel hebben meegenomen (waarbij de opties dan een nominale schaal met vier categorieën vormen) en analyses waarbij we twee categorieën hebben onderscheiden: of deelnemers factoren hadden gesignaleerd in een specifiek domein (“enkele” en “veel” samengenomen) of niet (“nee” en “onbekend” samengenomen).

In Tabel 5 is te zien dat de overeenstemming voor de meeste items laag was. Wanneer naar de overeenstemming voor twee categorieën wordt gekeken, bleek voor 23 items de alfa's lager dan .20 te zijn, voor 26 items lagen alfa's tussen .20 en .40, voor 14 items tussen .40 en .60, voor 5 items tussen .60 en .80 en slechts voor 4 items waren alfa's .80 of hoger. Veel bootstrap-betrouwbaarheidsintervallen bevatten de waarde 0, wat laat zien dat statistische onafhankelijkheid, ofwel niet meer dan toevalsovereenstemming, niet met 95% zekerheid kan worden uitgesloten. De betrouwbaarheidsintervallen zijn vaak ook breed, de schatting van de alfa's hebben dus een grote onzekerheidsmarge. In Tabel 5 staan ook percentages van hoe vaak er “ja” was gekozen per item (voor de gevalsbeschrijvingen samen), ofwel hoe vaak deelnemers gemiddeld dachten dat een factor aanwezig was. Deze percentages geven geen overeenstemming weer, maar geven aan of een factor relatief zeldzaam lijkt (percentage dichtbij 0, de factor wordt niet/nauwelijks geïdentificeerd door deelnemers) of veelvoorkomend (dichtbij 100, deelnemers identificeren de factor in (vrijwel) alle gevallen). Zowel relatief zeldzame als veelvoorkomende factoren blijken vaak brede betrouwbaarheidsintervallen en lage alfa's te hebben. De berekende overeenstemming bleek bij de signalen in de ouder-kindinteractie, bij de kindsignalen en bij de risicofactoren voor drie categorieën in de meeste gevallen lager dan voor twee. Voor de samenvattende conclusies bleek ook dat wanneer de categorie “onbekend” werd meegenomen bij het berekenen van de overeenstemming, de overeenstemming lager was. De overeenstemming over of er factoren waren gesignaleerd, dus zonder het onderscheid te maken tussen veel of enkele, viel nog lager uit.

We hebben niet voor alle items van de LIRIK de overeenstemming tussen beoordelaars berekend. De vraag over wie in het bijzonder een ernstig bedreiging vormt voor een kind, vragen naar de aanwezigheid van

andere, nader te specificeren signalen, risicofactoren en beschermende factoren, de vraag naar de aard van vermoedelijke of aangetoonde mishandeling, en naar het voortduren van aangetoonde mishandeling (als specificering van de hoofdconclusie over huidige veiligheid) hebben we buiten beschouwing gelaten. Te weinig deelnemers hadden deze vragen ingevuld om op betrouwbare wijze overeenstemming te kunnen bepalen.

Tabel 5. Krippendorff's alfa's voor interbeoordelaarsovereenstemming (met bootstrap 95%-betrouwbaarheidsintervallen) over de individuele items van de LIRIK (uitgezonderd de hoofdconclusies) en percentages van hoe vaak deelnemers vonden dat een factor aanwezig was, gemeten over cases en beoordelaars.

	2 Categorieën		3 Categorieën		4 Categorieën	Percentage "ja"
	α	(95%BI)	α	(95%BI)	α	(95%BI)
1. HUIDIGE VEILIGHEIDSSITUATIE						
A. Directe veiligheid						
ernstige bedreiging door handelen/nalaten ouder(s)/primaire opvoeder(s)	.40	(.19, .57)				30
ernstige bedreiging door jeugdige zelf	.16	(-.23, .48)				12
ernstige bedreiging door ander gezinslid/derde	.00	(-.88, .81)				3
<i>vermoedens/aanwijzingen levensbedreigende situatie of fysiek in gevaar</i>	.36	(.17, .54)	.32	(.19, .44)		35
B. Omgang ouder(s)-jeugdige						
zijn er feiten die wijzen op recent:						
lichamelijk geweld	.40	(.20, .60)	.27	(.19, .35)		30
psychisch geweld	.28	(.09, .48)	.27	(.19, .35)		49
lichamelijke verwaarlozing	.08	(-.73, .68)	.13	(.03, .24)		3
emotionele verwaarlozing	.26	(.06, .46)	.20	(.10, .30)		58
seksueel misbruik	.09	(-.87, .77)	.02	(-.09, .12)		2
getuige van huiselijk geweld	.41	(.21, .59)	.33	(.26, .41)		40
opvoeding en verzorging:						
bescherming en veiligheid	.35	(.13, .58)	.27	(.16, .37)		72
basale verzorging	.23	(-.08, .50)	.21	(.12, .29)		15
emotionele warmte (ondersteuning)	.37	(.15, .59)	.30	(.19, .41)		71
regels en grenzen	.31	(.08, .51)	.22	(.12, .32)		70
Stimulering	.20	(.01, .39)	.14	(.04, .24)		64
stabiliteit	.29	(.11, .48)	.21	(.13, .29)		55
<i>aanwijzingen voor bedreigend handelen of nalaten van de ouder(s)</i>	.05	(-.51, .52)	.29	(.17, .41)	.16	(.07, .24)
C. Jeugdige						

psychosociaal functioneren	.06	(-.42, .47)	.05	(-.19, .27)	91
lichamelijke gezondheid	.15	(-.15, .44)	.13	(.05, .21)	15
vaardigheden en cognitieve ontwikkeling	.27	(.07, .47)	.21	(.11, .30)	66
<i>kindsignalen die kunnen wijzen op kindermishandeling</i>	.12	(-.25, .47)	.29	(.16, .43)	.17 (.08, .26)
D. Risico- en beschermende factoren					
<i>Risicofactoren bij de ouder(s)</i>					
Functioneren als opvoeder					
eerder een kind mishandeld, verwaarloosd of misbruikt	.18	(-.12, .46)	.13	(.05, .21)	18
oantoreikende pedagogische kennis en/of vaardigheden	.09	(-.16, .33)	.07	(-.06, .19)	74
bagatelliseren/ontkennen aangetoonde kindermishandeling	.10	(-.16, .32)	.09	(.02, .17)	27
negatieve houding t.a.v. kind	.44	(.25, .63)	.29	(.21, .37)	40
Persoonlijk functioneren					
psychiatrische problematiek	.47	(.29, .65)	.26	(.19, .34)	32
verslavingsproblematiek	.94	(.85, 1.00)	.43	(.36, .51)	26
verstandelijke beperking	.79	(.63, .93)	.38	(.30, .46)	16
Beschikbaarheid voor de jeugdige					
fysiek afwezig	.32	(.14, .52)	.22	(.14, .30)	39
emotioneel afwezig	.24	(.04, .44)	.15	(.06, .24)	61
Voorgeschiedenis					
op jonge leeftijd (<18 jaar) ouder geworden	.19	(-.47, .73)	.40	(.31, .48)	4
zelf slachtoffer van kindermishandeling	.80	(.65, .91)	.53	(.45, .61)	22
eerder geweld gebruikt tegen personen	.45	(.22, .65)	.26	(.18, .35)	22
Problematische partnerrelatie	.51	(.35, .67)	.41	(.33, .49)	58
<i>Risicofactoren bij jeugdige, gezin en omgeving</i>					
jeugdige					
jong kind (<5 jaar)	.37	(-.06, .72)	.27	(.00, .53)	8
belaste voorgeschiedenis (bijv. prematuur)	.39	(.18, .61)	.26	(.18, .35)	21
(ernstige) ziekte, handicap of beperking	.65	(.44, .85)	.23	(.13, .33)	16
gedrags- en/of ontwikkelingsproblemen	.10	(-.38, .50)	.06	(-.14, .28)	91
moeilijk temperament	.04	(-.15, .22)	.05	(-.03, .13)	48
Ongewenst	.63	(.31, .87)	.21	(.12, .31)	9

gezin en omgeving						
laag opleidingsniveau	.62	(.43, .79)	.49	(.41, .56)		26
eenoudergezin, stiefgezin, groot gezin	.85	(.76, .94)	.78	(.72, .85)		43
veel conflicten	.47	(.28, .64)	.34	(.25, .43)		63
huiselijk geweld	.55	(.35, .69)	.37	(.30, .45)		42
instabiel, ongeregeld leven	.21	(-.01, .44)	.18	(.11, .25)		28
materiële/financiële problemen (werkloosheid, huisvesting)	.83	(.70, .94)	.59	(.52, .66)		43
ingrijpende levensgebeurtenissen	.31	(.08, .54)	.25	(.12, .37)		76
sociaal isolement/sociaal conflict	.31	(.12, .50)	.27	(.18, .37)		46
<i>risicofactoren kindermishandeling</i>	.15	(-.25, .54)	.33	(.20, .45)	.19	(.11, .27)
Beschermende factoren						
ouders						
gevoel van competentie, draagkracht	.22	(-.01, .44)				30
positief zelfbeeld	.59	(.40, .77)				27
ondersteunende partner	.29	(-.03, .57)				15
kan eigen jeugdervaringen hanteren	.14	(-.32, .54)				8
positieve jeugdervaringen	.55	(.29, .78)				16
kan steun vragen/profiteren van	.16	(-.07, .36)				32
emotionele beschikbaarheid	.33	(.01, .60)				15
Flexibiliteit	.17	(-.43, .76)				4
bereid en in staat om te veranderen	.49	(.29, .66)				32
gezin en omgeving						
steun informeel netwerk	.21	(.03, .41)				43
steun formeel netwerk	.09	(-.21, .38)				17
jeugdige						
sociaal vaardig	.40	(.14, .62)				17
positief zelfbeeld	.01	(-.87, .81)				3
bovengemiddelde intelligentie	.47	(.16, .78)				10
aantrekkelijk uiterlijk	.62	(.46, .79)				30
goede relatie belangrijke volwassene(n)	.15	(-.07, .35)				38
egoveerkracht (stressresistentie)	.00	(-.52, .41)				8
bereid en in staat om te veranderen	.27	(-.04, .59)				13

<i>beschermende factoren</i>	.11	(-.29, .45)	.23	(.09, .37)	.08	(-.03, .18)
2. RISICOTAXATIE						
A. Aanvullende risicofactoren bij mogelijke/aangetoonde kindermishandeling						
(vermoedelijke) pleger heeft direct toegang tot jeugdige	.31	(.12, .48)				50
derden hebben geen toezicht op jeugdige	.04	(-.27, .37)				17
B. Wat kan er gebeuren						
mogelijke risico's voor de jeugdige						
levensbedreigende situatie/direct fysiek in gevaar	.30	(.18, .42)				
voortduren/herhalen van kindermishandeling	.34	(.22, .47)				
ontstaan van kindermishandeling	.15	(-.02, .29)				
verwachte gevolgen voor de jeugdige	.20	(.06, .34)				
C. Beschermende factoren						
beschermende factoren die de risico's kunnen verminderen	.07	(-.33, .43)	.14	(-.05, .31)	.05	(-.01, .12)

Noot. Cursief gedrukte items zijn bedoeld om antwoorden op voorgaande items samen te vatten. Het "percentage 'ja'" geeft weer hoe vaak deelnemers de optie 'ja' kozen bij de items over de aanwezigheid van individuele factoren. Hierbij zijn alle cases samengenomen. Dit betreft dus geen percentage overeenstemming.

2.3.4 Evaluaties van deelnemers van de bruikbaarheid van de LIRIK.

Aan het eind van de tweede vragenlijst werden (in zowel de LIRIK als niet-LIRIK-versies) vragen gesteld over hoe deelnemers dachten over de bruikbaarheid van de LIRIK. Deze vragen werden door 48 tot 50 deelnemers beantwoord⁵. We vroegen in welke mate deelnemers vonden dat de LIRIK geschikt was voor de gestelde doelen: het ondersteunen van professionals bij de vorming van expliciete en onderbouwde oordelen over de veiligheid van kinderen en risico's in de nabije toekomst, door het systematisch beoordelen van beschikbare informatie aan de hand van relevante aandachtspunten (opties waren: zeer geschikt, deels, een beetje, nauwelijks, helemaal niet). Van de 50 deelnemers vonden er 17 de LIRIK hiervoor zeer geschikt (34%); 26 deels (52%); 7 een beetje (14%) en niemand nauwelijks of helemaal niet. De meeste deelnemers (85.7%) gaven aan dat ze de LIRIK niet gebruikten voor andere dan de gestelde doelen. Deelnemers die dit wel deden, zeiden de LIRIK ook te gebruiken bij oordelen over of uit huis geplaatste kinderen veilig terug konden keren, of om veiligheidsissues met ouders te bespreken.

We vroegen in welke mate deelnemers het eens waren met de stelling dat de LIRIK de hulpverlener goed in staat stelt om zijn of haar professionele oordeel als beroepskracht te geven (zeer eens, eens, geen mening, oneens, zeer oneens). De meeste deelnemers (71.4%) waren het hiermee eens of zeer eens, 22.4% had geen mening, 6.1% was het oneens. Verder was 75% het (zeer) eens met de stelling dat de LIRIK meerwaarde heeft ten opzichte van het eigen professionele oordeel zonder gebruikmaking van de LIRIK, 14.6% had geen mening, 10.4% was het hiermee oneens. De helft van de deelnemers was het eens met de stelling dat het gebruik van de LIRIK leidt tot een expliciet en gefundeerd oordeel over aard en ernst van kindermishandeling of onveilige opvoedsituatie, dat de basis kan vormen voor verschillende beslissingen, 29.2% had geen mening, 20.8% was het hiermee oneens. Ook werd gevraagd of deelnemers de vorm van de LIRIK, een vragenlijst, geschikt vonden voor het oordelen over kindveiligheid. De meeste deelnemers waren het hiermee eens (67.3%), 24.5% had geen mening, 8.2% was het hiermee oneens. Tenslotte vroegen we of deelnemers vonden dat het meerwaarde heeft de LIRIK samen met een collega/met collega's in te vullen. Hier was 88% het (zeer) mee eens, 12% had geen mening.

Deelnemers merkten op bij de verschillende vragen dat de LIRIK hen hielp om informatie te ordenen, een overzicht te krijgen van welke informatie nog nodig is, in gesprek te gaan met ouders over veiligheid en dat het gebruik van de LIRIK kan leiden tot meer objectieve oordelen. Deelnemers gaven aan dat ze door het gebruik van de LIRIK zich er meer bewust van waren dat zij steeds naar veiligheid en risico's moeten kijken. Aan de andere kant werd aangegeven dat het (te) veel tijd kost om de LIRIK volledig in te vullen en dat alleen het invullen van de LIRIK onvoldoende is. De items werden soms niet concreet genoeg bevonden en de ruimte voor nuanceringen te beperkt.

Op de vraag wat deelnemers belangrijk vonden voor een goed risicotaxatie-instrument noemden ze veel verschillende dingen. Veel deelnemers gaven aan dat een goed risicotaxatie-instrument objectief moet zijn, een volledig overzicht moet geven van relevante factoren, concrete en ondubbelzinnige factoren moet bevatten, beschermende factoren moet bevatten en ruimte moet geven voor nuanceringen en verklaringen en duidelijk onderscheid moet maken tussen feiten en vermoedens. Deelnemers vonden het ook belangrijk dat

⁵ Eerst vroegen we of deelnemers de LIRIK kenden. Alle deelnemers kenden de LIRIK, op een deelnemer na in de niet-LIRIK-conditie. Deze deelnemer werd na deze vraag doorgeleid naar de vraag over wat belangrijk is voor een goed risicotaxatie-instrument.

zij een instrument kunnen gebruiken samen met ouders en collega's. Tenslotte zou een goed instrument niet te lang moeten zijn en makkelijk bruikbaar in veel situaties.

2.4 Discussie

We hebben de interbeoordelaarsovereenstemming geanalyseerd voor conclusies van hulpverleners over veiligheid en risico's voor kinderen in gevalsbeschrijvingen, gegeven met en zonder gebruikmaking van de LIRIK. De overeenstemming over veiligheids- en risico-oordelen was onvoldoende, gezien het belang van deze oordelen (cf. Cichetti, 2001; Evers et al, 2010; Krippendorff, 2011). Deelnemers waren het in de minste mate eens over risico's in de nabije toekomst bij voorziene veranderingen. De overeenstemming voor de conclusies gegeven met de LIRIK bleek niet hoger dan die voor conclusies gegeven zonder dat de LIRIK was gebruikt. Verschillen in overeenstemming tussen de groepen (die oordeelden met of zonder de LIRIK) waren niet groot en wezen op hogere overeenstemming over directe veiligheid en lagere overeenstemming over risico's wanneer de LIRIK was gebruikt. We vonden ook geen hogere overeenstemming voor deelnemers die een training voor het gebruik van de LIRIK hadden gevolgd.

In vergelijkingen van veiligheids- en risico-inschattingen gemaakt met op consensus gebaseerde instrumenten en met actuariële instrumenten blijkt de interbeoordelaarsovereenstemming hoger voor actuariële instrumenten, maar nog altijd niet hoog (Austin et al., 2005; Baird et al., 1999; Shlonsky & Wagner, 2005). We kunnen nu een vergelijking toevoegen: die van een op consensus gebaseerd instrument en geen instrument. Hoewel studies lastig te vergelijken zijn, lijkt de door ons gevonden overeenstemming wat hoger dan gevonden is voor risicotaxaties gemaakt met andere op consensus gebaseerde instrumenten (Baird et al., 1999; Barber et al., 2007).

Dit onderzoek geeft ook meer inzicht in de betrouwbaarheid van het identificeren van individuele aanwijzingen voor (mogelijke) kindermishandeling en verwaarlozing. Andere studies hebben zich vaak beperkt tot analyses van de overeenstemming van conclusies die daarop zijn gebaseerd. Het identificeren van de meeste relevante aanwijzingen bleek niet erg betrouwbaar, wat een opvallende bevinding is als men ervan uitgaat dat een goed risicotaxatie-instrument overeenstemming over afzonderlijke items die aan de algemenere conclusies voorafgaan veronderstelt (Baird et al., 1999; Shlonsky & Wagner, 2005). Barber et al. (2007) vonden ook een lage overeenstemming voor afzonderlijke risicofactoren: Cohens kappa van lager dan .5 voor alle op één na van 22 onderzochte items. Orsi et al. (2014) vonden Fleiss' kappa lager dan .5 voor de meer dan de helft van de meer dan 100 onderzochte aanwijzingen. Voor sommige concrete, makkelijk vast te stellen aanwijzingen was de overeenstemming wel hoger (bijv. gezinssamenstelling), maar soms ook verrassend laag (bijv. leeftijd van het kind). Misschien hadden deelnemers de gevalsbeschrijvingen niet altijd zorgvuldig gelezen, hoewel zij hiernaar steeds terug konden kijken. Het is ook mogelijk dat hulpverleners minder op dergelijke statische risicofactoren letten, omdat deze weinig aanknopingspunten voor mogelijke interventies bieden. De overeenstemming was lager wanneer niet alleen werd gekeken of deelnemers een factor als aanwezig signaleerden, maar ook naar het onderscheid tussen "onbekend" en "nee". Deelnemers gebruikten deze laatste categorieën mogelijk niet allemaal volgens de instructie bij de LIRIK. De lage overeenstemming die we vonden voor de veiligheids- en risicoconclusies gevormd met de LIRIK is misschien te verwachten in het licht van het gebrek aan overeenstemming over de daaraan voorafgaande items. Orsi et al. (2014) verwijzen naar de moeilijkheid om goede voor kans gecorrigeerde overeenstemming te vinden, als te identificeren aanwijzingen weinig voorkomen. In dit onderzoek bleek de overeenstemming

vaak erg laag voor aanwijzingen die maar weinig werden gesignaleerd. Als de aanwezigheid (of afwezigheid) van een specifieke factor zeldzaam is, zouden professionals het er echter juist over eens moeten zijn wanneer deze dan aanwezig (of afwezig) is. Ditzelfde geldt voor conclusies over veiligheid: hoewel onveilige of risicovolle situaties relatief zeldzaam zullen zijn, is het cruciaal dat deze onveilige of risicovolle situaties correct geïdentificeerd worden. Overeenstemming over de afwezigheid van iets wat niet vaak voorkomt (of overeenstemming over de aanwezigheid van iets dat juist meestal voorkomt), is niet betekenisvol als beoordelaars het niet eens zijn over de situaties waarin het zeldzame toch voorkomt (of het veelvoorkomende afwezig is) (cf. Krippendorff, 2004).

Hoewel overeenstemming voor oordelen gemaakt met de LIRIK onvoldoende bleek, waren de deelnemers over het algemeen positief over de LIRIK. De meeste deelnemers geloofden dat de LIRIK geschikt is voor de gestelde doelen: om op systematische wijze tot expliciete, gefundeerde oordelen over kindveiligheid te komen. Er werd soms opgemerkt dat het invullen een extra administratieve belasting vormde, maar deelnemers gaven vaak aan dat zij geloofden dat de LIRIK nuttig was om te komen tot meer volledige, objectieve veiligheidsoordelen (zie ook Ten Berge & Meuwissen, 2013; Ten Berge & Van Rossum, 2009). Het gebruik van instrumenten die niet resulteren in (meer) betrouwbare en valide oordelen is feitelijk zinloos, en kan zelfs leiden tot overmatig vertrouwen in het oordeel (Regehr, Bogo, Shlonsky & LeBlanc, 2010). De LIRIK bevat factoren die naar voren zijn gekomen uit de empirische literatuur en de professionele praktijk, en die relevant zouden zijn om systematisch na te gaan om zo tot een weloverwogen veiligheidsinschatting te komen. Er wordt daarbij dus verondersteld dat professionals bewust (kunnen) redeneren over risico's. In de praktijk maken professionals misschien intuïtieve oordelen, zoals op basis van niet-pluis-gevoelens bij familiesituaties (Munro, 1999, 2005, 2008). Het is de vraag of het gebruik van een risicotaxatie-instrument het ongestructureerde oordeel kan verbeteren, door intuïties te controleren (Baumann et al., 2005; cf. Kahneman & Klein, 2009). Actuariële instrumenten voorzien, in tegenstelling tot de LIRIK en andere consensus-instrumenten, wel in een algoritme voor het wegvan de geïdentificeerde factoren. Dit kan leiden tot meer consistente risico-oordelen, die minder worden beïnvloed door intuïties van professionals.

Verschillen tussen professionals die dezelfde situaties beoordelen kunnen het gevolg zijn van verschillende vooronderstellingen, verwachtingen, of behoefte aan extra informatie (Mandel, Lehman & Yuille, 1994), verschillende persoonlijke beslisdrempels (Baumann, Dalglish, Fluke & Kern, 2011; Dalglish, 2000; Schuerman et al., 1999), of verschillende culturele contexten (Gold, Benbenishty, & Osmo, 2001). Enosh en Bayer-Topilsky (2014) vonden vooroordelen ten aanzien van gezinnen met lage SES en sociale minderheden in risico-inschattingen en uithuisplaatsingsbesluiten. Camasso en Jagannathan (2000) vonden dat risico-inschattingen vooral worden beïnvloed door de aard van de mishandeling. Deze laatste bevinding verklaren echter niet waardoor beoordelaars het oneens zijn over dezelfde casus.

