

De haalbaarheid en betrouwbaarheid van het screenen op een licht verstandelijke beperking met de SCIL via videoverbinding

Nienke Spaan, Matthijs Verzaal en Hendrien Kaal

Lectoraat LVB & Jeugdcriminaliteit, Kenniscentrum Samen Redzaam, Faculteit Sociaal Werk en Toegepaste Psychologie, Hogeschool Leiden

Onderzoeksverslag, april 2020

Samenvatting

AANLEIDING - De SCIL is een screeningsinstrument dat ontwikkeld is om te helpen een LVB in een vroeg stadium in de strafrechtketen of een hulpverleningstraject te herkennen. In situaties waarbij het lastig is om de cliënt of verdachte face-to-face op korte termijn te ontmoeten kan de gewenste vroege herkenning moeilijk zijn. Een mogelijke oplossing in dergelijke situaties is afname van de SCIL op afstand via een videoverbinding. De vraag die in dit onderzoek dan ook is gesteld is of de SCIL ook betrouwbaar is wanneer de afname niet face-to-face plaatsvindt maar via een videoverbinding.

METHODEN - In totaal is bij 89 respondenten de SCIL tweemaal afgenomen, eenmaal face-to-face en eenmaal via video-connectie, met een periode van tenminste 6 weken tussen de twee afnames. Hierbij werd gevarieerd in de volgorde van afname. Er is gebruik gemaakt van 'purposive sampling' door deels te werven via de zorg voor mensen met een LVB en deels via mbo-onderwijsinstellingen en onder beroepsgroepen, zoals facilitaire dienstverlening en woonbegeleiding, waar veel mensen met een mbo-opleiding werken. Voor afname via videoverbinding werd gebruik gemaakt van een laptop en skype-verbinding. Hierbij was een assistent aanwezig om ervoor te zorgen dat de respondenten geen technische handelingen hoefden te verrichten. Na de tweede afname van de SCIL is de respondenten een serie evaluatievragen voorgelegd.

RESULTATEN - Respondenten waren doorgaans tevreden met beide vormen van afname van de SCIL. Wel waren zij doorgaans positiever over de face-to-face afname dan over de afname via videoverbinding. Zij voelden zich bijvoorbeeld significant meer op hun gemak bij de face-to-face afname en hadden het gevoel dat deze een beter beeld gaf van hun vaardigheden. Er was hierin geen verschil tussen de

groep met een vermoeden van LVB en de groep zonder vermoedelijke LVB. De meeste respondenten zouden wel bereid zijn om in de toekomst een afname via videoverbinding te ondergaan, al is het enthousiasme hierover niet heel groot. Slechts een enkeling gaf de voorkeur aan afname per videoverbinding.

Gemiddeld behaalden respondenten een iets lagere totaalscore op de SCIL wanneer deze via videoverbinding werd afgenomen ($\mu = 16,31$, $SE=.77$) dan wanneer de afname face-to-face plaatsvond ($\mu = 16,94$, $SE .78$), $t(88)=2.47$, $p= .015$. Berekening van de betrouwbaarheid van het eindoordeel 'wel/niet vermoedelijk LVB' gaf een (lineair gewogen) Kappa van 0,771, $p=,000$, 95% BI: ,637-,904. Bij 10 van de respondenten (11%) werd gezien dat zij op basis van de afname via videoverbinding in een andere categorie vielen dan op basis van de face-to-face afname: vier mensen vielen op basis van de face-to-face afname onder de categorie 'vermoedelijk LVB' en op basis van de videoafname onder de categorie 'vermoedelijk geen LVB', bij zes mensen was dit precies andersom.

CONCLUSIE – De resultaten van dit onderzoek laten zien dat de SCIL met enige voorzichtigheid ook ingezet kan worden met afname via videoverbinding. Hierbij zal de SCIL met de oorspronkelijke afkapscore tot klein aantal extra ten onrechte als LVB gelabelde respondenten leiden.

1 Aanleiding

Europese richtlijnen stellen dat er rekening moet worden gehouden met de specifieke behoeften van zogenoemde kwetsbare personen in de justitiële keten, en zien mensen met een LVB als een van de groepen zogenaamde kwetsbare verdachten (artikel 28b lid 1 Sv). Om binnen de justitiële keten te kunnen afstemmen op de behoeften van verdachten van criminaliteit en hen beter toegang te geven tot hun rechten, is herkenning van LVB in een vroeg stadium van belang.

De SCIL, een screeningsinstrument voor LVB, wordt hiervoor al enige tijd ingezet, onder andere bij de reclassering en binnen penitentiaire inrichtingen. Ook in de context van de ZSM-werkwijze van het OM - waarbij nauw wordt samengewerkt met politie, reclassering, kinderscherming, slachtofferhulp en hulpverlening om zo tot een vlotte en passende afdoening van delicten te komen - wordt gezien dat het belangrijk is om een LVB te herkennen. Men zou er dan ook graag vaker de SCIL willen inzetten om deze herkenning te waarborgen. Een complicerende factor binnen de ZSM-werkwijze is echter dat men verdachten vaak niet face-to-face ziet maar dat eventueel contact doorgaans plaatsvindt via bijvoorbeeld telehoren. De verdachte, die zich op dat moment op het politiebureau bevindt, wordt dan in een cel met videoverbinding geplaatst en kan zo gehoord worden. Overwogen wordt om ook de SCIL via een dergelijke verbinding af te nemen. De gebruikelijke afname van de SCIL is echter tijdens face-to-face contact. Naar de betrouwbaarheid van de SCIL bij afname via een videoverbinding is niet eerder onderzoek gedaan. De vraag die in dit onderzoek dan ook is gesteld is:

Is de SCIL ook betrouwbaar wanneer de afname niet face-to-face plaatsvindt maar via een videoverbinding?

2 Achtergrond

2.1 Gebruik van de SCIL

De screener voor intelligentie en licht verstandelijke beperking (SCIL) is een screeningsinstrument om te beoordelen of iemand functioneert op het niveau van een LVB. Hiermee helpt de SCIL bij het identificeren van een beperking op grond van optredende symptomen, teneinde op dat moment adequate hulp te kunnen bieden. De SCIL leidt tot een vermoeden van het functioneren op het niveau van een LVB maar is geen diagnostisch instrument (Kaal, Nijman, & Moonen, 2015). Voor het daadwerkelijk vaststellen van een LVB is afname van een IQ-test en onderzoek naar het adaptieve vermogen noodzakelijk. Dit vereist echter specialistische kennis, vergt een relatief lange afnametijd en is vaak belastend voor de te testen persoon. Daardoor is dit praktisch vaak niet inzetbaar voor een vroege beoordeling van de mogelijke aanwezigheid van een LVB. Een snelle beoordeling en vroege signalering van een vermoedelijke LVB is van belang om zo tijdig passende bejegening te bieden en toegang tot rechten voor kwetsbare verdachten te waarborgen. Daadwerkelijke diagnostiek van een eventuele LVB kan dan, zo nodig, later in het traject worden ingezet.

De SCIL is relatief niet belastend, kent een korte afnametijd en kan – nadat zij hiertoe een training hebben gevolgd – ook worden afgenomen door bijvoorbeeld politieambtenaren en andere actoren in de justitiële keten. De SCIL screent op een IQ < 85 omdat dit als bovenwaarde van het IQ wordt gezien volgens de zogenoemde praktijkdefinitie van LVB in Nederland (Nijman, Kaal, van Scheppingen, & Moonen, 2018). Belangrijk is om hier te noemen dat de SCIL, evenals veel referentie-standaarden, niet 100% accuraat is (Gold et al., 2010). De SCIL is onder andere gevalideerd met een sample van 318 volwassenen waarbij het IQ recent was vastgesteld door middel van de Wechsler Adult Intelligence Scale (WAIS). Verdeling van de IQ score in deze sample was: 8% IQ 50-59; 30% IQ 70-84; 62% IQ > 85. Hierbij werd voor de SCIL een sensitiviteit van 82% en specificiteit van 89% vastgesteld in het voorstellen van een IQ < 85 (Nijman et al., 2018).

2.2 Inzet van audio-visuele middelen bij de afname van psychometrische tests

Voor dit onderzoek is onder andere gekeken naar literatuur uit medische vakgebieden. Contact tussen behandelaars en specialisten enerzijds en cliënten en patiënten anderzijds via bijvoorbeeld video- en audioverbinding wordt daar al langere tijd gebruikt en onderzocht. Op een wijze die vergelijkbaar is met de beoogde toepassing van een video- en audioverbinding voor het afnemen van de SCIL, wordt telecommunicatie binnen de medische vakgebieden bijvoorbeeld ingezet voor het stellen van een diagnose wanneer de behandelaar en de patiënt gescheiden zijn door grote afstand (Perednia & Allen, 1995 in Hersh et al., 2002). Al in 2002 voerden Hersh et al. een systematische analyse van literatuur uit waarin zij verschillende vormen van *telemedicine* definieerden. Situaties waarin sprake is van directe interactie tussen de deelnemende partijen wordt 'synchrone audio en video' of '*realtime telemedicine*' genoemd. Deze analyse van studies naar *telemedicine* liet zien dat hoewel deze manier van werken destijds al wijdverbreid was, het bewijs van de effectiviteit onder andere verschilde per doel; zo zou het bijvoorbeeld minder effectief zijn wanneer diagnose-stelling afhangt van moeilijk zichtbare kenmerken zoals bij reumatologie en voor sommige oogaandoeningen, terwijl het voor andere oog-aandoeningen juist effectiever kan zijn.

