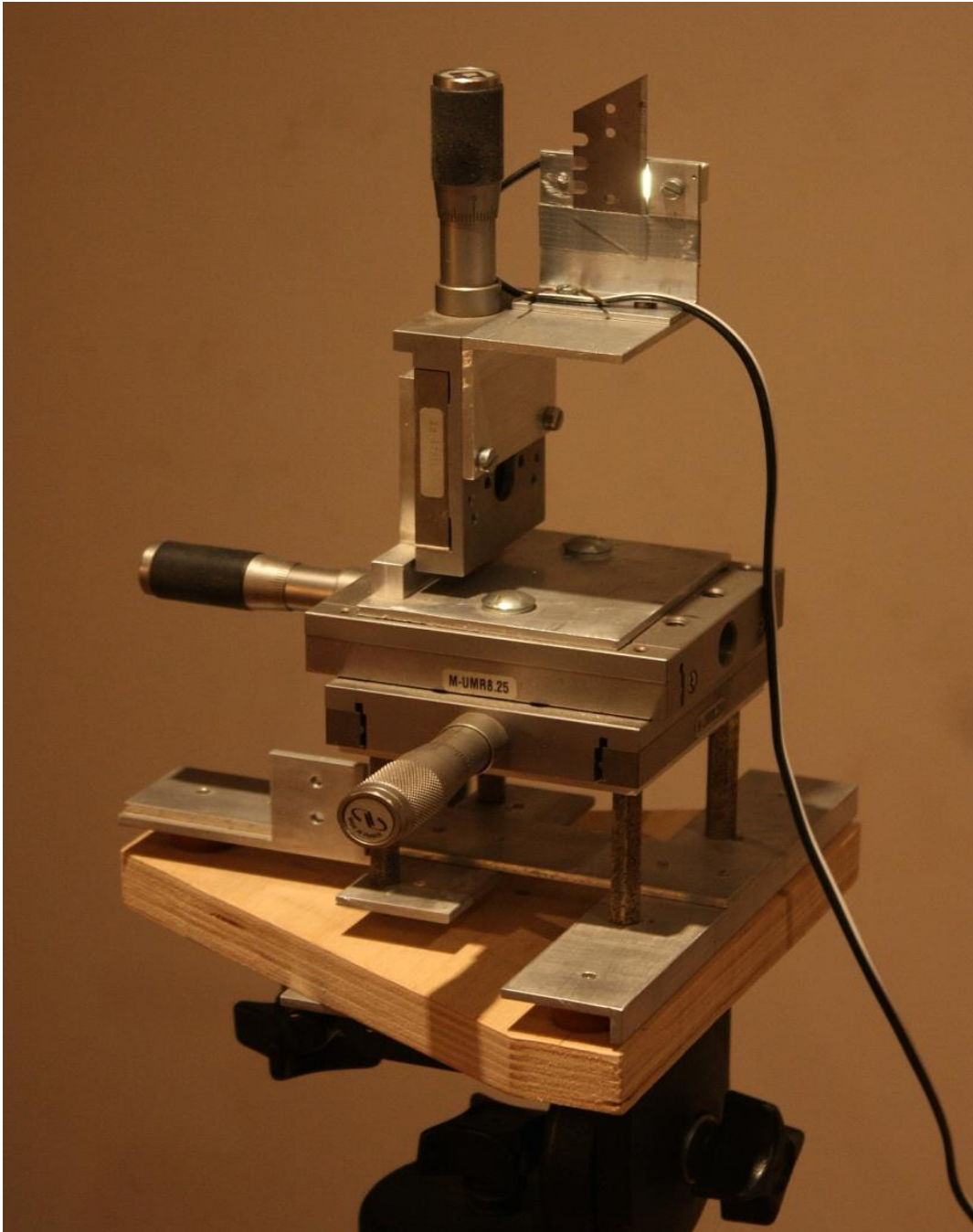
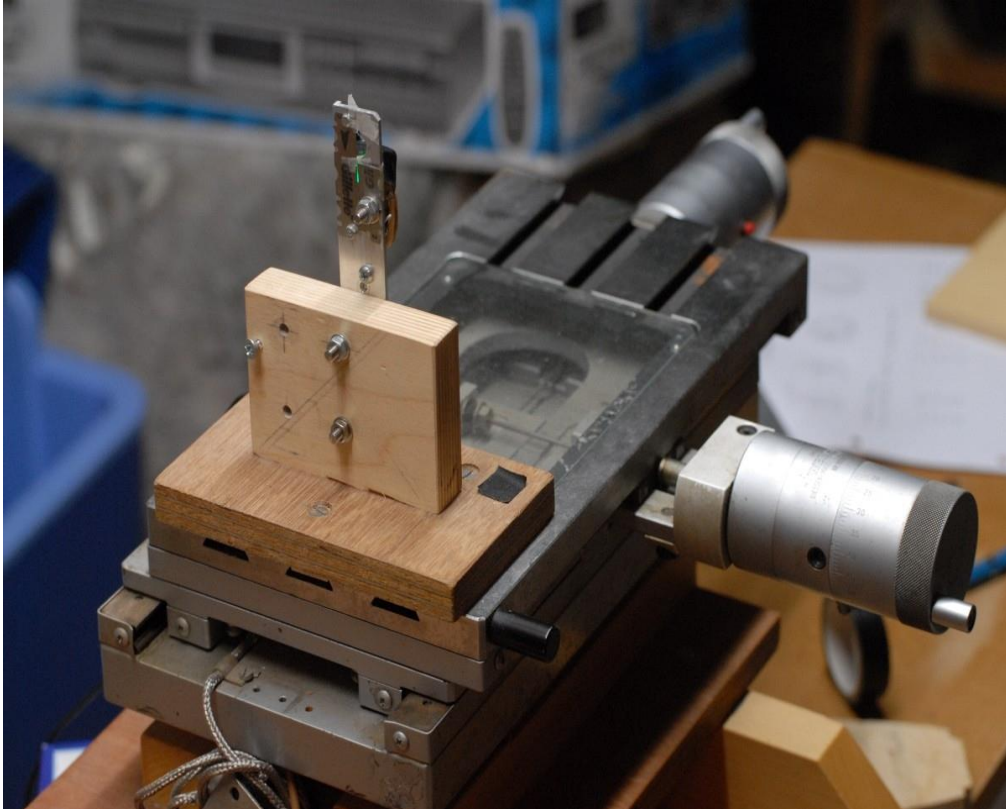


De Foucaulttest



Verslag van een onderzoek naar de Foucaulttest in Nederland en België

Jan van Gastel & Werkgroep Kijkerbouw VSRUG



Foucaulttester Tom Börger

Foucaulttestonderzoek

Opzet: Jan van Gastel

Organisatie: Werkgroep Kijkerbouw VSRUG (Guy en Jean-Pierre) en Jan van Gastel

Vragenlijst, dataverwerking en verslag: Jan van Gastel

Met dank aan:

Jimmy Oostvogels van Astromarket voor het belangeloos ter beschikking stellen van de twee spiegels

Arie Otte voor het kritisch doornemen van een eerdere versie van dit verslag

Xander van Gastel voor correcties van de Nederlandse taal

Alle testers die de moeite hebben genomen de spiegels te testen en het scoringsformulier in te vullen

Foto voorpagina: Foucaulttester Arjan te Marvelde

Inhoudsopgave

Par.	Onderwerp	Pag.
	Voorwoord	1
1	Opzet van het onderzoek	2
1.1	Aard van het onderzoek en pretentie	2
1.2	De vraagstelling van het onderzoek	2
1.3	De deelnemers	4
1.4	De spiegels	4
1.5	Het Coudermasker	4
1.6	Kwantitatieve en kwalitatieve tests	4
1.7	Gebruik van de tabellen 1 t/m 3	6
1.8	Statistisch pakket en significantie	6
1.9	Verklaring van enkele wetenschappelijke termen	7
2	Beoordeling van de spiegels op basis van kwantitatieve en kwalitatieve gegevens	8
2.1	Spiegel A	8
2.2	Spiegel B	11
3	Het onderzoek	16
3.1	Beschrijving van de onderzoeksresultaten: achtergrondinformatie	16
3.2	Beschrijving van het onderzoek: de testresultaten van de deelnemers	20
3.3	Verbanden tussen achtergrondvariabelen en kwaliteitsvariabelen	24
3.4	Verbanden tussen kwaliteitsmaten en eindoordeel van de testers over de spiegels	27
3.5	Kenmerken van de testers die de spiegels als goed of niet goed beoordeelden	28
4	Conclusies	31
4.1	Komen alle testers tot dezelfde resultaten?	31
4.2	Komen de testers tot dezelfde kwantitatieve resultaten als de interferometrische tests?	31
4.3	Overeenkomst van de kwalitatieve resultaten	34
4.4	Het eindoordeel van de testers over de spiegels	35
4.5	Het effect van ervaring	35
5	Adviezen	37
5.1	Astigmatisme	37
5.2	Ruwheid van het oppervlak	37
5.3	De kwaliteit van de rand	38
5.4	Het kwantitatieve meetresultaat	38
	Bijlagen	39

Voorwoord

Aanleiding voor het onderzoek was een discussie op de VVS-mailinglist en het Astroforum over de kwaliteit van telescoopspiegels. In die discussie werd gemeld dat in de VS in het verleden onderzoek is gedaan naar het testen van spiegels¹ en ². Ook in Duitsland, zo bleek later, is een onderzoek naar het testen van spiegels uitgevoerd³. Zo ontstond het idee om dat ook in Nederland en België te organiseren. We hebben ons beperkt tot de Foucaulttest en de resultaten daarvan vergeleken met een interferometrische test. Waarom (alleen) de Foucaulttest?

1. Het is de meest gebruikte test door amateurs die zelf spiegels maken.
2. Als iedereen dezelfde test uitvoert levert dat voor die test meer gegevens op dan wanneer verschillende tests worden gebruikt. Daardoor krijgen de verzamelde gegevens voor die test meer zeggingskracht.

We willen met dit onderzoek niet alleen bekijken hoe betrouwbaar en valide de resultaten van Foucaulttests zijn, maar we willen ook bevorderen dat met de veelgebruikte Foucaulttest test zo nauwkeurig mogelijke resultaten worden behaald, zowel kwantitatief als kwalitatief. Daartoe hebben we op basis van dit onderzoek adviezen opgesteld, die daar hopelijk een bijdrage aan kunnen leveren.

Jan van Gastel

(februari 2015)



Foucaulttestopstelling Ton Posthumus

¹ <http://web.archive.org/web/20130420122705/http://www.atmsite.org/contrib/Eason/RoundRobin/index.html>

² http://web.archive.org/web/20130420122514/http://www.atmsite.org/contrib/Rychnovsky/Round_Robin/RR%20Reports.html

³ <http://marty-atm.de/RoundRobin/RR.html>

1. De opzet van het onderzoek

1.1 Aard van het onderzoek en pretentie

Dit onderzoek is exploratief van aard. Er worden dus geen hypothesen getoetst en er liggen ook geen theorieën aan ten grondslag. Er zijn meer onderzoeken naar dit onderwerp uitgevoerd, maar ons zijn geen onderzoeken naar de Foucaulttest bekend waar zoveel – dertig- testers aan deelnamen. Het onderzoek leverde te veel gegevens op om ‘met de hand’ te verwerken. Daarom is er voor gekozen de gegevens te verwerken met een statistisch pakket ‘OpenStat’.

We hebben niet de pretentie uitspraken te doen die betrekking hebben op alle Foucaulttesters⁴ in de wereld. We gaan er wel van uit dat we met dit onderzoek wat zeggen over het Foucaulttesten in Nederland en België’. Op grond van flink wat ervaring met Foucaulttesten door onszelf en anderen en bestudering van enkele kleinschaliger onderzoeken⁵, zijn we wel van mening dat de geldigheid van veel van wat hier wordt gepresenteerd verder reikt dan de steekproef waar de analyses in dit rapport op zijn uitgevoerd. We hopen daarom, dat het onderzoek aanleiding is tot vergelijkbare onderzoeken met andere testers, liefst ook in andere landen, zodat kan worden nagegaan hoe ver de geldigheid van conclusies uit dit soort onderzoeken strekt. Ook zou het goed zijn meer onderzoek te doen naar andere tests, zoals interferometrische tests, de caustic test en kwalitatieve tests. Vergelijkingen tussen verschillende testtypen zijn ook erg belangrijk, maar zijn niet vaak gedaan en voor zover ze gedaan zijn kenden ze gewoonlijk weinig deelnemers. Ook hier valt zeker nog zinvol werk te verrichten.

1.2 De vraagstelling van het onderzoek

Vraagstelling van het onderzoek is: *hoe betrouwbaar en hoe valide zijn de resultaten van een Foucaulttest*. Het gaat hier om een ‘gewone’ visuele Foucaulttest, dus niet om speciale setups zoals Robo-Foucault, ‘Photoshop-Foucault, of anderszins een Foucaulttest waarbij camera’s of webcams worden gebruikt. Anders geformuleerd luidt de vraagstelling aldus: als testers met behulp van een Foucaulttest, eventueel met hulp van een Ronchitest, eenzelfde spiegel testen:

1. komen zij dan tot dezelfde testresultaten;
2. komen de *kwantitatieve* testresultaten dan overeen met die van een meer objectieve interferometrische test voor de kwantitatieve resultaten;
3. komen ze inzake *kwalitatieve* testresultaten overeen met ons in paragraaf 2.2.5 geformuleerde oordeel⁶;
4. Komen ze dan tot eenzelfde eindoordeel over de spiegel;
5. komen testresultaten van meer ervaren testers dan beter overeen met die van interferometrische tests dan die van minder ervaren testers⁷.

⁴ Hier in de betekenis van ‘personen die testen met behulp van de Foucaulttest’.

⁵ Zie voetnoten pagina 1.

⁶ Vanwege het ontbreken van een algemeen geaccepteerde standaard staat dit wel meer open voor discussie.

⁷ Voor zover het gaat om kwalitatieve maten is dat overigens lastig.

Om dit alles optimaal te kunnen onderzoeken, zouden eigenlijk alle deelnemers aan het onderzoek dezelfde spiegel met dezelfde tester in dezelfde omstandigheden, met hetzelfde Foucaultmasker moeten testen. Dit was helaas niet te organiseren. Wel is door alle testers hetzelfde Foucaultmasker gebruikt.

De gegevens zijn verzameld via een door de deelnemers in te vullen scoringsformulier (bijlage 1), dat vergezeld ging van een toelichting (bijlage 2). Hierin was vermeld dat alle deelnemers *individueel* moesten testen en *niet* met deelnemers over hun resultaten mochten spreken zolang zij of hun gesprekspartner de spiegel nog niet hadden getest. Voor zover we konden nagaan, heeft iedereen zich daaraan gehouden. Ook is gevraagd of zij geen vragen wilden overslaan. Aan dit laatste is helaas niet door iedereen voldaan. Waar relevant, komen we hierop terug.



Foucaulttester Michel Boers

1.3 De deelnemers

De deelnemers aan het onderzoek zijn geworven via de VVS-mailinglijst, via het Astroforum en via persoonlijke contacten. De enige eis om te kunnen meedoen aan de test was een zodanige ervaring met Foucaulttesten, dat men zelfstandig en alleen op grond van de instructies die door de organisatie werden toegezonden de test zou kunnen uitvoeren.