2.4.1 Beperkingen

Een beperking van dit onderzoek is dat we gevalsbeschrijvingen hebben gebruikt, die hoewel gebaseerd op echte casuïstiek, geen werkelijke en volledige gevallen waren, en dat deelnemers geen extra informatie konden vragen. Geschreven vignetten zijn niet helemaal representatief voor de werkelijke praktijk, hoewel vaak beter bruikbaar in onderzoek (Taylor, 2006). In de praktijk gebruiken hulpverleners mogelijk meer informatie dan zij nu beschikbaar hadden (bijvoorbeeld non-verbale informatie uit gesprekken en

huisbezoeken), wat zou kunnen leiden tot meer overeenstemmende oordelen (cf. Barber et al., 2007). Meer informatie hoeft echter niet te leiden tot betere oordelen, omdat dit ook meer complexe verwerking vraagt. Irrelevante informatie kan daarbij onterecht meegewogen worden (Dana, Dawes & Peterson, 2013; Grove et al., 2000). Deelnemers gaven meestal wel een conclusie over de veiligheid, hoewel zij steeds (ook) konden kiezen voor de optie dat ze te weinig informatie hadden. Dit hoeft niet te betekenen dat er genoeg informatie was, omdat professionals in de dagelijkse praktijk ook vaak tot inschattingen van de veiligheid moeten komen met onvolledige informatie.

Het aantal gevalsbeschrijvingen in dit onderzoek was beperkt. Waar veel studies naar interbeoordelaarsovereenstemming veel cases laten beoordelen door enkele beoordelaars, hadden wij relatief veel beoordelaars die enkele cases beoordeelden. Met meer gevalsbeschrijvingen en een meer systematische variatie van kenmerken zou men kunnen onderzoeken welke kenmerken samenhangen met meer of minder overeenstemming tussen beoordelaars. Onze gevalsbeschrijvingen leken representatief voor wat betreft de ernst van problemen. Deelnemers beoordeelden ze gemiddeld niet als ernstiger of minder ernstig dan wat zij tegenkomen in hun dagelijkse werk. Een vergelijking van de frequenties van de verschillende oordelen met die in een grote steekproef uit het validiteitsonderzoek suggereert echter dat in de gevalsbeschrijvingen in dit betrouwbaarheidsonderzoek sprake is van meer ernstige problematiek. Persoonlijke betrokkenheid bij cliënten zou de conclusies van hulpverleners kunnen beïnvloeden.

Deelnemers vormden geen representatieve steekproef van alle mogelijke LIRIK-gebruikers. De meeste deelnemers kwamen uit twee grote jeugdzorgorganisaties, er waren echter weinig deelnemers die op dezelfde locatie werkzaam waren. In beide condities (LIRIK en niet-LIRIK) waren deelnemers al bekend met de LIRIK. Het feit dat de meeste deelnemers al bekend waren met de LIRIK zou de vergelijkbare overeenstemming in de twee groepen (LIRIK, niet-LIRIK) kunnen verklaren, hoewel louter kennis van de LIRIK waarschijnlijk niet tot dezelfde systematische werkwijze voor het nagaan van de aanwezigheid van relevante factoren zal leiden als het werkelijk invullen van de lijst. De manier waarop deelnemers in de LIRIK-conditie tot hun oordelen kwamen lijkt misschien juist op de manier waarop deelnemers in de niet-LIRIK-conditie dat deden, omdat de LIRIK niet voorziet in een wegingsalgoritme voor de gevonden aanwijzingen en ook LIRIK gebruikers dus hun eigen wegeningen moesten maken.

3 Deelonderzoek 2. Predictieve validiteit

Cora Bartelink & Leontien de Kwaadsteniet

3.1 Inleiding

De kern van dit deelonderzoek draaide om de predictieve validiteit. De onderzoeksvragen hiervoor luiden:

1. Wat is de predictieve validiteit, de oordelen over de huidige veiligheid en toekomstige risico's van de jeugdige voor AMK-meldingen, Raadsmeldingen, ondertoezichtstellingen en/of geboden jeugdhulp?
2. Verschilt de predictieve validiteit van oordelen die met en zonder de LIRIK gemaakt zijn?
3. Wat is de predictieve validiteit van oordelen over de LIRIK-items voor AMK-meldingen, Raadsmeldingen, ondertoezichtstellingen en/of geboden jeugdhulp?

Aanvullend is gekeken naar de interne consistentie van de LIRIK. De onderzoeksvraag hierbij luidt:

4. Hoe is de samenhang tussen LIRIK-items en de oordelen over de huidige veiligheid en de toekomstige risico's?

Voorafgaand aan het onderzoek was de verwachting dat vaker AMK-meldingen, Raadsmeldingen, ondertoezichtstellingen en uithuisplaatsingen gedaan zouden worden, vaker jeugdhulp gericht op veiligheid geboden zou worden en zich vaker een crisis zou voordoen bij jeugdigen voor wie professionals concludeerden dat zij onveilig waren, of dat er grote risico's op onveiligheid voor de toekomst bestonden, dan bij jeugdigen voor wie professionals concludeerden dat zij op dit moment veilig waren en er geen risico's op onveiligheid voor de toekomst bestonden.

Ook was de verwachting voorafgaand aan het onderzoek dat de predictieve validiteit voor oordelen gemaakt met de LIRIK groter zou zijn dan voor het ongestructureerde professionele oordeel (= zonder LIRIK), omdat de LIRIK helpt om meer expliciet te beoordelen in hoeverre een jeugdige op dit moment veilig is en welke risico's er voor de nabije toekomst zijn, aan de hand van factoren waarvan onderzoek heeft aangetoond dat die een rol spelen in de oordeelsvorming over kindermishandeling.

Het is lastig de predictieve validiteit van de LIRIK en het ongestructureerde oordeel zuiver te bepalen, omdat verwacht mag worden dat professionals anders handelen wanneer zij concluderen dat het onveilig is of de risico's op onveiligheid groot zijn, dan wanneer zij een situatie als veilig inschatten en weinig risico's zien. Interventies die instellingen als Just en Spirit bieden kunnen leiden tot vermindering van onveiligheid en risico's. Het is ethisch niet verantwoord om een jeugdige waarvan gedacht wordt dat die in een onveilige of risicovolle situatie verblijft, daarin te laten blijven zonder iets te gaan doen om die situatie te verbeteren. Doordat er hulp geboden wordt maar ook door andere veranderingen in de gezinssituatie kan de predictieve validiteit van veiligheids- en risico-inschattingen lager lijken dan dat in werkelijkheid het geval was als er geen hulp geboden zou zijn. Het was daarom voorafgaand aan het onderzoek niet de verwachting dat de predictieve validiteit perfect zou zijn, maar we verwachtten wel dat in de situaties die als veilig of met klein risico beoordeeld zijn minder acties zouden zijn ondernomen dan in situaties die als onveilig of met groot risico beoordeeld zijn.

3.2 Methode

3.2.1 Onderzoeksgroep

In totaal zijn de gegevens van 370 jeugdigen verzameld onder drie instellingen (GGD Gelderland-Zuid, Juzt en Spirit). Bij 278 jeugdigen is de LIRIK ingevuld (LIRIK-groep), bij 92 jeugdigen is een ongestructureerd professioneel oordeel over de huidige veiligheid en toekomstige risico's voor de veiligheid gegeven (controlegroep) (zie voor de uitleg over de beide groepen paragraaf "Procedure"). Jeugdigen die op het moment van de veiligheids- en risicotaxatie in een residentiële instelling of een gesloten residentiële instelling verbleven werden van het onderzoek uitgesloten. In Tabel 1 zijn achtergrondgegevens per groep weergegeven.

In de onderzoeksgroep zaten 218 jongens en 149 meisjes. In beide groepen was de verhouding tussen het aantal jongens en meisjes gelijk ($\chi^2(1) = 0,26, p = .62$). De leeftijd van de jeugdigen op het moment dat de veiligheidsbeoordeling werd uitgevoerd, varieerde tussen 0 en 20 jaar. De gemiddelde leeftijd lag hoger in de controlegroep dan in de LIRIK-groep ($t(367) = 3,35, p = .001$).

In 265 gezinnen was slechts één jeugdige aangemeld voor hulp. In 70 gezinnen was meer dan één jeugdige aangemeld voor hulp. Wanneer meerdere jeugdigen waren aangemeld, varieerde het aantal aangemelde jeugdigen tussen twee en zes. In 21 dossiers bij Juzt en Spirit waarin één jeugdige aangemeld was, bleek uit de beschrijving van de situatie dat het om gezinsproblematiek ging (het ging hier om een zogenaamd gezinsdossier en niet om een dossier dat op naam van een individuele jeugdige was geregistreerd). De gegevens over het exacte aantal jeugdigen waren in deze gezinsdossiers niet altijd goed terug te vinden. De LIRIK-groep en de controlegroep verschilden niet in het aantal aangemelde jeugdigen in het gezin ($t(362) = -1,79, p = .07$).

De veiligheids- en risicotaxaties zijn tussen september 2013 en november 2014 gedaan (meetmoment T1). Na ongeveer zes maanden (meetmoment T2) hebben twee onderzoekers onafhankelijk de cliëntdossiers gescoord om informatie over de uitkomstvariabelen te verzamelen (periode tussen T1 en T2 varieerde tussen 146 en 360 dagen; gemiddelde 219 dagen en standaardafwijking 41 dagen). De tijd tussen de veiligheidsbeoordeling op T1 en het dossieronderzoek op T2 was voor de controlegroep met een verschil van 24 dagen significant lager dan voor de LIRIK-groep ($t(343) = -5,06, p = .000$) (zie Tabel 1). Voor de controlegroep vonden we echter niet minder uitkomstmaten (zie paragraaf 3.3).

Tabel 1. Kenmerken van de jeugdigen in de onderzochte dossiers in de LIRIK-groep en de controlegroep

	LIRIK	Ongestructureerd professioneel oordeel
Sekse		
Jongens	166	52
Meisjes	110	39
Leeftijd	9,6 jaar (SD = 5,1)	11,7 jaar (SD = 5,0)
Aantal kinderen aangemeld voor hulp	1,4 (SD = 0,91)	1,3 (SD = 0,73)
Instelling		
GGD-Gelderland-Zuid	28	16
Juzt	135	45
Spirit	115	31
Tijd tussen T1 en T2	225 dagen (SD = 34)	201 (SD = 52)

3.2.2 Procedure

Medewerkers maakten over jeugdigen die bij de instelling als cliënt kwamen een veiligheids- en risicotaxatie bij het openen van een zaak of gedurende het behandelingsproces (T1).

Medewerkers maakten drie inschattingen, met of zonder de LIRIK:

- Huidige veiligheid (antwoordmogelijkheden: er zijn geen aanwijzingen voor kindermishandeling; er is mogelijk sprake van kindermishandeling; er is sprake van feitelijk aangetoonde kindermishandeling; er is sprake van een levensbedreigende situatie of direct fysiek gevaar; onvoldoende informatie om een oordeel te vormen)
- Risico voor de jeugdige op dit moment (antwoordmogelijkheden: klein; reëel, groot, zeer groot)
- Risico voor de jeugdige bij voorziene veranderingen (antwoordmogelijkheden: klein; reëel, groot, zeer groot)

Het is niet bekend hoeveel medewerkers de veiligheids- en risicotaxaties hebben uitgevoerd in beide groepen en in hoeverre medewerkers in de LIRIK-groep ook de ongestructureerde veiligheids- en risicotaxatie in de controlegroep hebben uitgevoerd, omdat dit niet betrouwbaar te scoren was. Met name de LIRIK kon door een andere hulpverlener (en zelfs een andere instelling, namelijk Bureau Jeugdzorg) zijn ingevuld dan de uitvoerend hulpverlener. Aangezien alle beoordelingen met de LIRIK uit de periode van september 2013 tot en met maart 2014 en alle ongestructureerde oordelen die in de periode van maart tot en met november 2014 onder een brede groep medewerkers zijn verzameld in het onderzoek werden meegenomen, is het onwaarschijnlijk dat de resultaten zijn beïnvloed door een of enkele zeer uitzonderlijk beoordelende medewerkers. Medewerkers die de ongestructureerde veiligheids- en risicobeoordelingen uitvoerden kunnen ook met de LIRIK beoordelingen uitgevoerd hebben. Omdat de LIRIK organisatiebreed grotendeels was ingevoerd, was het niet mogelijk om de beoordelingen in de controlegroep te laten uitvoeren door alleen medewerkers die niet bekend waren met de LIRIK. De veiligheids- en risicobeoordelingen in de controlegroep kunnen dus beïnvloed zijn door kennis van de LIRIK.

Een deel van de jeugdigen werd met de LIRIK beoordeeld en een deel op ongestructureerde wijze. Naar het ongestructureerde oordeel werd gevraagd met de concluderende items inclusief de vraag om de conclusies toe te lichten, maar geen verdere items die op enige manier het professioneel oordeel van de werker kon beïnvloeden. In Bijlage 2 is het formulier dat medewerkers invulden voor het ongestructureerd professioneel oordeel te vinden.

Wanneer ouders met de instellingen te maken hebben, krijgen zij standaard een formulier waarop zij kunnen aangeven of zij wel of niet bereid zijn hun gegevens voor onderzoek ter beschikking te stellen. Over de jeugdigen waarvan ouders geen toestemming voor het gebruik van hun gegevens hebben gegeven, is geen informatie verzameld.

Na ongeveer zes maanden is een tweede meting uitgevoerd (T2). Onafhankelijke onderzoekers hebben de cliëntdossiers doorzocht op informatie over de aanwezigheid van:

- AMK-meldingen of meldingen bij Bureau Jeugdzorg (voor leesgemak verder AMK-meldingen genoemd),
- meldingen bij de Raad voor de Kinderbescherming (Raadsmeldingen),
- ondertoezichtstellingen (OTS),
- uithuisplaatsingen (UHP),
- interventies gericht op veiligheid,

Deze variabele werd gescoord als er in het cliëntdossier acties waren beschreven die expliciet bedoeld waren om te werken aan de veiligheid. Hierbij valt te denken aan de aanwezigheid van een veiligheidsplan, de beschrijving dat Signs of Safety is toegepast of dat in gesprekken met ouders veiligheid is besproken omdat daar zorgen over waren.

- crises.

Deze variabele werd gescoord wanneer beschreven stond in het cliëntdossier dat er een crisis geweest was en/of dat de crisisdienst van de instelling was ingeschakeld.

Het coderen van de cliëntdossiers gebeurde met behulp van een gestandaardiseerde checklist (zie Bijlage 3). In eerste instantie werden behalve bovenstaande uitkomstmaten ook aanvullende informatie gecodeerd die indicaties voor (on)veiligheid hadden moeten geven. Door gebrek aan informatie in de cliëntdossiers en omdat het onmogelijk bleek om de beschikbare informatie betrouwbaar te coderen hebben de onderzoekers hier verder van afgezien (zie Bijlage 8 voor verdere informatie hierover).

Twee onderzoekers hebben twee dagen proefgedraaid met de checklist en deze is aan de hand daarvan aangescherpt. Vervolgens hebben zij onafhankelijk veertig dossiers gescoord en is de interbeoordelaarsbetrouwbaarheid berekend. Tabel 2 geeft de interbeoordelaarsbetrouwbaarheid in percentages en kappa voor de uitkomstmaten weer. In eerste instantie varieerde Cohen's kappa tussen .48 en 1.00, wat matig tot goed is. Voor de items met een lage interbeoordelaarsbetrouwbaarheid zijn nieuwe afspraken over de scoring gemaakt. Vervolgens hebben de onderzoekers opnieuw 24 dossiers onafhankelijk gecodeerd. Cohen's kappa hiervoor varieerde tussen .65 en 1.00, wat als goed beschouwd mag worden (zie Bijlage 4 voor interpretatie van Cohen's kappa).

Tabel 2. Interbeoordelaarsbetrouwbaarheid in percentage en kappa over AMK-meldingen, OTS, hulp en uithuisplaatsing

	Antwoord categorieën	Ronde 1		Ronde 2	
		percentage	kappa	percentage	kappa
AMK melding	2	97	.79	100	-. ^A
Raadsmelding	2	100	1.00	100	1.00
OTS	3	89	.48	96	.65
UHP	2	97	.88	100	1.00
Veiligheidsinterventies	2	92	.53	96	.83
Crisis	2	95	.64	100	1.00

Noot. ^A Er was geen enkele AMK melding gedaan, waardoor kappa berekenen niet mogelijk is.

3.2.3 Analyses

In de eerste plaats zijn frequenties en percentages berekend voor hoe vaak verschillende oordelen ten aanzien van de huidige veiligheid en de risico's voor de toekomst zijn gegeven en van alle uitkomstmaten. Er is gecontroleerd met een χ^2 -toets of er significante verschillen waren tussen de LIRIK-groep en de controlegroep (ongestructureerd professioneel oordeel) in de veiligheids- en risicotaxatie en de uitkomsten op T2.

Vervolgens zijn voor beide groepen de veiligheids- en risico-inschattingen gekoppeld aan de uitkomstmaten op T2 om te zien of er samenhang is tussen de inschattingen op T1 en de uitkomsten op T2. Daarbij deden zich bij de huidige veiligheidsinschatting twee problemen voor. Het eerste probleem was dat medewerkers meerdere opties konden kiezen. Het andere probleem dat zich voordeed was dat de eerste vier antwoordmogelijkheden gezamenlijk een ordinale schaal vormden, maar dat de vijfde antwoordmogelijkheid (onvoldoende informatie) niet in deze ordening past. Om deze twee problemen op te lossen is zijn de vijf dichotome oorspronkelijke variabelen gehercodeerd tot nieuwe ordinale variabelen (vierpuntschaal). De antwoordmogelijkheid “onvoldoende informatie” is daarbij buiten beschouwing gelaten, omdat dit niet aangeeft hoe de medewerker de veiligheid beoordeelt. Als medewerkers naast deze antwoordmogelijkheid ook één van de vier andere opties hadden gekozen, dan werd die andere optie in de nieuwe variabele meegenomen. Wanneer medewerkers twee of meer van de antwoordmogelijkheden 1 tot en met 4 hadden gekozen, namen we ofwel de hoogste score (getransformeerde variabele “veiligheid-hoog”), de laagste score (getransformeerde variabele “veiligheid-laag”), of we maakten er een ontbrekende score van (getransformeerde variabele “veiligheid-missing”) (zie ook de analyse van interbeoordelaarsovereenstemming in 2.2.5). Uiteindelijk is ervoor gekozen om alleen verdere analyses uit te voeren met de getransformeerde variabele “veiligheid-missing”, omdat het om een klein aantal zaken (16) ging en het in de onderbouwing van de oordelen onduidelijk was of de medewerker nu vond dat de hoogste of de laagste score het beste bij de zaak paste⁶. Om interpretatie door de onderzoekers te vermijden is ervoor gekozen om alle zaken waar medewerkers meerdere oordelen over de huidige veiligheid hadden aangegeven buiten beschouwing te laten en in de analyses alleen die zaken mee te nemen waar medewerkers één oordeel over de huidige veiligheid hadden gegeven.

Om onderzoeksvraag 1 en 2 te beantwoorden zijn kruistabellen gemaakt waarin veiligheids- en risico-inschattingen gerelateerd zijn aan de uitkomsten op T2. In de tabellen is per antwoordcategorie op T1 weergegeven hoeveel van de uitkomsten op T2 aanwezig waren. Daarbij waren meer opties mogelijk (bijvoorbeeld AMK-melding en crisis).

In de LIRIK-groep bleek de conclusie dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar altijd gegeven te zijn in combinatie met een van de andere oordelen over de huidige veiligheid. Door de zaken waar medewerkers meerdere oordelen over de huidige veiligheid gegeven hadden weg te laten uit de analyses, kwam het oordeel dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar niet meer in de analyses over de samenhang voor. Om de LIRIK-groep goed met de controlegroep te kunnen vergelijken is besloten ook in de controlegroep de zaken waarin het oordeel was dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar niet in de analyses over de samenhang mee te nemen.

⁶ Overigens bleek uit eerste analyses die voor alle drie de hergecodeerde variabelen werden uitgevoerd dat de verschillen in uitkomsten maar klein waren.

Spearman's rangcorrelaties zijn berekend om de samenhang tussen de veiligheids- en risicotaxatie op T1 en de uitkomsten op T2 te onderzoeken. De interpretatie van correlaties is te vinden in Bijlage 4

Om de derde onderzoeksvraag te beantwoorden (predictieve validiteit van LIRIK-items) zijn Spearman's rangcorrelaties berekend.

Tot slot zijn eveneens Spearman's rangcorrelaties berekend om onderzoeksvraag 4 te beantwoorden (interne samenhang tussen LIRIK-items en oordelen over huidige veiligheid en risico's).

3.3 Resultaten voor de predictieve validiteit

Tabel 3 laat zien welke veiligheids- en risico-inschattingen de medewerkers van de drie instellingen gemaakt hebben (weergegeven in frequenties en in percentages van het totale aantal geïncludeerde jeugdigen). Bij de conclusie over de huidige veiligheid konden medewerkers meer opties kiezen, waardoor het aantal gekozen opties niet overeenkomt met het totale aantal jeugdigen. De percentages komen gezamenlijk niet altijd tot 100%, omdat er missings waren (medewerkers hadden in 70 zaken geen oordeel over de huidige veiligheid, in 53 zaken geen oordeel over de risico's voor de nabije toekomst en in 118 zaken geen oordeel over de risico's bij voorziene veranderingen ingevuld). Vrijwel alle missings bleken in de LIRIK-groep voor te komen (in de controlegroep ontbrak in drie zaken het oordeel over de huidige veiligheid, in één zaak het oordeel over de risico's voor de nabije toekomst en in vijf zaken het oordeel over de risico's bij voorziene veranderingen).

Om te bepalen of de groepen verschilden in hun veiligheids- en risico-inschattingen is een χ^2 -toets uitgevoerd. In deze toets is de antwoordmogelijkheid "onvoldoende informatie om een oordeel te vormen" buiten beschouwing gelaten. De χ^2 -toets laat zien dat er een significant verschil is tussen de LIRIK-groep en de controlegroep in hoe vaak verschillende conclusies over de huidige veiligheid en risico's worden gegeven. Gestandaardiseerde residuen geven aan in welke cellen de waargenomen frequentie significant hoger of lager ligt dan de verwachte frequentie. Wanneer gestandaardiseerde residuen boven 1,96 liggen komen die combinaties vaker voor dan je op basis van toeval mag verwachten. Cellen waar het omgekeerde het geval is en waar de gestandaardiseerde residuen onder -1,96 liggen, zijn combinaties die minder vaak voorkomen dan je op basis van toeval mag verwachten. De gestandaardiseerde residuen laten zien dat in de controlegroep medewerkers vaker de conclusie trokken dat er sprake was van een levensbedreigende situatie of direct fysiek gevaar dan in de LIRIK-groep. De χ^2 -toets voor de beide risico's voor de toekomst liet geen verschillen zien tussen de LIRIK-groep en de controlegroep voor hoe vaak de verschillende opties voor risico's op dit moment en bij voorziene veranderingen als even groot werden gekozen.