Psychiatrie werd aangewezen als een van de velden waarin *telemedicine* aantoonbaar werkzaam zou zijn voor met name diagnosestelling (Hersh et al., 2002). Zo concludeerde een studie naar

psychiatrische diagnosestelling middels semi-gestructureerde interviews door psychiaters, dat de betrouwbaarheid van de diagnosestelling na videoverbinding overeenkwam met die na face-to-face contact (Baigent et al., 1997; Hersh et al., 2002). Er werden wel enkele verschillen gevonden in de assessments maar die hadden geen invloed op de interpretatie van de situatie en uiteindelijk te stellen diagnose (Baigent et al., 1997). In Noorwegen werd onderzoek gedaan onder 32 normaalbegaafde jongeren waarbij meerdere neuropsychologische tests werden afgenomen via video en face-to-face (op dezelfde dag). Hierbij vonden de onderzoekers verschillende betrouwbaarheidscoëfficiënten in de range van 0,37-0,87. Geconcludeerd werd dat, ondanks kleine opgetreden verschillen, de prestatie via telecommunicatie consistent was met die bij conventionele methoden van testen. Statistisch significante verschillen werden vooral aangetoond bij de maten voor verbaal leren en auditieve aandacht; deze waren hoger bij afname via videoverbinding (Jacobsen, Sprenger, Anderson, & Krogstad, 2003). In een eerdere studie vonden Kirkwood, Peck & Bennie (2000) dat respondenten over het geheel genomen juist wat lager scoorden op tests die werden afgenomen middels telecommunicatie waarbij werd gesuggereerd dat dit een gevolg kon zijn van de kwaliteit van de videoverbinding. Op een ander onderdeel werd in dit onderzoek echter ook hoger gepresteerd bij afname per video, namelijk bij de “motor speed test of information processing” (Kirkwood, Peck, & Bennie, 2000). Jacobsen et al., (2003) maakten gebruik van een betere verbinding dan Kirkwood et al. en noemden dat een aantal respondenten had aangegeven zich minder afgeleid te voelen bij afname van tests middels telecommunicatie. Dit werd gezien als mogelijke verklaring voor de betere prestatie bij de video-afname van een aantal items van de tests. Ook hier werd geconcludeerd dat vanuit een klinisch standpunt, kleine verschillen in eindscores per afnamemethode waarschijnlijk geen invloed zouden hebben op de uiteindelijke interpretatie van neuropsychologische tests. Kirkwood et al. (2000) gaven verder aan dat afname middels videolink bijna 25% meer tijd kostte dan face-to-face afname.

Bij beoordeling van studies naar de werkzaamheid van *telemedicine* door Hersh et al. (2002) werd duidelijk dat veel studies naar de werkzaamheid van de inzet van *telemedicine* zwakheden vertoonden. De onderzoekers waarschuwden dat, om de werkzaamheid van de inzet van *telemedicine* voor diagnostische doeleinden en behandelbesluitvorming aan te tonen, het van belang was dat er grotere hoeveelheden respondenten en clinici werden geïnccludeerd in studies, dat er goed gekeken moest worden naar bijvoorbeeld inter-observator overeenstemming en dat vergelijkingen gemaakt zouden moeten worden met zogenaamde referentiestandaarden (Hersh et al., 2002).

De WHO Global Observatory for eHealth (2010) noemt diverse menselijke en culturele factoren die een uitdaging kunnen vormen voor het routinematig inzetten van *telemedicine* in de praktijk. Zo kan bij zowel patiënten als professionals weerstand tegen verandering deze vorm van contact in de weg staan. Ook een gebrek aan ICT-vaardigheden van de gebruikers heeft een negatieve invloed op het effectief kunnen inzetten van *telemedicine*. Daarnaast wordt door het WHO aangekaart dat het van belang is om juridische aspecten van hulp- en dienstverlening rondom bijvoorbeeld vertrouwelijkheid en privacy mee te wegen in de toepasbaarheid (WHO Global Observatory for eHealth, 2010). Het gebruik van *telemedicine* wordt doorgaans, indien haalbaar, echter als een goede kosteneffectieve optie beschouwd (Nord, Rising, Band, Carr, & Hollander, 2019). Wilson en Maeder (2015) concluderen op basis van hun literatuuronderzoek zelfs dat de toepassing van *telemedicine* inmiddels beschouwd kan worden als een essentieel onderdeel van de gezondheidszorg en meer specifiek op het gebied van psychiatrie en psychische gezondheid.

De toepasbaarheid van *telemedicine*, in combinatie met de inmiddels brede toegankelijkheid hiervan door technologische ontwikkelingen die relatief goedkoop en makkelijk inzetbaar zijn, leidde ertoe dat al in 2015 gesignaleerd werd dat deze vorm van zorgverlening steeds vaker, breder en binnen verschillende categorieën van zorg werd ingezet (Wilson & Maeder, 2015). De zogenaamde videoconferentie wordt het meest gebruikt en kon ook toen al met gewone 'consumentenproducten' en gebruik van infrastructuur aanwezig in de meeste huishoudens worden ingezet (Wilson & Maeder, 2015). Een goede implementatie is afhankelijk van juist gebruik van technologie. Zo noemen Fatehi, Armfield, Dimitrijevic, & Gray (2015) dat de ervaren kwaliteit en bruikbaarheid van videoconferenties afhankelijk zijn van technische factoren als frame rate, beeldgrootte, resolutie, mate van beeldcompressie en beschikbare telecommunicatiesnelheden. Met de huidige ontwikkelingen rondom goed toegankelijke technologie, zeker in Nederland, lijkt al snel aan de vereisten te worden voldaan.

Er zijn duidelijke parallellen aan te wijzen tussen de situatie waarbij justitie snel wil screenen op de mogelijke aanwezigheid van een LVB bij een verdachte, en *synchrone telemedicine* zoals dit bijvoorbeeld wordt gebruikt binnen de psychiatrie en spoedeisende hulpsituaties. Zo is er in beide situaties bijvoorbeeld sprake van een vergelijkbare noodzakelijkheid of wenselijkheid van snel contact tussen specifieke partijen voor assessment enerzijds en de onmogelijkheid om op elk gewenst moment die partijen elkaar in dezelfde fysieke ruimte te laten ontmoeten anderzijds. Literatuur laat zien dat *telemedicine* goed toepasbaar is in dit soort situaties waarin het van belang is dat op specifieke en kritieke momenten en op korte termijn, specialistische expertise beschikbaar is. Verder is het afnemen van psychologische assessments per video-connectie anno 2020 in een aantal situaties al routine. Zo maakt het bijvoorbeeld in Canada onderdeel uit van de dagelijkse praktijk, waarbij gebruik wordt gemaakt van getrainde coördinatoren zonder specifieke academische scholing in psychologische assessments (Goyal, Temple, Sawanas, & Brown, 2020; Temple, Drummond, Valiquette, & Jozsvai, 2010; persoonlijke communicatie met Temple, 2019).

Al met al is duidelijk dat de werkzaamheid van assessment van cognitieve vaardigheden en intelligentie via videoverbinding reeds lang beschreven wordt in de literatuur en dat dergelijke assessments dagelijks worden ingezet in de praktijk (Goyal, Temple, Sawanas, & Brown, 2020; Temple, Drummond, Valiquette, & Jozsvai, 2010). Toch is het van belang dat de SCIL niet op deze wijze wordt ingezet zonder nader inzicht in de invloed die afname via videoverbinding kan hebben op de betrouwbaarheid. De eerdergenoemde waarschuwingen uit de literatuur, onder andere bij systematische analyses naar de toepasbaarheid van *telemedicine*, met betrekking tot onderwerpen als inter-observator betrouwbaarheid, aantallen respondenten (bijvoorbeeld Hersh et al., 2002) en de invloed van factoren zoals opleiding, taal en fysieke en mentale mogelijkheden (bijvoorbeeld WHO Global Observatory for eHealth, 2010) onderstrepen het belang van onderzoek naar het gebruik van videoverbinding voor het afnemen van de SCIL alvorens de inzetbaarheid hiervan in de praktijk te overwegen.

3 Methoden

Om een antwoord te geven op de onderzoeksvraag is een onderzoek opgezet waarbij bij een groep respondenten twee keer de SCIL is afgenomen: één keer op de reguliere, face-to-face manier, en één

keer via een videoverbinding. Dit maakte het mogelijk om vast te stellen wat de mate van overeenstemming was tussen de twee afnames. Hieronder wordt de onderzoeksopzet nader besproken.

3.1 Studieontwerp

Voor dit onderzoek werd het niet haalbaar geacht om alle respondenten te onderwerpen aan een volledige IQ-test en/of diagnostisering van een LVB. Aangezien het onderzoek niet gericht was op het opnieuw valideren van de SCIL maar op het toetsen van de overeenstemming tussen twee methoden van afname, werd dit ook niet nodig geacht. Om ruis en vertekening door onbekende variabelen tegen te gaan is gekozen voor een onderzoeksdesign waarin elke respondent als zijn eigen controlepersoon functioneert. Er werd rekening gehouden met mogelijke invloed van de volgorde van de wijze van afname (eerst face-to-face of eerst videoverbinding) en een eventueel leereffect tussen afnames. Respondenten werden daarom zoveel mogelijk random verdeeld over twee groepen, waarbij de ene groep eerst een face-to-face afname onderging en de tweede groep eerste een afname via videoverbinding. De beoogde tijd tussen de twee afnames was 6-12 weken. Respondenten werden uitgelegd dat zij tweemaal bevroegd zouden worden middels een vragenlijst. Hierbij werd niet verteld dat het om identieke vragen ging; dit om te voorkomen dat zij extra hun best zouden doen om de vragen te onthouden. Na de tweede afname van de SCIL werd de respondenten een evaluatie-vragenlijst voorgelegd.