1.4 De spiegels

Er werden twee 20 cm f/5 spiegels getest, belangeloos ter beschikking gesteld door Astromarket (Jimmy Oostvogels). De spiegels waren veraluminiseerd, maar de coating mocht er worden afgehaald. Dit is gebeurd bij de Werkgroep Kijkerbouw VSRUG, in Gent.

Om een eventueel effect van het land van herkomst (België en Nederland) uit te sluiten, of in elk geval te kunnen berekenen, zijn de beide spiegels toen ongeveer de helft van de testers in beide landen een spiegel getest hadden 'de grens over' gestuurd, zodat beide spiegels zowel door testers uit Nederland als door testers uit België konden worden getest. Vanwege wat uitval van testers, is het niet gelukt beide spiegels in beide landen door evenveel personen te laten testen.

1.5 Het Coudermasker

Alle deelnemers hebben getest met behulp van door de organisatie van het onderzoek gemaakte 5-zonemaskers, gemaakt volgens de methode Carlin⁸. De reden hiervan was tweeledig:

- het uitsluiten van verschillen die eventueel ontstaan door verschillende maskers;
- voorkomen dat spiegels lang bij mensen thuis zouden liggen omdat mensen zelf nog een masker moesten maken, wellicht zelfs daarvoor nog spullen moeten kopen, etc.

De spiegels met masker werden via de post van de ene deelnemer naar de volgende gestuurd. De spiegel mocht men een week in bezit houden om te testen. De organisatoren verstrekten de benodigde adresgegevens voor het doorsturen. Elke deelnemer kreeg een eigen code, bestaande uit een letter (N voor Nederland, B voor België en een volgnummer. Deze code diende men op het scoringsformulier in te vullen.

1.6 Kwantitatieve en kwalitatieve tests

De testresultaten met betrekking tot *sferische aberratie* en (voor zover dat voorkomt) *astigmatisme* zijn interferometrisch *kwantitatief* verzameld. Voor beide fouttypen levert dat dus een getalsmatige uitkomst op. Ook de Foucaulttest levert een getal op voor de kwaliteit van de vorm van de spiegel, maar alleen voor *sferische aberratie*⁹ en niet voor astigmatisme. Met een interferometrische test kunnen de verschillende fouttypen uit elkaar worden getrokken en dus los van elkaar bekeken worden, met een Foucaulttester niet. De kwaliteit van de uiterste spiegelrand en de ruwheid van het oppervlak worden in dit

⁸ <http://web.telia.com/~u41105032/couder/zcalc.htm>

⁹ Hoewel het in de praktijk vaak geen glad verlopende sferische aberratie is, maar er andere fouten 'doorheen' zitten.

onderzoek *kwalitatief* vastgesteld, met Foucault- en Ronchitesters. Ook het vaststellen van *astigmatisme* met Foucault- en/of Ronchitester gebeurt kwalitatief. Dat laatste is een serieuze tekortkoming van deze test, want astigmatisme is een veelvoorkomende fout die vaak flinke nadelige gevolgen heeft voor de beeldkwaliteit. Met de Foucaulttest en Ronchitest is astigmatisme alleen te zien als het vrij veel is, te veel om nog van een goede spiegel te kunnen spreken. En zelfs als het niet in Foucault- of Ronchitest te zien is, kan het de Strehlratio nog aardig verlagen. Een PV-wavefrontfout voor astigmatisme van $1/6$ lambda bijvoorbeeld, verlaagt de Strehlratio al met 5% en wordt door vrijwel niemand opgemerkt in Foucault- en Ronchitests. De redenering: *'als je astigmatisme maar net kunt zien is maar weinig'* gaat dus helaas niet op. Ook niet voor de Ronchitest die de deelnemers als extra test mochten inzetten.

Het oordeel over hoe goed de rand is, hoe ruw het oppervlak is en of er veel of weinig astigmatisme in een spiegel zit is dus een *subjectieve inschatting*. Ook de kwantitatieve Foucaulttest kent een subjectieve component, omdat men 'op het oog' grijstinten moet vergelijken. Ten slotte zit er ook een subjectieve component in het bepalen welke getallen of observaties aangeven dat een spiegel 'slecht', 'matig' of 'goed' is.

Tabel 1 geeft een overzicht van hoe wij de kwaliteit van telescoopspiegels indelen.

classificatie	Strehlratio	PV-wavefrontfout ¹⁰
slecht	< 0.80	> 0.25
matig	0.80-0.85	0.25-0.21
Vrij goed	0.85-0.90	0.21-0.17
Goed	0.90-0.95	0.17-0.125
Zeer goed	0.95-1.00	< 0.125

Tabel 1 : classificatie kwantitatief gemeten spiegelkwaliteit

Merk op, dat in de door ons gehanteerde kwalificatie in tabel 1 'vrij goed' een middencategorie is tussen 'matig' en 'goed' en dus niet hetzelfde betekent als 'goed'.

Moeilijker is het om voor de kwalitatief gemeten aspecten van de kwaliteit van de spiegel: ruwheid en kwaliteit van de rand, zo'n classificatietabel te maken. Omdat we expliciet willen zijn over onze beoordeling van de bevindingen van de testers in het onderzoek, presenteren we hier toch tabellen met betrekking tot die fouttypen¹¹. Deze classificaties zijn gebaseerd op bestudering van discussies over ruwheid op internet. Ze zijn, omdat er niet gemeten wordt, *subjectiever* dan de bovenstaande en zullen dan ook *discutabeler* zijn in die zin, dat

Waargenomen	TDE/TUE	Rand is
Aan een kant heldere diffractierand, andere kant geheel donker	sterk	slecht
Verskil in helderheid diffractierand beide kanten	matig	matig
Geen zichtbaar verschil diffractierand beide kanten	geen	goed

Tabel 2: classificatie kwaliteit rand in Foucaulttester

¹⁰ Dit geldt voor sferische aberratie

¹¹ Deze kwalificaties sluiten aan bij hetgeen in het scoringsformulier werd gevraagd, maar zou verder verfijnd kunnen worden (zie hoofdstuk 5).

sommigen de kwalificaties in de tabellen 2 en 3 voor respectievelijk de *kwiteit van de rand* en de *mate van ruwheid* te streng of juist niet streng genoeg zullen vinden. We zullen hier in onze conclusies rekening mee houden. Algemeen geaccepteerde standaarden voor de beoordeling van de rand en de mate van ruwheid zouden welkom zijn. Ideaal gesproken zou daarvoor bekend moeten zijn hoeveel strooilicht door een bepaalde mate van ruwheid of minder goede rand wordt gegenereerd, omdat daardoor het contrast verlaagd wordt. Tot we dat weten, lijkt het handig *om streng om te gaan* met ruwheid en randkwaliteit, om de kans op negatieve effecten zo klein mogelijk te maken.

In tabel 2 op pagina 5 is de relatie aangegeven tussen het verschil in intensiteit van de diffractierand aan weerszijden van de spiegel en de beoordeling van de kwaliteit van de rand. In de Ronchitester vertaalt zich dat in de grootte en meer of minder duidelijke zichtbaarheid van de haakjes aan de uiteinden van de Ronchilijnen als het Ronchirooster buiten de ROC wordt geplaatst.

Bij de classificatie ‘*zeer ruw*’ (tabel 3 hieronder) is vaak sprake van ‘dog biscuit’: een blokjesachtig patroon over de gehele spiegel, in het Engels wel ‘primary ripple’ genoemd.

Waargenomen	ruwheid	oppervlak is
Direct in het oog springende ruwheid van het oppervlak	(zeer) ruw	slecht
Ruwheid goed te zien bij nauwkeurig kijken	licht ruw	matig
Ruwheid niet of heel moeilijk te zien ¹²	glad	goed

Tabel 3: classificatie mate van ruwheid spiegeloppervlak in Foucaulttester

Ook zeer duidelijk in het oog springende, van binnen naar buiten lopende ‘stralen’ vallen onder ‘(zeer) ruw’.

1.7 Gebruik van de tabellen 1 t/m 3

In de conclusies van dit verslag zullen de interferometrisch door Wolfgang Rohr (sferische aberratie en astigmatisme) en kwalitatief door ons verzamelde informatie (kwaliteit van de rand en mate van gladheid/ruwheid) beoordeeld worden aan de hand van de bovenstaande drie tabellen.

1.8 Statistische pakket en significantieniveau

De gegevens uit de ingevulde scoringsformulieren zijn ingevoerd in het open source statistisch programma ‘OpenStat’ en daarmee verwerkt. Als significantieniveau (zie paragraaf 1.9 voor uitleg en paragraaf 1.1 voor het kader) voor de analyses is een p-waarde van ten hoogste 0.05 gehanteerd. Bij een p-waarde tussen de 0.05 en 0.1 spreken we van een ‘tendens’. De p-waarden staan steeds in de tekst vermeld als over significantie wordt gesproken. Als testresultaten zijn in het onderzoek opgenomen:

1. de Peak-to-Valley (PV) wavefront fout;
2. de Strehlratio;

¹² De alleen met een Lyottester of geavanceerde professionele apparatuur zichtbaar te maken microruwheid laten we hier buiten beschouwing.

3. een subjectieve inschatting van de mate van astigmatisme in de spiegel;
4. een subjectieve inschatting van de mate van ruwheid van het spiegeloppervlak;
5. een subjectieve inschatting van de kwaliteit van de uiterste rand van de spiegel.

1.9 Verklaring van enkele in dit rapport gebruikte wetenschappelijke termen

Omdat wellicht niet iedereen gewend is verslagen van onderzoek te lezen, willen we hier enkele in het onderzoek voorkomende termen verhelderen.

1. Variabele(n)

In de tekst wordt steeds gesproken van 'een variabele' of 'variabelen'. Wat zijn dat? *Een variabele is 'een grootte die niet constant is, maar variabel'.* Zo is in dit onderzoek 'spiegel' een variabele, omdat die twee waarden kan aannemen: 1 en 2, namelijk voor spiegel A en spiegel B. Eenzelfde redenering geldt ook voor 'Land'. De twee waarden zijn in dit geval 'Nederland' en 'België', gecodeerd als 1 en 2. Uiteraard zijn er ook variabelen met meer dan twee waarden, bijvoorbeeld 'Strehl' (de Strehlratio). Deze kan in theorie alle waarden aannemen tussen 0 en 1. Een laatste voorbeeld: de variabele 'ervaring' kan in dit onderzoek de waarden 1, 2 en 3, ofwel geen ervaring, gemiddelde ervaring of veel ervaring aannemen.

We onderscheiden **onafhankelijke** en **afhankelijke** variabelen. *Onafhankelijke* variabelen zijn de variabelen waarvan we de invloed op de *afhankelijke* variabele willen berekenen. Voorbeeld: Als we willen weten op grond van welke kwaliteitsmaten de testers gekomen zijn tot hun eindoordeel (paragraaf 3.2.5) dan is 'oordeel' de afhankelijke variabele en zijn de variabelen 'Strehlratio', 'PV-fout', 'astigmatisme', 'ruwheid' en 'rand' de onafhankelijke variabelen.

2. Significant(ie)

Als er verbanden worden gevonden tussen verschillende variabelen kan dat een werkelijk verband zijn, maar het kan ook toeval zijn, bijvoorbeeld veroorzaakt doordat de onderzoeksgroep (hier de mensen die de spiegel getest hebben) op een bepaalde manier, gewild of ongewild, was samengesteld. Om te voorkomen dat besloten wordt tot verbanden die bij analyse toevallig in de onderzoeksgroep worden gevonden, maar in werkelijkheid niet bestaan, moet een verband een bepaalde sterkte hebben om als zodanig te worden geaccepteerd. De te accepteren kans dat het gevonden verband op 'toeval' berust wordt daarom vooraf vastgesteld. Het is vrij gebruikelijk om hiervoor *een waarde van 0.05* aan te houden. Je kunt dan zeggen dat *'de kans dat een bepaald verband tussen variabelen toevallig wordt gevonden 5% is'*. Uiteraard sluit je daarmee de kans op toeval niet geheel uit, maar je perkt die kans wel in. Je zou ook kunnen kiezen voor 1% of zelfs 0.1%. Hier hebben we echter gekozen voor de grens van 0.05. Als de gevonden kans, in de tekst genoemd als p-waarde (p staat voor probability) 0.05 is of kleiner, noemen we het verband 'significant' en gaan we er van uit dat het om een reëel verband gaat. Significant

betekent nog niet dat het verband ook inhoudelijk of theoretisch zinvol of relevant hoeft te zijn. Of dat zo is zal inhoudelijk beredeneerd moeten worden.

2. Beoordeling van de spiegels op basis van kwantitatieve en kwalitatieve gegevens

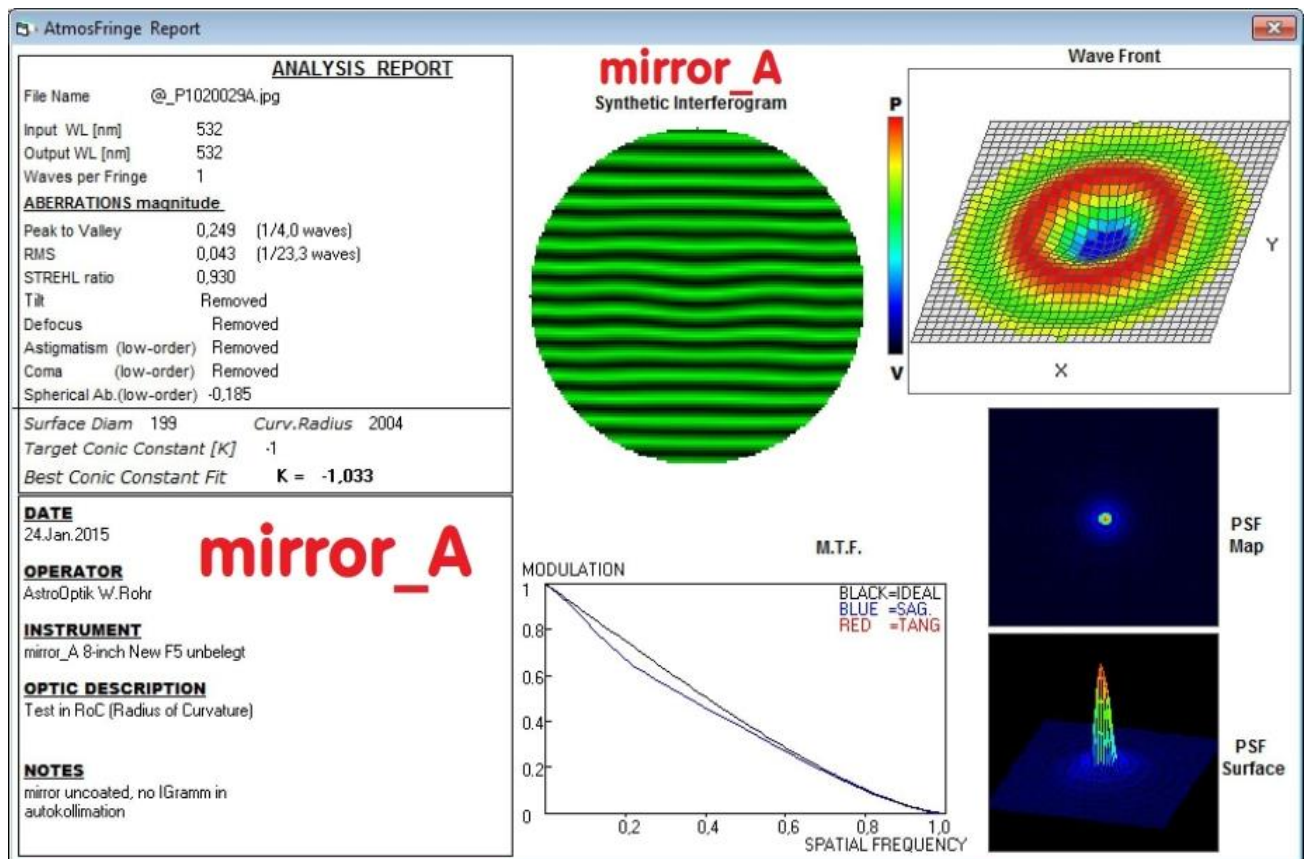
Zoals reeds vermeld, zijn er twee 20 cm f/5 spiegels getest. Om naast de Foucaulttestresultaten ook *objectieve* testresultaten ter vergelijking te kunnen laten zien, zijn de beide spiegels nadat iedereen had getest, opgestuurd naar Wolfgang Rohr, in Duitsland. Hem werd verzocht de volgende testinformatie te leveren:

1. de oppervlaktenauwkeurigheid – Strehlratio en PV-fout – van de spiegels;
2. de P-V wavefrontfout van het astigmatisme (indien aanwezig);
3. de P-V wavefrontfout van sferische aberratie;
4. de kwaliteit van de rand van de spiegel (kwalitatief);
5. de oppervlakteruwheid (kwalitatief) getest met Foucault- en/of Lyottest.