In Bijlage 5 zijn de frequenties en percentages van de overige LIRIK-items gerapporteerd.

Tabel 3. Frequenties en percentages van de veiligheids- en risico-oordelen gemaakt met de LIRIK en ongestructureerde oordelen

	LIRIK		Ongestructureerd professioneel oordeel		χ^2
	Frequentie (n=278)	Percentage	Frequentie (n=92)	Percentage	
Conclusie huidige veiligheidssituatie					10,52**
Er zijn geen aanwijzingen voor kindermishandeling	133	47,8	58	63,0	
Er is mogelijk sprake van	65	23,4	24	26,1	

kindermishandeling					
Er is sprake van feitelijk aangetoonde kindermishandeling	26	9,4	5	5,4	
Er is sprake van een levensbedreigende situatie of direct fysiek gevaar	3	1,1	4	4,3	
Onvoldoende informatie om een oordeel te vormen	34	12,1	3	3,3	
Conclusie risico's voor de jeugdige					
Conclusie risico's voor de jeugdige - op dit moment					0,87
Klein	138	49,6	60	65,2	
Reëel	57	20,5	21	22,8	
Groot	25	9,0	9	9,8	
Zeer groot	5	1,8	1	1,1	
Bij voorziene veranderingen in de nabije toekomst					5,67
Klein	90	32,4	59	64,1	
Reëel	41	14,7	18	19,6	
Groot	18	6,5	9	9,8	
Zeer groot	11	4,0	1	1,1	

** $p < .01$ (tweezijdig)

In Tabel 4 zijn de resultaten voor de uitkomsten op T2 gerapporteerd. Er is een χ^2 -toets uitgevoerd om na te gaan of er verschillen zijn tussen de LIRIK-groep en de controlegroep in de uitkomstvariabelen. In de controlegroep kwamen verhoudingsgewijs meer uithuisplaatsingen voor dan in de LIRIK-groep.

De uitkomstmaten sluiten elkaar deels uit (bijv. als er al een OTS is, dan zal er niet snel nog een AMK-melding worden gedaan), maar kunnen ook met elkaar in combinatie voorkomen (bijv. een OTS in combinatie met een uithuisplaatsing of veiligheidsinterventie). Daarom is ook de nieuwe variabele “onveilige uitkomst” berekend (nieuwe dichotome variabele). Daarin is geteld of tenminste één van de zes uitkomstmaten voor onveiligheid voorkwam bij een jeugdige.

In de LIRIK-groep werden in totaal bij 27,0% van de gevallen ($n = 75$) één of meerdere uitkomstmaten gerapporteerd, meestal was dit er één (46 van de 75). In de controlegroep werd bij 32,6% van de gevallen ($n = 30$) één of meerdere uitkomstmaten gerapporteerd, in de helft van de gevallen was dit er één (15 van de 30).

Tabel 4. Frequenties en percentages van de uitkomstmaten op T2 (AMK-meldingen, ondertoezichtstelling, hulp, crisis en uithuisplaatsing)

	LIRIK		Ongestructureerd professioneel oordeel		χ^2
	Frequentie (n=278)	Percentage	Frequentie (n=92)	Percentage	
AMK melding	17	6,1	3	3,3	1,10
Raadmelding	11	4,0	5	5,4	0,82
Nieuwe OTS op T2	2	0,7	2	2,2	
Nog aanwezige OTS op T2	26	9,4	14	15,2	2,46
Uithuisplaatsing	13	4,8	15	17,6	14,84**
Veiligheidsinterventies	34	12,2	9	9,8	0,40
Crisis/crisisdienst	20	7,2	8	8,7	0,22

Onveilige uitkomst	75	27,0	30	32,6	1,08
--------------------	----	------	----	------	------

* $p < .05$ (tweezijdig), ** $p < .000$ (tweezijdig)

In de zaken waarvoor de LIRIK is ingevuld is er tussen T1 en T2 in twee zaken een nieuwe ondertoezichtstelling uitgesproken. Op T2 waren er in totaal 26 ondertoezichtstellingen. In de controlegroep is er tussen T1 en T2 ook in twee zaken een nieuwe ondertoezichtstelling uitgesproken. Op T2 waren er in totaal 14 ondertoezichtstellingen. In de analyse van de samenhang tussen de veiligheids- en risicobeoordelingen op T1 en de uitkomsten op T2 zijn alle ondertoezichtstellingen die er op T2 aanwezig waren, meegenomen, omdat de aantallen voor de nieuwe ondertoezichtstellingen te klein waren voor zinvolle analyses en omdat het voortduren van een ondertoezichtstelling ook een indicatie voor de veiligheid is. In Bijlage 6 zijn nog enkele aanvullende gegevens over T1 en T2 te vinden.

3.3.1 *Samenhang tussen veiligheids- en risico-inschattingen en uitkomsten – rangcorrelaties*

Tabel 5 laat de Spearman's rangcorrelaties zien tussen de conclusies over huidige veiligheid en risico's enerzijds en anderzijds de uitkomsten op T2. Hogere scores voor de conclusies betekenen meer onveiligheid/grotere risico's. Dus de conclusie "er is geen sprake van kindermishandeling" kreeg score 1 en de conclusie "er is sprake van een levensbedreigende situatie" score 4. Datzelfde geldt voor de risicotaxatie: een klein risico kreeg score 1 en een groot risico score 4. Uitkomstmaten werden gecodeerd met score 1 (aanwezig) of score 0 (afwezig).

Omdat in de controlegroep meer onveilige situaties leken voor te komen dan in de groep waarin de LIRIK was gebruikt kunnen de rangcorrelaties eerder hoger uitvallen in de controlegroep. We hebben daarom ook berekend wat de maximaal mogelijke samenhang zou zijn, gegeven hoe vaak de verschillende oordelen werden gegeven en hoe vaak de verschillende uitkomstmaten voorkwamen. Zie hiervoor Bijlage 7.

Bij de zaken die met LIRIK beoordeeld werden, bleken ondertoezichtstellingen vaker voor te komen naarmate medewerkers oordeelden dat de huidige veiligheid slechter was en de risico's op dit moment en bij voorziene veranderingen als groter inschatten (zwakke correlaties). Ook bleken in de zaken met LIRIK vaker één van de uitkomstmaten voor te komen naarmate medewerkers oordeelden dat de huidige veiligheid slechter was of de risico's groter inschatten (zwakke tot matige correlaties).

In de zaken die op ongestructureerde wijze beoordeeld werden, bleek een sterke positieve samenhang tussen de veiligheids- en risico-inschattingen enerzijds en uithuisplaatsing anderzijds. Naarmate medewerkers de huidige veiligheid slechter vonden of de risico's op dit moment en bij voorziene veranderingen hoger inschatten, was er vaker een uithuisplaatsing. Ook werden er vaker veiligheidsinterventies en raadmeldingen uitgevoerd naarmate medewerkers oordeelden dat de huidige veiligheid slechter was en de risico's groter inschatten (matige correlatie). Ook bleek in de zaken die op ongestructureerde wijze beoordeeld werden, vaker één van de uitkomstmaten voor te komen naarmate de veiligheid meer in het geding was of de risico's groter inschat werden (matige tot sterke correlaties).

Tabel 5. Spearman's rangcorrelaties voor de samenhang tussen oordelen over huidige veiligheid en toekomstige risico's en uitkomsten op T2

	AMK-	Raads-	OTS	UHP	Veiligheids-	Crisis	Onveilige
--	------	--------	-----	-----	--------------	--------	-----------

	melding	melding			interventie		uitkomst ^a
LIRIK							
Conclusie huidige veiligheidssituatie	.18	.06	.39**	-.02	.10	.26*	.34**
Conclusie risico's voor de jeugdige - op dit moment	.27*	.10	.33**	.18	.29**	.22*	.41**
Bij voorziene veranderingen in de nabije toekomst	.12	-.06	.20*	.08	.03	.13	.19*
Ongestructureerd professioneel oordeel							
Conclusie huidige veiligheidssituatie	.19	.15	.34*	.56**	.50*	.31	.46**
Conclusie risico's voor de jeugdige - op dit moment	.29	.49*	.00	.38**	.67**	.30	.40**
Bij voorziene veranderingen in de nabije toekomst	.31	.30	.19	.55**	.40*	.36	.34**

Noot. ^a De variabele "onveilige uitkomst" is een dichotome variabele waarin gekeken is of één van de zes andere uitkomsten zich heeft voorgedaan.

* $p < .05$ (tweezijdig), ** $p < .000$ (tweezijdig)

3.3.2 Samenhang tussen veiligheids- en risico-inschattingen en uitkomsten – ordening naar niveau van onveiligheid en risico

Tabel 6 laat zien hoe vaak de verschillende uitkomsten die onveiligheid indiceren voorkomen gespecificeerd per niveau van beoordeelde onveiligheid ofwel in hoeverre bij stijging van de beoordeelde (on)veiligheid het aantal uitkomsten die onveiligheid indiceren, consequent toeneemt. In de tabel is eerst weergegeven welk percentage een veiligheidsoordeel uitmaakte van het totaal aantal oordelen per groep. Dus in 47,8% van alle gevallen in de LIRIK-groep werd geoordeeld dat de jeugdige veilig was (missings wel meegerekend). Vervolgens is per veiligheidsoordeel aangegeven bij welk percentage zich een specifieke uitkomstmaat voordeed. Dus in 3,1% van de zaken waar de medewerker op T1 met de LIRIK concludeerde dat er geen kindermishandeling speelde, bleek op T2 een AMK-melding gedaan te zijn. Daarnaast is in Tabel 6 de base rates gerapporteerd van uitkomsten: het percentage waarin een bepaalde uitkomstmaat zich in de totale groep (LIRIK of controlegroep) heeft voorgedaan. Dus in de LIRIK-groep is bij 4,6% van alle jeugdigen een AMK-melding gedaan. Naarmate de percentages van een uitkomstmaat meer oplopen bij oordelen die meer onveiligheid weergeven, worden met de oordelen beter meer onveilige van meer veilige situaties onderscheiden, ofwel is de predictieve validiteit van oordelen beter (Baird & Wagner, 2000). Daarnaast zouden bij oordelen van veiligheid de percentages voor hoe vaak een uitkomstmaat voorkomt lager moeten zijn dan de base rate, en bij oordelen van onveiligheid zouden de percentages hoger moeten zijn dan de base rate.

In de LIRIK-groep is voor ondertoezichtstellingen (OTS), crises en de onveilige uitkomsten samengenomen een consequente toename te zien naarmate geoordeeld wordt dat het onveiliger is. Bij het ongestructureerd professioneel oordeel is dit bij geen enkele uitkomst op T2 te zien.

Gekeken naar de verhoudingen met de base rates, blijkt dat de percentages uitkomstmaten in beide groepen onder de base rates liggen wanneer de conclusie werd getrokken dat er geen sprake was van

kindermishandeling, met uitzondering van het percentage uithuisplaatsingen in de LIRIK-groep. In situaties waarin geconcludeerd werd dat er mogelijk sprake was van kindermishandeling of kindermishandeling feitelijk aangetoond was, bleek in de meeste gevallen de resultaten boven de base rate te liggen. In de LIRIK-groep waren percentages in vergelijking met de base rates vaker zoals verwacht (hoger dan de base rate bij oordelen van (meer) onveiligheid) dan in de controlegroep. De LIRIK lijkt hier dus wat beter het onderscheid te maken tussen veilige en onveilige situaties dan het ongestructureerd professioneel oordeel.

Tabel 6. Percentages voor hoe vaak uitkomstmaten voor onveiligheid voorkomen per niveau van het oordeel over de huidige veiligheid

	Verdeling oordelen binnen de LIRIK- en controlegroep A	AMK-melding	Raads-melding	OTS	UHP	Veiligheids-interventie	Crisis	Onveilige uitkomst B
LIRIK								
Geen kindermishandeling	47,8	3,1	2,3	6,2	3,2	8,5	6,2	15,5
Mogelijk kindermishandeling	23,4	7,4	1,9	7,4	3,8	18,5	9,3	29,6
Feitelijk aangetoonde kindermishandeling	9,4	7,1	7,1	57,1	0,0	0,0	28,6	78,6
<i>Base rate</i> ^c		4,6	2,5	10,2	3,1	10,7	8,6	23,9
Ongestructureerd professioneel oordeel								
Geen kindermishandeling	63,0	1,7	1,7	8,6	3,8	1,7	5,2	15,5
Mogelijk kindermishandeling	26,1	4,5	4,5	31,8	38,1	18,2	18,2	63,6
Feitelijk aangetoonde kindermishandeling	5,4	0,0	0,0	0,0	33,3	0,0	0,0	33,3
<i>Base rate</i>		2,4	2,4	14,5	14,5	6,0	8,4	28,9

Noot. ^A De variabele “verdeling oordelen binnen de LIRIK- en controlegroep” geeft de percentages aan van de oordelen over de huidige veiligheid.

^B De variabele “onveilige uitkomst” is een dichotome variabele waarin gekeken is of één van de zes andere uitkomsten zich heeft voorgedaan.

^C De base rate is het percentage waarin een bepaalde uitkomstmaat zich in de totale groep heeft voorgedaan.

De zaken waarin medewerkers de conclusie trokken dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar zijn buiten beschouwing gelaten in de verdere analyses. Dat betekent niet dat in deze zaken geen acties zijn ondernomen op T2. In één van de drie zaken waar met de LIRIK de conclusie werd getrokken dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar, werd een raadmelding gedaan, in één was er een ondertoezichtstelling, in één was er een uithuisplaatsing, in twee

werd een veiligheidsinterventie gedaan, in één was er een crisis. In twee van de drie zaken kwam één of meerdere van de uitkomstmaten voor. In één van de zaken waar de conclusie op T1 met de LIRIK was dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar, is geen van de uitkomstmaten op T2 gevonden. In één van de vier zaken waar ongestructureerd de conclusie werd getrokken dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar, werd een AMK-melding gedaan, in twee een raadmelding, in één was er een ondertoezichtstelling, in drie een uithuisplaatsing, in twee was een veiligheidsinterventie uitgevoerd, in één was er een crisis geweest. In drie van de vier zaken was één of meerdere van de uitkomstmaten voor. In één van de zaken waar de conclusie op T1 met het ongestructureerd professioneel oordeel dat er sprake was van een levensbedreigende situatie/direct fysiek gevaar, is geen van de uitkomstmaten op T2 gevonden.

Tabel 7 en Tabel 8 laten zien hoe vaak de verschillende uitkomsten die onveiligheid indiceren voorkomen gespecificeerd per risiconiveau op dit moment (Tabel 7) of bij voorziene veranderingen (Tabel 8), ofwel in hoeverre bij stijging van de beoordeelde risico's het aantal uitkomsten die onveiligheid indiceren, consequent toeneemt. In de tabel is eerst weergegeven welk percentage een risico-oordeel uitmaakte van het totaal aantal oordelen per groep. Dus in Tabel 7 is te zien dat voor 49,6% van alle gevallen in de LIRIK-groep werd geoordeeld dat het risico voor de jeugdige op dit moment klein was; in Tabel 8 dat voor 32,1% van alle gevallen in de LIRIK-groep werd geoordeeld dat het risico voor de jeugdige bij voorziene veranderingen klein was. Vervolgens is per risico-oordeel aangegeven bij welk percentage zich een specifieke uitkomstmaat voordeed. Dus in 4,3% van de zaken waar de medewerker op T1 met de LIRIK concludeerde dat er een klein risico was op dit moment, bleek op T2 een AMK-melding gedaan te zijn. Daarnaast zijn in Tabel 7 en 8 ook de base rates gerapporteerd van uitkomsten: de percentages waarin een bepaalde uitkomstmaat zich in de totale groep (LIRIK of controlegroep) heeft voorgedaan. Dus in de LIRIK-groep is bij 6,7% van alle jeugdigen een AMK-melding gedaan. De base rates in Tabellen 6, 7 en 8 verschillen enigszins van elkaar doordat de percentages de verhouding weergeven tussen oordelen en uitkomsten, en er bij de verschillende oordelen (veiligheid, risico op dit moment en risico bij voorziene veranderingen) missings zijn (medewerkers hebben geen oordeel ingevuld).

Tabel 7. Percentages voor hoe vaak uitkomstmaten voor onveiligheid voorkomen per niveau van voor het oordeel over de risico's op dit moment

	Verdeling oordelen binnen de LIRIK- en controlegroep ^A	AMK-melding	Raadmelding	OTS	UHP	Veiligheidsinterventie	Crisis	Onveilige uitkomst ^B
LIRIK								
Klein	49,6	4,3	2,9	5,1	2,2	8,7	5,1	16,7
Reëel	20,5	7,0	5,3	8,8	10,9	15,8	14,0	38,6
Groot	9,0	16,0	4,0	28,0	0,0	28,0	12,0	72,0
Zeer groot	1,8	20,0	0,0	20,0	0,0	40,0	0,0	40,0
Base rate ^C		6,7	3,6	8,9	4,1	13,3	8,0	28,9
Ongestructureerd professioneel								

oordeel								
Klein	65,2	1,7	1,7	15,0	11,3	1,7	5,0	20,0
Reëel	22,8	9,5	14,3	19,0	19,0	19,0	19,0	57,1
Groot	9,8	0,0	11,1	11,1	44,4	44,4	11,1	55,6
Zeer groot	1,1	0,0	0,0	0,0	100	0,0	0,0	100
Base rate		3,3	5,5	15,4	17,9	9,9	8,8	33,3

Noot. ^A De variabele “verdeling oordelen binnen de LIRIK- en controlegroep” geeft de percentages aan van de oordelen over de huidige veiligheid.

^B De variabele “onveilige uitkomst” is een dichotome variabele waarin gekeken is of één van de zes andere uitkomsten zich heeft voorgedaan.

^C De base rate is het percentage waarin een bepaalde uitkomstmaat zich in de totale groep heeft voorgedaan.

In beide groepen is voor vrijwel geen van de uitkomstmaten een consequente toename te zien naarmate geoordeeld wordt dat de risico's groter zijn. In de LIRIK-groep is alleen voor meer AMK-meldingen een consequente toename te zien naarmate geoordeeld wordt dat de risico's groter zijn. In de controlegroep was een consequente toename te zien voor uithuisplaatsingen.

Gekeken naar de verhoudingen met de base rates, blijkt dat de percentages uitkomstmaten in beide groepen onder de base rates liggen wanneer de conclusie werd getrokken dat het risico klein was. De percentages uitkomstmaten liggen meestal boven de base rates wanneer de conclusie getrokken werd dat het risico reëel, groot of zeer groot was. In de LIRIK-groep waren percentages in vergelijking met de base rates vaker zoals verwacht (hoger dan de base rate bij oordelen van (meer) risico) dan in de controlegroep. De LIRIK lijkt hier dus wat beter het onderscheid te maken tussen veilige en onveilige situaties dan het ongestructureerd professioneel oordeel.

Tabel 8. Percentages voor hoe vaak uitkomstmaten voor onveiligheid voorkomen per niveau van voor het oordeel over de risico's bij voorziene veranderingen

	Verdeling oordelen binnen de LIRIK- en controlegroep ^A	AMK-melding	Raadsmelding	OTS	UHP	Veiligheids-interventie	Crisis	Onveilige uitkomst ^B
LIRIK								
Klein	32,1	5,6	4,4	7,8	3,4	12,2	6,7	24,4
Reëel	19,2	7,3	2,4	12,2	4,9	14,6	17,1	29,3
Groot	7,0	11,1	0,0	22,2	5,9	27,8	0,0	50,0
Zeer groot	2,7	9,1	9,1	18,2	0,0	0,0	18,2	45,5
Base rate ^C		6,9	3,8	11,2	3,8	13,8	9,4	30,0
Ongestructureerd professioneel oordeel								
Klein	64,1	1,7	3,4	11,9	7,5	5,1	5,1	23,7
Reëel	19,6	11,1	11,1	16,7	33,3	22,2	16,7	44,4
Groot	9,8	0,0	11,1	33,3	44,4	11,1	22,2	66,7
Zeer groot	1,1	0,0	0,0	0,0	100	100	0,0	100
Base rate		3,4	5,7	14,9	28,5	10,3	9,2	33,3

Noot. ^A De variabele “verdeling oordelen binnen de LIRIK- en controlegroep” geeft de percentages aan van de oordelen over de huidige veiligheid.

^B De variabele “onveilige uitkomst” is een dichotome variabele waarin gekeken is of één van de zes andere uitkomsten zich heeft voorgedaan.

^C De base rate is het percentage waarin een bepaalde uitkomstmaat zich in de totale groep heeft voorgedaan.

In beide groepen is voor geen enkele uitkomstmaat een consequente toename te zien naarmate geoordeeld wordt dat de risico’s bij voorziene veranderingen groter zijn.

Gekeken naar de verhoudingen met de base rates, blijkt dat de percentages uitkomstmaten in beide groepen meestal onder de base rates liggen wanneer de conclusie werd getrokken dat het risico klein was, met uitzondering van het percentage raadsmeldingen in de LIRIK-groep. De percentages uitkomstmaten liggen meestal boven de base rates wanneer de conclusie getrokken werd dat het risico reëel, groot of zeer groot was. In beide groepen vielen de percentages in vergelijking met de base rates even vaak onder de base rate, terwijl deze boven de base rates hadden moeten liggen.

3.3.3 Samenhang tussen de LIRIK-items en uitkomsten op T2

Vervolgens is ook voor de items van de LIRIK die vooraf gaan aan de drie hoofdconclusies gekeken in hoeverre deze samenhangen met de uitkomsten op T2. In Tabel 9 zijn Spearman’s rangcorrelaties weergegeven, waarbij voor de aanwezigheid van zorgen of een risico- of beschermende factor een score 1 was gegeven en bij afwezigheid een score 0, evenals als voor de uitkomstmaten (zie 3.3.1). De samenhang tussen enerzijds zorgen en risicofactoren en anderzijds de uitkomsten op T2 zouden positief moeten zijn, voor beschermende factoren worden negatieve correlaties verwacht, op basis van de literatuur.