3.2 Respondenten

Om een voldoende diverse populatie te includeren is gebruik gemaakt van *purposive sampling*. De helft van de respondenten is geworven via de zorg voor mensen met een LVB. Hiervoor zijn diverse zorgorganisaties benaderd, die cliënten met een (vermoede) LVB hebben gevraagd of zij bereid waren mee te werken. Als zij hierin toestemden werden hun gegevens met de onderzoekers gedeeld. De andere helft van de respondenten bestond uit mensen met werk of een opleiding op mbo-niveau: zij werden deels geworven via mbo-onderwijsinstellingen en deels onder beroepsgroepen, zoals facilitaire dienstverlening en woonbegeleiding, waar veel mensen met een mbo-opleiding werken. Bij de werving is ook gestreefd naar een evenwichtige verdeling tussen mannen en vrouwen. Alle respondenten waren ouder dan 18 jaar. Verder zijn als inclusiecriteria de criteria voor de validiteit van de SCIL gehanteerd. Dit houdt in dat de respondenten de Nederlandse taal voldoende machtig waren om de opdrachten en instructies te kunnen begrijpen en dat er geen sprake was van acute psychiatrische problematiek.

Beoogd werd om in totaal 120 respondenten te includeren waarbij de SCIL tweemaal werd afgenomen. Het uiteindelijke aantal gerealiseerde tweede afnames is lager, deels doordat het lastig was om mensen bereid te vinden mee te doen aan het onderzoek, maar grotendeels ook omdat 23 respondenten niet konden deelnemen aan een tweede afname door maatregelen rondom de uitbraak van het COVID-19 virus in Nederland. Verder waren 3 respondenten na de eerste afname niet meer bereikbaar voor een tweede meting en gaven 7 respondenten om uiteenlopende redenen aan niet deel te willen nemen aan de tweede afname. Dit heeft er ook voor gezorgd dat de verdeling van de respondenten wat betreft de volgorde van afnames uiteindelijk niet 50/50 is. In tabel 1 wordt beschreven wat de kenmerken zijn van de respondenten die uiteindelijk zijn geïncludeerd in de analyses.

Tabel 1: Kenmerken van respondenten waarbij 2 afnames van de SCIL hebben plaatsgevonden (N=89)

Geslacht	Man: n=42 (47%) / vrouw: n=47 (53%)
Leeftijd	Mediaan 30, Q ₁ : 23,9 ; Q ₃ :46,1, (min-max: 18-63)
Populatie	LVB*: n=45 (51%) / mbo**: n=44 (49%)
Tijd tussen afnames (in dagen)	Mediaan: 46, Q ₁ : 42 ; Q ₃ : 116 (min-max: 40-226)
Methode eerste afname	Face to face: n=38 (43%) / videoverbinding: n=51 (57%)

* Werving van respondent via zorginstantie voor mensen met een LVB

* Werving van respondent via opleidingsinstellingen mbo of mbo beroepsgroepen

3.3 Afname SCIL

Afhankelijk van of bij de potentiële respondent de eerste afname face-to-face of via videoverbinding zou worden afgenomen, vond het informatieve gesprek tussen de respondent en de veldonderzoeker en het bespreken van het informed consent plaats per video of face-to-face. Na ondertekening van het formulier voor informed consent vond de afname van de SCIL plaats conform de instructies in de handleiding van de SCIL. Er was in dit opzicht geen verschil tussen de face-to-face afname en de afname via videoverbinding. Alle afnames werden bovendien gedaan door dezelfde onderzoeker.

De tijdsperiode tussen twee afnames was ten minste zes weken. Hoewel bij de opzet van het onderzoek een maximaal tijdsbestek werd beoogd van 3 maanden, bleek dit in de praktijk niet haalbaar, met name door de planning van de respondenten. Hierdoor varieert de tijd tussen afname van 6 weken tot 7,5 maand (zie tabel 1); bij 34 respondenten was de tijdsperiode tussen 2 afnames groter dan 3 maanden, waarvan bij 15 groter dan 5 maanden.

3.4 Afname SCIL per videoverbinding

Bij afname per video bevond de onderzoeker die de SCIL afnam zich in een andere ruimte dan de respondent. De respondent werd verwelkomd door een assistent die de laptop voor videocontact klaar had gezet en beschikbaar was voor eventuele technische vragen. Van de respondent werd geen kennis van computergebruik gevraagd; alles werd in orde gemaakt door de assistent zodat de respondent zich alleen met de communicatie met de testafnemer hoefde bezig te houden. Indien zich technische problemen voordeden werden deze opgelost door de assistent. Links en rechts van de laptop die gebruikt werd voor de videoconnectie waren een bruine en een witte envelop geplaatst met in een van de enveloppen het SCIL-opgavenformulier (waarop de respondent enkele taken zelf moet uitvoeren) en in de andere envelop, indien het de tweede afname betrof, het evaluatieformulier. Indien het de eerste afname betrof lag er een consentformulier op tafel. Op deze manier kon de afnemer van de SCIL makkelijk verwijzen naar de verschillende formulieren. Daar waar dit door praktische omstandigheden niet lukte, gaf de aanwezige assistent de benodigde formulieren aan de respondent.

3.5 Gebruikte techniek voor afname SCIL via videoverbinding

Omdat aannemelijk is dat in de praktijk een eventuele afname van de SCIL per video gebonden is aan relatief makkelijk toegankelijke middelen, is in dit onderzoek gebruik gemaakt van laptops met een schermdiameter van 35,5 cm zonder extra microfoon of externe speakers. Beide laptops maakten gebruik van Windows 10. Voor de videoverbinding werd gebruik gemaakt van een verbinding via Skype.

Daar waar mogelijk werd gebruik gemaakt van aanwezige wifi, indien dit niet toegankelijk was maakten zowel de laptop van de respondent als die van de afnemer van de SCIL gebruik van een zogenaamde hotspot, opgezet met een mobiele telefoon.

3.6 Het scoren van de SCIL

Na afname van de SCIL werden de door de respondent gegeven antwoorden aan de hand van de handleiding gescoord. De meeste items krijgen hierbij 0, 1 of 2 punten, bij enkele items kan alleen 0 of 2 punten toegekend worden. In de praktijk wordt dit scoren doorgaans gedaan door één beoordeelaar. Om eventuele systematische bias door de manier van scoren te verminderen zijn de ingevulde SCIL-formulieren voor dit onderzoek onafhankelijk gescoord door twee verschillende beoordelaars. De interbeoordelaarsbetrouwbaarheid is berekend per item door middel van een Kappa maat, weergegeven in tabel 2. Daar waar een verschil zat in beoordeling is het betreffende item nagekeken door een derde, onafhankelijke, beoordeelaar die op basis van de handleiding de eindbeoordeling, zoals gebruikt voor de data van dit onderzoek, heeft gegeven.

Tabel 2: Interbeoordelaarsbetrouwbaarheid bij het scoren van de SCIL (N=178)

SCIL	n verschil	% overlap	K _w	p	95% BI
Vraag 1* - speciaal onderwijs	0	100%	1,000	,000	(1,00 – 1,00)
Vraag 2** - opleiding	8	95,5%	,947	,000	(,911 - ,984)
Vraag 3* - hulpverlening	0	100%	1,000	,000	(1,00 – 1,00)
Vraag 4* - steun	0	100%	1,000	,000	(1,00 – 1,00)
Vraag 5* - geldrekenen	2	98,9%	,966	,000	(,919 – 1,01)
Vraag 6* - tijdrekenen 1	0	100%	1,000	,000	(1,00 – 1,00)
Vraag 7* - tijdrekenen 2	2	98,9%	,977	,000	(,945 – 1,01)
Vraag 8** - achteruit spellen	6	96,6%	,949	,000	(,901 - ,996)
Vraag 9* - lezen	10	94,4%	,884	,000	(,814 - ,954)
Vraag 10* - spreekwoord	11	93,8%	,875	,000	(,803 - ,946)
Vraag 11** - rekenen	0	100%	1,000	,000	(1,00 – 1,00)
Vraag 12** - schrijven	30	83,1%	,855	,000	(,796 - ,915)
Vraag 13** - voorlezen	6	96,6%	,950	,000	(,900 – 1,00)
Vraag 14** - klok tekenen	28	84,3%	,867	,000	(,807 - ,926)
Vermoeden LVB*	5	97,2%	,943	,000	(,893 - ,992)

* Dichotome variabele; waarde Kappa is lineair gewogen Kappa.

** 0-2 punten per antwoord, waarde Kappa is kwadratisch gewogen Kappa.

3.7 Evaluatievragen

Na de tweede afname van de SCIL werd een evaluatie-vragenlijst voorgelegd aan de deelnemers. Hierop stonden de volgende negen stellingen:

1. Bij de afname met de computer kon ik de afnemer goed genoeg zien
2. Bij de afname met de computer kon ik de afnemer goed genoeg horen
3. Ik voelde me op mijn gemak bij de afname met de computer
4. Ik voelde me op mijn gemak bij de face-to-face afname
5. De afname met de computer gaf een goed beeld van mijn vaardigheden
6. De face-to-face afname gaf een goed beeld van mijn vaardigheden
7. Bij de afname met de computer voelde ik me vrij te zeggen / vragen wat ik wilde
8. Bij de face-to-face afname voelde ik me vrij te zeggen / vragen wat ik wilde.
9. Als ik in de toekomst zo'n test zou moeten doen, zou ik dat wel via de computer willen doen.