Omdat de spiegels niet veraluminiseerd waren kon Wolfgang Rohr geen Foucault- en Lyottest gebruiken (te weinig licht), waardoor punt 4 en 5 niet door hem konden worden gerealiseerd. Daarom hebben we dat zelf gedaan met behulp van de Foucaulttest. Aan de hand van de foto's is onze beoordeling van de rand en het spiegeloppervlak door elke lezer van dit rapport te controleren.

2.1 Spiegel A

2.1.1 Interferometrische testresultaten



Figuur 1: Strehlratio en P-V wavefrontfout sferische aberratie van spiegel A (Rapport Wolfgang Rohr)

Uit de interferometrische analyse van Rohr (zie figuur 1) komt een Strehlratio van 0.93 en een P-V-wavefrontfout van $1/5.4 \lambda$. Dat is alleen sferische aberratie, want eventuele andere fouttypen zijn in de analyse niet meegenomen. Rohr had op kwalitatieve gronden¹³ geconcludeerd dat er 'geen significant' astigmatisme in de spiegel aanwezig was. Dit bleek in overeenstemming met onze eigen kwantitatieve interferometrische testresultaten.

2.1.2 Resultaten kwalitatieve Foucault- en Ronchitest

Om een goed beeld van de uiterste rand van de spiegel en van de mate van gladheid/ruwheid van het oppervlak te krijgen, hebben we de spiegel door een Ronchitester en een Foucaulttester bekeken. De resultaten daarvan zullen hieronder besproken worden.

Foto 1, genomen door de Ronchitester, is genomen buiten de radius of curvature (ROC). Op de foto is het volgende de te zien:

1. Een enigszins afgevalven rand (TDE). Dat is te zien aan de iets naar buiten buigende haakjes aan de uiteinden van de gebogen lijnen.
2. Lichte ruwheid van het spiegeloppervlak. Dit blijkt uit de wat rafelig aandoende randen van de ronchilijnen.

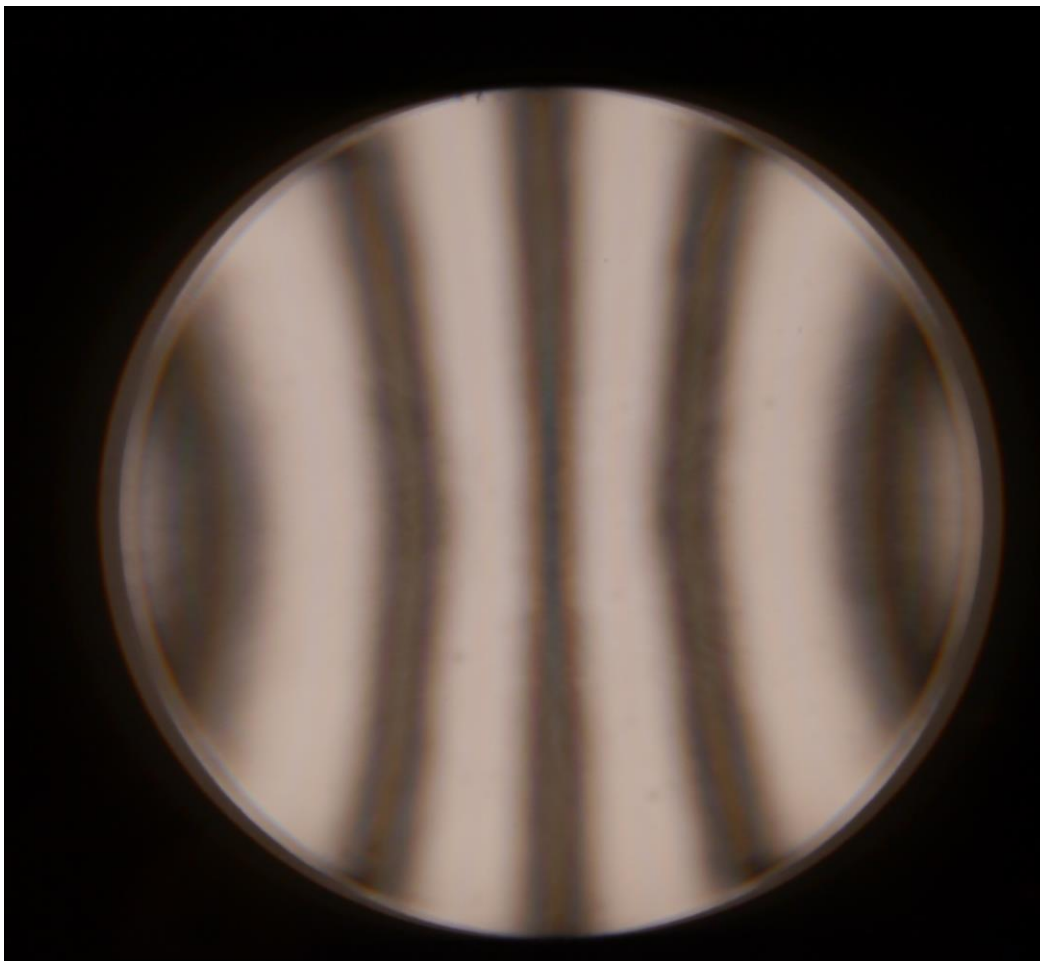


Foto 1: Ronchitester spiegel A (Jan van Gastel)

¹³ Hoewel wij om kwantitatieve gegevens hadden gevraagd

1. Een te ver uitgediept centrum van de spiegel. Dit is te zien aan het feit dat de ronchilijnen (vanaf links geteld) 2 en 4 in het centrum van de spiegel sterker gebogen zijn dan buiten het centrum en aan het feit dat de middelste Ronchilijn in het centrum veel dunner is dan verder naar buiten toe.

Alvorens dit te interpreteren kijken we naar foto 2, genomen door een Foucaulttester, die op een andere manier dezelfde fouten in beeld brengt. We zien het volgende:

1. Een licht afgevallen rand. Dat is te zien aan het verschil tussen de lichte lijn aan de linkerkant van de spiegel en de lichte lijn aan de rechterkant van de spiegel. Als de rand goed is, zijn beide lijnen even helder. Als de afgevallen rand zeer sterk is, is de rand aan de linkerzijde nog helderder en valt het lichte randje aan de rechterkant helemaal weg.
2. Ruwheid van het spiegeloppervlak die we willen plaatsen tussen 'licht ruw' en '(zeer) ruw' (tabel 3) in zit. We zouden dit 'matig ruw' kunnen noemen. Het is te zien over de gehele spiegel (alleen in het donkere deel valt het weg in de schaduw).
3. Een te ver uitgediept centrum.

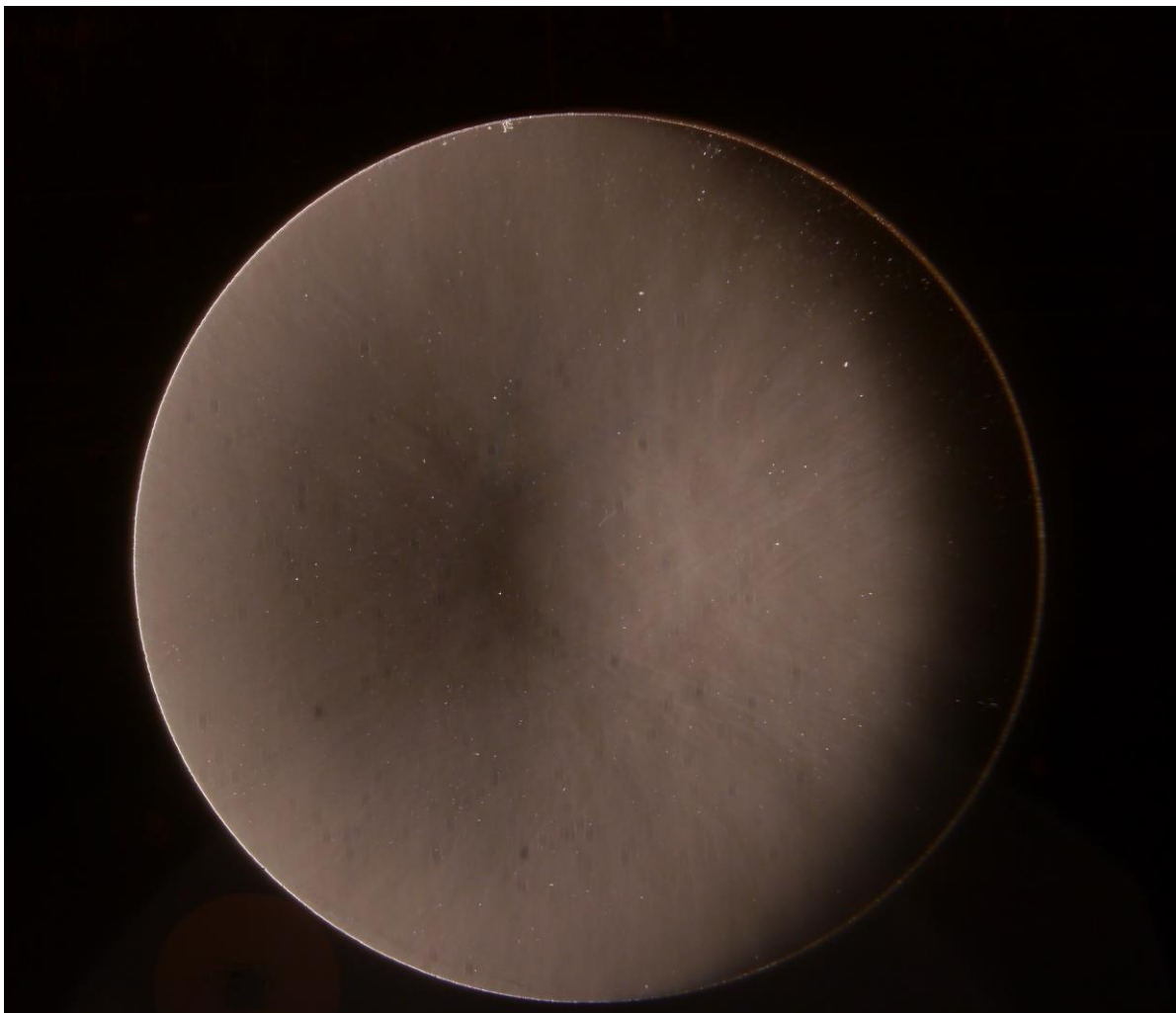


Foto 2: Foucaulttest spiegel A (Jan van Gastel)

2.1.4 Hoe ernstig zijn deze defecten?

De Ronchi- en Foucaulttest leveren hetzelfde beeld op: een afgevalven rand, een ruw oppervlak en een 'gat' in het centrum.

Om met het laatste defect – het te ver uitgediepte centrum - te beginnen: dit is niet van groot belang. Het te ver uitgediepte deel is niet groot (qua diameter) en zal voor het grootste deel – bij een wat grotere vangspiegel zelfs helemaal - schuil gaan onder de vangspiegel, waardoor het effect op de beeldvorming verwaarloosbaar is.

Het deel van de rand dat is afgevalven (sterk naar beneden gebogen) ligt in de orde van grootte van 1 millimeter of iets minder. Een randje van 1 mm vormt 2% van het oppervlak van de spiegel. Het verstrooit licht en kan daardoor verlies aan contrast veroorzaken.

Het oppervlak van de spiegel is 'matig ruw'. Ook hierdoor wordt het licht verstrooid en ook dit veroorzaakt enig contrastverlies.

2.1.5 Conclusie ten aanzien van spiegel A

Op grond van de Strehlratio van de interferometrische analyse alleen, zouden we moeten besluiten dat de spiegel 'goed' is (zie tabel 1). Op grond van de *licht afgevalven rand* en het *matig ruwe oppervlak* zullen we dit oordeel echter moeten bijstellen. Het 'gat' in het centrum speelt geen rol, want dat gaat zo goed als geheel schuil onder de schaduw van de vangspiegel. Alles bij elkaar genomen classificeren we deze spiegel daarom een klasse lager dan goed. Daarmee komt deze spiegel terecht in de middencategorie 'vrij goed', zoals gedefinieerd in tabel 1 van dit verslag.