Tabel 9. Spearman’s rangcorrelaties voor de samenhang tussen scores op LIRIK-items en uitkomsten op T2

	AMK- melding	Raads- melding	OTS	UHP	Veiligheids- interventie	Crisis	Onveilige uitkomst
1. Huidige veiligheidssituatie							
A. Directe veiligheid							
Bedreiging door ouder(s)/primaire opvoeder(s)	.06	-.03	.15*	.05	.00	.04	.07
Ernstige kindermishandeling	-.04	.20**	.11	.18**	.08	.13*	.11
Onvoldoende bescherming	.01	.05	.20**	.08	.08	.08	.20**
Onvoldoende basiszorg	.07	.06	.15*	-.02	.10	.00	.21**
Opvoeder niet beschikbaar	.03	.08	.06	.01	.05	.01	.12*
Ouder(s) belemmeren zicht op / toegang tot jeugdige?	.05	-.04	-.06	.06	.00	.12*	.04
Bedreiging door jeugdige zelf	-.07	.02	.02	.15*	-.10	-.07	.01
Bedreiging door ander gezinslid/derde	.09	-.05	.05	-.05	.18**	.00	.13*
Zijn er vermoedens dat de jeugdige op dit moment fysiek in gevaar is?	-.05	.05	.05	.02	.23**	.01	.08
B. Omgang ouder(s) - jeugdige							
Lichamelijk of psychisch geweld	.01	.14*	.05	.15*	.13*	.14*	.14*
Verwaarlozing	.07	.03	.21*	-.04	.08	.13	.23**
Seksueel misbruik ^A	-	-	-	-	-	-	-

Getuige van huiselijk geweld	.21**	.21**	.09	.05	.26**	.10	.21**
Bescherming en veiligheid	.05	.07	.20**	.10	.20**	.11	.28**
Basale verzorging	.08	.12	.16*	.04	.11	.06	.24**
Emotionele warmte (ondersteuning)	.05	.12	.06	.15*	.02	-.05	.16*
Regels en grenzen	.10	-.02	.08	.13	.12	.13	.21**
Stimulering	.12	.02	.09	.18**	.03	.03	.19**
Stabiliteit	.13	.12	.04	.16*	.19**	.18**	.31**
Ouder geeft geen inzicht in eigen handelen	.06	.03	.02	.01	.08	.10	.08
Onvoldoende informatie beschikbaar	.13*	-.02	-.01	-.03	.13*	.11	.07
Zijn er aanwijzingen voor bedreigend handelen of nalaten van de ouder(s)?	.03	.04	.12	-.06	.16*	.02	.16*
C. Jeugdige(n)							
Psychosociaal functioneren	.12	.08	.12	.14*	.07	.12	.22**
Lichamelijke gezondheid	.16*	.14*	.12	.09	.05	.32**	.22**
Vaardigheden en cognitieve ontwikkeling	.07	.15*	.08	.03	.01	.03	.09
Zijn er kindsignalen die wijzen op een onveilige opvoedingssituatie of kindermishandeling?	.11	.11	.15*	.05	.15*	.16**	.25**
D. Risicofactoren							
<i>Risicofactoren bij de ouders</i>							
Eerder een kind mishandeld, verwaarloosd of misbruikt	.09	.22**	.20**	.08	.28**	.24**	.26**
Ontoereikende pedagogische kennis en/of vaardigheden	.05	-.02	.12	.13	.10	.07	.20**
Problemen in de ouder-kind interactie	.12	.00	.11	.19**	.05	.13	.21**
Bagatelliseren/ontkennen aangetoonde kindermishandeling	.10	-.01	.05	-.02	.10	.01	.12
Negatieve houding t.a.v. het kind	.03	.06	.05	.15*	.06	.05	.15*
Psychiatrische problematiek	-.05	.05	-.01	-.09	.25**	-.00	.09
Verslavingsproblematiek	.01	.04	.04	-.06	.26**	.07	.08
Verstandelijke beperking	-.04	-.04	.15*	-.03	.03	.07	.16*
Fysieke beschikbaarheid	.01	.02	.05	-.14*	.05	-.00	-.02
Emotionele beschikbaarheid	.07	.11	-.03	-.06	.07	.00	.05
Op jonge leeftijd (< 18 jaar) ouder geworden	.02	-.05	.07	-.04	.09	.08	.09
Zelf slachtoffer van kindermishandeling	-.03	.12	.22*	.13	.19*	.22*	.25**
Eerder geweld gebruikt tegen personen	.04	.10	.25**	.04	.30**	.15	.12
Problematische partnerrelatie	.10	.17*	.08	.03	.23**	.08	.21**
Anders	.20	.20	.02	.21	.29*	.18	.31*
<i>Risicofactoren bij de jeugdige</i>							
Jong kind (< 5 jaar)	.16*	.00	.10	-.09	.23**	.13	.17*
Belaste voorgeschiedenis (Ernstige) ziekte of handicap	.11	.21**	.20**	.09	.16*	.37**	.28**
Gedrags- en/of ontwikkelingsproblemen	.03	-.04	.10	.11	.05	.09	.08
Moelijk temperament	.04	.04	-.07	-.02	.02	.08	.04
Ongewenst	-.04	.12	.17*	.14*	.23**	.24**	.20**
Anders	.03	.03	.06	-.07	.03	.13	.07
<i>Risicofactoren bij gezin en/of omgeving</i>							
Eenoudergezin, stiefgezin, groot gezin	.12	.04	.14*	.10	.12	.06	.26**
Veel conflicten	.10	.16*	.08	.21**	.13*	.05	.27**

Huiselijk geweld	.09	.10	.22**	.21**	.18*	.25**	.19**
Instabiel, ongeregeld leven	-.01	.10	.28**	.07	.20**	.04	.25**
Materiële/financiële problemen	.12	.04	.22**	.08	.23**	.17*	.30**
Ingrijpende levensgebeurtenissen	.08	.11	.21**	.09	.21**	.13	.28**
Sociaal isolement/conflict	.02	-.04	-.01	.07	.09	-.04	.08
Anders	.06	-.08	.02	-.08	.14	.10	.08
Zijn er risicofactoren voor kindermishandeling?	.08	.02	.08	.10	.19**	.16*	.22**
2. Risicotaxatie							
Aanvullende risicofactoren bij vermoedelijke of aangetoonde kindermishandeling							
(Vermoedelijke) pleger heeft direct toegang tot de jeugdige	.06	-.01	.05	.03	.26**	-.01	.13*
Derden hebben geen zicht op de jeugdige	-.03	-.03	.07	.12	.05	.08	.13*
Wat kan er gebeuren?							
Mogelijke risico's: Levensbedreigende situatie of direct fysiek gevaar	.09	.03	.07	.00	.13	.06	.23**
Mogelijke risico's: Ontstaan van kindermishandeling	.11	-.04	.14*	.08	.10	-.00	.23**
Mogelijke risico's: Voortduren/herhaling van kindermishandeling	.14*	.07	.25**	.10	.15*	.10	.33**
Mogelijke risico's: Anders	.09	-.17	-	.07	.47**	.30*	.40**
Verwachte gevolgen voor de jeugdige	.12	.02	.24**	.09	.12	.17*	.34**
Beschermende factoren							
Zijn er beschermende factoren die de risico's kunnen verminderen?	-.24**	-.07	-.07	-.06	-.16*	-.14	-.31**
<i>Beschermende factoren bij de ouders</i>							
Ouder 1							
Gevoel van competentie, draagkracht	.05	.03	-.03	-.07	.02	.04	-.03
Positief zelfbeeld	.02	.05	-.05	-.01	-.07	.06	-.04
Ondersteunende partner	-.05	-.10	.07	-.05	.07	-.02	.01
Kan eigen jeugdervaringen hanteren	.03	-.03	-.07	-.05	.09	-.07	.03
Positieve jeugdervaringen	-.01	-.00	-.03	-.06	.05	-.06	.02
Kan steun vragen/profiteren van	.01	-.02	.02	-.02	.15*	.11	.09
Emotionele beschikbaarheid	.03	-.02	.01	-.08	.03	.03	-.03
Flexibiliteit	-.01	.11	-.03	-.07	.09	.06	.04
Bereid en in staat om te veranderen	.06	.03	-.09	-.08	.13*	.01	.05
Ouder 2							
Gevoel van competentie, draagkracht	-.04	.00	-.02	-.11	.08	-.02	-.05
Positief zelfbeeld	-.01	.03	-.06	-.09	.01	.01	-.08
Ondersteunende partner	-.09	-.01	-.04	-.07	.06	-.00	-.07
Kan eigen jeugdervaringen hanteren	-.05	-.02	-.04	-.03	.03	-.06	-.05

Positieve jeugdervaringen	-.08	-.00	-.02	-.02	.03	.00	-.05
Kan steun vragen/profiteren van	-.06	-.02	-.02	-.08	.04	-.01	-.07
Emotionele beschikbaarheid	-.02	-.01	-.01	-.11	.09	.03	-.02
Flexibiliteit	-.11	.02	-.07	.00	.03	-.04	-.07
<i>Beschermende factoren bij de jeugdige</i>							
Jeugdige 1							
Sociaal vaardig	-.07	-.01	-.05	-.03	.07	-.02	-.03
Positief zelfbeeld	-.00	.05	-.05	-.06	-.00	-.02	-.03
Bovengemiddelde intelligentie	-.04	-.01	-.11	-.02	-.02	-.01	-.08
Aantrekkelijk uiterlijk	-.01	-.03	.02	-.09	.04	-.10	.00
Goede relatie belangrijke volwassene(n)	-.09	-.05	-.01	-.08	.02	-.04	-.04
Egoveerkracht (stressresistentie)	-.06	.07	-.03	-.09	-.03	.04	-.05
Bereid en in staat om te veranderen	-.06	-.04	-.11	-.02	-.01	.01	-.05
Jeugdige 2							
Sociaal vaardig	.00	.03	-.03	-.06	.19**	-.07	.09
Positief zelfbeeld	.02	.05	.05	-.05	.25**	-.06	.15*
Bovengemiddelde intelligentie	.08	-.03	-.04	-.03	.12	-.04	.04
Aantrekkelijk uiterlijk	.05	-.06	-.09	-.06	.12	-.08	-.00
Goede relatie belangrijke volwassene(n)	-.00	.03	.02	.01	.23**	-.07	.15*
Egoveerkracht (stressresistentie)	.07	.10	-.05	-.03	.10	-.04	.08
<i>Beschermende factoren bij gezin en omgeving</i>							
Steun informeel netwerk	.13*	-.00	.00	.04	.12*	.11	.16**
Steun formeel netwerk	.07	.02	.06	.06	.12*	.14*	.16**

Noot. ^A Het was niet mogelijk een correlatie voor het item “seksueel misbruik” te berekenen, omdat medewerkers in geen enkele zaak de conclusie trokken dat er sprake was van seksueel misbruik.

* $p < .05$ (tweezijdig), ** $p < .000$ (tweezijdig)

Veel correlaties bleken laag, wat duidt op geen of heel weinig samenhang met de uitkomsten op T2. Een klein aantal items bleek zwak maar statistisch significant samen te hangen met de uitkomsten op T2. Wat opvalt, is dat per uitkomstmaat verschillen bestaan in de items die ermee samenhangen. Items die met een uitkomstmaat samenhangen, hingen niet per se ook met andere uitkomstmaten samen. Er zijn maar enkele LIRIK-items die samenhangen met meerdere uitkomstmaten, namelijk:

- Deel 1A: ernstige kindermishandeling,
- Deel 1B: lichamelijk of psychisch geweld, getuige van huiselijk geweld, bescherming en veiligheid, stabiliteit,
- Deel 1C: lichamelijke gezondheid, aanwezigheid van kindsignalen,
- Deel 1D: eerder kind mishandeld, zelf slachtoffer van kindermishandeling, eerder geweld gebruikt, problematische partnerrelatie, jong kind, belaste voorgeschiedenis, ongewenst, veel conflicten, huiselijk geweld, instabiel/ongeregeld leven, materiële/financiële problemen, ingrijpende levensgebeurtenissen, aanwezigheid van risicofactoren,
- Deel 2: voortduren/herhaling van kindermishandeling, verwachte gevolgen, aanwezigheid van beschermende factoren, steun informeel netwerk.

Hoewel de tussentijdse conclusie of er beschermende factoren zijn wel negatief samenhang met de uitkomsten op T2, bleken de afzonderlijke beschermende factoren geen samenhang te hebben met de uitkomsten op T2, met uitzondering van steun vanuit het informele en formele netwerk. Deze laatste twee variabelen bleken echter tegen de verwachting in niet negatief, maar met positief samen te hangen met enkele uitkomstmaten (met name crisis en alle uitkomsten).

Tenslotte is er gekeken of er samenhang is tussen het aantal zorgen, risicofactoren en beschermende factoren op de verschillende domeinen enerzijds en de uitkomsten op T2 anderzijds. Hiertoe is per domein de som berekend van het aantal keer dat er zorgen, risicofactoren of beschermende factoren aanwezig waren. Tabel 10 toont de rangcorrelaties.

Tabel 10. Spearman's rangcorrelaties voor de samenhang tussen het aantal zorgen/risico- en beschermende factoren enerzijds en anderzijds uitkomsten op T2

	AMK- melding	Raads- melding	OTS	UHP	Veiligheids- interventie	Crisis	Onveilige uitkomst
Aantal zorgen over omgang ouder – jeugdige	.11	.12*	.13*	.17**	.15*	.11	.32**
Aantal zorgen over functioneren jeugdige	.12*	.21**	.15*	.11	.08	.17**	.27**
Aantal risicofactoren ouders	.09	.14*	.13*	.08	.22**	.15*	.26**
Aantal risicofactoren jeugdige	.12*	.12*	.14*	.04	.17**	.24**	.24**
Aantal risicofactoren gezin en omgeving	.11	.15**	.24**	.18**	.22**	.17**	.36**
Aantal beschermende factoren ouders	.02	.01	-.04	-.08	.13*	.06	.05
- Ouder 1	.04	.01	-.03	-.07	.12	.05	.05
- Ouder 2	-.07	-.01	-.06	-.05	.06	-.00	-.06
Aantal beschermende factoren jeugdige	-.07	-.02	-.07	-.08	.03	-.05	-.04
Aantal beschermende factoren gezin en omgeving	.11	.01	.03	.05	.14*	.15*	.18**

* $p < .05$ (tweezijdig), ** $p < .000$ (tweezijdig)

De meeste correlaties zijn zwak tot zeer zwak. Als naar de variabele “onveilige uitkomst” gekeken wordt, is de samenhang zwak tot matig voor de zorgen en risicofactoren. Opvallend is dat het aantal beschermende factoren meestal niet samenhangt met de uitkomsten op T2, met uitzondering van de beschermende factoren in gezin en omgeving (sociale steun informeel en formeel netwerk). In tegenstelling tot de verwachtingen werd daar echter een positieve samenhang gevonden.

3.4 Resultaten voor de interne consistentie

Aanvullend is onderzocht in hoeverre scores op de LIRIK-items voorafgaand aan de hoofdconclusies samenhangen met de hoofdconclusies. Hiertoe zijn Spearman's rangcorrelaties berekend, waarbij voor de aanwezigheid van zorgen of een risico- of beschermende factor een score1 was gegeven en bij afwezigheid een

score 0, en conclusies waren gescoord met 1 tot 4 waarbij hogere scores oordelen weergeven van meer onveiligheid/hoger risico (zie 3.3.3 en 3.2).

Tabel 11. Spearman's rangcorrelaties voor de samenhang tussen scores op de LIRIK-items over zorgen, risicofactoren en beschermende factoren en conclusies over huidige veiligheid en risico's

	Veiligheidstaxatie	Risico dit moment	Risicotaxatie Risico bij voorziene veranderingen
1. Huidige veiligheidssituatie			
A. Directe veiligheid			
Bedreiging door ouder(s)/primaire opvoeder(s)	.30**	.28**	.21**
Ernstige kindermishandeling	.18*	.11	.11
Onvoldoende bescherming	.43**	.41**	.25**
Onvoldoende basiszorg	.40**	.35**	.25**
Opvoeder niet beschikbaar	.26**	.29**	.14
Ouder(s) belemmeren zicht op / toegang tot jeugdige?	.17*	.12	.10
Bedreiging door jeugdige zelf	.19**	.20**	.19*
Bedreiging door ander gezinslid/derde	.18*	.11	.19*
Zijn er vermoedens dat de jeugdige op dit moment fysiek in gevaar is?	.24**	.37**	.19*
B. Omgang ouder(s) - jeugdige			
Lichamelijk of psychisch geweld	.38**	.34**	.15
Verwaarlozing	.56**	.53**	.34**
Seksueel misbruik ^A	-	-	-
Getuige van huiselijk geweld	.29**	.30**	.33**
Bescherming en veiligheid	.63**	.59**	.34**
Basale verzorging	.30**	.37**	.25**
Emotionele warmte (ondersteuning)	.47**	.47**	.25**
Regels en grenzen	.48**	.43**	.18*
Stimulering	.43**	.40**	.13
Stabiliteit	.47**	.48**	.35**
Anders	.30	.29	.40*
Ouder geeft geen inzicht in eigen handelen	.22**	.23**	.22**
Onvoldoende informatie beschikbaar	.15*	.09	-.02
Zijn er aanwijzingen voor bedreigend handelen of nalaten van de ouder(s)?	.67**	.64**	.34**
C. Jeugdige(n)			
Psychosociaal functioneren	.25**	.35**	.11
Lichamelijke gezondheid	.11	.19**	.07
Vaardigheden en cognitieve ontwikkeling	.23**	.24**	.26**
Anders	.33*	.38**	.52**
Zijn er kindsignalen die wijzen op een onveilige opvoedingssituatie of kindermishandeling?	.66**	.57**	.26**
D. Risicofactoren			
<i>Risicofactoren bij de ouders</i>			
Eerder een kind mishandeld, verwaarloosd of misbruikt	.29**	.18*	.16

Ontoereikende pedagogische kennis en/of vaardigheden	.58**	.52**	.17
Problemen in de ouder-kind interactie	.31**	.31**	.25**
Bagatelliseren/ontkennen aangetoonde kindermishandeling	.48**	.46**	.36**
Negatieve houding t.a.v. het kind	.28**	.24**	.13
Psychiatrische problematiek	.27**	.30**	.12
Verslavingsproblematiek	.32**	.38**	.30**
Verstandelijke beperking	.14	.25**	.13
Fysieke beschikbaarheid	.12	.12	.06
Emotionele beschikbaarheid	-.06	.03	.02
Op jonge leeftijd (< 18 jaar) ouder geworden	.04	.13	.14
Zelf slachtoffer van kindermishandeling	.35**	.25**	.04
Eerder geweld gebruikt tegen personen	.27**	.25**	.17
Problematische partnerrelatie	.21**	.34**	.27**
Anders	.38**	.35*	.09
<i>Risicofactoren bij de jeugdige</i>			
Jong kind (< 5 jaar)	.05	.02	.09
Belaste voorgeschiedenis	.06	.09	.26**
(Ernstige) ziekte of handicap	.11	.16*	.06
Gedrags- en/of ontwikkelingsproblemen	.10	.09	.19*
Moeilijk temperament	.10	.12	.04
Ongewenst	.08	.05	.07
Anders	.26	.13	-.14
<i>Risicofactoren bij gezin en/of omgeving</i>			
Eenoudergezin, stiefgezin, groot gezin	.36**	.41**	.37**
Veel conflicten	.29**	.33**	.21*
Huiselijk geweld	.31**	.26**	.21*
Instabiel, ongeregeld leven	.24**	.39**	.22**
Materiële/financiële problemen	.33**	.41**	.27**
Ingrijpende levensgebeurtenissen	.44**	.32**	.20*
Sociaal isolement/conflict	.27**	.28**	.23**
Anders	.36**	.28*	.09
Zijn er risicofactoren voor kindermishandeling?	.74**	.60**	.40**
2. Risicotaxatie			
Aanvullende risicofactoren bij vermoedelijke of aangetoonde kindermishandeling			
(Vermoedelijke) pleger heeft direct toegang tot de jeugdige	.34**	.31**	.10
Derden hebben geen zicht op de jeugdige	.07	.10	-.01
Wat kan er gebeuren?			
Mogelijke risico's: Levensbedreigende situatie of direct fysiek gevaar	.50**	.58**	.42**
Mogelijke risico's: Ontstaan van kindermishandeling	.70**	.71**	.39**
Mogelijke risico's: Voortduren/herhaling van kindermishandeling	.70**	.70**	.43**
Mogelijke risico's: Anders	.54**	.65**	.41**
Verwachte gevolgen voor de jeugdige	.76**	.78**	.49**
Beschermende factoren			
Zijn er beschermende factoren die de risico's kunnen verminderen?	-.36**	-.37**	-.18*

Beschermende factoren bij de ouders

Ouder 1

Gevoel van competentie, draagkracht	-.14*	-.21**	-.07
Positief zelfbeeld	-.10	-.19**	-.09
Ondersteunende partner	-.10	-.15*	-.08
Kan eigen jeugdervaringen hanteren	-.09	-.12	.06
Positieve jeugdervaringen	-.12	-.16*	-.02
Kan steun vragen/profiteren van	.01	-.14*	-.07
Emotionele beschikbaarheid	-.20**	-.27**	-.02
Flexibiliteit	-.10	-.12	-.04
Bereid en in staat om te veranderen	-.01	-.11	.01

Ouder 2

Gevoel van competentie, draagkracht	-.24**	-.26**	-.21**
Positief zelfbeeld	-.25**	-.24**	-.20**
Ondersteunende partner	-.17*	-.28**	-.24**
Kan eigen jeugdervaringen hanteren	-.16*	-.16*	-.05
Positieve jeugdervaringen	-.16*	-.16*	-.16*
Kan steun vragen/profiteren van	-.24**	-.31**	-.15
Emotionele beschikbaarheid	-.21**	-.30**	-.20**
Flexibiliteit	-.24**	-.22**	-.16*

Beschermende factoren bij de jeugdige

Jeugdige 1

Sociaal vaardig	.05	-.02	.04
Positief zelfbeeld	.05	-.04	.04
Bovengemiddelde intelligentie	.06	-.06	-.19*
Aantrekkelijk uiterlijk	.15*	.05	.00
Goede relatie belangrijke volwassene(n)	-.03	-.07	-.15*
Egoveerkracht (stressresistentie)	.10	.01	-.01
Bereid en in staat om te veranderen	.12	-.05	-.16*

Jeugdige 2

Sociaal vaardig	-.12	-.18**	.04
Positief zelfbeeld	-.14	-.18**	-.02
Bovengemiddelde intelligentie	-.07	-.06	.11
Aantrekkelijk uiterlijk	-.14	-.14*	.05
Goede relatie belangrijke volwassene(n)	-.15*	-.18**	-.04
Egoveerkracht (stressresistentie)	-.10	-.13	-.04

Beschermende factoren bij gezin en omgeving

Steun informeel netwerk	.06	-.01	-.09
Steun formeel netwerk	.15*	.06	.15

Noot. ^AEen correlatie berekenen tussen “seksueel misbruik” en de conclusie over de huidige veiligheid en de risicotaxatie was niet mogelijk, omdat in geen van de zaken de conclusie was getrokken dat er sprake was van seksueel misbruik.