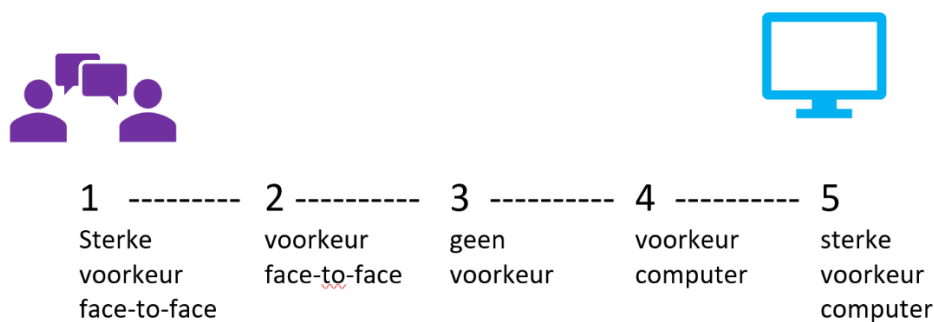
Respondenten gaven de mate van instemming aan op een 7-punts schaal, waarbij 1 'sterk mee eens' was, en 7 'sterk mee oneens'. Deze schalen werden visueel ondersteund met gekleurde duimpjes (zie fig. 1). Ook werd met behulp van symbolen aangegeven of de vragen bijvoorbeeld over de face-to-

face afname of de afname via videoverbinding ging. Tenslotte werd de volgende vraag gesteld, met antwoorden op een vijf-punts-schaal (zie fig. 2):

10. Aan welke manier van afnemen geeft u de voorkeur?



Figuur 1: 7-punts Likertschaal gebruikt bij vragen 1-9 van de evaluatie-vragenlijst



Figuur 2: 5-punts schaal gebruikt bij vraag 10 van de evaluatie-vragenlijst

Bovenstaande vragen zijn gebaseerd op de evaluatievragen gebruikt door Schutte et al. (2015), die een vergelijkbaar onderzoek deden naar de bruikbaarheid en de betrouwbaarheid van de afname via videoverbinding van de ADOS module 4, een autisme assessment instrument voor volwassenen. Schutte et al. stelden de evaluatievragen na de computerafname, ongeacht of dit de eerste of de tweede afname was. Zodoende was in hun studie maar een deel van de respondenten in staat om een vergelijking te maken tussen de twee manieren van afname. Omdat wij het belangrijk vonden die vergelijking wel te maken, is er in dit onderzoek voor gekozen de vragen van Schutte et al. enigszins aan te passen en de lijst bij iedereen voor te leggen na de tweede afname.

3.8 Data analyse

Voor analyse van de data is gebruikt gemaakt van IBM SPSS Statistics versie 23. Literatuur over onderzoeken naar de betrouwbaarheid van testafname via videoverbinding betreft over het algemeen relatief kleine groepen respondenten en laat een discussie zien over de juiste wijze van data-analyse. De discussie over betrouwbaarheidsonderzoeken wordt in de literatuur ook breder gevoerd. Zo wordt bepleit dat zowel samenhang als overeenstemming onderdeel uitmaken van replicerbaarheid en rapportage van betrouwbaarheidsintervallen van belang is (Bujang & Baharum, 2017; Hersh et al., 2002; Koo & Li, 2016; Watson & Petrie, 2010). In eerste instantie was de opzet dan ook om de data van onderhavig onderzoek te analyseren door middel van een Two-way Mixed ICC om zowel inzicht in samenhang tussen de twee vormen van afname als inzicht in de mate van overeenstemming te verkrijgen (zie bijv. Koo & Li, 2016). De verkregen data van totaalscores op de SCIL per afname-methode bleken echter niet normaal verdeeld. Gezien de scoremogelijkheden op de SCIL, met een maximaal

haalbare score van 28 punten en afkapscore gehanteerd voor de beoordeling of er bij de respondent een vermoedelijke LVB is van 19,5, is dit begrijpelijk. Dit staat echter de toepassing van dit ICC-model in de weg.

In plaats van een ICC-analyse is er daarom voor gekozen om een gewogen Kappa-waarde te berekenen voor ieder item en voor de eindbeoordeling (ja/nee vermoeden LVB) van de SCIL. Op basis van de eerdergenoemde discussie in de literatuur is voor de items die 3 scoremogelijkheden hadden (0, 1 of 2 punten) een kwadratisch gewogen Kappa berekend. Deze vorm van gewogen Kappa kan toegepast worden wanneer er sprake is van drie of meer ordinale categorieën en weegt ook de mate van afwijkingen in de scores mee. Kortgezegd: grotere afwijkingen wegen zwaarder dan kleine afwijkingen, hierdoor wordt naast de mate van overeenstemming dus ook de mate van verschil in scores meegewogen. Deze kwadratisch gewogen Kappa kan gezien worden als een equivalent van de ICC (Warrens, 2011; Watson & Petrie, 2010). Voor de dichotome data (i.e. de vragen van de SCIL waarop 2 scoremogelijkheden zijn (0 of 2 punten) en voor de conclusie (wel/niet vermoedelijke LVB) is een zogenaamde lineair gewogen Kappa berekend. Bij alle Kappa's is ook het 95% betrouwbaarheidsinterval gerapporteerd.

Voordat op basis van de SCIL tot de dichotome conclusie wordt gekomen of er wel of niet sprake is van een vermoedelijke LVB wordt er een totaalscore berekend die kan variëren van 0-28 punten. De gemiddelde scores per afnamemethode zijn vergeleken met een gepaarde t-test (data voldeed aan assumpties). Hierbij is gebruik gemaakt van Bias corrected and accelerated (BCa) bootstrapping (5000). Dit houdt o.a. in dat er bij schatting van het betrouwbaarheidsinterval gebruik wordt gemaakt van stratificeren waarbij de variabele 'volgorde van afname' (eerst face to face of eerst video) is meegewogen. Voor visuele inspectie van de overeenstemming (*agreement*) van variantie is gebruik gemaakt van een zogenaamde Bland Altman plot voor repeated measures (Bland & Altman, 1999).

Bij de analyse van de evaluatievragenlijst is in eerste instantie gekeken naar het gemiddelde, de mediaan en de modus. Vervolgens is middels een t-toets gekeken of er verschil is in de ervaring tussen respondenten waarbij op basis van de face-to-face afname van de SCIL sprake is van een vermoeden van LVB, en respondenten waar vermoed wordt dat dit niet zo is. Ook is, middels de pearson correlatie, gekeken of de ervaringen met de verschillende manieren van afname samenhangen met de leeftijd van de respondent. Om te kijken of men tevredener is over de face-to-face afname of de afname via videoverbinding is voor vragenparen 3-4, 5-6 en 7-8 een verschilscore berekend. Hierbij geeft een positieve score een voorkeur voor face-to-face afname aan, en een negatieve een voorkeur voor afname via videoverbinding. Middels een t-test kon worden gekeken of er verschil in voorkeur was tussen mensen met vermoedelijke LVB en mensen zonder vermoedelijke LVB. Middels een Pearson's correlatie werd bekeken of er samenhang was tussen de voorkeur en de leeftijd van de respondent.

3.9 Ethische waarborgen

Het lectoraat LVB en jeugdcriminaliteit maakt onderdeel uit van Hogeschool Leiden. Het onderzoeksvoorstel is daarom ter beoordeling voorgelegd aan de onderzoeksadviescommissie van de Faculteit Sociaal Werk en Toegepaste Psychologie van Hogeschool Leiden. Omdat respondenten met een (vermoedelijke) LVB grotendeels werden geworven via zorginstanties, is het onderzoeksvoorstel ook beoordeeld door de interne onderzoekscommissie, de ethische commissie en de juridische afdeling van een van de grotere instellingen die medewerking verleenden. Uiteindelijke toestemming is verleend door zowel de commissie van de Hogeschool als door de instellingscommissies. Voor het omgaan met

en bewaren van persoonsgegevens zijn de richtlijnen opgevolgd van de privacydesk van Hogeschool Leiden. Nadat potentiële respondenten aangaven deel te willen nemen aan het onderzoek werd eerst een gesprek gevoerd tussen de onderzoeker en de potentiële respondent over wat deelname precies zou inhouden en welke tijdsinvestering werd gevraagd. Als de respondent na dit gesprek nog steeds wilde deelnemen aan het onderzoek werd het *informed consent*-formulier gezamenlijk doorgesproken en ondertekend door de respondent. Iedere respondent kreeg een cadeaubon naar keuze ter waarde van € 10 na de tweede afname van de SCIL.

4 Resultaten

4.1 Betrouwbaarheid SCIL-afname via videoverbinding

Kappa-waardes berekend voor de 14 items van de SCIL vallen tussen de 0,355 en 0,954 (zie tabel 3). Vier items hadden een $K > 0,81$ (bijna perfect); zes items een $K = 0,61-0,80$ (voldoende tot goed); drie items een $K = 0,41-0,60$ (redelijk) en één item een $K = 0,21-0,40$ (matig). Het valt op dat bij enkele vragen respondenten wat vaker lager scoren bij video-afname dan bij face-to-face afname dan andersom.

Gemiddeld behaalden respondenten een iets lagere totaalscore op de SCIL wanneer deze via videoverbinding werd afgenomen ($\mu = 16,31$, $SE = .77$) dan wanneer de afname face-to-face plaatsvond ($\mu = 16,94$, $SE = .78$), $t(88) = 2.47$, $p = .015$. Het gemiddelde verschil is daarmee 0,63 ($sd = 2,4$), BCa 95% $CI = (0,124 - 1.135)$.

De SCIL leidt uiteindelijk tot een eindoordeel 'wel/niet vermoedelijk LVB'. Berekening van de betrouwbaarheid van dit eindoordeel gaf een (lineair gewogen) Kappa van 0,771, $p = ,000$, 95% $BI: ,637- ,904$. Bij 10 van de respondenten (11%) werd gezien dat zij op basis van de afname via videoverbinding in een andere categorie vielen dan op basis van de face-to-face afname: vier mensen vielen op basis van de face-to-face afname onder de categorie 'vermoedelijk LVB' en op basis van de videoafname onder de categorie 'vermoedelijk geen LVB', bij zes mensen was dit precies andersom.