2.2 Spiegel B

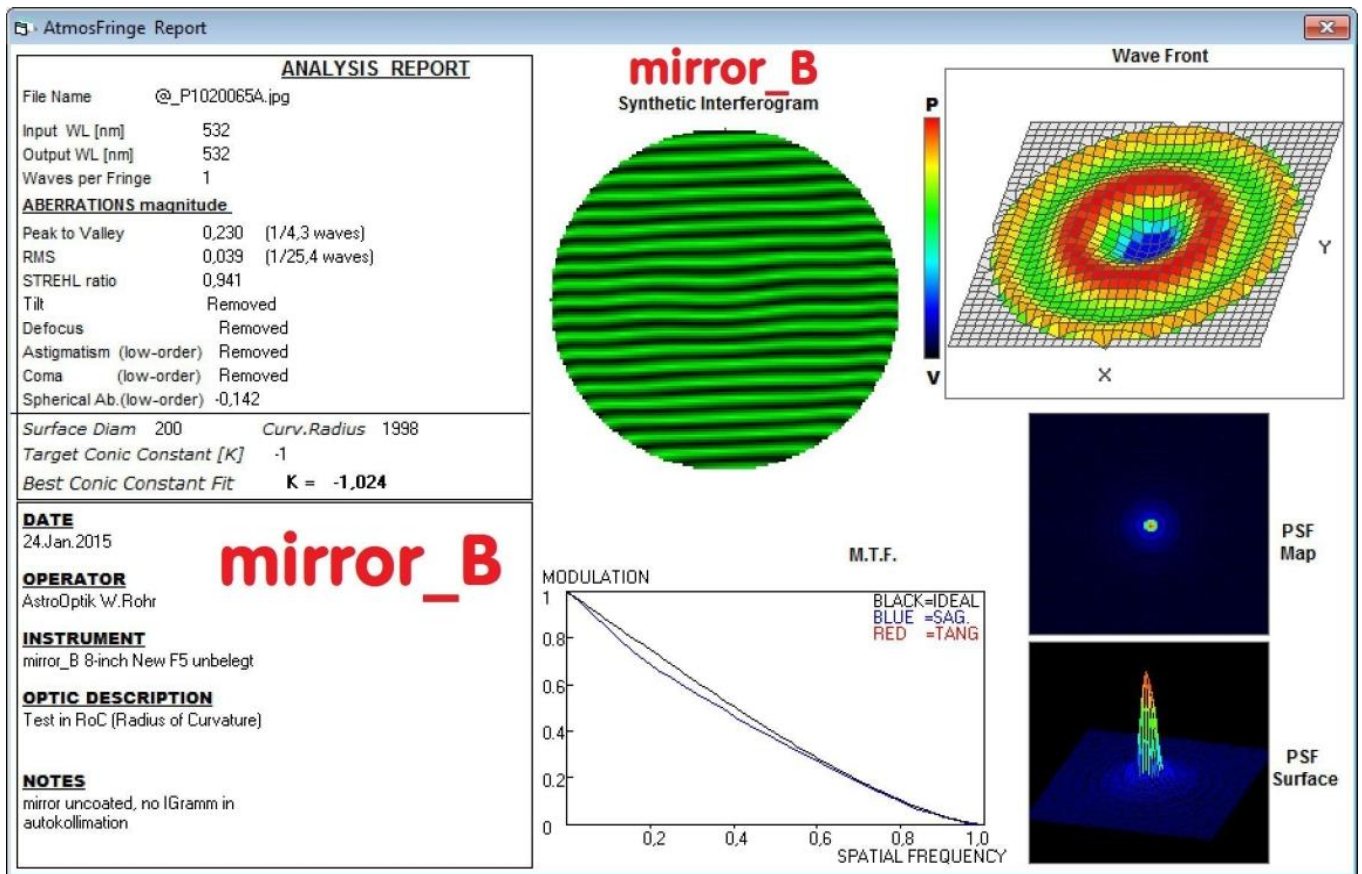
2.2.1 Interferometrische testresultaten

De interferometrische testresultaten (figuur 2 hieronder) van Wolfgang Rohr leveren een Strehlratio op van 0.94 en een P-V wavefrontfout van $1/7$ lambda. Het gaat ook hier alleen om *sferische aberratie*, omdat andere fouttypen niet waren meegenomen in de analyse. Volgens de *kwalitatieve* test¹⁴ voor astigmatisme van Rohr zat er ook in deze spiegel geen astigmatisme. Omdat wij graag een *kwantitatieve* meting van astigmatisme wilden hebben, is Rohr verzocht een kwantitatieve analyse aan te leveren. Deze werd daarna ook door hem aangeleverd en daaruit bleek dat er wel degelijk astigmatisme in de spiegel aanwezig was¹⁵ en wel ongeveer $1/1.5$ ofwel 0.67 lambda PV-wavefront¹⁶ ligt (figuur 3). In figuur 3 is aan de blauwe en rode kleuren goed te zien dat er inderdaad astigmatisme aanwezig is.

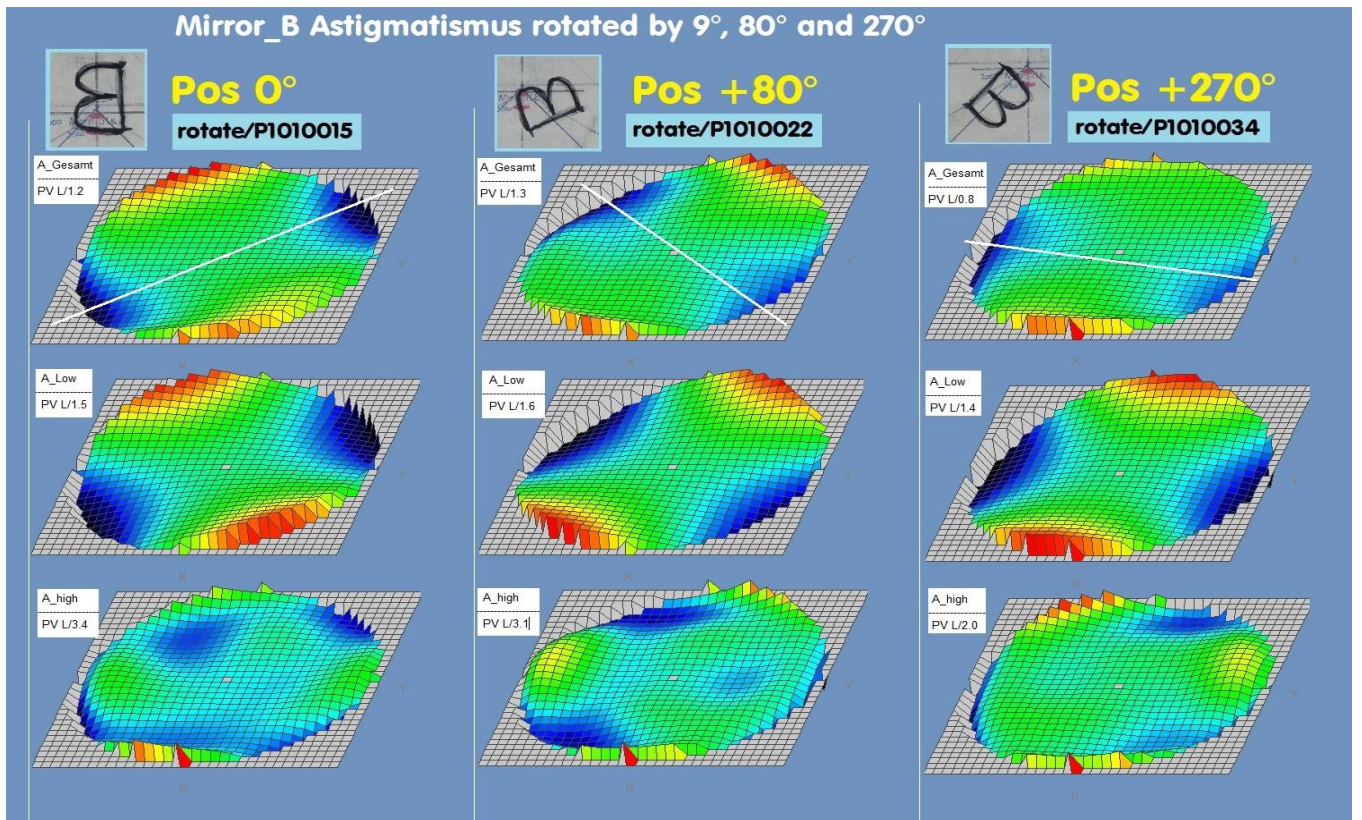
¹⁴ Inschatten van de mate van 'rondheid' van de 'bull's-eye's van de interferogrammen.

¹⁵ Waaruit blijkt dat we voor het vaststellen van astigmatisme heel voorzichtig moeten zijn met dit kwalitatief te doen. Luchtbewegingen en andere omgevings- en instrumentele zaken leiden gemakkelijk tot een verkeerde interpretatie, terwijl die er met een interferometer kwantitatief goed uit te filteren zijn.

¹⁶ Twee andere interferometrische analyses (met verschillende interferometers) leverden vergelijkbare waarden, namelijk $1/1.6$ en $1/1.7$ lambda astigmatisme op.



Figuur 2: Strehlratio en De P-V wavefrontfout sferische aberratie van spiegel B (Rapport Wolfgang Rohr).



Figuur 3: Astigmatisme spiegel B, getest op $\pm 1/1.5$ lambda P-V wavefront (Rapport Wolfgang Rohr).

2.2.2 Resultaten kwalitatieve Foucault- en Ronchitest

We hebben ook deze spiegel met de Foucaulttester en Ronchitester bekeken om de mate van gladheid/ruwheid van het oppervlak en de kwaliteit van de uiterste rand te bekijken (zie de twee onderstaande foto's).

Foto 3, genomen door de Ronchitester, is genomen buiten de radius of curvature (ROC). Op de foto is het volgende de te zien:

1. Een afgevalen rand (TDE). Dat is te zien aan de naar buiten buigende haakjes aan de uiteinden van de gebogen lijnen.
2. Ruwheid van het spiegeloppervlak. Dit blijkt met name uit de wat rafelig aandoende randen van de ronchilijnen.

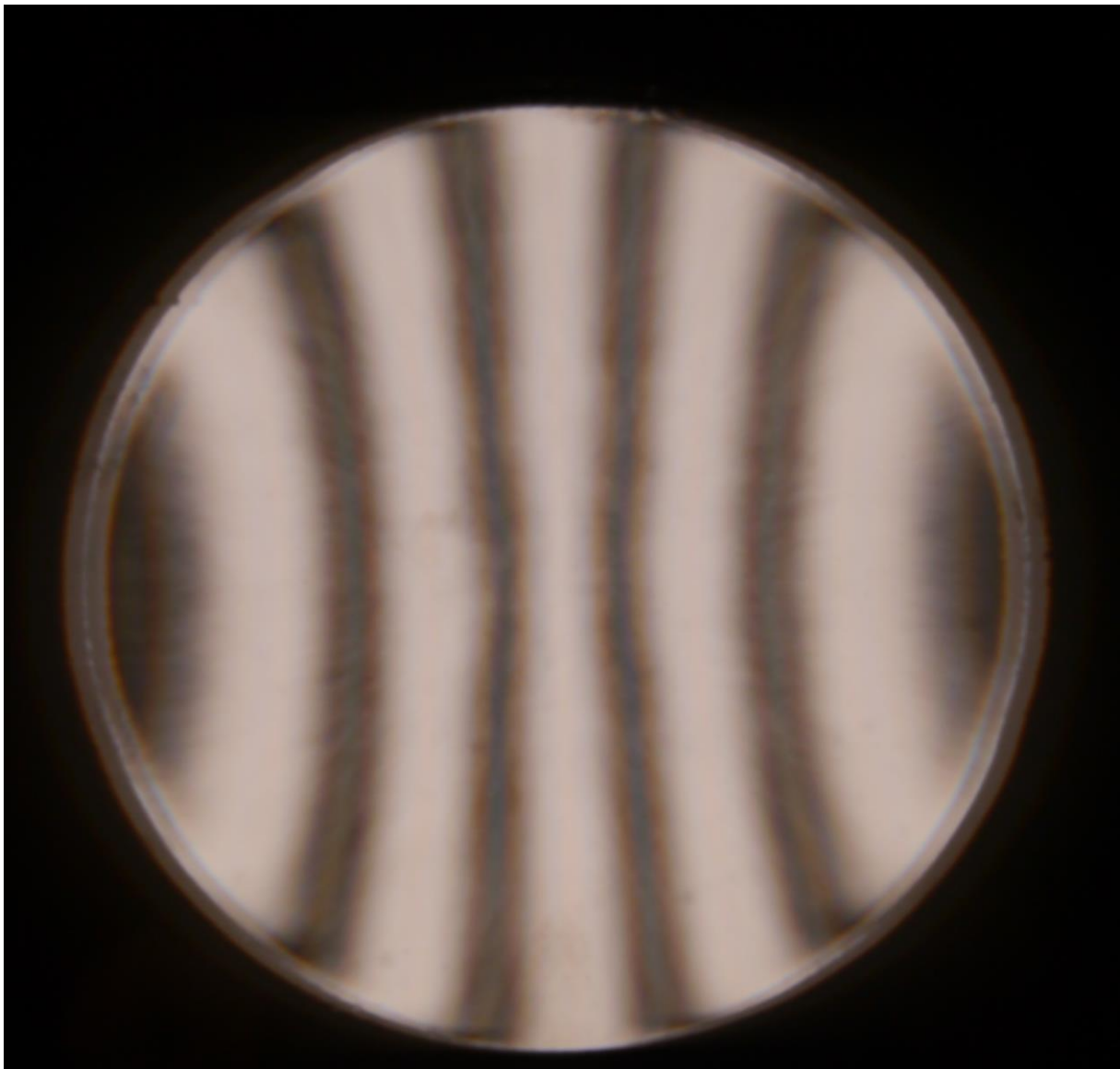


Foto 3: Ronchifoto spiegel B (jan van Gastel)

3. Een te ver uitgediept centrum van de spiegel. Dit is te zien aan het feit dat de twee binnenste ronchilijnen in het centrum van de spiegel sterker gebogen zijn dan buiten het centrum.

Alvorens dit te interpreteren, gaan we net als bij spiegel A naar de foto kijken die door een Foucaulttester is gemaakt en die op een andere manier dezelfde fouten in beeld brengt: foto 4.

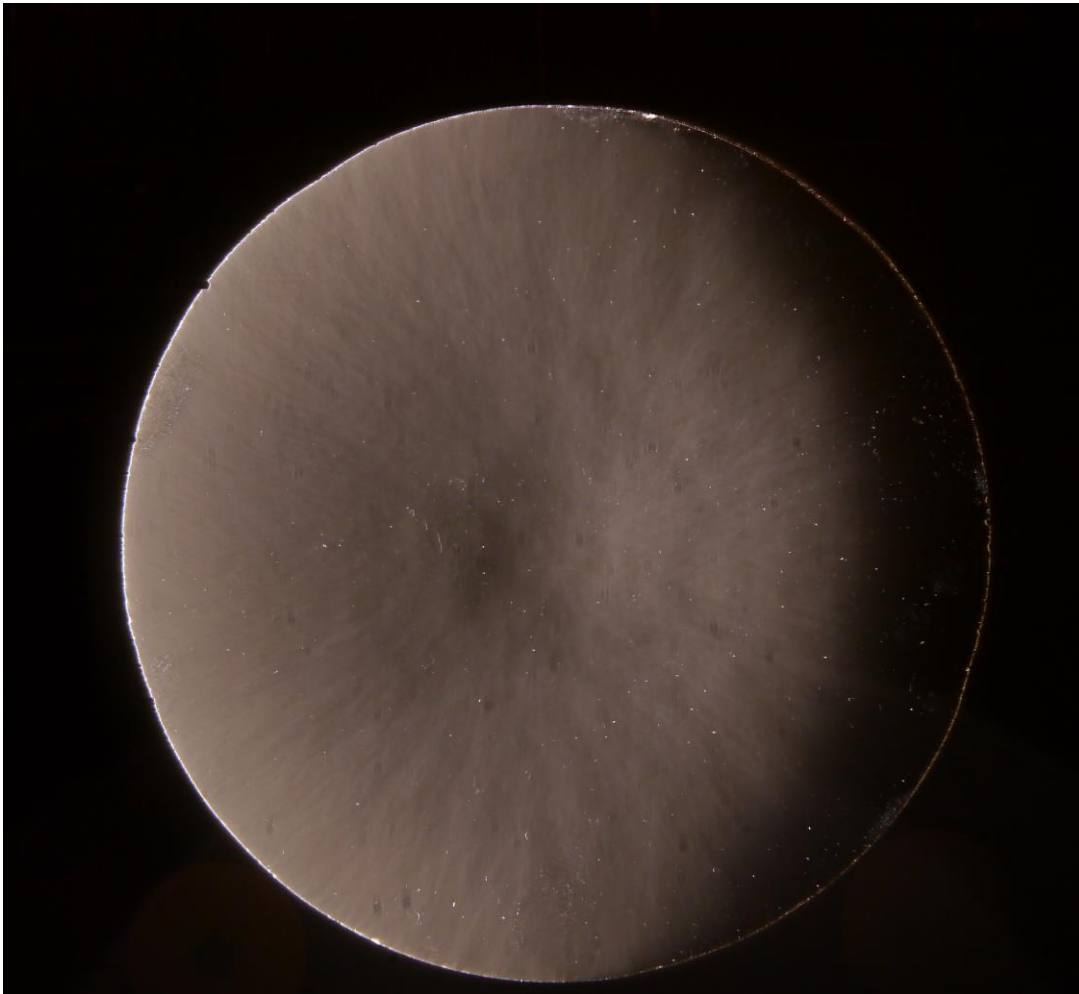


Foto 4: Foucaulttest spiegel B (Jan van Gastel)

Op foto 4 zien we het volgende:

1. Een matig (zoals gedefinieerd in tabel 2) afgevalen rand, dus iets meer dan bij spiegel A het geval is.
2. Een flinke ruwheid van het spiegeloppervlak, te classificeren als 'ruw' (zie tabel 3 voor de definiëring) is te zien over de gehele spiegel (alleen in het donkere deel valt het weg in de schaduw).
3. Een te ver uitgediept centrum.