* $p < .05$ (tweezijdig), ** $p < .000$ (tweezijdig)

Tabel 11 laat zien dat het varieert hoe sterk scores op de items van de LIRIK samenhangen met de conclusies over de huidige veiligheid en risico's. De tussenconclusies of er sprake is van bedreigend handelen of nalaten van ouders, of er kindsignalen zijn en of er risicofactoren voor kindermishandeling zijn, hingen sterk tot zeer sterk samen met de drie hoofdconclusies. Scores op de items over de directe veiligheid, de omgang tussen ouders en jeugdige, de kindsignalen en de meeste risicofactoren bij de ouders en bij gezin en omgeving hingen matig sterk samen met de hoofdconclusies. Scores voor de risicofactoren bij de jeugdige en de

beschermende factoren (met uitzondering van de beschermende factoren bij ouder 2) hingen niet of nauwelijks samen met de hoofdconclusies. De samenhang tussen scores op de items en de risico's bij voorziene veranderingen was lager dan de samenhang met de conclusies over de huidige veiligheid en risico's op dit moment.

Ook is gekeken in hoeverre er een samenhang is tussen het aantal zorgen, risico- of beschermende factoren enerzijds en de getrokken conclusies over de huidige veiligheid en risico's anderzijds. Daarvoor zijn somscores berekend voor het aantal zorgen over de omgang tussen ouders en jeugdige, het aantal zorgen over het functioneren van de jeugdige en het aantal risico- en beschermende factoren dat medewerkers aangevinkt hebben. In Tabel 12 zijn de Spearman's rangcorrelaties weergegeven.

Tabel 12. Spearman's rangcorrelaties voor de samenhang tussen het aantal zorgen/risicofactoren of beschermende factoren enerzijds en conclusies over huidige veiligheid en risico's anderzijds

	Veiligheidstaxatie	Risicotaxatie	
		Risico op dit moment	Risico bij voorziene veranderingen
Aantal zorgen over omgang ouder – jeugdige	.69**	.66**	.40**
Aantal zorgen over functioneren jeugdige	.31**	.40**	.31**
Aantal risicofactoren ouders	.43**	.46**	.31**
Aantal risicofactoren jeugdige	.18**	.19**	.26**
Aantal risicofactoren gezin en omgeving	.52**	.53**	.40**
Aantal beschermende factoren ouders	-.15*	-.31**	-.11
- Ouder 1	-.12	-.25**	-.06
- Ouder 2	-.22**	-.31**	-.20**
Aantal beschermende factoren jeugdige	.12	-.05	-.11
Aantal beschermende factoren gezin en omgeving	.12	.03	.04

* $p < .05$ (tweezijdig), ** $p < .000$ (tweezijdig)

Tabel 12 laat zien dat er een zwakke tot sterke positieve relatie is tussen het aantal zorgen over de omgang, het functioneren van de jeugdige, risicofactoren bij de ouders en in gezin en omgeving enerzijds en anderzijds de veiligheids- en risicotaxatie. Er is geen of een zeer zwakke relatie met het aantal risicofactoren bij de jeugdige en met de beschermende factoren.

3.5 Discussie

De eerste hoofdvraag van dit deelonderzoek was in hoeverre oordelen over de huidige veiligheid, risico op kindermishandeling op dit moment en risico op kindermishandeling bij voorziene veranderingen gemaakt met de LIRIK verschillende uitkomsten die onveiligheid indiceren voorspellen (AMK-melding, raadsmelding, ondertoezichtstelling, uithuisplaatsing, interventie gericht op veiligheid en crisis(interventie)). De tweede hoofdvraag was of oordelen gemaakt met de LIRIK deze uitkomsten beter voorspellen dan oordelen waarbij geen gebruik is gemaakt van de LIRIK (ongestructureerde oordelen).

De samenhang tussen de oordelen over de huidige veiligheid en de risico's voor de nabije toekomst enerzijds en de uitkomsten op T2 anderzijds bleek beperkt. Bij de meeste uitkomstmaten bleek geen consequente toename van uitkomsten die onveiligheid indiceren bij oordelen van meer onveiligheid of hoger

risico. Wel verhielden deze uitkomsten zich per veiligheids- of risiconiveau redelijk logisch ten opzichte van de base rates, dus bij oordelen van veiligheid/laag risico waren er verhoudingsgewijs minder uitkomsten van onveiligheid dan overall, bij oordelen van onveiligheid/hoog risico meer. Deze resultaten komen overeen met uitkomsten van eerder onderzoek naar andere consensus-based instrumenten (Baird & Wagner, 2000; D'Andrade, Benton, & Austin, 2005).

We vonden geen duidelijke verschillen tussen de LIRIK-groep en de controlegroep, oordelen gemaakt met de LIRIK voorspelden niet beter dan het ongestructureerd professioneel oordeel. Op het terrein van risicotaxatie bij kindermishandeling zijn nauwelijks vergelijkingen gemaakt tussen het gebruik van een instrument en het ongestructureerd professioneel oordeel. Een onderzoek van Baumann, Law, Sheets, Reid, en Graham (2005), waarin zij actuariële instrumenten vergelijken met een ongestructureerd oordeel, laat zien dat ook een actuariële instrument niet altijd beter de uitkomsten voorspelt dan het ongestructureerd oordeel. Tijdens het dossieronderzoek is het de onderzoekers opgevallen dat de veiligheids- en risicotaxaties (al dan niet met LIRIK) vaak los leken te staan van het verdere hulpverleningsproces. Conclusies over veiligheid en risico's waren nauwelijks in de cliëntdossiers zelf terug te vinden en nog minder kwam het voor dat conclusies hierover beschreven werden in combinatie met acties die ondernomen waren of zouden worden. De conclusies die medewerkers trekken lijken dus niet direct te leiden tot bepaalde acties of dit is op zijn minst niet duidelijk terug te vinden in de dossiers. Dit kan verklaren waarom er zo weinig samenhang is met de uitkomsten op T2.

De hulpverleners die de beoordelingen met de LIRIK gaven waren niet steeds dezelfde als de uitvoerende hulpverlener (het ongestructureerde oordeel was altijd door de uitvoerende hulpverlener uitgevoerd). Dit kan ook verklaren waarom in dit onderzoek oordelen over onveiligheid of risico's niet altijd leidden tot een bepaalde actie. Dagleish en anderen hebben laten zien dat er de factoren die het oordeel over wat er in een situatie aan de hand is beïnvloeden, verschillen van de factoren die maken dat mensen actie ondernemen (zie Baumann, Dagleish, Fluke, & Kern, 2011). Mensen kunnen het eens zijn over de mate van onveiligheid van een jeugdige, maar toch verschillen in de actie die zij ondernemen. Als een conclusie over de veiligheid en risico's van de ene persoon naar de andere persoon wordt overgedragen, kan dit tot een verschil van mening over de conclusie, maar ook over de benodigde actie leiden. Onderzoek heeft ook steeds weer laten zien dat beslissingen over hulp van andere factoren afhankelijk zijn dan casusspecifieke factoren, bijvoorbeeld persoonlijke overtuigingen en kennis over beschikbaar hulpaanbod (Berben, 2000; Lekkerkerker, Bartelink, & Eijgenraam, 2011; Ten Berge, 1998). Dat kan van invloed zijn op de predictieve validiteit.

De vraag naar de predictieve validiteit van een risicotaxatie-instrument voor kindermishandeling is lastig te beantwoorden, omdat conclusies over onveiligheid en mogelijke risico's aanleiding voor werkers zijn om actie te ondernemen om de veiligheid te verbeteren. Dat kan zijn door interventies die als uitkomstmaat in dit onderzoek zijn gehanteerd, maar kan ook op andere manieren, bijvoorbeeld tijdens de contacten die er met een gezin zijn (al dan niet in de vorm van hulpverlening). Er kunnen ook dingen veranderd zijn in de gezinssituatie waardoor de veiligheid groter of juist kleiner is geworden. De gemeten predictieve validiteit kan door deze factoren lager uitvallen. We hebben geprobeerd om deze factoren te meten in het dossieronderzoek. Er bleek hierover echter weinig informatie in de dossiers te vinden, die bovendien niet betrouwbaar te coderen was.

De derde onderzoeksvraag was in hoeverre de LIRIK-items de uitkomstmaten voorspelden. Wat betreft de samenhang tussen de LIRIK-items en de uitkomsten op T2 is gebleken dat maar een beperkt aantal items met de uitkomsten samenhangt en dat die samenhang meestal maar zwak tot matig was. Wel is gebleken dat er een zwakke tot matige samenhang was tussen het aantal zorgen/ risicofactoren en de meeste uitkomstmaten.

Dat de afzonderlijke items slechts zwak tot matig de uitkomsten voorspellen is logisch aangezien het bij kindermishandeling om complexe problematiek en een samenspel tussen verschillende factoren gaat die uiteindelijk bijdragen aan onveiligheid en toekomstige risico's en daarmee aan de beslissing over de verdere aanpak (Ten Berge et al., 2012). Eerder onderzoek bevestigt dat casusspecifieke factoren vaak maar een kleine rol spelen in de beslissing en dat andere factoren, bijvoorbeeld persoonlijke overtuigingen en kennis over beschikbaar hulpaanbod, een grotere rol spelen in de beslissing (Berben, 2000; Lekkerkerker, Bartelink, & Eijgenraam, 2011; Ten Berge, 1998). Welke LIRIK-items samenhangen met de uitkomsten hing nogal van de uitkomstmaat af. Veel items bleken wel met één uitkomstmaat samen te hangen, maar niet met meerdere uitkomstmaten. Slechts een klein aantal items bleek met meerdere uitkomstmaten samen te hangen. Een mogelijke verklaring hiervoor is dat de uitkomstmaten niet exact een objectieve bepaling van kindermishandeling vormen, maar slechts een benadering ervan. Daardoor kunnen het ook verschillende factoren zijn die het vóórkomen van verschillende uitkomsten bepalen.

Opvallend is dat de samenhang van de uitkomsten op T2 met veel factoren in de omgang tussen ouders en jeugdige en risicofactoren bij de ouders ontbreekt, terwijl uit wetenschappelijke literatuur wel gebleken is dat dit belangrijke voorspellers voor kindermishandeling zijn (Alink et al., 2011; Mulder, 2014; Stith, Liu, Davies, Boykin, Alder & Dees, 2009; Ten Berge et al., 2012). Mogelijk komt dit door de lage overeenstemming over de items (zie Hoofdstuk 2). Wanneer items niet betrouwbaar worden ingevuld, kan de predictieve validiteit ervan nooit hoog zijn. Onder andere bleek dat professionals de items “fysieke beschikbaarheid” en “emotionele beschikbaarheid” verschillend scoren. Sommige scoren ze als ouders juist wel fysiek en emotioneel beschikbaar zijn, anderen juist als ze niet beschikbaar zijn (dit laatste is correct). Een andere mogelijke verklaring voor het ontbreken van een samenhang van de uitkomsten met factoren in de omgang tussen ouders en jeugdige en risicofactoren bij de jeugdige is dat professionals onvoldoende kennis hebben hoe deze factoren de opvoeding kunnen beïnvloeden en daardoor onvoldoende hebben meegewogen in hun beslissingen.

De laatste onderzoeksvraag ging in op de interne samenhang van de LIRIK. De samenhang van de scores op LIRIK-items met de oordelen over de huidige veiligheid en risico's voor de nabije toekomst en bij voorziene veranderingen bleek meestal groter te zijn dan de samenhang met de uitkomsten, hoewel er veel variatie was. Ook bleek dat naarmate professionals meer zorgen en risicofactoren zagen, zij ook vaker oordeelden dat er vaker sprake was van kindermishandeling of een risicovolle situatie. De tussenconclusies of er sprake is van bedreigend handelen of nalaten van ouders, of er kindsignalen zijn en of er risicofactoren voor kindermishandeling zijn, hingen het sterkst samen met de drie hoofdconclusies. Dit is goed nieuws, omdat het om tussenconclusies gaat die bijdragen aan het uiteindelijke oordeel over de huidige veiligheid en de risico's voor de nabije toekomst. Dat LIRIK lijkt dus redelijk consistent ingevuld te worden. Dat scores op de afzonderlijke items minder sterk samenhangen met de conclusies is te verwachten, aangezien het de combinatie van factoren is die uiteindelijk leidt tot onveiligheid en risico's voor de jeugdige (Ten Berge et al.,

2012). De scores op de risicofactoren bij de jeugdige en de beschermende factoren hingen nauwelijks samen met de hoofdconclusies. Uit literatuur blijkt echter dat het met name factoren in de omgang tussen ouders en jeugdige en factoren in het functioneren van de ouders zijn die bepalen of kindermishandeling zich voordoet (Alink et al., 2011; Mulder, 2014; Stith, Liu, Davies, Boykin, Alder & Dees, 2009; Ten Berge et al., 2012). Het resultaat dat risicofactoren bij de jeugdige minder sterk samenhangen met de oordelen over de huidige veiligheid en risico's sluit dus aan bij de huidige beschikbare literatuur. Wat betreft de beschermende factoren kan het zijn dat die in situaties die zeer onveilig zijn nauwelijks een bufferfunctie hebben. In de literatuur over beschermende factoren bij kindermishandeling bestaat onduidelijkheid over welke factoren mogelijk beschermend zijn en hoe de aanwezigheid van bepaalde factoren beschermend kan werken (Development Service Group, Washington, 2013; Harper Browne, 2014). Het kan ook zijn dat professionals de risicofactoren bij de jeugdige en de beschermende factoren onvoldoende meewegen in hun oordeel over de veiligheid en risico's.

3.5.1 Beperkingen

Dit onderzoek kent een aantal beperkingen, naast de eerder genoemde beperking dat tussentijdse interventies anders dan de uitkomstmaten en veranderingen in gezinssituaties de uitkomsten kunnen beïnvloeden. Een andere beperking is dat medewerkers die met het ongestructureerd professioneel oordeel de veiligheids- en risicotaxatie uitvoerden, bekend waren en ook werkten met de LIRIK, omdat de LIRIK binnen de organisaties al geïmplementeerd was. Mogelijk dat hierdoor de gevonden verschillen tussen groepen kleiner waren dan zij zouden zijn als een controlegroep geen kennis van en ervaring met de LIRIK zou hebben. In zowel de LIRIK-groep als de controlegroep was de predictieve validiteit echter beperkt. De dossiers bevatten vaak weinig informatie, ook wanneer je meer specifieke informatie zou verwachten gezien wat er wel instaat. Hierdoor was het lastig om naast de harde uitkomstmaten zoals in de resultaten gerapporteerd, ook meer zachte uitkomstmaten voor veiligheid te scoren op basis van de cliëntdossiers. De termijn tussen de beoordelingen op T1 en de uitkomsten op T2 was beperkt tot zes maanden en soms korter, doordat het onderzoek in een beperkte tijd moest plaatsvinden. Het is mogelijk dat een aantal uitkomsten zich pas later hebben voorgedaan, waardoor de samenhangen van oordelen met een uitkomst niet gevonden kon worden in deze beperkte tijdspanne. Met het bekijken van uitkomsten bijvoorbeeld na een jaar zouden mogelijk meer uitkomsten geregistreerd worden.

Er bleken enkele verschillen in de beoordeelde situaties van jeugdigen tussen de LIRIK-groep en de controlegroep. In de controlegroep bleek het aantal jeugdigen waarvan medewerkers meenden dat er sprake was van een levensbedreigende situatie of direct fysiek gevaar hoger te zijn. Ook waren jeugdigen in de controlegroep vaker uit huis geplaatst dan in de LIRIK-groep. Het lijkt er daarmee op dat de problematiek in de controlegroep ernstiger was dan in de LIRIK-groep. Wanneer de situatie ernstiger is, zijn mensen het over het algemeen vaker eens dat een interventie moet plaatsvinden en zullen er ook meer interventies plaatsvinden. Dit kan verklaren waarom de samenhang tussen de oordelen en de uitkomsten op T2 in de controlegroep hoger lagen dan in de LIRIK-groep.

Gezien het zeer kleine aantal zaken in de steekproef waar geconcludeerd werd dat er sprake was van een levensbedreigende situatie of direct fysiek gevaar, is het niet goed mogelijk om betrouwbare conclusies te trekken over de validiteit van de LIRIK voor deze categorie.

4 Conclusies, discussie en aanbevelingen

Cora Bartelink, Leontien de Kwaadsteniet, Ingrid ten Berge, Mariëtte Blik, Mariëtte van Brandenburg, Willemijn van Gastel, Mariëlle Hornstra, Carolien Konijn & Cilia Witteman

4.1 Inleiding

Om kinderen en ouders verantwoorde zorg te kunnen verlenen is het noodzakelijk een goede inschatting te maken van de huidige veiligheidssituatie van de jeugdige, factoren die daar in positieve of negatieve zin aan bijdragen en de risico's voor de jeugdige in de toekomst. De LIRIK is een checklist gericht op gestructureerde oordeelsvorming en ontwikkeld om hulpverleners bij deze inschatting te ondersteunen.

In dit onderzoek is nagegaan in welke mate oordelen gevormd met de LIRIK betrouwbaar en valide zijn. De onderzoeksvragen waren:

1. Worden met behulp van de LIRIK betrouwbare en valide oordelen gevormd over de veiligheid van kinderen?
2. Is het oordeel met behulp van de LIRIK meer betrouwbaar en valide dan het ongestructureerde professionele oordeel van de hulpverlener?

Het onderzoek bestond uit twee deelonderzoeken: een vignetonderzoek en een dossieronderzoek. In het vignetonderzoek stond de vraag naar de interbeoordelaarsovereenstemming centraal. Om de interbeoordelaarsovereenstemming vast te stellen beoordeelden hulpverleners een aantal vignetten op de huidige veiligheid en toekomstige risico's voor de jeugdige. Een deel van de hulpverleners deed dit met behulp van de LIRIK, een ander deel zonder een instrument te gebruiken. We vergeleken in welke mate hulpverleners het eens waren over de conclusies over de huidige veiligheid en risico's.

In het dossieronderzoek stond de vraag naar de predictieve validiteit centraal. Om de predictieve validiteit vast te stellen voerden hulpverleners bij gezinnen in hun caseload een veiligheids- en risicotaxatie uit. Deels gebeurde dit met de LIRIK, deels zonder de LIRIK. Na een half jaar werd een dossieronderzoek uitgevoerd om na te gaan in hoeverre er AMK-meldingen, Raadsmeldingen, ondertoezichtstellingen, uithuisplaatsingen, crises en interventies gericht op veiligheid waren geweest. We hebben onderzocht in hoeverre er samenhang was tussen de conclusies over de veiligheid en risico's enerzijds en de uitkomsten na een half jaar anderzijds.

4.2 Resultaten samengevat

De resultaten van de twee deelonderzoeken worden in aparte subparagrafen besproken.

4.2.1 **Betrouwbaarheid: Interbeoordelaarsovereenstemming**

In het vignetonderzoek bleek de overeenstemming over de oordelen over de huidige veiligheid en de risico's voor de toekomst onvoldoende, wanneer de LIRIK werd gebruikt. Professionals waren het het minst eens over de risico's in de nabije toekomst bij voorziene veranderingen. Verschillen in overeenstemming tussen de professionals die oordeelden met de LIRIK en zonder de LIRIK waren niet groot en wezen op iets hogere

overeenstemming over directe veiligheid en iets lagere overeenstemming over risico's wanneer de LIRIK was gebruikt. Ook de overeenstemming voor de meeste items van de LIRIK die voorafgaan aan de conclusies over de huidige veiligheid en risico's voor de nabije toekomst was laag. De professionals waren wel redelijk positief over de bruikbaarheid van de LIRIK voor het vormen van oordelen over veiligheid en risico's.

4.2.2 Predictieve validiteit

Uit het dossieronderzoek bleken zowel ongestructureerde oordelen als oordelen gemaakt met de LIRIK slechts heel beperkt de veiligheidsuitkomsten na een half jaar te voorspellen. De samenhang tussen oordelen en uitkomsten was zwak en varieerde voor de verschillende uitkomsten en tussen de groepen (LIRIK-groep en controlegroep). Slechts enkele items van de LIRIK die voorafgaan aan de conclusies over veiligheid en risico's bleken enigszins samen te hangen met de uitkomsten een half jaar later.

We vonden een duidelijkere samenhang tussen de LIRIK-items en de oordelen over de huidige veiligheid en risico's, wat erop wijst dat professionals hun conclusies wel baseerden op daaraan voorafgaande items.

4.3 Verklaringen voor bevindingen

Lage overeenstemming op itemniveau. Uit het vignetonderzoek bleek dat de overeenstemming op itemniveau laag was. Zelfs bij concrete items als de risicofactor "jong kind" bleek soms dat professionals het niet met elkaar eens waren. Overeenstemming op itemniveau lijkt noodzakelijk om overeenstemming over conclusies en daaruit volgende beslissingen over ingrijpen te bereiken, zeker omdat in het dossieronderzoek is gevonden dat de conclusies die professionals trekken samenhangen met de scores die zij op voorafgaande items geven.

Lage overeenstemming over conclusies. Een lage overeenstemming over oordelen over veiligheid en risico's tussen professionals impliceert een beperkte validiteit van deze oordelen. Wanneer hulpverleners verschillend oordelen over dezelfde situatie, kunnen deze oordelen de uitkomsten niet goed voorspellen. Een gebrek aan overeenstemming kan leiden tot meer vals negatieven (jeugdigen die mishandeld of verwaarloosd worden of een hoog risico daarop hebben die niet opgemerkt worden) en meer vals positieven (valse beschuldigingen van mishandeling of verwaarlozing) (Camasso & Jagannathan, 2000; 2013). Als niet eerst de betrouwbaarheid van veiligheidsoordelen wordt verbeterd, zal nooit een goede validiteit voor deze oordelen gerealiseerd kunnen worden.

Ruimte voor individuele oordelen. We vonden een lage overeenstemming, zowel over de oordelen over de huidige veiligheid en risico's voor de nabije toekomst als voor de items die hieraan voorafgaan. Er ontbreken duidelijke criteria voor het scoren van specifieke LIRIK-items en er is voor gekozen geen precieze criteria te geven voor het aantal zorgen of risicofactoren dat ertoe moet leiden dat professionals een bepaalde conclusie moeten trekken over of een situatie onveilig of risicovol is. Ook is er geen weging van de verschillende factoren aangegeven. Zo is er ruimte voor het professionele oordeel in specifieke situaties. Deze ruimte impliceert ook dat individuele professionals ruimte hebben om subjectieve afwegingen te maken en eigen normen te hanteren.

Ander onderzoek laat zien dat persoonlijke ervaring, kennis en beslisdrempels grote invloed op oordeelsvorming en besluitvorming hebben (Berben, 2000; De Kwaadsteniet et al., 2010; Munro, 1999).

Hoewel de verwachting was dat door de LIRIK, door relevante aandachtspunten aan te dragen, deze oordeelsvorming minder afhankelijk zou zijn van de individuele professional, hebben we daar in dit onderzoek geen bevestiging voor gevonden.