Tabel 3 Analyse overeenstemming methoden van afnemen SCIL (face to face en per video-link) (N=89)

SCIL	K_w	p	95% BI	n verschil ¹	FtF>Vid ²	FtF<Vid ³
Vraag 1* - speciaal onderwijs	,954	,000	(,892-1,017)	2	1	1
Vraag 2** - opleiding	,846	,000	(,754-,938)	11	5	6
Vraag 3* - hulpverlening	,854	,000	(,741-,967)	6	2	4
Vraag 4* - steun	,690	,000	(,404-,976)	4	2	2
Vraag 5* - geldrekenen	,557	,000	(,343-,770)	13	7	6
Vraag 6* - tijdrekenen 1	,579	,000	(,606-,886)	15	10	5
Vraag 7* - tijdrekenen 2	,745	,000	(,606-,885)	11	8	3
Vraag 8** - achteruit spellen	,552	,000	(,375-,728)	27	15	12
Vraag 9* - lezen	,695	,000	(,543-,848)	13	7	6
Vraag 10* - spreekwoord	,819	,000	(,700-,939)	8	5	3
Vraag 11** - rekenen	,756	,000	(,634-,877)	22	14	8
Vraag 12** - schrijven	,768	,000	(,686-,850)	29	18	11
Vraag 13** - voorlezen	,666	,000	(,531-,802)	27	17	10
Vraag 14** - klok tekenen	,355	,001	(,162-,547)	36	19	17
Vermoeden LVB*	,771	,000	(,637-,904)	10	4***	6***

* 0 of 2 punten per antwoord mogelijk. Waarde Kappa is lineair gewogen Kappa.

** 0-3 punten per antwoord, waarde Kappa is kwadratisch gewogen Kappa.

***score LVB=1, NB=0; dus 4 mensen scoorden op FtF LVB en bij video NB; 6 mensen scoorden bij FtF NB en bij video LVB.

¹ Absoluut aantal respondenten met andere score op deze vraag bij de verschillende afnames

² Absoluut aantal respondenten dat hoger scoorde bij face to face afname dan bij video

³ Absoluut aantal respondenten dat lager scoorde bij face to face afname dan bij video

Visuele inspectie van de variantie middels een Bland Altman plot liet geen systematische verschillen zien tussen de methodes van afname. Ruim 95% van de variantie viel binnen de zogenaamde ‘grenzen van overeenstemming’. Er lijkt ook geen relatie te zijn tussen de verschillen tussen de twee soorten afnames en de ‘werkelijke’ score op de SCIL.

4.2 Resultaten evaluatievragen

Wanneer we kijken naar de evaluatievragen zien we dat mensen doorgaans tevreden zijn met beide vormen van afname van de SCIL. In tabel 4 is te zien dat de modus 1 is voor vraag 1 – 8; dit betekent dat de meeste respondenten het met deze stellingen eens zijn. Zij vonden de afnemer via de videoverbinding goed te zien en te horen, voelden zich bij beide vormen van afname op hun gemak, hadden de indruk dat beide vormen van afname een goed beeld gaf van hun vaardigheden, en voelde zich bij beide manieren van afname vrij om te zeggen en vragen wat zij wilden. Er zijn geen significante verschillen in de beoordeling van afnames tussen de respondenten waarbij op basis van de SCIL een LVB vermoed wordt en de respondenten bij wie hier vermoedelijk geen sprake van is. Tabel 4 laat wel een significante samenhang zien wanneer gekeken wordt naar de correlatie tussen de antwoorden op de tevredenheidsvragen over de vormen van afname en leeftijd: oudere respondenten hebben iets minder dan jongeren het gevoel dat de face-to-face afname een goed beeld gaf van hun vaardigheden. Zij leken zich bij face-to-face afname ook wat minder vrij te voelen te zeggen wat ze wilden dan jongeren. Bij videoafname werd geen verschil tussen jongeren en ouderen gevonden.

Omdat bij vragen 3 – 8 steeds twee keer dezelfde vraag is gesteld voor de face-to-face afname en de afname via videoverbinding is het mogelijk om te kijken of men op de ene vraag positiever of negatiever antwoordde dan op de andere. In tabel 5 zien we dat mensen doorgaans positiever zijn over de face-to-face afname dan over de afname via videoverbinding. Zij voelden zich significant meer op hun gemak bij de face-to-face afname en hadden het gevoel dat deze een beter beeld gaf van hun vaardigheden. Er was hierin geen verschil tussen de groep met een vermoeden van LVB en de groep zonder vermoedelijke LVB. Er werd voor een verschilscore wel een samenhang gevonden met leeftijd: hoe lager de leeftijd, hoe meer men het gevoel had dat face-to-face een beter beeld gaf van de vaardigheden.

Items 9 en 10 in tabel 4 laten tenslotte zien of men in de toekomst een afname via videoverbinding zou willen doen en waar de voorkeur van de respondent naar uitgaat. De meeste respondenten zouden wel bereid zijn om in de toekomst een afname via videoverbinding te ondergaan. Toch is met een mediaan van 3 het enthousiasme niet heel groot. Bij de laatste vraag wordt dan ook een milde voorkeur voor de face-to-face afname gevonden. Er is bij beide vragen geen verschil tussen mensen met en zonder een vermoedelijke LVB. Een samenhang met leeftijd wordt ook niet gevonden.

5 Discussie

5.1 Beperkingen van het onderzoek

Studiepopulatie - Bij het studieontwerp is er met opzet voor gekozen om de respondenten te betrekken uit groepen mensen met een (waarschijnlijke) LVB en mensen die mbo-opgeleid zijn. Een meer diverse onderzoekspopulatie, dus inclusief hoger opgeleide respondenten, zou wellicht een betere afspiegeling geven van de maatschappij, maar geeft minder inzicht in het functioneren van afname van de SCIL. De SCIL is immers het meest betrouwbaar bij scores die verder van de afkapscore afliggen. Door expliciet te focussen op populaties die naar verwachting dichterbij de afkapscore zouden scoren, is gekozen voor een opzet waarbij eventuele problemen met afname met videoverbinding maximaal

Tabel 4: Post-SCIL afname evaluatievragen: resultaten totaal, uitgesplitst naar vermoeden LVB/normaal begaafd, en samenhang met leeftijd.

ITEM	Descriptives										Verschil tussen LVB/NB				Samenhang leeftijd		
	Mean					Std.					LVB		NB		T-test		
	Mean	Std.	Med	Mod	N	Mean	Std.	Med	Mod	N	Mean	Std.	Mean	Std.	Sign	Pear. Corr.	Sign
1. Bij de afname met de computer kon ik de afnemer goed genoeg zien	1,70	1,03	1	1	89	1,60	1,09	1,60	1,09	89	1,82	0,97	1,82	0,97	,323	-,013	,213
2. Bij de afname met de computer kon ik de afnemer goed genoeg horen	1,99	1,39	2	1	88	2,00	1,58	2,00	1,58	88	1,97	1,14	1,97	1,14	,930	-,014	,177
3. Ik voelde me op mijn gemak bij de afname met de computer	2,20	1,33	2	1	89	2,20	1,53	2,20	1,53	89	2,21	1,06	2,21	1,06	,985	,038	,726
4. Ik voelde me op mijn gemak bij de face-to-face afname	1,76	1,20	1	1	89	1,78	1,27	1,78	1,27	89	1,74	1,12	1,74	1,12	,888	,087	,416
5. De afname met de computer gaf een goed beeld van mijn vaardigheden	2,24	1,27	2	1	89	2,10	1,33	2,10	1,33	89	2,41	1,19	2,41	1,19	,255	-,070	,517
6. De face-to-face afname gaf een goed beeld van mijn vaardigheden	1,92	1,19	2	1	89	1,74	1,18	1,74	1,18	89	2,15	1,18	2,15	1,18	,104	,215	,043
7. Bij de afname met de computer voelde ik me vrij te zeggen / vragen wat ik wilde	1,93	1,26	2	1	89	1,86	1,25	1,86	1,25	89	2,03	1,29	2,03	1,29	,541	-,043	,692
8. Bij de face-to-face afname voelde ik me vrij te zeggen / vragen wat ik wilde.	1,64	,93	1	1	89	1,58	1,03	1,58	1,03	89	1,72	0,79	1,72	0,79	,492	,208	,050
9. Als ik in de toekomst zo'n test zou moeten doen, zou ik dat wel via de computer willen doen.	3,02	1,84	3	1	89	3,08	2,05	3,08	2,05	89	2,95	1,55	2,95	1,55	,732	-,007	,536
10. Aan welke manier van afnemen geeft u de voorkeur?	2,46	1,03	3	3	89	2,50	1,17	2,46	1,17	89	2,41	0,85	2,41	0,85	,687	0,16	,142

Lagere getallen geven een sterkere instemming aan, behalve bij het laatste item. Daar is een lager getal een voorkeur voor face-to-face afname.

Tabel 5: Post-SCIL afname evaluatievragen: verschil in ervaring face-to-face afname versus afname via videoverbinding

ITEM	Score op item										Verschilscore "computer minus face-to-face"					
	Computer		Face-to-face		T-test		NB		LVB		T-test		leeftijd			
	Mean	Std.	Mean	Std.	Sig*	Mean	Std.	Mean	Std.	Mean	Std.	Sig*	Pear. Corr.	Sig*		
Ik voelde me op mijn gemak bij de ... afname	2,20	1,33	1,76	1,20	,000	,46	1,02	,42	1,50	,883	-,04	,699				
De ... afname gaf een goed beeld van mijn vaardigheden	2,24	1,27	1,92	1,19	,020	,26	,99	,36	1,43	,700	-,28	,009				
Bij de ... afname voelde ik me vrij te zeggen / vragen wat ik wilde	1,93	1,26	1,64	,93	,026	,31	1,40	,28	1,07	,916	-,20	,056				

* (2-tailed); een positieve verschilscore geeft een voorkeur voor face-to-face afname aan.

zichtbaar worden. De gekozen manier van werven van respondenten leidt niet tot een steekproef die representatief is voor een populatie. Dit werd ook niet van groot belang geacht omdat de respondenten hun eigen match waren. Toch zou een studie met een andere populatie tot andere resultaten kunnen leiden. We vermoeden dat de studie in zijn huidige opzet de populaties waarvoor de screening ingezet gaat worden op het gebied van opleidingsniveau dicht benadert. Bij hogere opleidingen is de kans dat bij een tweede afname tot een afwijkend eindoordeel wordt gekomen alleen maar kleiner.