De - kwalitatieve - Ronchi- en Foucaulttest tests leveren dus voor deze spiegel hetzelfde beeld op als voor spiegel A: een afgevalen rand, een ruw oppervlak en een 'gat' in het centrum. Wel zijn ruwheid en afgevalen rand ernstiger dan bij spiegel A het geval was.

2.2.3 Hoe ernstig zijn deze defecten?

Het te ver uitgediepte centrum is ook hier niet van groot belang. Het te ver uitgediepte deel is niet groot (qua diameter) en zal voor het grootste deel schuil gaan onder de vangspiegel, waardoor het effect op de beeldvorming verwaarloosbaar is.

Het deel van de rand dat is afgevalen (sterk naar beneden gebogen) is klein, geschat in de orde van grootte van 1-2 millimeter. Een randje van 1 mm vormt bij een 20 cm spiegel al 2% van het oppervlak van de spiegel. Het verstrooit het licht en kan daardoor verlies aan contrast veroorzaken.

Het spiegeloppervlak is ruw. Ook hierdoor wordt het licht verstrooid en ook dit veroorzaakt contrastverlies.

2.2.4 Conclusie met betrekking tot spiegel B.

De vormnauwkeurigheid zoals berekend in de analyse van Rohr, met een Strehlratio van 0.94, zou 'goed' zijn, ware het niet dat de spiegel zoals hierboven vermeld, meer dan $1/2$ lambda PV-wavefront astigmatisme bevat. Dit feit alleen al, maakt de spiegel tot een slechte spiegel. Inclusief het astigmatisme zakt de Strehlratio van de spiegel van 0.94 naar onder de 0.60. Daar komt dan nog het effect van de *matig afgevalen rand* en het *ruwe* oppervlak bij. Eindconclusie is dan ook dat deze spiegel niet verder komt dan de in tabel 1 gedefinieerde kwalificatie 'slecht'.

2.2.5 Kwaliteitsgegevens en eindoordeel met betrekking tot beide spiegels samengevat

spiegel	strehl	sferische aberratie	astigmatisme	rand	ruwheid	oordeel
1	0.93	$1/5.4 \lambda$	geen	licht TDE	matig ruw	vrij goed
2	0.94	$1/7 \lambda$	$1/1.5 \lambda$	matig TDE	ruw	slecht

Tabel 4: Samenvatting kwaliteitsgegevens spiegel A en B

In tabel 4 zijn de conclusies van beide spiegels samengevat. Verderop in dit verslag zullen we de bevindingen van de deelnemers aan het onderzoek vergelijken met de conclusies uit de tabel hierboven.

3. Het onderzoek

3.1 Beschrijving van de onderzoeksresultaten: achtergrondinformatie

3.1.1 De deelnemers: land van herkomst en geteste spiegel

Land	Spiegel A	Spiegel B	totaal
Nederland	6	6	12
België	5	13	18
totaal	11	19	30

Tabel 5: land van herkomst deelnemers Foucaulttestonderzoek

Er namen 12 testers uit Nederland en 18 uit België deel aan het onderzoek. In totaal waren er dus 30 deelnemers. Spiegel A werd door 11, spiegel B door 19 deelnemers getest.

3.1.2 De ervaring van de deelnemers

In het onderzoek is gevraagd naar de ervaring van de deelnemers met het testen van spiegels. Dit werd als volgt geoperationaliseerd:

- Weinig ervaring: niet meer dan een spiegel getest.
- Gemiddelde ervaring: twee spiegels getest of een spiegel getest en daarnaast anderen geholpen met testen of samen met iemand spiegels getest.
- Veel ervaring: meer dan twee spiegels getest, of twee spiegels getest en anderen geholpen met testen of anderen leren testen.

Land	Weinig ervaring	Gemiddelde ervaring	Veel ervaring	totaal
Nederland	2	3	7	12
België	7	8	3	18
totaal	9	11	10	30

Tabel 6: testervaring van de deelnemers.

Over het geheel genomen is, zoals in tabel 6 is te zien, de ervaring tamelijk evenredig verdeeld over de drie categorieën. Wel was het ervaringsniveau van de Nederlandse testers significant hoger dan dat van de Belgische testers ($p=0.03$). We zullen verderop nagaan of dit in verband staat met de testresultaten.

3.1.3 Waar heeft men de spiegel getest?

Niet alle testers beschikten zelf over een Foucaulttester, maar hebben bij iemand anders thuis of op een sterrenwacht de test uitgevoerd.

Land	thuis	Bij sterrenwacht	Bij iemand anders	totaal
Nederland	10	2	0	12
België	4	9	5	18
totaal	14	11	5	30

Tabel 7: waar heeft men de spiegel getest?

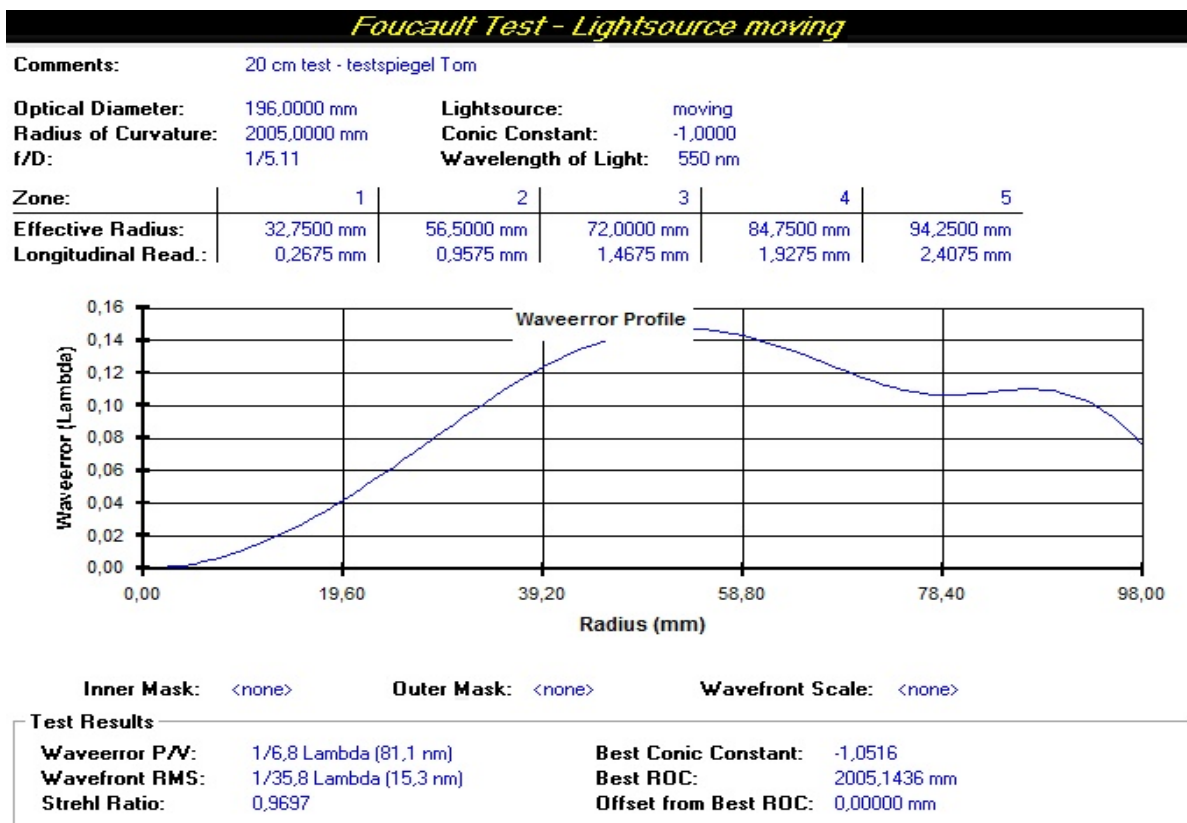
In tabel 7 zien we, dat in België de meeste deelnemers bij een sterrenwacht hebben getest en in Nederland de meeste mensen thuis. Alleen in België heeft een aantal mensen bij iemand anders thuis getest: zes personen, die allemaal bij een van hen thuis hebben getest.

3.1.4 De gebruikte Foucaulttester: fixed source of moving source?

Land	Fixed source	Moving source	totaal
Nederland	1	11	12
België	16	2	18
totaal	17	13	30

Tabel 8: Fixed source of moving source

In tabel 8 zien we dat de Belgische deelnemers op twee na allemaal een 'fixed source' tester gebruikten, de Nederlanders op een na allemaal een 'moving source' tester.



Figuur 4: Profielweergave spiegel A, Foucaulttest Analysis (foto Tom Börger)

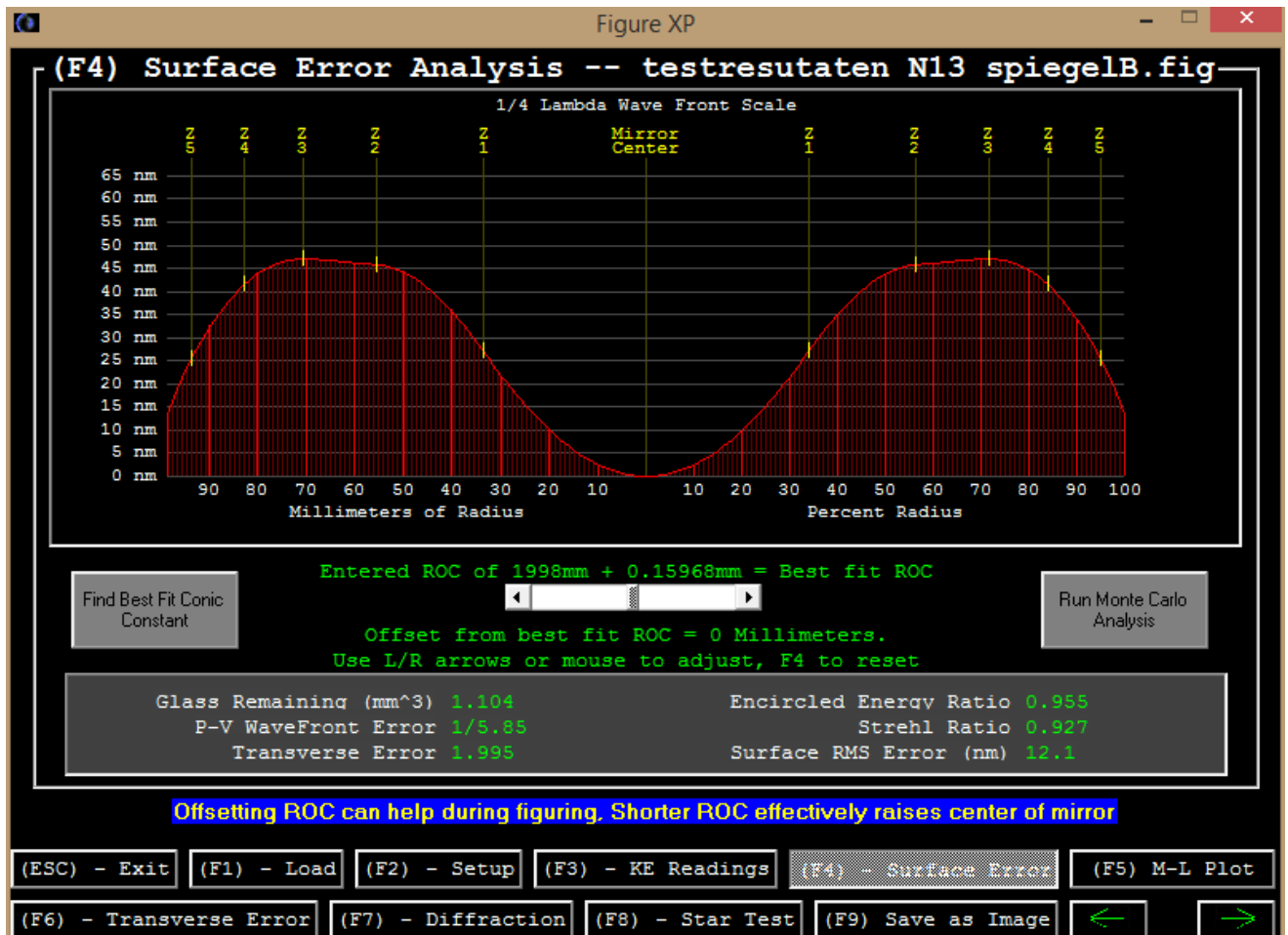
3.1.5 De gebruikte software

Voor het verwerken van de metingen met de Foucaulttester wordt gewoonlijk FigureXP of Foucaulttest Analysis gebruikt. Tabel 9 geeft aan of dat in Nederland en België verschilde.

Land	Figure XP	Foucaulttest analysis	totaal
Nederland	10	2	12
België	9	9	18
totaal	19	11	30

Tabel 9: FigureXP of Foucaulttest Analysis

We zien in deze tabel dat Belgische testers vaker Foucaulttest Analysis dan Nederlandse testers.



Figuur 5: Profielweergave spiegel B, FigureXP (foto Jan van Gastel)

3.1.6 kleur van de lichtbron

Land	wit	groen	blauw	rood	totaal
Nederland	5	7	0	0	12
België	16	0	1	1	18
totaal	21	7	1	1	30

Tabel 10: De kleur van de lichtbron van de testapparatuur

In tabel 10 zien we dat de meeste testers een witte lichtbron gebruikten. Blauw en rood werden slechts zelden gebruikt.

3.1.7 Gebruik van een Ronchitester

De testers mochten naast de Foucaulttester een Ronchitester gebruiken als zij van mening waren dat bepaalde defecten in een spiegel daardoor beter te zien of te interpreteren zouden zijn.

Land	Geen Ronchi	Wel Ronchi	totaal
Nederland	7	5	12
België	5	13	18
totaal	12	18	30

Tabel 11: Gebruik van een Ronchitester

Voor beide spiegels werd ongeveer even vaak naast de Foucaulttester een Ronchitester gebruikt. Belgische testers gebruikten iets vaker een Ronchitester dan Nederlandse testers.

3.1.8 Over hoeveel assen heeft men de spiegel getest en hoeveel metingen¹⁷ verrichtte men?

Vaak test men een spiegel over meer dan een as, om een beter kwantitatief beeld te krijgen van de spiegel. Aan de deelnemers is daarom gevraagd over hoeveel assen men de spiegel heeft getest.

Land	1 as	2 assen	3 assen	3+ assen	totaal
Nederland	3	5	2	2	12
België	10	4	4	0	18
totaal	13	9	6	2	30

Tabel 12: Aantal assen waarover men heeft getest

In tabel 12 zien we dat Nederlandse deelnemers iets vaker over meer assen getest hebben dan de Belgische deelnemers, die dus vaker over een as testten. De Nederlandse testers gebruikten *beduidend meer metingen*, namelijk gemiddeld 5.6 vergeleken met 2.4 voor de Belgische deelnemers aan het onderzoek ($p=0.03$). Er is ook bekeken of ervaren testers vaker over meer dan een as testen en meer metingen verrichten dan minder ervaren testers. Dit bleek beide niet het geval te zijn ($p=0.33$ en 0.17).