Complexe problematiek. Kindermishandeling is een complex en dynamisch probleem. Gezinnen waar kindermishandeling speelt, hebben vaak te maken met allerlei met elkaar samenhangende problemen. Onderzoek laat zien dat er allerlei risicofactoren voor kindermishandeling zijn, , maar er is geen één op één relatie tussen de aanwezigheid van risicofactoren en het zich voordoen van kindermishandeling. Wat nu precies maakt dat in het ene gezin waarin een bepaalde risicofactor voorkomt kinderen wel mishandeld worden, maar in een ander gezin niet, is onbekend, waardoor kindermishandeling in individuele gezinnen lastig te voorspellen is (Gambrell & Shlonsky, 2000; Hermanns, Öry, & Schrijvers, 2005; Rycus & Hughes, 2003). Dit gebrek aan wetenschappelijke kennis maakt het moeilijk om betrouwbare oordelen te vormen, met als gevolg dat overeenstemming over oordelen over veiligheid en toekomstige risico's laag is. Daar komt bij dat er geen precieze definities en criteria voor kindermishandeling zijn. Er is geen eenduidig antwoord op de vraag wanneer je nu van kindermishandeling moet spreken en professionals hebben daar verschillende opvattingen over. Professionals die moeten oordelen over de veiligheid en risico's op kindermishandeling in een gezin, hebben dus te maken met een grote mate van onzekerheid in de oordeelsvorming. Daarnaast beschikken zij vaak niet over volledige en objectieve informatie die nodig is om tot een valide oordeel te komen. Bovendien zijn interventies die geschikt zijn voor gezinnen waar mishandeling of verwaarlozing speelt maar beperkt beschikbaar en van vele is onduidelijk of ze effectief zijn (Davidson-Arad, 2005, 2010; Doyle, 2007; Pinto & Maia, 2013; Rubin, 2012). Een instrument dat weinig sturing geeft aan de weging en oordeelsvorming en aan daarop volgende besluitvorming zal dan ook onvermijdelijk tot verschillen in conclusies leiden en tot verschillende keuzes voor interventies.

Implementatie van de LIRIK. Uit observaties in het dossieronderzoek bleek dat items van de LIRIK soms verkeerd werden ingevuld. Een voorbeeld hiervan was de risicofactor “fysieke beschikbaarheid”. Uit toelichting van enkele medewerkers bleek dat ze die aankruisten op het moment dat een ouder wel fysiek beschikbaar was, terwijl anderen dit item aankruisten als een ouder niet beschikbaar was. Een ander voorbeeld is dat medewerkers die in Deel 1A van de LIRIK aangaven dat er sprake was van direct gevaar of een levensbedreigende situatie voor de jeugdige, dit niet altijd bij de conclusies over de huidige veiligheid aangaven. Blijkbaar was hen onvoldoende helder dat deze items feitelijk hetzelfde waren en gingen zij inconsistent met de LIRIK om. Uit het vignetonderzoek bleek dat niet iedereen die de LIRIK gebruikte, ook getraind was in het gebruik ervan. Dat geeft aanleiding voor zorgen over de implementatie van de LIRIK. Ander onderzoek laat zien dat een goede implementatie van instrumenten en methoden een belangrijke voorwaarde voor succes ervan is (Daamen, 2013; Prins, 2011).

4.4 Beperkingen van het onderzoek

Dit onderzoek geeft een beeld van hoe de betrouwbaarheid en validiteit van veiligheids- en risicotaxaties uitgevoerd met de LIRIK in de dagelijkse praktijk zijn. Er zijn enkele kanttekeningen bij het onderzoek te plaatsen, doordat het in de praktijk is uitgevoerd.

Voor beide deelonderzoeken geldt dat de controlegroep (ongestructureerd professioneel oordeel, zonder gebruikmaking van de LIRIK) geen optimale controlegroep was. Medewerkers in de controlegroep waren bekend met en werkten meestal al met de LIRIK. Hun oordelen gemaakt zonder de LIRIK kunnen dus beïnvloed zijn door hun kennis over de LIRIK. In zowel de LIRIK-groep als de controlegroep waren de betrouwbaarheid en validiteit van de oordelen echter onvoldoende.

In het eerste deelonderzoek is gebruik gemaakt van gevalsbeschrijvingen, niet van echte casuïstiek (hoewel wel daarop gebaseerd), zodat onzeker is hoe de overeenstemming zou zijn bij echte casuïstiek. Gezien de consistente bevindingen ten aanzien van overeenstemming op dit gebied, is het echter onwaarschijnlijk dat deze veel beter zou zijn wanneer dezelfde beoordelaars dezelfde casus beoordelen in de praktijk.

Het tweede deelonderzoek was voor de dataverzameling van de uitkomstmaten op het tweede meetmoment afhankelijk van de kwaliteit van de dossiervorming. Geconstateerd werd dat professionals niet altijd alle gegevens over jeugdigen en gezinnen zorgvuldig vastlegden. Hierdoor kunnen bepaalde uitkomsten zich wel hebben voorgedaan, waar we in dit onderzoek niet van op de hoogte waren. Daarnaast kunnen andere instellingen meldingen hebben, waarvan de instellingen waar we het onderzoek uitvoerden niet op de hoogte waren. We beschikten niet over de mogelijkheid om gegevens over AMK-meldingen, raads meldingen, ondertoezichtstellingen, uithuisplaatsingen en andere interventies rechtstreeks bij Bureau Jeugdzorg of andere instellingen op te vragen.

Het tweede deelonderzoek liet zien dat de situaties die de controlegroep en de LIRIK-groep beoordeelden waarschijnlijk niet helemaal vergelijkbaar waren wat betreft ernst van de problematiek. In de controlegroep oordeelden professionals vaker dat er sprake was van een levensbedreigende situatie of direct fysiek gevaar en was vaker sprake van een uithuisplaatsing. Mogelijk was in de zaken van de controlegroep dus sprake van ernstiger problematiek dan in de zaken waarvoor de LIRIK is ingevuld. Verondersteld kan worden dat het gemakkelijker is om een hogere validiteit te bereiken wanneer problematiek in gezinnen meer divers is qua ernst, omdat samenhang tussen oordelen en uitkomsten beter tot uitdrukking kan komen bij meer spreiding op deze variabelen. In beide groepen werd relatief vaak geoordeeld dat een situatie veilig is en werden relatief weinig uitkomsten die onveiligheid indiceren waargenomen (wat ook in de dagelijkse praktijk het geval is). Noodgedwongen – zoals in elk onderzoek naar risicotaxatie bij kindermishandeling – moest in het onderzoeken van de predictieve validiteit gekozen worden voor uitkomstmaten waaruit onveiligheid afgeleid kan worden. Dat er een AMK-melding is gedaan hoeft nog niet te betekenen dat er ook feitelijk kindermishandeling plaatsvindt. Een jeugdige kan ook om andere redenen dan kindermishandeling uit huis geplaatst worden.

4.5 Aanbevelingen

Het geheel overziend concluderen we dat de betrouwbaarheid en validiteit van oordelen waarbij gebruik is gemaakt van de LIRIK onvoldoende zijn en niet beter dan voor ongestructureerde oordelen. Omdat het ongestructureerde oordeel echter ook in belangrijke mate subjectief gebleken, wat de aanleiding was voor het ontwikkelen van de LIRIK, is het niet zomaar aan te bevelen de LIRIK niet langer te gebruiken. Voor andere risicotaxatie-instrumenten is ook niet aangetoond dat zij leiden tot meer objectieve oordelen.

Professionals hebben ook redenen aangedragen waarom het zinvol zou kunnen zijn om de LIRIK te blijven gebruiken. De LIRIK zou ervoor kunnen zorgen dat professionals meer gestructureerd werken en aandacht schenken aan factoren die relevant zijn bij kindermishandeling (zie ook Bartelink, De Kwaadsteniet & Ten

Berge, 2011). Daarnaast ervaren professionals het zelf ook als zinvol, omdat het hen helpt om met een brede blik naar de veiligheid van jeugdigen te kijken.

Wanneer de LIRIK wordt gebruikt, blijven oordelen over veiligheid en risico's blijven dus in belangrijke mate subjectief. Daarvan moeten professionals zich bewust zijn, ook wanneer zij meerwaarde in het gebruik van de LIRIK zien. Het is belangrijk dat zij helder over deze subjectiviteit communiceren met ouders en jeugdigen. We doen aanbevelingen voor (onderzoek naar) wat naast het gebruik van risicotaxatie-instrumenten gedaan kan worden om oordeelsvorming over veiligheid en mogelijke risico's in de praktijk te verbeteren en adviseren we om na te gaan hoe de LIRIK zelf verder verbeterd kan worden.

4.5.1 Verbeteringen in de praktijk

Het is belangrijk dat professionals in de praktijk zich realiseren dat de LIRIK geen objectief meetinstrument is, maar dat ze het wel als ondersteuning kunnen gebruiken bij het systematisch kijken naar veiligheid en risico's. Wanneer professionals zich niet bewust zijn van de grote mate van subjectiviteit van hun oordelen, die er ook is wanneer zij deze op gestructureerde wijze zoals met de LIRIK vormen, zal dit tot overmatig vertrouwen leiden in hun oordelen, en mogelijk ook in de daarop volgende besluiten. Professionals moeten er dus steeds rekening mee houden dat zij een onveilige situatie mogelijk onterecht als veilig inschatten, en dat de inschatting dat een situatie onveilig is onterecht kan zijn, en steeds alert blijven op signalen voor (on)veiligheid. De LIRIK moet gezien worden als hulpmiddel in de bredere context van besluitvorming. Er moet verder gekeken worden naar wat professionals helpt om een zo zorgvuldig, objectief mogelijk veiligheids- en risicotaxatie te maken.

Omdat de betrouwbaarheid onvoldoende bleek, is het aan te bevelen om steeds samen met collega's en/of een gedragswetenschapper de veiligheid en risico's te beoordelen. Intercollegiaal en interdisciplinair overleg kan inzicht geven in de verschillende inschattingen die professionals maken en leidt wellicht tot meer eenduidigheid in de oordeelsvorming. Belangrijk aandachtspunt daarbij is dit vooral ook samen met collega's te doen die hierin een verschillend standpunt innemen, zodat verschillende perspectieven op veiligheid en risico's goed naar voren kunnen komen.

Het is ook belangrijk dat professionals de onzekerheid waarmee en onderbouwing op basis waarvan zij hun conclusies trekken helder kunnen uitleggen aan gezinnen, maar waar mogelijk ook samen de veiligheid in het gezin bespreken. Hulpverleners moeten openstaan voor de meningen en informatie van ouders of jeugdige. Dit levert een breder en mogelijk meer betrouwbaar beeld van de veiligheid en mogelijke risico's op dan wanneer zij als professionals dit zonder hun inbreng doen. Daarnaast is het belangrijk dat professionals feedback krijgen op hun oordeelsvorming. Door regelmatig gericht te evalueren en te reflecteren op de eigen oordeelsvorming en de uitkomsten bij jeugdigen kan een leerproces ontstaan. Daarbij gaat het om de vraag hoe een professional de veiligheid en risico's beoordeeld heeft en wat er na verloop van een half jaar of een jaar met een jeugdige is gebeurd. Is een jeugdige bij een Veilig Thuis gemeld? Is er een ondertoezichtstelling uitgesproken of is een jeugdige uit huis geplaatst, terwijl de professional vond dat het veilig was? Of is na een Veilig Thuis melding gebleken dat een kind veilig was? Feedback kan helpen om te onderzoeken welke factoren invloed hebben op uitkomsten en hoe factoren bijdragen aan de oordeelsvorming en de beslissingen. Op basis daarvan kunnen professionals hun oordelen in toekomstige zaken aanpassen. Noors onderzoek laat zien dat feedback ertoe kan leiden dat professionals vaker juiste oordelen vormen en beslissingen nemen (Finnila, Santtila, Mattila, & Niemi, 2012).

Gezien de lage samenhang tussen conclusies en latere uitkomsten en de observatie dat de conclusies over de veiligheids- en risicotaxatie nauwelijks in de cliëntdossiers zijn terug te vinden (de meeste ingevulde LIRIKs zaten in een ander digitaal systeem dan het cliëntdossier), is het noodzakelijk dat professionals veel meer aandacht schenken aan wat hun conclusies over de veiligheid en risico's betekenen voor hun handelen, dit kunnen uitleggen en ook expliciet in de cliëntdossiers vermelden. Wat gaan ze doen en waarom gezien hun conclusies over de veiligheid van de jeugdige en het risico op kindermishandeling?

4.5.2 Verbetering van de LIRIK

Ten eerste is het van belang de betrouwbaarheid van de LIRIK-items te vergroten. Zeker voor de items die voorafgaan aan de tussenconclusies en conclusies is het noodzakelijk dat hulpverleners duidelijker weten wanneer zij die als zorg- of risicofactor moeten zien. Dit kan door in de digitale versie explicieter aan te geven wat de betekenis van de items is en misschien ook door items in de vragende vorm te formuleren, met duidelijke criteria.

De betrouwbaarheid en validiteit van oordelen gemaakt met actuariële risicotaxatie-instrumenten blijken steeds wat beter dan die van oordelen gemaakt met op consensus gebaseerde instrumenten. Sommige LIRIK-items blijken niet samen te hangen met de conclusies over de veiligheid en risico's en ook niet met de uitkomsten op T2. De literatuur over risico- en beschermende factoren bij kindermishandeling laat over sommige van deze aspecten ook niet altijd een duidelijk beeld zien. Mogelijk kunnen sommige items uit de LIRIK verwijderd worden. Hiervoor zou een vergelijking van verschillende nationale en internationale onderzoeken gedaan moeten worden om na te gaan welke factoren wel en niet bijdragen aan de voorspelling van onveiligheid en risico's. Verder zou dan het toevoegen van een expliciet wegingsalgoritme voor geïdentificeerde factoren de overeenstemming over veiligheids- en risico-oordelen mogelijk vergroten, en (daarmee) mogelijk ook de predictieve validiteit.

Overwogen kan worden of de LIRIK in een andere vorm beter bruikbaar is en minder de schijn van een objectief instrument wekt. De vorm van een checklist zorgt er mogelijk voor dat professionals verwachten een objectief antwoord te krijgen op de vraag hoe veilig en hoe risicovol een situatie voor een jeugdige is. Wellicht dat andere vormen, bijvoorbeeld een opsomming van de aandachtspunten en risicofactoren, professionals beter helpen om systematisch na te lopen wat belangrijke aandachtspunten zijn en de veiligheid en mogelijke risico's voor jeugdigen te beoordelen.

Ook blijkt onder professionals vraag naar een hulpmiddel dat zij samen met ouders kunnen gebruiken om de veiligheid en risico's in kaart te brengen. Gezien de subjectiviteit van veiligheids- en risico-oordelen gemaakt met (en zonder) de LIRIK zouden deze misschien altijd samen met de betrokken gezinnen gemaakt moeten worden. Hoewel sommige professionals de LIRIK samen met ouders invullen, zijn er ook professionals die de LIRIK hiervoor niet goed bruikbaar vinden. Gezien de ontwikkeling dat besluitvorming steeds vaker samen met ouders gebeurt, bijvoorbeeld in werkwijzen als Signs of Safety, is het aan te bevelen om na te gaan hoe de LIRIK daarvoor het beste vormgegeven kan worden. Ook zouden jeugdigen zelf in dit proces veel meer betrokken moeten worden (Arad-Davidson & Benbenishty, 2008; Cashmore, 2002; De Kinderombudsman, 2014b).

4.5.3 Verder onderzoek en theorievorming

Verder onderzoek is nodig om veiligheids- en risicotaxatie te verbeteren. Er is nog weinig bekend over de weging van aanwezige zorgen en risicofactoren en hoe die bijdragen aan het ontstaan en herhaling van kindermishandeling. Mogelijk dat het aangeven van een weging van zorgen en risicofactoren tot betrouwbaardere en meer valide oordelen kan leiden. Hoe die weging van factoren zou moeten plaatsvinden, zou verder onderzocht moeten worden.

Gezien ook de beperkte betrouwbaarheid en validiteit van andere risicotaxatie instrumenten zou daarnaast ook gezocht moeten worden naar andere mogelijkheden om de veiligheids- en risicotaxatie te verbeteren. Daarbij zijn wellicht intercollegiaal overleg en de betrokkenheid van ouders en jeugdigen in het beslisproces van belang. Welk effect heeft het wanneer professionals samen met collega's en gezinnen systematisch (dat wil zeggen aan de hand van de LIRIK of een ander instrument) tot hun oordeel over de veiligheid en risico's komen? Leidt dat tot betrouwbaardere en meer valide oordelen?

Tenslotte is het cruciaal dat investeringen in verbeteringen van de professionele oordeels- en besluitvorming er uiteindelijk voor zorgen dat jeugdigen veiliger kunnen opgroeien en dat waar er sprake is van onveiligheid of risico's de juiste hulp wordt ingezet. Voor alle voorstellen voor aanpassingen zal moeten worden nagegaan of deze daadwerkelijk tot verbeteringen leiden in de geboden zorg.

Referenties

- Alink, L.R.A., IJzendoorn, M.H. van, Bakermans-Kranenburg, M.J., Pannebakker, F., Vogels, T. & Euser, S. (2011). *Kindermishandeling in Nederland Anno 2010: De Tweede Nationale Prevalentiestudie Mishandeling van Kinderen en Jeugdigen (NPM-2010)*. Leiden: Casimir Publishers.
- Arad-Davidzon, B. & Benbenishty, R. (2008). The role of workers' attitudes and parent and child wishes in child protection workers' assessments and recommendation regarding removal and reunification. *Children and Youth Services Review*, 30, 107-121.
- Austin et al. (2005)**
- Baird, D. & Wagner, D. (2000). The relative validity of actuarial- and consensus-based risk assessment systems. *Children and Youth Services Review*, 22, 839-871.
- Baird, D., Wagner, D., Healy, T. & Johnson, K. (1999). Risk assessment in child protective services: consensus and actuarial model reliability. *Child welfare*, LXXVIII, 723-748.
- Barber et al. (2007)**
- Barlow, J., Fisher, J. D. & Jones, D. (2012). Systematic review of models of analysing significant harm. Oxford: Oxford University.
- Bartelink, C. & Kooijman, K. (2013). Inschatten van veiligheid en kans op kindermishandeling: noodzaak, instrumenten en ontwikkelingen. *Tijdschrift voor Sociale Geneeskunde*, 91 (7), 391-393.
- Bartelink, C., Kwaadsteniet, L. de & Berge, I. J. ten (2011). *Beter beslissen over kindermishandeling. Onderzoek naar de effecten van gestructureerde besluitvorming middels ORBA*. Utrecht: Nederlands Jeugdinstituut.
- Baumann, D. J., Dagleish, L., Fluke, J., & Kern, H. (2011). The decision-making ecology. Washington, DC: American Humane Association.
- Berben, E.G.M.J. (2000). *Als iedereen hetzelfde was... Indicietelling in de jeugdzorg*. Maastricht: Shaker Publishing B.V.
- Berge, I. J. ten (1998). *Besluitvorming in de kindbescherming. de ontwikkeling en evaluatie van een checklist voor de beoordeling van meldingen bij de raad voor de kindbescherming*. Delft: Eburon.
- Berge, I.J. ten (2008). *Instrumenten voor risicotaxatie in situaties van (vermoedelijke) kindermishandeling*. Utrecht: MO groep.
- Berge, I. ten, Addink, A., Baat, M. de, Bartelink, C., Rossum, J. van, & Vinke, A. (2012). Stoppen en helpen: Een adequaat antwoord op kindermishandeling. Utrecht/Amsterdam: Nederlands Jeugdinstituut/SWP.
- Berge, I.J. ten, Bartelink, C. & Kwaadsteniet, L. de (2011). *Beter beslissen over kindermishandeling: Onderzoek naar de effecten van gestructureerde oordeelsvorming middels ORBA*. Utrecht: Nederlands Jeugdinstituut.
- Berge, I. ten, Eijgenraam, K. & Bartelink, C. (2014). *Licht Instrument Risicotaxatie Kindveiligheid*. Utrecht: Nederlands Jeugdinstituut.

- Berge, I. ten & Meuwissen, I. (2013). *Bruikbaarheid en mogelijke aanpassingen van de LIRIK voor de toepassing in de (L)VB-sector: Bevindingen van de pilot augustus 2012 – oktober 2013*. Utrecht: Nederlands Jeugdinstituut.
- Berge, I. ten & Rossum, J. van (2009). *Evaluatie en bijstelling GCT en LIRIK. Samenvatting resultaten en aanpassingen*. Utrecht: Nederlands Jeugdinstituut.
- Bruyn, E. E. J. de, Ruijsenaars, A. J. J. M., Pameijer, N. K., & Aarle, E. J. M. van (2003). *De diagnostische cyclus: Een praktijkleer*. Leuven/Leusden: Acco.
- Camasso, M. J. & Jagannathan, R. (2000). Modelling the reliability and predictive validity of risk assessment in child protective services. *Children and Youth Services Review*, 22, 873-896.
- Camasso & Jagannathan (2013). Decision making in child protective services: A risky business? *Risk Analysis*, 33, 1636-1649.
- Cashmore, J. (2002). Promoting the participation of children and young people in care. *Child Abuse and Neglect*, 26, 837-847.
- Cicchetti (2001)**
- Dagleish (2000)**
- Dana, Dawes & Peterson (2013)**
- Davidson-Arad, B. (2005). Fifteen month follow-up of children at risk: Comparison of the quality of life of children removed from home and children remaining at home. *Child and Youth Services Review*, 27, 1-20.
- Davidson-Arad, B. (2010). Four perspectives on the quality of life of children at risk kept at home and removed from home in Israel. *British Journal of Social Work*, 40, 1719-1735.
- De Kinderombudsman (2014b). *Kinderrechtenmonitor 2014*. Den Haag: De Kinderombudsman.
- Development Services Group, Inc. (2013). *Protective factors for populations served by the administration on children, youth and families: A literature review and theoretical Framework*.
<http://www.dsgonline.com/acyf/DSG%20Protective%20Factors%20Literature%20Review%202013%20Exec%20Summary.pdf>
- Doyle, J. (2007). Child protection and child outcomes: Measuring the effects of foster care. *The American Economic Review*, 97, 1583-1608.
- Elst, M. van der, Sondeijker, F., Vogel, I., Jansen, W. & Hermanns, J. (2012). *Veiligheidsrisicotaxatie bij opvoedhulp en opgroeihulp aan gezinnen met kinderen van 0-12 jaar*. GGD Rotterdam-Rijnmond en Van Montfoort/Collegio.
- Enosh & Bayer-Topilsky (2014)**
- Evers et al. (2010)**
- Faber, L.L. (2012). De LIRIK: validiteitsonderzoek naar het Licht Instrument Risicotaxatie Kindermishandeling. Masterthese Orthopedagogiek, Vrije Universiteit Amsterdam.
- Finnila, K., Santtila, P., Mattila, J., & Niemi, P. (2012). The effects of experience, outcome feedback, and cognitive feedback on decision-making in child sexual abuse cases: A simulation study. *Nordic Psychology*, 64, 242-257.

Gambrill, E. & Shlonsky, A. (2000). Risk assessment in context. *Children and Youth Services Review*, 22, 813-837.

Gold Benbenishty & Osmo (2001)

Grove et al. (2000)

Harper Browne, C. (2014). *The strengthening families approach and protective factors framework: Branching out and reaching deeper*. Washington, DC: Center for the Study of Social Policy.