Toewijzing aan methode eerste afname – Idealiter zouden respondenten in het onderzoek willekeurig zijn toegewezen aan een van de twee groepen ‘eerste afname face-to-face’ of ‘eerste afname video’. Dit was praktisch gezien echter niet haalbaar. Omdat bij de videoafname een assistent nodig was, was het soms praktisch om de SCIL bij een aantal respondenten op een locatie op een dag op dezelfde manier af te nemen. Ook waren de onderzoekers soms afhankelijk van de beschikbare ruimtes. De volgorde van afname (eerst video of eerst face to face) is dan ook veelal gebaseerd geweest op praktische overwegingen. Belangrijk hierbij is dat de indeling niet heeft plaatsgevonden op basis van kenmerken van de respondenten, en dat dit gold voor zowel respondenten geworven uit de mbo-opgeleide populatie als respondenten uit de LVB-populatie.

Beperkingen met betrekking tot het aantal respondenten - Tijdens het studieontwerp was de opzet om 120 respondenten te includeren voor het berekenen van een robuuste ICC waarde waarbij het risico op zowel een type I als type II fout laag was. Dit aantal respondenten is met name vanwege de maatregelen als gevolg van de Coronapandemie niet behaald. Omdat juist de gepaarde metingen van belang zijn, is alleen de data van respondenten die beide afnames van de SCIL afgerond hebben geïnccludeerd. Het aantal geïnccludeerde respondenten was dus lager dan beoogd. Indien geen verschil tussen gemiddelde scores zou zijn gevonden dan had er twijfel kunnen zijn of er ten onrechte werd aangenomen dat er inderdaad geen verschil was. Omdat er op basis van deze 89 respondenten een klein en significant verschil is aangetoond, mag op basis van de resultaten verwacht worden dat de afname via videoverbinding daadwerkelijk leidt tot een iets lagere score.

Technische problemen bij de afname via videoverbinding - Tijdens de afnames via videoverbinding is het een paar keer voorgekomen dat zich problemen voordeden met de verbinding. In een geval betrof dit een situatie waarbij de verbinding meerdere keren kort vastliep. Dit had als praktisch gevolg dat er wat vaker pauzes tussen vragen plaatsvonden. De aanwezige assistent heeft geassisteerd bij de verbindingproblemen en deze verholpen. In één geval lukte het de assistent niet om de skype-verbinding tot stand te brengen waardoor de onderzoeker zelf kort de ruimte is ingegaan voor herstel van het probleem.

Rol van de assistent – Alleen bij afname via videoverbinding werd gebruik gemaakt van de inzet van een assistent. Zoals in methoden aangegeven was diens rol voornamelijk faciliterend. De assistent ontving de respondent, leidde deze naar de kamer waarin de videoverbinding tot stand werd gebracht, loste eventuele problemen op en reikte waar nodig formulieren aan op verzoek van de SCIL-afnemer die via de videoverbinding communiceerde met de respondent (en de assistent). Omdat het in veel gevallen praktisch niet haalbaar was voor de assistent om buiten de kamer waar de respondent aan het onderzoek deelnam te wachten, is de assistent gedurende afname veelal aanwezig geweest in dezelfde ruimte. Omdat de respondentengroep met een (vermoedelijke) LVB veelal deelnam binnen de organisatie waarvan zij zorg ontvingen, is de assistent bij deze groep relatief vaker in dezelfde

kamer aanwezig geweest. In andere gevallen, bijvoorbeeld op een mbo-school, vervulde de assistent regelmatig ook de rol van poortwachter: er kon dan alleen voldoende rust geboden worden als de assistent net voor de deur van de ruimte “de wacht hield” om te voorkomen dat andere mensen per ongeluk de ruimte binnenliepen. Hoewel de aanwezigheid van een assistent leidde tot andere omstandigheden dan bij een reguliere face-to-face afname, lijkt dit geen invloed te hebben gehad op het verloop van afname van de SCIL of de score. Op basis van de ervaringen kan wel geconcludeerd worden dat een assistent die zorgdraagt voor de technische kant van de afname, maar ook respondenten kan geruststellen waar nodig, bij in ieder geval een deel van de onderzoekspopulatie noodzakelijk was. Ook omdat bekend is dat stress invloed heeft op prestaties, is het belangrijk die stress zo veel mogelijk te reduceren.

5.2 Discussie van de resultaten

De afgelopen jaren is er een toegenomen aandacht voor het belang van herkenning van LVB in de strafrechtketen en daarbuiten. Om herkenning van LVB te verbeteren is de SCIL ontwikkeld waarmee in relatief korte tijd tot een vermoeden van (het functioneren op het niveau van) een LVB kan worden gekomen. Dit screeningsinstrument is gevalideerd voor afname in een face-to-face setting. Soms blijkt face-to-face contact in de praktijk echter niet haalbaar; dit riep de vraag op of afname via een videoverbinding ook tot betrouwbare resultaten zou kunnen leiden.

De resultaten van deze studie laten zien dat respondenten de afname via videoverbinding doorgaans positief waarden. Tweënnegentig procent van de respondenten kon de testafnemer goed zien, 86 procent kon de afnemer goed horen. Drieëntachtig procent voelde zich op zijn of haar gemak bij afname via videoverbinding, 82 procent had het idee dat de videoafname een goed beeld gaf van zijn of haar vaardigheden, en 89 procent voelde zich vrij te zeggen wat hij of zij wilde. Vierenvijftig procent zou in de toekomst een dergelijke test wel weer via de computer willen doen, nog eens 27 procent staat hier neutraal in; dat betekent dat slechts 19 procent hier negatief tegenover staat. Bijna de helft (47%) van de respondenten heeft echter wel een voorkeur voor face-to-face afname, nog eens 43 procent staat hier neutraal tegenover; 10 procent geeft de voorkeur aan een afname met de computer.

Voor een klein aantal respondenten lijkt een afname per videoverbinding dus positiever dan een afname face-to-face. Zo was er een respondent die aangaf de afname via videoverbinding fijner te vinden omdat er meer afstand was tot de testafnemer. Een andere respondent gaf aan dat er bij de afname via videoverbinding minder afleiding was. Daarentegen moet ook opgemerkt worden dat voor sommige respondenten de afname via videoverbinding veel stress opriep. Zo was er een respondent die voor het eind van de SCIL-afname van de computer wegliep als gevolg van de stress die werd ervaren bij afname middels video. Een oudere respondent die aangaf niet digivaardig te zijn was van te voren erg angstig en moest uitgebreid gerustgesteld worden door de testafnemer alvorens mee te willen doen aan de videoafname. Eén potentiële respondent gaf meteen bij het zien van de laptop aan niet meer mee te willen werken aan het onderzoek.

De statistische analyse van de SCIL-scores laat zien dat de Kappa-waardes voor de individuele items wisselend zijn: op sommige items is de overeenstemming bijna perfect, op andere items is de overeenstemming redelijk of zelfs matig. Zoals eerder beschreven is uit de literatuur bekend dat bepaalde vaardigheden betere resultaten kunnen laten zien wanneer deze face-to-face getest worden dan per

video, of juist beter tot uiting komen wanneer dit getest wordt per video dan face to face (zie bijvoorbeeld: Jacobsen et al., 2003; Kirkwood et al., 2000). In onderhavige onderzoek leken respondenten bij sommige items, met name de items waarbij ze iets moeten doen zoals tijdrekenen, sommetjes maken, dicteezinnen schrijven of een tekst moeten lezen vaker hoger te scoren bij de face-to-face afname dan bij de videoafname dan andersom.

Uiteindelijk zijn de (gewogen) Kappawaarden per item niet erg belangrijk. Het is immers bekend dat geen van de items alleen voorspellend is voor een LVB; de items staan niet op zichzelf maar vormen samen de SCIL. Het uiteindelijk oordeel van de SCIL wordt bepaald op basis van de totaalscore. Analyse van de totaalscore op SCIL toont dan ook een klein maar significant verschil tussen de afnamemethodes waarbij de gemiddelde totaalscore na afname per video 0,63 punt lager lag dan wanneer de afname face-to-face plaatsvond. Gezien het feit dat een ander antwoord op één vraag al twee punten verschil kan maken, is dit verschil niet erg groot.

Desalniettemin verdient de significant lagere gemiddelde score bij de afname via videoverbinding onze nadere aandacht. Het gevonden verschil in de score op de videoafname versus de face-to-face afname zou theoretisch gezien kunnen liggen aan een eventueel leereffect. De groep mensen bij wie eerst een video-afname plaatsvond was immers iets groter dan de groep waar bij de eerste afname face-to-face was. De testafnemer registreerde tijdens het veldwerk wel dat bijna alle respondenten bij de tweede afname een keer een opmerking maakten dat ze een vraag herkenden van de eerste afname. Deze herkenning trad vaak nog niet op tijdens de inleiding maar bij de vierde of vijfde vraag van de SCIL. Tegelijkertijd leek deze herkenning vervolgens het antwoord op die vraag niet of nauwelijks te beïnvloeden, wat impliceert dat herkenning niet hetzelfde is als het optreden van een leereffect. Om uit te sluiten dat de gevonden verschillen tussen de face-to-face afname en de afname via videoverbinding het gevolg zijn van een leereffect is eerst gekeken of respondenten op de tweede afname daadwerkelijk structureel hoger scoorden dan op de eerste afname. Gemiddeld werd op de tweede afname inderdaad iets hoger gescoord ($\mu = 16,89$, $SE=.78$) dan op de eerste afname ($\mu = 16,37$, $SE=.76$), $t(88)=2.01$, $p= .048$. Zes respondenten vielen bij de eerste meting in de categorie 'vermoedelijk LVB' en bij de tweede meting in de categorie 'geen vermoeden LVB'; bij vier respondenten was dit andersom. Hoewel er dus een klein, ternauwernood significant, leereffect lijkt te zijn, lijkt dit niet tot een duidelijke bias in het eindoordeel te leiden. Een analyse op een gewogen dataset, waarbij gecorrigeerd werd voor het grotere aandeel respondenten waarbij de eerste afname via videoverbinding plaatsvond, laat zien dat ook dan nog steeds een, klein maar significant, verschil gevonden wordt tussen de face-to-face afname ($\mu = 16,85$, $SE=.78$) en de videoverbinding ($\mu = 16,28$, $SE=.77$), $t(88)=2.20$, $p= .030$.