3.1.9 De randondersteuning van de spiegel tijdens het testen

Een spiegel kan op de teststand op verschillende manieren ondersteund worden. Het meest gebruikt zijn een sling (draagband) en een ondersteuning met twee pennen, of daarop gelijkende ondersteuning.

Land	sling	2 pennen	totaal
Nederland	1	11	12
België	0	18	18
totaal	1	29	30

Tabel 13: Ondersteuning van de spiegel op de teststand

Het is (tabel 13) duidelijk dat voornamelijk 'twee pennen' of iets wat daarop lijkt als ondersteuning voor de spiegel worden gebruikt.

3.1.10 Moeilijk uit te lezen zones

Niet alle zones in de Foucaulttest zijn gemakkelijk af te lezen. Het moeilijkst worden gewoonlijk de eerste zone (vanaf het centrum) en de randzone gevonden. Ook in dit onderzoek hebben we gevraagd welke zones men moeilijk vond om te beoordelen. Men kon de twee moeilijkst te beoordelen zones noemen.

¹⁷ Omdat er wel tot 12 metingen werden verricht hebben we hier geen tabel van gemaakt maar volstaan met het geven van gemiddelden.

zone	moeilijkst	Op een na moeilijkst
1	10	3
2	0	0
3	0	1
4	6	2
5	12	11

Tabel 14: Welke zones vond met het moeilijkst te beoordelen

Tabel 14 laat zien, dat de testers vooral zone 5 en zone 1 lastig uit te lezen vonden, maar ook zone 4 werd aardig wat keren genoemd, vooral als moeilijkste zone.

3.2. Beschrijving van de onderzoekresultaten: de testresultaten van de deelnemers

Ter beoordeling van de kwaliteit van de spiegel is het niet voldoende om alleen de kwantitatieve, door de gebruikte software geleverde Peak-to-Valley-maat of Strehlratio te gebruiken. Ook kwalitatieve informatie over de spiegel, zoals de kwaliteit van de uiterste rand, het al dan niet aanwezig zijn van astigmatisme en de ruwheid van het oppervlak



Foucaulttester Rik Terhorst

dienen in het uiteindelijke oordeel een rol te spelen. Over al deze zaken is de testers gevraagd om te rapporteren. We geven eerst de resultaten *per spiegel* weer in beschrijvende zin en gaan daarna bekijken of er verbanden zijn met hierboven besproken gegevens.

3.2.1 De peak-to-Valley-wavefrontfout en de Strehlratio

We hebben in de vragenlijst zowel gevraagd naar de PV-wavefrontfout als naar de Strehlratio, die we hier beide zullen behandelen. De correlatie tussen beide maten is -0.92. Eigenlijk wordt voor het betrouwbaar berekenen van de Foucaulttest op te weinig punten getest, maar het is toch interessant om na te gaan op welke punten de uitkomsten verschillen van de PV-waarden.

In tabel 15 hieronder zijn de resultaten met betrekking tot de PV-wavefrontfout samengevat voor spiegel A en spiegel B:

PV-fout	Spiegel A		Spiegel B	
	aantal	perc	aantal	perc
0.03 – 0.12	2	18	6	32
0.12 – 0.21	6	55	8	42
0.21 - 0.29	3	27	2	11
0.29 – 0.38	0	0	1	5
0.38 – 0.46	0	0	2	11

Tabel 15: De PV wavefrontfout voor spiegel A

In tabel 15 zien we, dat de door de testers gevonden PV-waarden van deze spiegels een flinke spreiding vertonen, vooral die van spiegel B, met als uiterste grenzen waarden van $1/30$ en ongeveer $1/2$ lambda. De gemiddelde PV-fout bedraagt 0.167 voor spiegel A en 0.169 voor spiegel B, ofwel ongeveer $1/6$ lambda voor beide spiegels. De spreiding bedraagt 0.051 voor spiegel A en 0.116 voor spiegel B. De spreiding voor spiegel B is dus ruim twee maal die van spiegel A. Het 0.95 betrouwbaarheidsinterval van het gemiddelde ligt tussen $0.133 - 0.201$ lambda voor spiegel A en tussen 0.114 en 0.225 voor spiegel B.

In tabel 16 zijn de door de testers gevonden Strehlratio's voor beide spiegels samengevat. Gemiddeld vonden de testers voor spiegel A een Strehlratio van 0.94 en voor spiegel B een Strehlratio van 0.92. De spreiding is 0.032 voor spiegel A en 0.099 voor spiegel B.

Strehlratio	Spiegel A		Spiegel B	
	aantal	perc.	aantal	perc.
0.66 – 0.73	0	0	2	10
0.73 – 0.80	0	0	0	0
0.80 - 0.86	0	0	1	5
0.86 – 0.93	4	36	4	21
0.93 – 0.96	2	18	3	16
0.96 – 1.00	5	46	9	47

Tabel 16: De Strehlratio voor spiegel B

Van de testers van spiegel A vond 46% een Strehlratio van 0.93 of hoger, voor spiegel B is dit 47%. Voor spiegel A is het 95% betrouwbaarheidsinterval van het gemiddelde $0.92 - 0.96$ en voor spiegel B $0.87 - 0.97$.

Samengevat: de gemiddeld door de testers gevonden PV-waarde van beide spiegels – 0.167 en 0.169 - verschillen zeer weinig van elkaar, de spreiding van de PV-waarden van spiegel A is minder dan de helft van die van spiegel B. Ook de gemiddelde Strehlratio's verschillen met resp. 0.94 en 0.92 weinig van elkaar. De spreiding van de Strehlratio van spiegel A is ongeveer 1/3 van die van spiegel B. Geen van de verschillen is significant, wat betekent dat de twee spiegels op PV-waarde en Strehlratio als kwalitatief gelijk werden beoordeeld ($p=0.95$).

3.2.2 Astigmatisme

Een fout die, indien niet te klein, op kwalitatieve wijze met de Foucaulttester en Ronchitester zichtbaar kan worden gemaakt is astigmatisme. 'Kwalitatief' betekent, dat alleen maar kan worden gezegd of er veel of weinig (met eventueel gradaties) astigmatisme in een spiegel zit, en niet hoeveel. Met een *gewone* Foucaulttester onder gunstige omstandigheden is astigmatisme *met veel moeite* te zien tot een PV-waarde van ongeveer 1/6 lambda PV-wavefront, doch de meeste testers halen dit (lang) niet. Van alle vragen in verband met kwaliteit, is de vraag naar astigmatisme in dit onderzoek het meest onbeantwoord gebleven. Bij het testen van spiegel B heeft zelfs bijna 30% deze vraag niet beantwoord. Sommigen gaven aan niet (goed) te weten op welke manier ze astigmatisme in de Foucaulttester zichtbaar moesten maken. Dat is jammer, want het is een veelvoorkomende fout zowel in zelfgemaakte als in commerciële spiegels, die vaak

astigmatisme	Spiegel A		Spiegel B	
	aantal	perc.	aantal	perc.
geen	5	45	2	11
weinig	6	55	9	50
veel	0	0	2	11
zeer veel	0	0	0	0
niet beantwoord	0	0	5	28

Tabel 17: Astigmatisme waargenomen in spiegel A en B

zichtbaar negatieve gevolgen heeft op de beeldkwaliteit. In tabel 17 zijn de antwoorden op de vraag naar astigmatisme samengevat voor spiegel A en B. In spiegel A werd blijkens tabel 17 geen tot weinig astigmatisme waargenomen door de meeste testers. Ook voor spiegel B is dat zo, maar met minder testers in de categorie '*geen astigmatisme*'. In spiegel B zeiden 2 personen veel astigmatisme te hebben aangetroffen. Er is een duidelijke tendens om in spiegel B vaker astigmatisme waar te nemen dan in spiegel A ($p=0.06$).

3.2.3 De gladheid of ruwheid van het spiegeloppervlak

De mate van ruwheid van het spiegeloppervlak is in een Foucaulttester goed te zien. Vergelijken we spiegel A en B in tabel 18 dan zien we dat spiegel B als ruwer wordt beoordeeld dan spiegel 1. Dit net niet significant, maar wel een tendens in de antwoorden van de testers ($p=0.09$).

Gladheid/ruwheid	Spiegel A		Spiegel B	
	aantal	perc.	aantal	perc.
Zeer glad	0	0	0	0
Redelijk glad	5	45	2	11
Ietsje ruw	5	45	15	78
Zeer ruw	1	10	2	11

Tabel 18: Mate van gladheid/ruwheid zoals door de testers beoordeeld in spiegel A

Duidelijk is verder, dat niemand een van de twee spiegel zeer glad vond en dat beide spiegels voornamelijk als 'redelijk glad' tot 'ietsje ruw' beoordeeld werden. Weinig testers vonden de spiegel die ze testten 'zeer ruw'.

3.2.4 De beoordeling van de uiterste rand van de spiegels

In tabel 19 is aangegeven hoe de testers de rand van de spiegels beoordeelden. De meeste testers van beide spiegels gaven aan 'ietsje TDE' te hebben waargenomen.

De uiterste rand	Spiegel A		Spiegel B	
	aantal	perc.	aantal	perc.
Goed	3	25	1	6
Ietsje TDE	7	58	13	72
Veel TDE	0	0	3	17
Ietsje TUE	2	17	1	6
Veel TUE	0	0	0	0

Tabel 19: kwaliteit van de rand van spiegel A zoals beoordeeld door de testers van deze spiegel

We kunnen de door de testers verstrekte resultaten ook anders ordenen, om na te gaan in welke mate men de rand 'goed' dan wel 'matig' dan wel 'slecht' vond. Zowel TDE als TUE verminderen de kwaliteit van de rand, dus we kunnen 'weinig TDE' en 'weinig TUE' en 'veel TDE' en 'veel TUE' in respectievelijk de categorieën 'matig' en 'slecht' onderbrengen. Het resultaat van deze herordening zien we voor beide spiegels in tabel 20.

De uiterste rand	Spiegel A		Spiegel B	
	aantal	perc.	aantal	perc.
goed	3	25	1	6
matig	9	75	14	78
slecht	0	0	3	16

Tabel 20: kwaliteit rand spiegel A herordend

We zien in tabel 20, dat de rand van spiegel B als slechter beoordeeld wordt dan de rand van spiegel A. Dit verschil in beoordeling is significant ($p=0.04$) en klopt ook met de werkelijkheid zoals we eerder hebben gezien.

3.2.5 Het eindoordeel van de testers over de spiegels

Aan het eind van de test werd aan de deelnemers gevraagd hoe hun eindoordeel over de spiegel luidde. In dit oordeel kon men dus alles meenemen wat men kwantitatief had gemeten en kwalitatief had beoordeeld. Hieronder bespreken we de oordelen per spiegel. Hoewel het hier en daar in tabel 21 lijkt of er verschil is tussen het eindoordeel over beide

spiegels, is dat niet zo ($p= 0.73$). De conclusie luidt dus, dat beide spiegels als *even goed of even slecht* beoordeeld werden door de testers. Als we per spiegel naar het oordeel kijken *per land* zien we ook geen verschil tussen Nederlandse en Belgische beoordelaars ($p= 0.40$ en 0.63 voor respectievelijk spiegel A en spiegel B).

Eindoordeel	Spiegel A		Spiegel B	
	aantal	perc	aantal	perc
Slecht	1	9	1	5
Matig	3	27	10	53
Vrij goed	6	55	5	26
Goed	1	9	3	16
Zeer goed	0	0	0	0
totaal	11	100	19	100

Tabel 21: Eindoordeel van de testers over de spiegels

3.3 Verbanden tussen achtergrondvariabelen en kwaliteitsvariabelen

Deze paragraaf gaat over verbanden tussen achtergrondvariabelen en kwaliteitsmaten. Zo is het bijvoorbeeld interessant na te gaan of ervaren testers tot een ander oordeel kwamen over de kwaliteit van de spiegel (PV-waarde, astigmatisme, rand, ruwheid, eindoordeel). Interessant is bijvoorbeeld ook of testers die naast de Foucaulttest de Ronchitester gebruikten, tot een ander oordeel kwamen inzake de kwaliteit van het oppervlak of de spiegelrand en of een groter aantal metingen, of het meten over meer dan een as, van invloed zijn op de gevonden PV-wavefrontfout. In de linker kolom staan steeds de achtergrondvariabelen.

3.3.1 Verband tussen achtergrondvariabelen en de door de testers gevonden PV-wavefrontfout en Strehlratio

In tabel 22 gaan we na, of de variabelen in de linker kolom invloed hebben op de door de testers gevonden PV-wavefrontfout. We doen dat voor beide spiegels samen.

	f-waarde	p-waarde	Significant?
ervaring	0.18	0.84	Nee
plaats	1.16	0.33	Nee
source	0.00	0.99	Nee
Assen	0.13	0.72	Nee
Metingen	0.05	0.83	Nee

Tabel 22: verbanden tussen PV-waarde en een aantal onafhankelijke variabelen

In tabel 22 komt geen enkel verband in de buurt van een significante p-waarde van 0.05. We concluderen dus dat *geen enkele achtergrondvariabele* van invloed was op de door de testers gevonden PV-fout. We hebben deze analyse ook gedaan voor de *Strehlratio*, met hetzelfde resultaat.

3.3.2 Verband tussen achtergrondvariabelen op het waargenomen astigmatisme

In tabel 23 zijn de verbanden tussen achtergrondvariabelen en het oordeel over het aanwezig zijn van astigmatisme samengevat. Omdat geen van de p-waarden gelijk is aan of

kleiner is dan het significantieniveau van 0.05. concluderen we, dat er geen verband is tussen achtergrondvariabelen en het waarnemen door de testers van astigmatisme in de spiegels.

	f-waarde	p-waarde	Significant?
Ronchi	1.25	0.28	Nee
ervaring	0.24	0.79	Nee
plaats	1.07	0.36	Nee
source	0.04	0.84	Nee

Tabel 23: verband tussen het waarnemen van astigmatisme en onafhankelijke variabelen

3.3.3 Verband tussen achtergrondvariabelen en oordeel inzake ruwheid van het oppervlak

In tabel 24 zijn de analyses voor de relatie van onafhankelijke variabelen en de aangetroffen ruwheid van het spiegeloppervlak samengevat.

	f-waarde	p-waarde	Significant?
Ronchi	0.15	0.70	Nee
Ervaring	0.06	0.94	Nee
plaats	0.29	0.75	Nee
source	0.66	0.42	Nee

Tabel 24: Verband tussen de gladheid/ruwheid van het oppervlak onafhankelijke variabelen

Geen van de p-waarden in tabel 24 is kleiner of gelijk aan 0.05, dus er is geen significant verband gevonden tussen de mate waarin men de spiegel ruw of glad vond en de in de linker kolom van tabel 24 genoemde variabelen.

3.3.4 Verband tussen achtergrondvariabelen en oordeel inzake kwaliteit van de rand

	f-waarde	p-waarde	Significant?
Ronchi	4.56	0.04**	Ja
ervaring	3.17	0.06*	Nee
plaats¹⁸	9.32	0.00**	Ja
source	9.60	0.00**	Ja

Tabel 25: Verband tussen de kwaliteit van de rand en een aantal onafhankelijke variabelen

Op een na (ervaring) hangen volgens tabel 25 alle in de linker kolom genoemde variabelen significant samen met de mate waarin men de rand goed of minder goed vindt. Hoe kunnen we dat interpreteren?