Hayes & Krippendorff (2007)

Hermanns, J., Öry, F., & Schrijvers, G. (2005). *Helpen bij opgroeien en opvoeden: eerder, sneller en beter. Een advies over vroegtijdige signalering en interventies bij opvoed- en opgroei problemen*. Utrecht: Inventgroep.

Kahneman & Klein (2009).

Knoke, D. & Trocmé, N. (2005). Reviewing the evidence on assessing risk for child abuse and neglect. *Brief Treatment and Crisis Intervention*, 5, 310-327.

Krippendorff (2004)

Krippendorff (2011)

Lahlah, E., Knaap, L.M. van der & Bogarts, S. (2013). Dangerous boys or boys in danger. *European Journal of Criminology*

Lekkerkerker, L., C. Bartelink, & K. Eijgenraam (2011). *De indicatiestelling bij de Brabantse Bureaus Jeugdzorg nader bekeken. Een onderzoek naar de kwaliteit van het indicatieproces en de betrouwbaarheid van het indicatiebesluit*. Utrecht: Nederlands Jeugdinstituut.

Lindsey, D. (1992). Reliability of the foster care placement decision: A review. *Research on Social Work Practice*, 2, 65-80.

Mulder, T.M. (2014). *Risk factors for child abuse and neglect: A meta-analytic review*. University of Amsterdam, Unpublished Master thesis, Forensic Child and Youth Care Sciences.

Munro, E. (1999).

Munro, E. (2005).

Munro, E. (2008). *Effective child protection*. London: Sage.

Orsi et al. (2014)

Osmo, R. & Benbenishty, R. (2004). Children at risk: rationales for risk assessments and interventions. *Children and Youth Services Review*, 26, 1155-1173.

Pameijer, N. K. & Beukering, J. T. E. van (2004). *Handelingsgerichte diagnostiek : Een praktijkmodel voor diagnostiek en advisering bij onderwijsleerproblemen*. Leuven/ Voorburg: Acco.

Pinto, R. J. & Maia, A. C. (2013). Psychopathology, physical complaints and health risk behaviours among youths who were victims of childhood maltreatment: A comparison between home and institutional interventions. *Children and Youth Services Review*, 35, 603-610.

Regehr, C, Bogo, M., Shlonsky, A. & LeBlanc, V. (2010). Confidence and professional judgment in assessing children's risk of abuse. *Research on Social Work Practice*, 20, 621-628.

Rossi, P., Schuerman, J. & Budde, S. (1996). *Understanding child maltreatment decisions and those who make them*. Chicago, University of Chicago.

- Rubin, A. (Ed.; 2012). *Programs and interventions for maltreated children and families at risk*. Hoboken, NJ: John Wiley and Sons (Clinician's guide to evidence-based practice series).
- Ruiter, C. de, Hildebrand, M. & Hoorn, S. van der (2012). Risicotaxatie bij kindermishandeling: De Child Abuse Risk Evaluation – Nederlandse versie (CARE-NL). In: H.P.B. Lodewijks & L. Van Domburg (red). *Instrumenten voor risicotaxatie: Kinderen en jeugdigen* (pp. 169-190). Amsterdam: Pearson.
- Ruiter, C. de & Jong, E.M. de (2006). *CARE-NL: Richtlijn voor gestructureerde beoordeling van het risico van kindermishandeling*. Enschede: FEBO Druk.
- Ruiter, C. de, Jong, E.M. de, & Reus, M. (2013). Risicotaxatie van kindermishandeling in teamverband: een experimenteel onderzoek. *Kind en Adolescent*, 34, 30-44.
- Rycus, J. S. & Hughes, R. C. (2003). *Issues in Risk Assessment: Policy White Paper*. Columbus Ohio: North American Resource Center for Child Welfare.
- Schuerman et al (1999)**
- Shlonsky, A. & Wagner, D. (2005). The next step: Integrating actuarial risk assessment and clinical judgment into an evidence-based practice framework in CPS case management. *Children and Youth Services Review*, 27, 409-427.
- Stith, S.M., Liu, T., Davies, L.C., Boykin, E.L., Alder, M.C., Harris, J.M., Som, A., McPherson, M., & Dees, J. (2009). Risk factors in child maltreatment: A meta-analytic review of the literature. *Aggression and Violent Behavior*, 14, 13-29.

Taylor (2006)

Bijlagen

Bijlage 1. Licht Instrument Risicotaxatie Kindveiligheid (LIRIK)

Bijlage 2. Ongestructureerd professioneel oordeel

Bijlage 3. Checklist dossieronderzoek

Bijlage 4. Interpretatie van overeenstemming en samenhang

Bijlage 5. Uitkomsten LIRIK

Bijlage 6. Gegevens over T1 en aanvullende gegevens over T2

Bijlage 7. Oorspronkelijke en ideale rangcorrelaties

Bijlage 8. Zachte uitkomsten



Licht Instrument Risicotaxatie Kindveiligheid (LIRIK)

Herziene versie 2014

Ingrid ten Berge, Karin Eijgenraam & Cora Bartelink © NJi

download van www.nji.nl/lirik

LIRIK

Naam jeugdige/gezin:

Geboortedatum en sekse jeugdige:

man vrouw

Datum:

Ingevuld door:

1. Huidige veiligheid

A. Directe veiligheid

Ernstige bedreiging door handelen/nalaten ouder(s)/primaire opvoeder(s):

- Ernstige kindermishandeling
- Onvoldoende bescherming
- Onvoldoende basiszorg
- Opvoeder niet beschikbaar
- Ouder(s) belemmeren zicht op / toegang tot jeugdige

Ernstige bedreiging door jeugdige zelf

Ernstige bedreiging door ander gezinslid / derde, namelijk:

Zijn er vermoedens of aanwijzingen dat de jeugdige op dit moment in een levensbedreigende situatie verkeert of fysiek in gevaar is? ja nee onbekend

Toelichting:

B. Omgang ouder(s) - jeugdige

Zijn er feiten die wijzen op recent:

Lichamelijk geweld ja nee onbekend

Psychisch geweld ja nee onbekend

Lichamelijke verwaarlozing ja nee onbekend

Emotionele verwaarlozing ja nee onbekend

Seksueel misbruik ja nee onbekend

Getuige van huiselijk geweld ja nee onbekend

Opvoeding en verzorging:

Bescherming en veiligheid zorgen adequaat onbekend

Basale verzorging zorgen adequaat onbekend

Emotionele warmte (ondersteuning) zorgen adequaat onbekend

Regels en grenzen zorgen adequaat onbekend

Stimulering zorgen adequaat onbekend

Stabiliteit zorgen adequaat onbekend

Anders:

Als één of meer onbekend:

Ouder geeft geen inzicht in eigen handelen

Onvoldoende informatie beschikbaar

Zijn er aanwijzingen voor bedreigend handelen of nalaten van de ouder(s)? veel enkele geen onbekend

Toelichting:

C. Jeugdige

Psychosociaal functioneren zorgen adequaat onbekend

Lichamelijke gezondheid zorgen adequaat onbekend

Vaardigheden en cognitieve ontwikkeling zorgen adequaat onbekend

Anders:

Zijn er kindsignalen die kunnen wijzen op kindermishandeling? veel enkele geen onbekend

Toelichting:

D. Risico- en beschermende factoren

Risicofactoren bij de ouder(s)

Functioneren als opvoeder

Eerder een kind mishandeld, verwaarloosd of misbruikt	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Ontoereikende pedagogische kennis en/of vaardigheden	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Bagatelliseren/ontkennen aangetoonde kindermishandeling	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Negatieve houding t.a.v. kind	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Persoonlijk functioneren				
Psychiatrische problematiek	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Verslavingsproblematiek	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Verstandelijke beperking	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Beschikbaarheid voor de jeugdige				
Fysiek afwezig	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Emotioneel afwezig	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Voorgeschiedenis				
Op jonge leeftijd (< 18 jaar) ouder geworden	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Zelf slachtoffer van kindermishandeling	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Eerder geweld gebruikt tegen personen	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Problematische partnerrelatie	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Anders:				
Risicofactoren bij jeugdige, gezin en omgeving				
Jeugdige				
Jong kind (< 5 jaar)	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Belaste voorgeschiedenis (bijv. prematuur)	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
(Ernstige) ziekte, handicap of beperking	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Gedrags- en/of ontwikkelingsproblemen	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Moeilijk temperament	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Ongewenst				
Anders:				
Gezin en omgeving				
Laag opleidingsniveau	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Eenoudergezin, stiefgezin, groot gezin	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Veel conflicten	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Huiselijk geweld	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Instabiel, ongeregeld leven	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Materiële/financiële problemen (werkloosheid, huisvesting)	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Ingrijpende levensgebeurtenissen	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Sociaal isolement/sociaal conflict	<input type="checkbox"/> ja	<input type="checkbox"/> nee	<input type="checkbox"/> onbekend	
Anders:				
Zijn er risicofactoren voor kindermishandeling?	<input type="checkbox"/> veel	<input type="checkbox"/> enkele	<input type="checkbox"/> geen	<input type="checkbox"/> onbekend
Toelichting:				
Beschermende factoren				
Ouder(s)	Gezin en omgeving	Jeugdige		
<input type="checkbox"/> Gevoel van competentie, draagkracht	<input type="checkbox"/> Steun informeel netwerk	<input type="checkbox"/> Sociaal vaardig		
<input type="checkbox"/> Positief zelfbeeld	<input type="checkbox"/> Steun formeel netwerk	<input type="checkbox"/> Positief zelfbeeld		
<input type="checkbox"/> Ondersteunende partner		<input type="checkbox"/> Bovengemiddelde intelligentie		
<input type="checkbox"/> Kan eigen jeugdervaringen hanteren		<input type="checkbox"/> Aantrekkelijk uiterlijk		
<input type="checkbox"/> Positieve jeugdervaringen		<input type="checkbox"/> Goede relatie belangrijke volwassene(n)		
<input type="checkbox"/> Kan steun vragen/profiteren van		<input type="checkbox"/> Egoveerkracht (stressresistentie)		
<input type="checkbox"/> Emotionele beschikbaarheid		<input type="checkbox"/> Bereid en in staat om te veranderen		
<input type="checkbox"/> Flexibiliteit				
<input type="checkbox"/> Bereid en in staat om te veranderen				
Anders:				
Zijn er beschermende factoren?	<input type="checkbox"/> veel	<input type="checkbox"/> enkele	<input type="checkbox"/> geen	<input type="checkbox"/> onbekend
Toelichting:				
Conclusie huidige veiligheidssituatie				
<input type="checkbox"/> Er lijkt op dit moment geen sprake van kindermishandeling	Toelichting:			
<input type="checkbox"/> Er is mogelijk sprake van kindermishandeling				

er lijkt sprake van bedreigend handelen of nalaten van de ouder(s)

er zijn kindsignalen die daarop wijzen

er zijn risicofactoren bij de ouder(s)

er zijn risicofactoren bij jeugdige, gezin en/of omgeving

ouder geeft geen informatie of inzicht in eigen handelen

Er is sprake van feitelijk aangetoonde kindermishandeling

op dit moment

recent, maar gestopt

Er is sprake van een levensbedreigende situatie of direct fysiek gevaar

Onvoldoende informatie om een oordeel te vormen

2. Risicotaxatie

A. Aanvullende risicofactoren bij mogelijke of aangetoonde kindermishandeling

(Vermoedelijke) pleger heeft direct toegang tot de jeugdige

Derden hebben geen zicht op de jeugdige

B. Wat kan er gebeuren?

	Aard	Kans	
Mogelijke risico's voor de jeugdige	<input type="checkbox"/> Levensbedreigende situatie of direct fysiek in gevaar	<input type="checkbox"/> Groot	<input type="checkbox"/> Klein
	<input type="checkbox"/> Voortduren/herhaling van kindermishandeling	<input type="checkbox"/> Groot	<input type="checkbox"/> Klein
	<input type="checkbox"/> Ontstaan van kindermishandeling	<input type="checkbox"/> Groot	<input type="checkbox"/> Klein
	<input type="checkbox"/> Anders:	<input type="checkbox"/> Groot	<input type="checkbox"/> Klein
Verwachte gevolgen voor de jeugdige	<input type="checkbox"/> Zeer ernstig	<input type="checkbox"/> Ernstig	<input type="checkbox"/> Gering

C. Beschermende factoren

Zijn er beschermende factoren die de risico's kunnen verminderen? veel enkele geen onbekend

Toelichting:

Conclusie risico's voor de jeugdige

	laag	<input type="checkbox"/> 1	<input type="checkbox"/> 2	<input type="checkbox"/> 3	<input type="checkbox"/> 4	<input type="checkbox"/> 5	Zeer hoog
Op dit moment							
Toelichting:							
Bij voorziene veranderingen in de nabije toekomst							
Toelichting:							

Bijlage 2. Ongestructureerd professioneel oordeel

Dossiernummer:

Naam cliënt:

Naam medewerker:

Functie medewerker:

Afdeling/hulpvorm:

Datum van invullen:

1. Wat is de conclusie t.a.v. de huidige veiligheidssituatie?

- Er is sprake van een levensbedreigende situatie of direct fysiek gevaar
- Er is sprake van feitelijk aangetoonde kindermishandeling
- Er is mogelijk sprake van kindermishandeling
- Er zijn geen aanwijzingen voor kindermishandeling
- Onvoldoende informatie om een oordeel te vormen

Toelichting:

2. Wat is de conclusie t.a.v. risico's voor de jeugdige?

2a. Op dit moment: zeer groot groot reëel klein

Toelichting:

2b. Bij voorziene veranderingen in de nabije toekomst: zeer groot groot reëel klein

Toelichting:

Bijlage 3. Checklist dossieronderzoek

Datum van invullen T2:

Uitvoerder T2: Cora Raoul Floor

Dossiernummer:

Geboortedatum jeugdige:

Geslacht: jongen meisje

Organisatie: Spirit Juzt GGD Gelderland-Zuid

Naam medewerker:

Achtergrondgegevens gezin

1. Heeft de jeugdige broers en zussen die ook voor hulp zijn aangemeld?

Ja

Nee

Nee, maar het betreft opvoedings- of gezinsproblematiek

Zo ja, vermeld de geboortedatum en sekse van elk kind in het gezin

Kind	Geboortedatum	Sekse
Kind 1		
Kind 2		
Kind 3		
Kind 4		
Kind 5		
Kind 6		
Kind 7		
Kind 8		

2. Gezinssamenstelling:

Eenoudergezin

Tweeoudergezin

Samengesteld gezin

Anders, namelijk

Type hulp bij T1

Met dit onderdeel van de vragenlijst beogen we inzicht te krijgen in de hulp die de jeugdige/het gezin krijgt bij T1 (rond aanmelding of bij eerste veiligheidsinschatting).

3. A. Is er sprake van een ondertoezichtstelling?

Ja; startdatum OTS:

Nee

B. Is er sprake van een uithuisplaatsing?

- ja, in het netwerk van het gezin
- ja, in een pleeggezin
- ja, in een woon- of behandelgroep
- ja, in een gezinshuis
- ja, in gesloten jeugdzorg/jeugdzorg plus
- nee

Startdatum UHP:

4. Reden van aanmelding (aard & ernst van de problematiek)

.....
.....
.....
.....

5. Wordt er al hulp geboden bij T1?

- Ja, namelijk [naam interventie of beschrijving van de hulp]
- Nee (ga naar vraag 10)

6. Wat waren de doelen van de hulp?

.....
.....
.....
.....

7. Is al eerder hulp geboden voor jeugd- of opvoedingsproblemen?

- Ja, namelijk [naam interventie of beschrijving van de hulp]
- Nee

Actuele stand van zaken (T2)

Bij T2 (half jaar na veiligheidsbeoordeling T1) gaat het om de vraag hoe de veiligheid van de jeugdige is veranderd tussen T1 en T2. Dit wordt gemeten door te kijken naar AMK- en Raadsmeldingen, interventies gericht op de veiligheid en uithuisplaatsing.

8. Is de interventie of hulp die op T1 geboden werd afgesloten?

- Ja (ga naar vraag 9)
- Nee (ga naar vraag 10)

9. Zijn de doelen van de hulp behaald? (doelrealisatie)

- Ja, volledig behaald
- Ja, deels behaald

- Nee
- Onbekend
- N.v.t.

9A. Is de OTS die er op T1 was beëindigd?

- Ja
- Nee (OTS loopt nog of is verlengd)

Datum beëindiging OTS:

9B. Is de uithuisplaatsing die er op T1 was beëindigd?

- Ja
- Nee

Datum beëindiging uithuisplaatsing:

10. Zijn/worden er op T2 interventies uitgevoerd (anders dan op T1)?

- Ja, namelijk [naam interventie of beschrijving van de hulp]
- Nee

11. Zijn/worden er interventies/technieken specifiek gericht op de veiligheid uitgevoerd (bijv. Signs of Safety / of aandacht voor veiligheid in gesprek met ouders)?

- Ja (geef beschrijving van wat er gedaan is)

.....

.....

.....

.....

- Nee

12. Is de gezinssamenstelling gewijzigd?

- Ja; beschrijf wat gewijzigd is:
- Nee

13. Is het gezin bij het AMK gemeld?

- Ja (ga naar vraag 14)
- Nee (ga naar vraag 15)

14. Is kindermishandeling bevestigd in het AMK onderzoek?

- Ja
- Nee
- N.v.t.

15. Is het gezin bij de Raad voor de Kinderbescherming gemeld?

- Ja (ga naar vraag 16)
- Nee (ga naar vraag 18)

16. Is een verzoek tot ondertoezichtstelling bij de kinderrechter aangevraagd?

- Ja
- Nee
- N.v.t.

17. Is een ondertoezichtstelling uitgesproken?

- Ja
- Nee
- N.v.t.

18. Is de crisisdienst ingeschakeld / Heeft zich een crisissituatie voorgedaan?

- Ja
- Nee

19. Is het gezin verwezen naar een andere hulpverleningsinstelling?

- Ja, namelijk..... [naam interventie of beschrijving van de hulp die aan het gezin bij deze instelling geboden wordt]
- Nee

20. Is het kind uit huis geplaatst?

- Ja
 - in het netwerk van het gezin
 - in een pleeggezin
 - in een woon- of behandelgroep
 - in een gezinshuis
 - in gesloten jeugdzorg/jeugdzorg plus
- Nee

Ontwikkelingen in de 6 maanden tussen T1 en T2

Tussen T1 en T2 kunnen zich veranderingen in de gezinssituatie hebben voorgedaan die de veiligheid kunnen beïnvloeden, ten positieve en ten negatieve.

21. Is na T1 opnieuw een inschatting van de veiligheid en risico's gemaakt?

- ja (ga naar vraag 22 en 23)
- nee (ga naar vraag 24)

22. Wat is de conclusie t.a.v. de huidige veiligheidssituatie?

- Er lijkt op dit moment geen sprake van kindermishandeling

- Er is mogelijk sprake van kindermishandeling
- Er is sprake van feitelijk aangetoonde kindermishandeling
- Er is sprake van een levensbedreigende situatie of direct fysiek gevaar
- Onvoldoende informatie om tot een oordeel te komen

23. Wat is de conclusie t.a.v. risico's voor de jeugdige?

23a. Op dit moment: zeer groot groot reëel klein

23b. Bij voorziene veranderingen
in de nabije toekomst: zeer groot groot reëel klein

24. Staan in het dossier notities waaruit blijkt dat de volgende items met betrekking tot de omgang tussen ouder(s) en jeugdige verbeterd, gelijk gebleven of verslechterd zijn:

	Verbeterd	Gelijk gebleven	Verslechterd	Instabiel	Onbekend/n.v.t.
Bescherming en veiligheid					
Basale verzorging					
Emotionele warmte (ondersteuning)					
Regels en grenzen					
Stimulering					
Stabiliteit					

25. Staan in het dossier notities waaruit blijkt dat er veranderingen zijn in de volgende risicofactoren:

	Verbeterd	Gelijk gebleven	Verslechterd	Instabiel	Onbekend/n.v.t.
Psychiatrische problematiek bij ouder(s)					
Verslavingsproblematiek bij ouder(s)					
Fysieke afwezigheid van de ouder(s)					
Emotionele afwezigheid van de ouder(s)					
Conflicten					
Huiselijk geweld					
Materiële/financiële problemen					
Sociaal isolement/sociaal conflict					
Ingrijpende gebeurtenissen					

26. Staan in het dossier notities waaruit blijkt dat er veranderingen zijn in de volgende beschermende factoren:

	Verbeterd	Gelijk gebleven	Verslechterd	Instabiel	Onbekend/n.v.t.
Ouder(s)					
Gevoel van competentie					
Ondersteunende partner					
Steun vragen/profiteren van					
Emotionele beschikbaarheid					
Flexibiliteit					
Gezin en omgeving					
Steun informeel netwerk					
Steun formeel netwerk					
Jeugdige					

27. Staan in het dossier aanwijzingen voor/vermoedens van kindermishandeling vermeld?

ja (neem de beschrijving over)

.....
.....
.....
.....

nee

28. Is er andere informatie die relevant is voor de veiligheidssituatie die je hierboven niet hebt kunnen benoemen?

ja (neem de beschrijving over)

.....
.....
.....
.....

nee

Beoordeling veiligheid en risico's op T1

Uit het dossier nemen we de conclusies over de huidige veiligheid en risico's over.

Datum van invullen veiligheidsbeoordeling:

29. Wat is de conclusie t.a.v. de huidige veiligheidssituatie?

- Er lijkt op dit moment geen sprake van kindermishandeling
- Er is mogelijk sprake van kindermishandeling
- Er is sprake van feitelijk aangetoonde kindermishandeling
- Er is sprake van een levensbedreigende situatie of direct fysiek gevaar
- Onvoldoende informatie om tot een oordeel te komen

30. Wat is de conclusie t.a.v. risico's voor de jeugdige?

- 2a. Op dit moment: zeer groot groot reëel klein
- 2b. Bij voorziene veranderingen
in de nabije toekomst: zeer groot groot reëel klein

31. Is de LIRIK gebruikt om een inschatting van de veiligheid en risico's te maken?

- Ja (LIRIK in zijn geheel overnemen)
- Nee (conclusies en onderbouwing die wel in het dossier beschreven zijn overnemen)

Bijlage 4. Interpretatie van overeenstemming en samenhang

Tabel 1 geeft aan hoe de verschillende overeenstemmingsmaten geïnterpreteerd worden (zie Cohen, 1960; Landis & Koch, 1977).

Tabel 1. Kwalificatie overeenstemming

Kwalificatie	Overeenstemming	
	percentage	kappa
Slecht (no or slight agreement)	0 – 40	-1.00 - .20
Matig (fair)	41 – 60	.21 – .40
Redelijk (moderate agreement)	61 – 80	.41 – .60
Goed (substantial to almost perfect agreement)	81 of hoger	.61 – 1.00

Tabel 2 geeft weer hoe correlaties geïnterpreteerd worden.