Hoewel we in dit onderzoek, op basis van aangehaalde discussie in literatuur over vergelijkbare studies, hebben gekozen voor rapportage van (gewogen) Kappawaarden en een gepaarde t-toets, werd in het oorspronkelijke test-hertest onderzoek als onderdeel van een validatiestudie een Pearson correlatie gerapporteerd: 0.92 ($n = 28$) (Nijman, Kaal, van Scheppingen, & Moonen, 2018). Om een vergelijking te kunnen maken tussen de resultaten van de verschillende onderzoeken, hebben we ook voor het onderhavige onderzoek de Pearson correlatie berekend tussen de scores voor afname via videoverbinding en de scores voor afname face-to-face: $r = ,95$ ($n=89$), BCa 95% betrouwbaarheidsinterval ($,92 - ,97$). Hieruit valt op te maken dat de in dit onderzoek gevonden Pearson correlatie nog iets hoger uitvalt dan in het test-hertest onderzoek waarbij de onderwaarde van het betrouwbaarheidsinterval overeenkomt met de eerder gerapporteerde r .

Waar het in de praktijk uiteindelijk om draait is de eindconclusie die aan de SCIL-afname verbonden wordt, met betrekking tot het wel of niet aanwezig zijn van een vermoedelijke LVB. Ter interpretatie van Kappa worden in principe categorieën gebruikt in termen: 'slecht' ($k < 0$); 'gering' (0-0,20); 'matig' (0,21-0,40); 'redelijk' (0,41-0,60); 'voldoende tot goed' (0,61-0,80) en 'bijna perfect' (0,81-1,00) waarbij de afkappunten eigenlijk arbitrair gekozen zijn (zie bijvoorbeeld "Cohen's Kappa - WikiStatistiek," n.d.; Landis & Koch, 1977). Bij interpretatie is het daarom van belang om verder dan alleen de Kappa-waarde te kijken, bijvoorbeeld naar het doel en de consequenties van een te lage mate van overeenstemming. De in dit onderzoek gerapporteerde lineair gewogen Kappa van 0,77 met 95% BI (.637-.904) valt in wat betreft classificatie dus onder de bovenkant van de categorie 'voldoende tot goed' waarbij de onderkant van het BI valt aan de onderkant van deze categorie en de bovenwaarde van het BI onder de categorie 'bijna perfect' valt. Er is dus een duidelijke overeenstemming tussen de score bij de face-to-face afname en de score bij de afname via videoverbinding. Uiteindelijk valt bij 11 procent (10 van de 98) van de respondenten het eindoordeel bij de ene meting anders uit dan bij de andere meting. Dit is niet significant hoger ($\chi^2(1)=0,071$, $p=,791$) dan bij het oorspronkelijke test-hertest onderzoek waarbij 7 procent (2 van 28) in een andere categorie viel (Nijman, Kaal, van Scheppingen, & Moonen, 2018). De uiteindelijke impact van de afnamemethode lijkt dus niet heel groot te zijn.

Gezien het statistische significante kleine verschil tussen gemiddelde totaalscore op de SCIL is gekeken wat de invloed zou zijn van het verleggen van het afkappunt op onder andere de groep van tien respondenten die van beoordelingscategorie 'vermoedelijk LVB' of 'vermoedelijk geen LVB' wisselden. Daarbij werd de vraag gesteld wat het oplevert wanneer de afkapscore bij een videoafname op 18,5 gesteld wordt in plaats van op 19,5 punten. Het aantal respondenten die van categorie wisselen op basis van de dichotome beoordeling vermoedelijk wel of geen LVB neemt dan af van 10 naar 7 respondenten. Voor drie respondenten geldt dat zij op basis van de verlaagde afkapscore bij beide afnames tot de groep 'vermoedelijke geen LVB' behoren. Daarmee neemt de (lineair gewogen) Kappa tussen face-to-face afname waarbij het originele afkappunt gehanteerd wordt en afname per video met een aangepaste afkapscore van 18,5 punten krijgt toe: $\kappa_w = ,841$, 95% BI (.727 - ,954). Er waren in dit onderzoek geen respondenten die door het verlagen van de afkapscore 'ten onrechte' in de groep 'vermoedelijk niet LVB' werden geplaatst. In theorie is dit echter wel mogelijk. In dat geval zou deze persoon voor hem/haar passende maatregelen mislopen en verkeerd bejegend kunnen worden. Daar komt bij dat in alle gevallen, juist bij mensen die rondom het afkappunt scoren, rekening gehouden moet worden met een mogelijke misclassificatie. Bovendien is in de context van de strafrechtsketen, waarvoor dit onderzoek primair is opgezet, een iets grotere sensitiviteit te verkiezen boven een iets grotere specificiteit. Hoewel we ons realiseren dat hierover discussie mogelijk is, wordt het verleggen van de afkapscore vooralsnog door ons niet geadviseerd. Wel wordt geadviseerd om in de toekomst waar mogelijk, als afname per videoverbinding plaatsvindt, data te verzamelen over de behaalde SCIL-scores en resultaten van uitgevoerde diagnostiek.

Een identieke totaalscore op de SCIL per respondent tijdens test en hertest werd niet verwacht. De bij validering van de SCIL gerapporteerde sensitiviteit en specificiteit bij de gekozen afkappunten betekent dat er bij een aantal mensen die boven de afkapscore uitkomen toch sprake is van een LVB en vice versa (Kaal et al., 2015; Nijman, Kaal, van Scheppingen, & Moonen, 2018). Er zijn diverse veranderlijke factoren die invloed hebben op de prestatie van de respondent en die niet per keer meetbaar in kaart te brengen zijn. Zo is aangetoond dat factoren als gebrek aan slaap invloed hebben op geheugen, concentratie op niet emotionele stimuli en affect (Gobin, Banks, Fins, & Tartar, 2015) en daarmee net als stress en preoccupatie met omstandigheden en financiële problemen (Mullainathan & Shafir,

2014) negatieve invloed kunnen hebben op IQ-score. Ook werd, zoals eerder vermeld, bij eerder onderzoek naar de test-hertest betrouwbaarheid van de SCIL geconstateerd dat er een kleine, maar acceptabele, afwijking was tussen twee opvolgende metingen (Nijman, Kaal, van Scheppingen, & Moonen, 2018). Visuele inspectie van de data in het onderhavige onderzoek liet zien dat de totaalscores voor beide SCIL-afnames bij zeven van de tien respondenten waarbij de tweede afname tot een ander eindoordeel leidde beide keren om en nabij het afkappunt lagen. Er was dus maar een klein verschil in de antwoorden nodig om tot een ander eindoordeel te komen. Daarbij werden ook in dit onderzoek factoren waargenomen die mogelijk het verschil in score tussen de eerste en tweede meting konden verklaren: zo werd er bij een respondent een sterke wietgeur waargenomen bij de meting waarop hij aanzienlijk lager scoorde dan de keer daarna. Bij een andere respondent was bij de eerste afname sprake van net gestarte medicatie die duidelijk slaperig maakte. Een derde respondent was bij de tweede afspraak een nieuw hulpverleningstraject gestart, waardoor deze respondent op de vraag naar hulpverleningservaring anders scoorde. De respondent met het grootste verschil in eindscore (7 punten) gaf bij de tweede afname al bij item 3 aan de vragen te herkennen, waardoor bij deze respondent vermoed kan worden dat er sprake was een (sterk) leereffect. Deze factoren zijn allemaal factoren die ook in de dagelijkse praktijk een rol kunnen spelen. Wanneer we kijken wat deze tien respondenten antwoorden op de evaluatievragen zien we slechts bij een respondent dat deze negatiever antwoordde op evaluatievragen over bijvoorbeeld het kunnen zien of horen via de videoverbinding en bij de afname via videoverbinding ook daadwerkelijk een lagere eindscore behaalde. Bij de andere respondenten was hier geen sprake van of juist van het tegenovergestelde; er werd aangegeven dat de onderzoeker via video goed te zien en te horen was terwijl er bij afname middels videoverbinding een lagere SCIL score werd behaald.

6 Conclusie

De resultaten van dit onderzoek laten zien dat het screeningsinstrument SCIL, dat is ontwikkeld voor face-to-face afname, met enige voorzichtigheid ook ingezet kan worden met afname via videoverbindingen. Hoewel respondenten doorgaans een voorkeur lijken te hebben voor face-to-face afname reageren zij ook positief op vragen over afname via videoverbinding. Daarbij zijn er geen belangrijke verschillen tussen mensen met en zonder vermoedelijke LVB (zoals vastgesteld met de SCIL) en tussen oudere en jongere respondenten. Slechts een klein percentage heeft zoveel moeite met afname via videoverbinding dat hij of zij dat echt liever niet ondergaat.