Voor het beoordelen van de rand is niet van belang of de gebruikte Foucaulttester een 'moving' of een 'fixed' source apparaat is. In beide typen is het aan- of afwezig zijn van een diffractierand aan weerszijden van de spiegel en/of het helderheidsverschil in de diffractierand aan beide kanten van de spiegel even goed te beoordelen. Bovendien hangt 'source' significant samen met 'ervaring' ($p = 0.01$) en ook met 'plaats' ($p = 0.000$). Tenslotte is het niet aannemelijk dat het feit of men thuis, op een sterrenwacht of bij iemand anders test invloed heeft op het al dan niet (kunnen) waarnemen van een minder goede rand. Het is aannemelijk - en dat blijkt ook uit een variantieanalyse ($F = 9.42$, $p = 0.001$) met 'plaats' en

¹⁸ Testte men thuis, bij iemand anders of op een sterrenwacht.

'ervaring' als variabele dat mensen die thuis geen tester hadden maar elders testen *minder ervaring* hebben met testen dan mensen die thuis zelf over een tester beschikten. Dit alles overziend, lijkt het dus plausibel dat van de genoemde variabelen 'ervaring' de (of in elk geval een belangrijke) achterliggende variabele is met betrekking tot de verbanden van 'source' en 'plaats' met het oordeel inzake de spiegelrand. Als we 'source' en 'plaats' om die reden weglaten, houden we nog twee onafhankelijke variabelen over: 'ervaring' en 'Ronchi' (het wel of niet gebruikt hebben van een Ronchitester) om in een regressievergelijking te brengen. De analyse levert het volgende op:

variabele	Beta	Standaardfout	T-waarde	P-waarde
Ervaring	0.41	0.11	5.59	0.023
Ronchi	-0.33	0.17	-1.91	0.07

Tabel 26: Resultaat regressieanalyse 'kwaliteit van de rand'

De bijdrage van beide onafhankelijke variabelen is significant, zoals in tabel 26 in de kolom met p-waarden – allemaal beduidend lager dan 0.05 - te zien is. Tenslotte krijgen we met deze variabelen de volgende samenvatting van de regressieanalyse (tabel 27):

R	R2	F-waarde	P-waarde
0.56	0.31	5.59	0.01

Tabel 27: Verklaarde variantie en significantieniveau regressieanalyse kwaliteit rand'

R geeft de gecombineerde samenhang (correlatie) van alle drie de onafhankelijke variabelen met 'rand'. Adjusted R2 geeft de hoeveelheid van de variantie in 'rand' die verklaard wordt door de drie onafhankelijke variabelen samen. Dat is dus 31%. De p-waarde van 0.01 geeft aan dat de resultaten significant zijn. Inhoudelijk betekent deze analyse dat in het algemeen:

- ervaren testers de rand van de door hen geteste spiegel als *beter* beoordeelden dan onervaren testers;
- testers die naast de Foucaulttest ook een Ronchitester gebruikten, de rand *slechter* vonden dan testers die geen Ronchitester gebruikten.

3.3.5 Conclusies met betrekking tot de relatie achtergrondvariabelen en kwaliteitsmaten

De conclusies met betrekking tot de in de analyses gebruikte achtergrondvariabelen met kwaliteitsmaten luiden als volgt:

- a. Er is geen verband tussen achtergrondvariabelen en de *P-V-wavefrontfout* en *Strehlratio*.
- b. Er is een tendens om in spiegel B meer astigmatisme waar te nemen dan in spiegel A.
- c. Er is een tendens om het oppervlak van spiegel B als ruwer te beoordelen dan dat van spiegel A.
- d. Er is significant verband tussen achtergrondvariabelen en de waargenomen *kwaliteit van de spiegelrand*: ervaren testers vonden de rand beter dan onervaren testers, testers die ook een Ronchitest gebruikten vonden de rand slechter dan testers die geen Ronchitester gebruikten.

3.4 Verband tussen kwaliteitsmaten en het eindoordeel van de testers over de spiegel

3.4.1 Analyse van de relatie tussen kwaliteitsmaten en eindoordeel

Er is de testers gevraagd naar verschillende kwaliteitsmaten te kijken bij het testen: *de PV-wavefrontfout, de Strehlratio, het astigmatisme, de mate van ruwheid van het spiegeloppervlak en de kwaliteit van de uiterste rand van de spiegel*. Ook is gevraagd wat hun eindoordeel is over de spiegel (tabel 21). Het is interessant na te gaan op welke kwaliteitsmaten het oordeel van de testers is gebaseerd. We doen dit door middel van een 'best fit'-regressieanalyse, met 'oordeel' als afhankelijke en de bovengenoemde kwaliteitsmaten als onafhankelijke variabelen. Omdat we geen vooropgezet idee hebben over hoe het eindoordeel ontstaat, laten we het statistiekprogramma zelf bepalen hoe de situatie in elkaar zit. Wel nemen we voor de zekerheid ook de variabele 'ervaring' mee in de analyse, vanwege de mogelijkheid dat ervaren testers op grond van vergelijkbare gegevens een ander oordeel geven over een spiegel dan minder ervaren testers. We hebben de variabele 'astigmatisme' niet meegenomen, omdat 20% van de testers geen antwoord op de vraag naar astigmatisme had ingevuld en daardoor uit de gehele analyse zouden vallen¹⁹. Deze 'best fit'-regressieanalyse levert de volgende uitslag op:

variabele	Beta	R2 incr.	T-waarde	P-waarde
ruwheid	-0.46	0.20	-1.41	0.17
PV	-0.24	0.07	-2.67	0.01
rand	0.22	0.05	1.27	0.22

Tabel 28: Resultaat regressieanalyse 'eindoordeel'

De kolom ' R^2 incr.' in tabel 28 geeft aan hoeveel procent (extra) wordt verklaard door elke onafhankelijke variabele, dus respectievelijk 20, 7 en 5% voor 'ruwheid', 'PV-wavefrontfout' en *kwaliteit van de rand*. De grootste bijdrage wordt dus geleverd door 'ruwheid'.

De samenvatting van de invloed van 'ruwheid', 'PV-fout' en 'rand' op 'oordeel' zien we in tabel 29 hieronder:

R	Adjusted R2	F-waarde	P-waarde
0.559	0.32	3.72	0.025

Tabel 29: Verklaarde variantie en significantieniveau regressieanalyse 'eindoordeel'

Aan de p-waarde van 0.025 zien we dat de bijdrage van 'ruwheid', 'PV' en 'rand' samen significant is. Dat betekent, dat de testers zich in hun eindoordeel hebben laten leiden door de mate van gladheid of ruwheid van het oppervlak, door de PV-wavefrontfout en door de kwaliteit van de rand van de spiegel.

De waarde van 'R' in tabel 29 geeft de correlatie aan tussen het 'oordeel' enerzijds en 'ruwheid', 'PV' en 'rand' anderzijds'. De term 'adjusted R²' geeft aan hoeveel procent van de variantie in 'oordeel' wordt verklaard door deze drie variabelen samen. Dit is dus 32%. Er blijft dus nog 68% te verklaren over, maar met de huidige onderzoeksgegevens kan dat niet.

¹⁹ Gezien de lage correlatie van -0.16 tussen astigmatisme en eindoordeel zou astigmatisme ook geen rol in de analyse hebben gespeeld zo bleek uit een eerdere analyse. Dit neemt niet weg dat sommige testers de mate van astigmatisme die zij meenden te zien wel kunnen hebben meegenomen in hun oordeel.

Voor sommige testers zal de verklaring in een andere kwaliteitsmaat zitten, zoals astigmatisme, maar voor de totale groep testers gaat dit niet op, want dan zouden ook die andere variabelen een significante bijdrage hebben geleverd. We zien ook, dat 'ervaring' geen rol blijkt te spelen in het eindoordeel over de spiegel als totaal, wat impliceert dat weinig ervaren testers tot eenzelfde eindoordeel over de door hen geteste spiegel zijn gekomen als ervaren testers.

3.4.2 Conclusies met betrekking tot de relatie tussen kwaliteitsmaten en eindoordeel

Concluderend stellen we vast dat de testers, naarmate hun oordeel over de spiegel gunstiger was:

- a. Het spiegeloppervlak minder ruw vonden;
- b. Een lagere (dus gunstiger) waarde voor de P-V-wavefrontfout vonden;
- c. De rand van de spiegel beter vonden.

Dit ligt in de lijn van de verwachtingen. Het maakt verder duidelijk dat de testers zich niet door slechts één kwaliteitsaspect hebben laten leiden in hun oordeel over de spiegel, maar daarin een aantal kwaliteitsaspecten hebben meegenomen. Het zegt *niet* dat ze ook gelijk hebben in hun oordeel over de spiegel.

3.5 Kenmerken van testers die een spiegel 'goed' of 'niet goed' beoordeelden

We kunnen ook op een iets andere manier naar het oordeel van de testers, door ze in twee groepen in te delen: testers die de spiegel als 'vrij goed of goed' beoordeelden en mensen die de spiegel als 'slecht of matig' beoordeelden. We combineren de vier categorieën van de variabele 'oordeel' dus tot twee categorieën, door de categorieën '*slecht en matig*' en de categorieën '*vrij goed en goed*' te combineren²⁰. Vervolgens voeren we een zogenaamde discriminantanalyse uit. We stellen hierbij niet als eis dat elke variabele afzonderlijk een significante bijdrage moet leveren.

We voeren hier weer dezelfde kwaliteitsvariabelen in plus ervaring. Ook nemen we de variabele 'assen' (over hoeveel assen heeft men getest) op en de variabele 'Ronchi' (maakte men gebruik van een Ronchitester), omdat deze zaken bij het bepalen van de kwaliteit en het eindoordeel relevant kunnen zijn. In de eerste stap van de analyse bleek, dat 'Ronchi' niets toevoegde, reden waarom het verder uit de analyse is weggelaten. Dit geldt ook voor '*astigmatisme*' dat we in eerste instantie hadden meegenomen in de analyse en ook voor '*metingen*', ofwel of men een of meer metingen heeft verricht. De resultaten, samengevat in tabel 30, zijn de volgende.

- a. De correlatie van alle bij punt b (tabel 30) genoemde onafhankelijk variabelen samen met '*oordeel*' is 0.75. Dat is significant ($p=0.003$) en betekent een verklaarde variantie van 75^2 ofwel 56%. De p-waarde voor de analyse is 0.003. Hoe hoger het getal (in absolute zin) in de rechterkolom in tabel 30, hoe belangrijker de bijdrage van de betreffende variabele.

²⁰ Eigenlijk hoort hier ook 'zeer goed' bij, maar dit antwoord werd door geen van de testers gegeven.

- b. Om de bijdrage van alle variabelen te berekenen, creëert de analyse een functie (root) waarmee de gebruikte variabelen samenhangen. Hoe hoger het getal in tabel 30 (in absolute zin) hoe sterker het verband van de variabelen met de functie.

variabele	Correlatie met discriminantfunctie
Ervaring	0.36
PV-wavefrontfout	-0.37
Strehlratio	0.40
Ruwheid	-0.49
Aantal assen waarover gemeten	0.71
Kwaliteit rand	0.47

Tabel 30: correlaties van variabelen met discriminantfunctie

Inhoudelijk kunnen we de in tabel 30 samengevatte resultaten als volgt interpreteren: testers die de spiegel die ze testten 'vrij goed' of 'goed' van kwaliteit vonden:

- hadden meer ervaring;
- vonden een lager PV-waarde en een hogere Strehlratio;
- beoordeelden het spiegeloppervlak als gladder;
- testten over meer assen;
- beoordeelden de rand als beter.

Dit alles in vergelijking tot de groep van testers die de spiegel 'slecht' of 'matig' vonden. Ook in deze analyse zien we dus dat de testers zich door *verscheidene kwaliteitsaspecten* laten leiden om te komen tot een eindoordeel. Daarnaast spelen blijkbaar ook de achtergrondvariabelen: *ervaring*' en - vooral - het over een of over meer dan een as hebben getest een rol. De rol van ervaring moeten we hier overigens niet overschatten. Laten we *'ervaring'* weg, dan daalt de verklaarde variantie minder dan 1% en wordt er slechts 1 persoon extra in de verkeerde groep geplaatst (zie punt c hieronder).

Ook deze analyse zegt alleen iets over hoe de testers de spiegels beoordeelden op grond van hun eigen kwantitatieve en kwalitatieve testresultaten en zegt dus niet dat de testresultaten of het eindoordeel van een van de twee beoordelingsgroepen beter in overeenstemming was met de werkelijke kwaliteit van de spiegels.

- c. Op grond van deze analyse plaatst het programma de testers in een van de twee door de analyse gevormde beoordelingsgroepen, om door vergelijking met de groep waartoe de personen op grond van hun werkelijke antwoorden behoren te beoordelen of de keuze van de onafhankelijke variabelen een juiste is. De cijfers 1 en 2 (tabel 31) voor de *werkelijke* en *voorspelde* groep staan respectievelijk voor '*slecht tot matig*' en voor '*vrij goed tot goed*'.

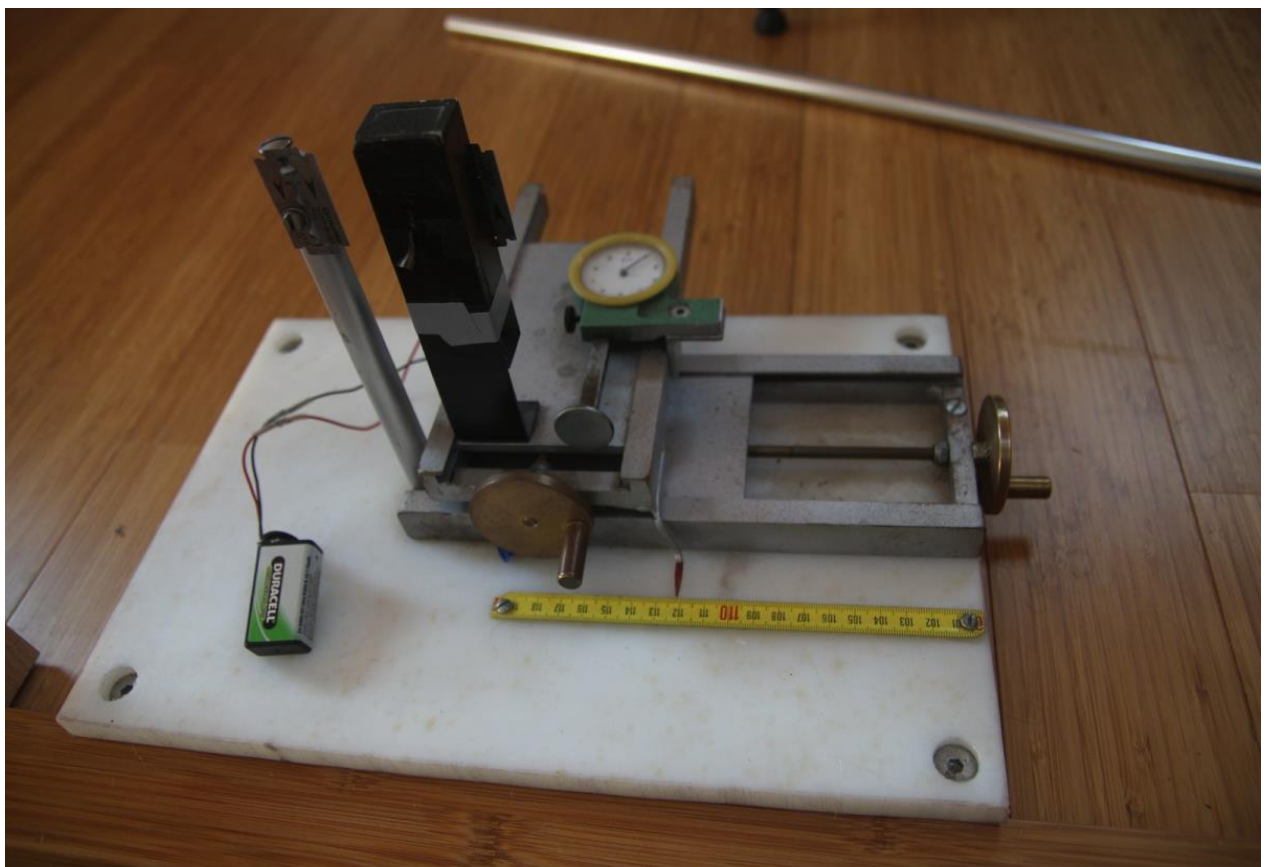
Werkelijke groep	Voorspelde groep		totaal
	1	2	
1	14	1	15
2	2	12	14
totaal	16	13	29

Tabel 31: Aantal testers in werkelijke en voorspelde groep

In tabel 31 staat horizontaal de groep waartoe de testers *werkelijk* behoren en verticaal de groep waarin het programma hen *middels de discriminantanalyse plaatst*. We zien het volgende:

- Van de 15 testers die in werkelijkheid in groep 1 (spiegel slecht of matig) zitten, worden er 14 in de juiste groep geplaatst.
- Van de 14 testers die in werkelijkheid tot groep 2 (spiegel vrij goed of goed) behoren, worden er door het programma 12 in de juiste groep geplaatst.