Tabel 2. Interpretatie van correlaties

Kwalificatie	Correlatie
Zeer sterke positieve samenhang	+ .70 or higher
Sterke positieve samenhang	+ .40 to + .69
Matig sterke positieve samenhang	+ .30 to + .39
Zwakke positieve samenhang	+ .20 to + .29
Geen of zeer zwakke positieve samenhang	+ .01 to + .19
Geen of zeer zwakke negatieve samenhang	- .01 to - .19
Zwakke negatieve samenhang	- .20 to - .29
Matig sterke negatieve samenhang	- .30 to - .39
Sterke negatieve samenhang	- .40 to - .69
Zeer sterke negatieve samenhang	- .70 or higher

Bijlage 5. Uitkomsten LIRIK

Tabel 3. Frequenties en percentages van de LIRIK-items

	Frequentie (n=278)	Percentage
1. Huidige veiligheidssituatie		
A. Directe veiligheid		
Bedreiging door ouder(s)/primaire opvoeder(s)	41	14,7
Ernstige kindermishandeling	7	2,5
Onvoldoende bescherming	31	11,2
Onvoldoende basiszorg	27	9,7
Opvoeder niet beschikbaar	37	13,3
Ouder(s) belemmeren zicht op / toegang tot jeugdige?	8	2,9
Bedreiging door jeugdige zelf	18	6,5
Bedreiging door ander gezinslid/derde	13	4,7
Zijn er vermoedens dat de jeugdige op dit moment fysiek in gevaar is?		
Ja	13	4,7
Nee	242	87,1
Onbekend	14	5,0
B. Omgang ouder(s) - jeugdige		
Lichamelijk of psychisch geweld		
Ja	32	11,5
Nee	194	69,8
Onbekend	13	4,7
Verwaarlozing		
Ja	36	12,9
Nee	185	66,5
Onbekend	19	6,8
Seksueel misbruik		
Ja	0	0
Nee	217	78,1
Onbekend	19	6,9
Getuige van huiselijk geweld		
Ja	22	7,9
Nee	183	65,8
Onbekend	35	12,6
Bescherming en veiligheid		
Geen zorgen	164	59,0
Zorgen	72	25,9
Onbekend	10	3,6
Basale verzorging		
Geen zorgen	202	72,7
Zorgen	28	10,1
Onbekend	12	4,3
Emotionele warmte (ondersteuning)		
Geen zorgen	157	56,5
Zorgen	67	24,1
Onbekend	16	5,8
Regels en grenzen		
Geen zorgen	116	41,7

Zorgen	110	39,6
Onbekend	19	6,8
Stimulering		
Geen zorgen	164	59,0
Zorgen	56	20,1
Onbekend	24	8,6
Stabiliteit		
Geen zorgen	123	44,2
Zorgen	96	34,5
Onbekend	20	7,2
Anders		
Geen zorgen	49	17,6
Zorgen	8	2,9
Onbekend	9	3,2
Ouder geeft geen inzicht in eigen handelen	17	6,1
Onvoldoende informatie beschikbaar	34	12,2
Zijn er aanwijzingen voor bedreigend handelen of nalaten van de ouder(s)?		
Ja	70	25,2
Nee	175	62,9
Onbekend	17	6,1
C. Jeugdige(n)		
Psychosociaal functioneren		
Geen zorgen	82	29,5
Zorgen	148	53,2
Onbekend	16	5,8
Lichamelijke gezondheid		
Geen zorgen	193	69,4
Zorgen	32	11,5
Onbekend	14	5,0
Vaardigheden en cognitieve ontwikkeling		
Geen zorgen	124	44,6
Zorgen	100	36,0
Onbekend	21	7,6
Anders		
Geen zorgen	37	13,3
Zorgen	19	6,8
Onbekend	6	2,2
Zijn er kindsignalen die wijzen op een onveilige opvoedingssituatie of kindermishandeling?		
Ja	58	20,9
Nee	191	68,7
Onbekend	20	7,2
D. Risicofactoren		
<i>Risicofactoren bij de ouders</i>		
Eerder een kind mishandeld, verwaarloosd of misbruikt		
Ja	26	9,4
Nee	164	59,0
Onbekend	61	21,9
Ontoereikende pedagogische kennis en/of vaardigheden		
Ja	79	28,4

Nee	131	47,1
Onbekend	44	15,8
Problemen in de ouder-kind interactie		
Ja	112	40,3
Nee	93	33,5
Onbekend	33	11,9
Bagatelliseren/ontkennen aangetoonde kindermishandeling		
Ja	22	7,9
Nee	190	68,3
Onbekend	34	12,2
Negatieve houding t.a.v. het kind		
Ja	39	14,0
Nee	184	66,2
Onbekend	27	9,7
Psychiatrische problematiek		
Ja	51	18,3
Nee	139	50,0
Onbekend	62	22,3
Verslavingsproblematiek		
Ja	15	5,4
Nee	196	70,5
Onbekend	39	14,0
Verstandelijke beperking		
Ja	6	2,2
Nee	202	73,0
Onbekend	40	14,4
Fysieke beschikbaarheid		
Ja	119	42,8
Nee	116	41,7
Onbekend	15	5,4
Emotionele beschikbaarheid		
Ja	108	38,8
Nee	109	39,2
Onbekend	35	12,6
Op jonge leeftijd (< 18 jaar) ouder geworden		
Ja	11	4,0
Nee	211	75,9
Onbekend	29	10,4
Zelf slachtoffer van kindermishandeling		
Ja	27	9,7
Nee	92	33,1
Onbekend	133	47,8
Eerder geweld gebruikt tegen personen		
Ja	20	7,2
Nee	131	47,1
Onbekend	92	33,1
Problematische partnerrelatie		
Ja	87	31,3
Nee	114	41,0
Onbekend	39	14,0
Anders		

Ja	18	6,5
Nee	42	15,1
Onbekend	20	7,0
<i>Risicofactoren bij de jeugdige</i>		
Jong kind (< 5 jaar)		
Ja	64	23,0
Nee	177	63,7
Onbekend	3	1,1
Belaste voorgeschiedenis		
Ja	37	13,3
Nee	157	56,5
Onbekend	53	19,1
(Ernstige) ziekte of handicap		
Ja	12	4,3
Nee	221	79,5
Onbekend	8	2,9
Gedrags- en/of ontwikkelingsproblemen		
Ja	162	58,3
Nee	67	24,1
Onbekend	18	6,5
Moeilijk temperament		
Ja	101	36,3
Nee	121	43,5
Onbekend	21	7,6
Ongewenst		
Ja	5	1,8
Nee	191	68,7
Onbekend	48	17,3
Anders		
Ja	18	6,5
Nee	50	18,0
Onbekend	10	3,6
<i>Risicofactoren bij gezin en/of omgeving</i>		
Eenoudergezin, stiefgezin, groot gezin		
Ja	133	47,8
Nee	115	41,4
Onbekend	4	1,4
Veel conflicten		
Ja	108	38,8
Nee	121	43,5
Onbekend	26	9,4
Huiselijk geweld		
Ja	29	10,4
Nee	177	63,7
Onbekend	49	17,6
Instabiel, ongeregeld leven		
Ja	49	17,6
Nee	179	64,4
Onbekend	24	8,6
Materiële/financiële problemen		
Ja	95	34,2

Nee	132	47,5
Onbekend	25	9,0
Ingrijpende levensgebeurtenissen		
Ja	118	42,4
Nee	98	35,3
Onbekend	34	12,2
Sociaal isolement/conflict		
Ja	40	14,4
Nee	175	62,9
Onbekend	32	11,5
Anders		
Ja	11	4,0
Nee	54	19,4
Onbekend	8	2,9
Zijn er risicofactoren voor kindermishandeling?		
Ja	69	24,8
Nee	141	50,7
Onbekend	21	7,6
Conclusie huidige veiligheidssituatie		
<i>Er is sprake van een levensbedreigende situatie of direct fysiek gevaar</i>	3	1,1
<i>Er is sprake van feitelijk aangetoonde kindermishandeling</i>	26	9,4
op dit moment	11	4,0
in het verleden (nu gestopt)	20	7,2
<i>Er is mogelijk sprake van kindermishandeling</i>	65	23,4
er lijkt sprake van bedreigend handelen of nalaten van de ouder(s)	50	18,0
er zijn kindsignalen die daarop wijzen	51	18,3
er zijn risicofactoren bij de ouder(s)	70	25,2
er zijn risicofactoren bij jeugdige, gezin en/of omgeving	63	22,7
ouder geeft geen informatie of inzicht in eigen handelen	18	6,5
<i>Er zijn geen aanwijzingen voor kindermishandeling</i>	133	47,8
<i>Onvoldoende informatie om een oordeel te vormen</i>	34	12,2
2. Risicotaxatie		
Aanvullende risicofactoren bij vermoedelijke of aangetoonde kindermishandeling		
(Vermoedelijke) pleger heeft direct toegang tot de jeugdige	30	10,8
Derden hebben geen zicht op de jeugdige	4	1,4
Wat kan er gebeuren?		
Mogelijke risico's: Levensbedreigende situatie of direct fysiek gevaar		
Groot	2	0,7
Reëel	22	7,9
Klein	107	38,5
Geen	62	22,3
Mogelijke risico's: Ontstaan van kindermishandeling		
Groot	8	2,9
Reëel	52	18,7
Klein	83	29,9
Geen	55	19,8
Mogelijke risico's: Voortduren/herhaling van kindermishandeling		

Groot	10	3,6
Reëel	42	15,1
Klein	81	29,1
Geen	56	20,1
<i>Mogelijke risico's: Anders</i>		
Groot	4	1,4
Reëel	2	0,7
Klein	21	7,6
Geen	17	6,1
<i>Verwachte gevolgen voor de jeugdige</i>		
Zeer ernstig	8	2,9
Ernstig	39	14,0
Matig ernstig	33	11,9
Gering	105	37,8
Beschermende factoren		
<i>Zijn er beschermende factoren die de risico's kunnen verminderen?</i>		
Geen	9	3,2
Enkele	95	34,2
Veel	65	23,4
Onbekend	15	5,4
<i>Beschermende factoren bij de ouders</i>		
<i>Ouder 1</i>		
Gevoel van competentie, draagkracht	107	38,5
Positief zelfbeeld	71	25,5
Ondersteunende partner	91	32,7
Kan eigen jeugdervaringen hanteren	68	24,5
Positieve jeugdervaringen	52	18,7
Kan steun vragen/profiteren van	140	50,4
Emotionele beschikbaarheid	113	40,6
Flexibiliteit	85	30,6
Bereid en in staat om te veranderen	133	47,8
<i>Ouder 2</i>		
Gevoel van competentie, draagkracht	50	18,0
Positief zelfbeeld	38	13,7
Ondersteunende partner	57	20,5
Kan eigen jeugdervaringen hanteren	33	11,9
Positieve jeugdervaringen	27	9,7
Kan steun vragen/profiteren van	61	21,9
Emotionele beschikbaarheid	56	20,1
Flexibiliteit	41	14,7
<i>Beschermende factoren bij de jeugdige</i>		
<i>Jeugdige 1</i>		
Sociaal vaardig	106	38,1
Positief zelfbeeld	50	18,0
Bovengemiddelde intelligentie	30	10,8
Aantrekkelijk uiterlijk	118	42,4
Goede relatie belangrijke volwassene(n)	132	47,5
Egoveerkracht (stressresistentie)	41	14,7
Bereid en in staat om te veranderen	76	27,3
<i>Jeugdige 2</i>		

Sociaal vaardig	16	5,8
Positief zelfbeeld	12	4,3
Bovengemiddelde intelligentie	5	1,8
Aantrekkelijk uiterlijk	19	6,8
Goede relatie belangrijke volwassene(n)	17	6,1
Egoveerkracht (stressresistentie)	6	2,2
<i>Beschermende factoren bij gezin en omgeving</i>		
Steun informeel netwerk	127	45,7
Steun formeel netwerk	112	40,3
Conclusie risico's voor de jeugdige		
Conclusie risico's voor de jeugdige - op dit moment		
Klein	138	49,6
Reëel	57	20,5
Groot	25	9,0
Zeer groot	5	1,8
Bij voorziene veranderingen in de nabije toekomst		
Klein	90	32,4
Reëel	41	14,7
Groot	18	6,5
Zeer groot	11	4,0

Bijlage 6. Gegevens over T1 en aanvullende gegevens over T2

Bij het dossieronderzoek zijn ook gegevens over de aanwezigheid van (eerdere) hulp, een ondertoezichtstelling of uithuisplaatsing verzameld. In Tabel 4 is weergegeven hoe vaak deze zaken voorkwamen op T1.

Tabel 4. Frequenties en percentages over overige informatie over T1

	LIRIK		Ongestructureerde professioneel oordeel	
	Frequentie (n=278)	Percentage	Frequentie (n=108)	Percentage
OTS T1	28	10,1	15	16,3
UHP T1	19	6,8	10	10,9
Hulp T1	264	95,0	79	85,9
Eerder hulp	228	82,0	62	67,4

Bij het dossieronderzoek is nog een aantal aanvullende variabelen informatie verzameld. Deze bleken bij nader inzien niet geschikt voor verdere analyses, vanwege het kleine aantal zaken of omdat ze te weinig zeiden over de veiligheid van de jeugdige. In Tabel 5 is weergegeven hoe vaak deze zaken voorkwamen.

Tabel 5. Nog aanvullend verzamelde informatie over T2

	LIRIK		Ongestructureerd professioneel oordeel	
	Frequentie (n=278)	Percentage	Frequentie (n=92)	Percentage
Hulp T1 afgesloten	144	51,8	44	47,8
Doelen behaald				
Ja	100	36,0	25	27,2
Nee	29	10,4	11	12,0
Onbekend	11	4,0	8	8,7
N.v.t. ⁷	138	49,6	48	52,2
OTS afgesloten				
Ja	4	1,4	3	3,3
Nee	24	8,6	12	13,0
UHP T1 afgesloten				
Ja	8	2,9	2	2,2
Nee	11	4,0	8	8,7
Hulp T2	112	40,3	49	53,3
Kindermishandeling bevestigd	0	0	0	0
Verzoek OTS				
Ja	2	0,7	3	3,3
Nee	4	1,4	2	2,2
Verwijzing andere hulp	74	26,6	26	28,3

In Tabel 6 is weergegeven hoe hoog de interbeoordelaarsbetrouwbaarheid van de T1- en overige T2-variabelen was. Na een verbeterslag bleek de interbeoordelaarsbetrouwbaarheid zeer goed te zijn.

⁷ We hebben n.v.t. gescoord als er geen doelen in het dossier beschreven waren of als er geen hulp geboden was.

Tabel 6. Interbeoordelaarsbetrouwbaarheid over T1-variabelen en overige T2-variabelen

	Antwoord categorieën	Ronde 1		Ronde 2	
		percentage	kappa	percentage	kappa
OTS T1	2	100	1.00	100	1.00
UHP T1	2	100	1.00	100	1.00
Hulp T1	2	97	.87	100	-.8
Eerder hulp	2	84	.33	100	1.00
Hulp T1 afgesloten	2	97	.94	100	1.00
Doelen behaald	5	85	.79	71	.61
OTS afgesloten	2	66	.36	100	1.00
UHP T1 afgesloten	2	71	-.05	100	1.00
Hulp T2	2	79	.57	100	1.00
Kindermishandeling bevestigd	3	94	.48	100	-.9
Verzoek OTS	3	74	.23	100	1.00
Verwijzing andere hulp	2	73	.39	92	.83

⁸ Er was altijd hulp aanwezig, waardoor kappa berekenen niet mogelijk is.

⁹ In geen enkel dossier was vastgelegd dat kindermishandeling werd bevestigd na een AMK onderzoek.

Bijlage 7. Oorspronkelijke rangcorrelaties (daadwerkelijke en ideale rangcorrelaties)

Omdat er bij de oordelen vier keuzemogelijkheden zijn en bij de uitkomsten twee (wel of niet aanwezig) is het niet mogelijk de maximale correlatie van 1.00 te behalen. Daarom is onderzocht wat gezien de verzamelde gegevens de maximaal te behalen correlatie was. Dit is gedaan door zowel de oordelen van laag naar hoog te ordenen als de uitkomstmaten. Dit is gedaan in twee aparte bestanden en vervolgens zijn de variabelen in hetzelfde databestand gezet, zodat de optimale situatie ontstond waarbij in de zaken waar het oordeel het onveiligst of het risicovolst was ook de uitkomstmaten voorkwamen en in de zaken waar het oordeel was dat er geen sprake was van kindermishandeling of de risico's klein geen uitkomstmaat aanwezig was. Vervolgens zijn over deze ideale situatie rangcorrelaties berekend. Deze ideale rangcorrelaties zijn tussen haakjes toegevoegd aan Tabel 7. Met uitzondering van de variabele "alle uitkomstmaten" bleken de maximale rangcorrelaties aanzienlijk lager dan 1.00 te liggen.

Tabel 7. Oorspronkelijke (en maximaal haalbare) rangcorrelaties voor de oordelen over de huidige veiligheid en risico's en de uitkomsten op T2

	AMK- melding	Raads- melding	OTS	UHP	Veiligheids- interventie	Crisis	Onveilige uitkomst
LIRIK							
Conclusie huidige veiligheidssituatie	.10 (.55)	.03 (.47)	.24** (.62)	-.01 (.51)	.07 (.69)	.15* (.58)	.31** (.91)
Conclusie risico's voor de jeugdige - op dit moment	.14* (.51)	.04 (.41)	.21** (.63)	.08 (.44)	.20** (.69)	.12 (.54)	.37** (.90)
Bij voorziene veranderingen in de nabije toekomst	.07 (.58)	-.03 (.49)	.14 (.70)	.04 (.52)	.02 (.76)	.08 (.62)	.17* (.90)
Ongestructureerde professioneel oordeel							
Conclusie huidige veiligheidssituatie	.16 (.38)	.25* (.48)	.23* (.69)	.51** (.71)	.36** (.60)	.18 (.58)	.48** (.96)
Conclusie risico's voor de jeugdige - op dit moment	.10 (.35)	.22* (.45)	.00 (.70)	.27** (.72)	.41** (.61)	.17 (.57)	.38** (.96)
Bij voorziene veranderingen in de nabije toekomst	.11 (.36)	.14 (.47)	.14 (.73)	.41** (.75)	.25* (.63)	.21 (.59)	.32** (.93)

Bijlage 8. Zachte uitkomsten

Interbeoordelaarsbetrouwbaarheid “subjectieve” uitkomstmaten

In eerste instantie was een aantal softe of subjectieve uitgangsmaten meegenomen bij T2 om na te gaan of er mogelijk verbetering of verslechtering zichtbaar was in de dossiers zonder dat er sprake was van een van de harde uitkomstmaten (AMK-meldingen, OTS, uithuisplaatsing).

Ook over deze items is de interbeoordelaarsbetrouwbaarheid berekend. Voor het grootste deel bleek deze laag te zijn. Omdat de verwachting was dat deze hoe dan ook niet betrouwbaar te scoren waren, zijn de voor de rest niet meer in het onderzoek meegenomen.

Tabel 8. Interbeoordelaarsbetrouwbaarheid over de omgang ouder en jeugdige

	Kappa
Bescherming en veiligheid	.49
Basale verzorging	.49
Emotionele warmte (ondersteuning)	.37
Regels en grenzen	.27
Stimulering	.35
Stabiliteit	.53

Tabel 9. Interbeoordelaarsbetrouwbaarheid over de risicofactoren

	Kappa
Psychiatrische problematiek bij ouder(s)	.46
Verslavingsproblematiek bij ouder(s)	.32
Fysieke afwezigheid van de ouder(s)	.20
Emotionele afwezigheid van de ouder(s)	.20
Conflicten	.52
Huiselijk geweld	.28
Materiële/financiële problemen	.32
Sociaal isolement/sociaal conflict	.48
Ingrijpende gebeurtenissen	.23

Tabel 13. Beschermende factoren

	Kappa
Ouder(s)	
Gevoel van competentie	.38
Ondersteunende partner	.20
Steun vragen/profiteren van	.29
Emotionele beschikbaarheid	.34
Flexibiliteit	-
Gezin en omgeving	
Steun informeel netwerk	.20
Steun formeel netwerk	.37
Jeugdige	
Sociaal vaardig	.47
Positief zelfbeeld	.18
Goede relatie met belangrijke volwassene(n)	.19

Uitkomsten “subjectieve” items

In 226 dossiers waren de soft- of subjectieve uitkomstmaten gescoord. In onderstaande tabellen zijn de uitkomsten daarvan te vinden.

Tabel 11. Frequentie van veranderingen in de omgang tussen ouder(s) en jeugdige (n=226)

	Verbeterd	Gelijk gebleven	Verslechterd	Instabiel	Onbekend/n.v.t.
Bescherming en veiligheid	12	9	2	1	202
Basale verzorging	8	9	2	1	206
Emotionele warmte (ondersteuning)	16	17	2	1	189
Regels en grenzen	37	17	1	2	169
Stimulering	23	14	0	0	169
Stabiliteit	37	12	1	9	167

Tabel 12. Frequentie van veranderingen in de risicofactoren (n=226)

	Verbeterd	Gelijk gebleven	Verslechterd	Instabiel	Onbekend/n.v.t.
Psychiatrische problematiek bij ouder(s)	4	13	3	0	206
Verslavingsproblematiek bij ouder(s)	4	3	0	0	219
Fysieke afwezigheid van de ouder(s)	11	15	5	0	195
Emotionele afwezigheid van de ouder(s)	8	11	4	1	201
Conflicten	29	21	2	3	171
Huiselijk geweld	6	3	2	1	214
Materiële/financiële problemen	6	17	2	0	201
Sociaal isolement/sociaal conflict	3	4	0	0	219
Ingrijpende gebeurtenissen	5	4	5	2	210

Tabel 13. Frequentie van veranderingen in beschermende factoren (n=226)

	Verbeterd	Gelijk gebleven	Verslechterd	Instabiel	Onbekend/n.v.t.
Ouder(s)					
Gevoel van competentie	37	23	6	4	156
Ondersteunende partner	28	20	6	3	169
Steun vragen/profiteren van	10	23	1	3	189
Emotionele beschikbaarheid	14	12	1	0	199
Flexibiliteit	7	5	0	0	213
Gezin en omgeving					
Steun informeel netwerk	12	26	1	4	182
Steun formeel netwerk	14	113	9	6	82
Jeugdige					
Sociaal vaardig	27	17	2	3	176
Positief zelfbeeld	20	15	3	2	186
Goede relatie met belangrijke volwassene(n)	34	18	5	1	167
Egoveerkracht	20	12	0	3	188

Bijlage 9. Klankbordgroep

Deelnemers aan de klankbordgroep waren:

Mascha Kamphuis	Jeugdarts en onderzoeker bij TNO, voorzitter Artsen Jeugdgezondheidszorg Nederland (AJN)
Karin Kleine	Bestuurslid van het vermogensfonds SOD en Fie van der Hoop fonds
Carlo Schuengel	Hoogleraar Ontwikkelingspedagogiek Vrije Universiteit Amsterdam
Ignace Vermaes	Directeur Kwaliteit van Zorg & Innovatie Pluryn