Ook de resultaten van de herhaalde afname laten zien dat de afname via videoverbinding redelijk betrouwbaar is: de Kappa-waarde behorende bij de vergelijking van de eindscores bij de twee vormen van afname laat een adequate/goede samenhang zien, en de test-hertest betrouwbaarheid is niet slechter dan in de originele rapportage. Wel is er sprake van een iets lagere gemiddelde score bij de afname via videoafname. Onderzoekers zijn van mening dat het niet verstandig is om hiervoor te corrigeren door de afkapscore te veranderen, omdat het van groot belang is zo min mogelijk mensen met een vermoedelijke LVB missen. Ook zal een correcte bejegening niet in het geding komen bij het hanteren van de originele afkapscore terwijl het andersom verstrekkende gevolgen kan hebben.

Een zekere marge van onzekerheid rondom beoordeling op basis van de SCIL is onvermijdelijk; zeker gezien de sensitiviteit en specificiteit van de SCIL en de test-hertest betrouwbaarheid zoals gezien bij het oorspronkelijke valideringsonderzoek lijkt een percentage van elf procent afwijkende eindoordeelen niet verontrustend. Wel bevestigt dit onderzoek dat niet blindgevaren kan worden op de score op de SCIL, onafhankelijk van de wijze waarop deze wordt afgenomen. Van belang is dat er aandacht is voor mogelijke invloeden op het niveau van presteren, zoals stressverhogende of concentratievermindende omstandigheden.

De SCIL is geen diagnostisch instrument: om uitsluitel te bieden over aanwezigheid van een LVB is een uitgebreide IQ-test en onderzoek naar adaptief vermogen nodig. Juist wanneer een score wordt gevonden rond het afkappunt is verdere diagnostiek onontbeerlijk (Kaal et al., 2015). Desalniettemin is gebruik van de SCIL belangrijk en praktisch goed haalbaar. Hoe wordt gehandeld naar aanleiding van een SCIL score – of er wordt gekozen voor het wel of niet toepassen van diagnostiek bijvoorbeeld - zal afhankelijk zijn van doelstelling van afname, of de SCIL nou face-to-face afgenomen of middels een videoverbinding.

In de praktijk zal moeten worden ervaren wat de reacties van respondenten zijn op afname via videoverbinding in de beoogde context. Dit onderzoek suggereert dat afname via videoverbinding betrouwbaar is in een onderzoekssetting. Er is in de onderzoekssetting steeds sprake geweest van een assistent bij de afname. Dit betekent dat uit dit onderzoek niet zonder meer geconcludeerd kan worden dat afname plaats kan vinden door het leggen van een skype-verbinding met een respondent thuis. Dit onderzoek geeft ook geen inzicht in de impact die de extra stress of weerstand die afname in een justitiële context kan hebben. In de onderzoekssetting was een enkele keer sprake van weigering van deelname aan afname via videoverbinding. Gezien het feit dat respondenten vrijwillig deelnamen aan dit onderzoek en op hoofdlijnen wisten wat hen te wachten stond, kan het zijn dat weigeringen in de praktijk vaker zullen voorkomen. Gezien het bovenstaande verdient het aanbeveling om bij toekomstig gebruik van de SCIL via videoverbinding in de praktijk bij te houden welke praktische problemen en twijfels dit oproept, zodat hier een beter beeld van ontstaat. Ook is het van belang om in de praktijk rekening te houden met eventuele weerstand en bijbehorende emoties over afname via videoverbinding bij verdachten en hier aandacht aan te besteden door bijvoorbeeld voldoende uitleg te geven over wat er (niet) van de respondent verwacht wordt.

7 Literatuur

Baigent, M. F., Lloyd, C. J., Kavanagh, S. J., Ben-Tovim, D. I., Yellowlees, P. M., Kalucy, R. S., & Bond, M. J. (1997). Telepsychiatry: “Tele” yes, but what about the “psychiatry”? *Journal of Telemedicine and Telecare*, 3(SUPPL. 1), 3–5. <https://doi.org/10.1258/1357633971930346>

Bland, J. M., & Altman, D. G. (1999). Measuring agreement in method comparison studies. *Statistical Methods in Medical Research*, 8(2), 135–160. <https://doi.org/10.1177/096228029900800204>

Bujang, M. A., & Baharum, N. (2017). A simplified guide to determination of sample size requirements for estimating the value of intraclass correlation coefficient: a review. *Archives of Orofascial Sciences*, 12(1), 1–11. Retrieved from http://www.dental.usm.my/aos/docs/Vol_12/aos-article-0246.pdf

- Cohen's kappa - WikiStatistiek. (n.d.). Retrieved April 2, 2020, from https://wikistatistiek.amc.nl/index.php/Cohen%27s_kappa#Referenties
- Fatehi, F., Armfield, N. R., Dimitrijevic, M., & Gray, L. C. (2015). Technical aspects of clinical videoconferencing: a large scale review of the literature. *Journal of Telemedicine and Telecare*, 21(3), 160–166. <https://doi.org/10.1177/1357633X15571999>
- Gobin, C. M., Banks, J. B., Fins, A. I., & Tartar, J. L. (2015). Poor sleep quality is associated with a negative cognitive bias and decreased sustained attention. *Journal of Sleep Research*, 24(5), 535–542. <https://doi.org/10.1111/jsr.12302>
- Goyal, S., Temple, V., Sawanas, C., & Brown, D. (2020). Cognitive profile of adults with intellectual disabilities from indigenous communities in Ontario, Canada. *Journal of Intellectual & Developmental Disability*, 45(1), 59–65. <https://doi.org/10.3109/13668250.2018.1470160>
- Hersh, W., Helfand, M., Wallace, J., Kraemer, D., Patterson, P., Shapiro, S., & Greenlick, M. (2002). A systematic review of the efficacy of telemedicine for making diagnostic and management decisions. *Journal of Telemedicine and Telecare*, 8(4), 197–209. <https://doi.org/10.1258/135763302320272167>
- Jacobsen, S. E., Sprenger, T., Anderson, S., & Krogstad, J.-M. (2003). Neuropsychological assessment and telemedicine: A preliminary study examining the reliability of neuropsychology services performed via telecommunication. *Journal of the International Neuropsychological Society*, 9(3), 472–478. <https://doi.org/10.1017/s1355617703930128>
- Kaal, H. L., Nijman, H. L. I., & Moonen, X. M. H. (2015). Identifying offenders with an intellectual disability in detention in The Netherlands. *Journal of Intellectual Disabilities and Offending Behaviour*, 6(2), 94–101. <https://doi.org/10.1108/JIDOB-04-2015-0008>
- Kaal, H., Nijman, H., & Moonen, X. (2015). Screener voor intelligentie en licht verstandelijke beperking (SCIL) - stand van zaken 1 juli 2015. Retrieved from <https://www.hsleiden.nl/binaries/content/assets/hsl/lectoraten/lvb-en-jeugdcriminaliteit/factsheetjuli2015.pdf>
- Kirkwood, K. T., Peck, D. F., & Bennie, L. (2000). The consistency of neuropsychological assessments performed via telecommunication and face to face. *Journal of Telemedicine and Telecare*, 6(3), 147–151. <https://doi.org/10.1258/1357633001935239>
- Koo, T. K., & Li, M. Y. (2016). A Guideline of Selecting and Reporting Intraclass Correlation Coefficients for Reliability Research. *Journal of Chiropractic Medicine*, 15(2), 155–163. <https://doi.org/10.1016/j.jcm.2016.02.012>
- Landis, J. R., & Koch, G. G. (1977). The Measurement of Observer Agreement for Categorical Data. Source: *Biometrics* (Vol. 33).
- Mullainathan, S., & Shafir, E. (2014). Freeing up Intelligence. *Nature Publishing Group* (Vol. 25). <https://doi.org/10.1038/scientificamericanmind0114-58>
- Nijman, H., Kaal, H., van Scheppingen, L., & Moonen, X. (2018). Development and Testing of a Screener for Intelligence and Learning Disabilities (SCIL). *Journal of Applied Research in Intellectual Disabilities*, 31(1), e59–e67. <https://doi.org/10.1111/jar.12310>
- Nord, G., Rising, K. L., Band, R. A., Carr, B. G., & Hollander, J. E. (2019). On-demand synchronous audio video telemedicine visits are cost effective. *American Journal of Emergency Medicine*, 37(5), 890–894. <https://doi.org/10.1016/j.ajem.2018.08.017>

- Schutte, J. L., McCue, M. P., Parmanto, B., McGonigle, J., Handen, B., Lewis, A., ... Saptono, A. (2015). Usability and Reliability of a Remotely Administered Adult Autism Assessment, the Autism Diagnostic Observation Schedule (ADOS) Module 4. *Telemedicine and E-Health*, 21(3), 176–184. <https://doi.org/10.1089/tmj.2014.0011>
- Temple, V., Drummond, C., Valiquette, S., & Jozsvai, E. (2010). A comparison of intellectual assessments over video conferencing and in-person for individuals with ID: Preliminary data. *Journal of Intellectual Disability Research*, 54(6), 573–577. <https://doi.org/10.1111/j.1365-2788.2010.01282.x>
- Warrens, M. J. (2011). Cohen's linearly weighted Kappa is a weighted average of 2 x 2 Kappas. *Psychometrika*, 76(3), 471–486. <https://doi.org/10.1007/S11336-011-9210-Z>
- Watson, P. F., & Petrie, A. (2010, June 1). Method agreement analysis: A review of correct methodology. *Theriogenology*. Elsevier. <https://doi.org/10.1016/j.theriogenology.2010.01.003>
- WHO Global Observatory for eHealth. (2010). *Telemedicine: opportunities and developments in Member States: report on the second global survey on eHealth*. WHO Global Observatory for eHealth. <https://doi.org/ISSN 2220-5462>
- Wilson, L. S., & Maeder, A. J. (2015). Recent Directions in Telemedicine: Review of Trends in Research and Practice. *Healthcare Informatics Research*, 21(4), 213. <https://doi.org/10.4258/hir.2015.21.4.213>