In totaal worden dus slechts drie testers in een verkeerde groep geplaatst. Dit betekent dat het programma op grond van de bovengenoemde zes variabelen 26 van de 29 testers, ofwel 90%, in de juiste groep plaatst. We kunnen dus zeggen dat, als we de scores (antwoorden op de betreffende vragen) weten van de testers op de zes in tabel 31 genoemde variabelen, we met 90% zekerheid kunnen zeggen of ze de spiegel als *slecht of matig*, dan wel als *vrij goed of goed* beoordeeld hebben, ongeacht of hun oordeel en hun testresultaten waarop dat oordeel is gebaseerd, juist zijn of niet.



Foucaulttester Philip Mollet

4. Conclusies

In dit hoofdstuk willen we de uitkomsten van het onderzoek terugkoppelen naar de vraagstelling uit hoofdstuk 1, die luidde:

Als testers met behulp van een Foucaulttest, eventueel met hulp van een Ronchitest, eenzelfde spiegel testen:

1. komen zij dan tot dezelfde testresultaten;
2. komen de *kwantitatieve* testresultaten dan overeen met die van een meer objectieve interferometrische test voor de kwantitatieve resultaten;
3. komen ze inzake *kwantitatieve* testresultaten overeen met ons in paragraaf 2.2.5 geformuleerde oordeel (zoals eerder gezegd discutabel vanwege het ontbreken van een algemeen geaccepteerde standaard);
4. Komen ze dan tot eenzelfde eindoordeel over de spiegel;
5. komen testresultaten van meer ervaren testers dan beter overeen met die van interferometrische tests dan die van minder ervaren testers.

4.1 Komen de testers tot dezelfde resultaten

Over het antwoord op deze vraag kunnen we kort zijn: de testers komen op geen enkele kwaliteitsmaat tot eenzelfde resultaat, zo blijkt uit de resultaten die zijn samengevat in de tabellen 15 t/m 20. Ook de eindoordelen over de spiegel lopen nogal uiteen (tabel 21). De spreiding is over het algemeen vrij groot, dus over het algemeen liggen de resultaten van de deelnemers niet dicht bij elkaar. Dit is in overeenstemming met de resultaten van de in het voorwoord (zie voetnoten pagina 1) genoemde Amerikaanse²¹ en Duitse onderzoeken. De Foucaulttest zoals gewoonlijk gebruikt, blijkt dus niet erg betrouwbaar te zijn.

4.2 Komen de testers tot kwantitatieve resultaten vergelijkbaar met interferometrische tests

4.2.1 Sferische aberratie

Om na te gaan in welke mate de *PV-wavefrontfout* en de *Strehlratio* die de testers vonden overeenkwamen met de interferometrisch gevonden waarde hebben we twee nieuwe variabelen gecreëerd: '*PV-verschil*' en '*Strehl-verschil*'. Dit is het absolute verschil van de door de testers gevonden PV-waarden en Strehlratio's en de interferometrisch gevonden waarden. De interferometrisch gevonden waarden vatten we hier op als indicatoren voor de werkelijke kwaliteit van de spiegels. Met behulp van deze variabelen, waarop elke tester een score kreeg, namelijk zijn afwijking van de het interferometrische testresultaat, gaan we na hoe dicht de resultaten van de testers bij die van de interferometrische analyse lagen.

We moeten nu eerst vaststellen *wat we een 'goede overeenkomst' vinden*. Daarbij gaan we uit van de breedte van de in tabel 1 samengevatte kwaliteitscategorieën. Voor de PV-fout komen we dan op 0.02 lambda. Dat lijkt een smalle marge, maar omdat het gaat om het

²¹ Spreiding in P-V resultaat voor drie spiegels in USA onderzoek: 0.10, 0.07 en 0.04. Spreiding is 0.05 en 0.12 voor spiegel A en B in ons onderzoek (zie par. 3.2.1). Van Duits onderzoek is geen totaalrapportage gemaakt.

gebied van -0.02λ tot $+0.02 \lambda$, is het een range van 0.04λ . Alle PV-waarden tussen $1/4.9 - 1/6.1 \lambda$ voor spiegel A en alle PV-waarden tussen $1/6.1$ en $1/8.1 \lambda$ voor spiegel B vallen binnen deze marge. Voor de Strehlratio leggen we het verschil op 2.5%. Omdat ook dit gezien de absolute waarde van de variabel ‘*Strehl-verschil*’ zowel boven als onder de interferometrisch gevonden Strehlratio kan zitten, is dit dus een range van 5% en daarmee ook precies de range van een kwaliteitscategorie uit tabel 1. Wijkt het Foucaulttestresultaat meer van de in de interferometrische analyse gevonden PV-waarde en Strehlratio af dan wat binnen deze ranges valt, *dan spreken we niet meer van een ‘goede overeenkomst’*.

In tabel 32 zijn de aantallen en percentages testers vermeld die testwaarden binnen de genoemde grenzen vonden.

	abs. verschil	aantal	%
PV-verschil	≤ 0.02	9	30
Strehlverschil	≤ 0.025	9	30

Tabel 32: overeenkomst PV-fout en Strehlratio met interferometer tests

Onze conclusie uit tabel 32 luidt dat slechts *een minderheid* van testers binnen de grenzen valt waarbinnen we spreken van een goede overeenkomst met de interferometrische tests, namelijk 30% voor zowel de PV-waarde als de Strehlratio. Zelfs als we voor *PV-verschil* de waarde van 0.02λ twee keer zo groot maken: 0.04λ ofwel een range van 0.08λ scoren nog maar 14 testers (47%) binnen deze aangepaste range. Er is dus geen goede overeenkomst tussen interferometrische en Foucaulttest resultaten²². Van de niet binnen de marges voor de ‘*PV-verschil*’ en ‘*Strehl-verschil*’ scorende testers vond respectievelijk 52% en 63% een te gunstige en respectievelijk 48% en 37% een te ongunstige waarde (zie bijlage 3) in vergelijking tot de interferometrische testresultaten. Met andere woorden: de kwaliteit van de spiegels werd ongeveer *even vaak onderschat als overschat* als we de PV-waarden als uitgangspunt nemen en iets *vaker overschat* als we kijken naar de Strehlratio’s. Het is bekend dat de Strehlratio’s die volgen uit een Foucaulttest gewoonlijk aan de hoge kant zijn en dat zien we ook hier terug. Kunnen we nu zeggen dat de 30% testers die binnen de in tabel 32 genoemde marges testten ‘*de goede testers*’ zijn? Nee, dat kunnen we niet. Dit onderzoek staat geen conclusies toe over wie wel of niet een goede tester is. De kans is groot, dat in een ander, vergelijkbaar onderzoek andere testers tot resultaten komen die binnen de marges vallen. Uit tabel 32 blijkt, dat – binnen ons uitgangspunt dat de interferometrische test een juiste uitslag geeft - de Foucaulttest²³ *geen erg valide test* ²⁴ is.

Op grond van de gegevens van tabel 22 concludeerden we in paragraaf 3.3.1, dat geen enkele achtergrondvariabele van invloed was op de door de testers gevonden PV-fout. Het is nu interessant om na te gaan of deze achtergrondvariabelen wel van invloed zijn op de variabelen ‘*PV-verschil*’ en ‘*Strehl-verschil*’. Testers die hierop een *lage score* hebben, zitten

²² Er was geen verschil tussen de beide spiegels

²³ Zoals eerder vermeld (par. 1.2, pag.2) hebben we het steeds over de *visuele* Foucaulttest

²⁴ Geen test die gewoonlijk de kwaliteit van de spiegel goed bepaalt.

namelijk dichter bij de interferometrische score dan testers die hier relatief hoog op scoren. In tabel 33 zijn de resultaten van deze analyse samengevat. Het verband tussen ‘assen’ en

	f-waarde	p-waarde	Significant?
ervaring	0.41	0.67	Nee
plaats	0.70	0.51	Nee
source	1.23	0.28	Nee
Assen	3.57	0.07*	Nee
Metingen	0.82	0.37	Nee

Tabel 33: verbanden tussen PV-verschil en onafhankelijke variabelen

‘PV-verschil’, zo zien we in tabel 33 (vorige pagina), is net niet significant ($p=0.07$), maar gaat wel duidelijk in die richting. Inhoudelijk zouden we kunnen zeggen dat er een *tendens* is, dat testers die over meer dan een as testen met hun resultaat dichter in de buurt van het interferometrisch resultaat kwamen. We hebben deze analyse ook voor ‘Strehl-verschil’ uitgevoerd. Dat leverde hetzelfde resultaat op.

We vonden verder ook, dat Nederlandse testers significant ($p=0.04$) dichter bij de interferometrisch gevonden waarden zaten dan Belgische²⁵. Hier kunnen we in het onderzoek geen verklaring voor vinden.

4.2.2 Astigmatisme

Hier vergelijken we een *kwantitatief* resultaat van een interferometrische test met een *kwantitatief* resultaat van de Foucaulttest en eventueel Ronchitest.

In spiegel A kwam *geen astigmatisme* voor. Dat betekent (zie tabel 17) dat 5 testers (42%) het bij het rechte eind hadden. Er waren 6 testers die ‘*weinig astigmatisme*’ waarnamen in spiegel A. Dat is dus *onjuist* en het is gezien het afwezig zijn van astigmatisme ook niet mogelijk dat het is waargenomen (of het zou een bijeffect van de betreffende metingen moeten zijn).

Spiegel B bevatte zowel volgens de interferometrische analyse van Wolfgang Rohr als volgens onze eigen interferometrische analyse meer dan $\frac{1}{2}$ lambda PV-wavefront astigmatisme. Dat is (*zeer*) *veel*. In tabel 17 zien we, dat 2 personen (11%) *veel astigmatisme* constateerden in spiegel B en geen enkele tester zeer veel. Twee testers van spiegel B gaven aan dat er ‘*geen*’ astigmatisme in de spiegel zat en 9 personen meldden ‘*weinig*’ astigmatisme. Dat mensen ‘*weinig astigmatisme*’ melden is onjuist, doch begrijpelijk. In een Foucaulttest en Ronchitest lijkt het niet veel, maar dat heeft te maken met het feit dat deze tests *niet erg gevoelig* zijn voor astigmatisme. Als je het zonder veel moeite²⁶ kunt zien, is het al gauw $\frac{1}{4}$ lambda of meer.

De conclusie met betrekking tot astigmatisme luidt, dat dit door het overgrote deel van de testers *onjuist* is beoordeeld

²⁵ Gemiddelde van Belgische en Nederlandse deelnemers resp. 0.088 en 0.034 lambda, standaardafwijking 0.084 en 0.029 lambda.

²⁶ Dit is bedoeld zoals het er staat. Astigmatisme dat zelfs met *enige* moeite pas gezien kan worden ligt al gauw tussen de $\frac{1}{2}$ en $\frac{1}{4}$ lambda.

4.3 Overeenkomst van de kwalitatieve resultaten

4.3.1 De mate van gladheid/ruwheid van het oppervlak

In paragraaf 2.1.1 en 2.2.1, waar we de foto's bespraken die de ruwheid van de spiegels lieten zien, beoordeelden wij het oppervlak van spiegel A als *'matig ruw'* en dat van spiegel B als *'(zeer) ruw'*, hierbij uitgaande van de in tabel 3 geformuleerde classificatie (zie ook tabel 4). In tabel 18 werd spiegel A door een tester en spiegel B door twee testers als zeer ruw beoordeeld. Voor spiegel A klopt dit niet, maar voor spiegel B is dit een juist oordeel, mede gezien het feit dat men moest kiezen tussen *'iets ruw'* en *'zeer ruw'* (een tussencategorie *'ruw'* was er niet). Er zijn voor beide spiegels veel testers die de spiegel als *'iets ruw'* beoordeeld hebben. Voor spiegel B is eigenlijk een te mild oordeel, maar ook dit antwoord kan voor deze spiegel (mede) zijn ingegeven door het feit dat men bij het waarnemen van ruwheid van het spiegeloppervlak moest kiezen tussen *'iets ruw'* en *'zeer ruw'*. Dit gecombineerd met het praktijkgegeven dat er spiegels zijn die nog (veel) ruwer zijn dan de in dit onderzoek geteste, kan aanleiding zijn geweest de spiegels niet als *'zeer ruw'* maar als *'iets ruw'* te beoordelen. Zouden we om die reden voor beide spiegels de *beoordeling 'iets ruw'* als goed aanmerken, dan zou spiegel A door 45% en spiegel B door 78% van de testers juist zijn beoordeeld. Rekenen we ook *'zeer ruw'* goed voor spiegel B, dan heeft 89% deze spiegel qua ruwheid juist beoordeeld. Spiegel A werd door 45% van de testers als *'redelijk glad'* beoordeeld. Meer ervaren testers gaven niet vaker een juist antwoord dan minder ervaren testers. Een redenering enigszins vergelijkbaar met bovenstaande voor spiegel B zouden we ook voor spiegel A kunnen volgen, omdat *'iets ruw'* en *'redelijk glad'* moeilijk van elkaar te onderscheiden zijn. Als we zowel *'redelijk glad'* als *'iets ruw'* goedkeuren, heeft 90% van de testers het oppervlak van deze spiegel goed beoordeeld. De conclusie luidt dan ook:

- a. Dat beide spiegels wat ruwheid betreft objectief gezien wellicht iets *te gunstig* zijn beoordeeld;
- b. Dat we op *methodologische gronden* (te weinig gespecificeerde antwoordmogelijkheden) besluiten dat beide spiegels qua oppervlakte ruwheid door de meeste testers goed zijn beoordeeld. Om dit in de toekomst te voorkomen is het handig om een betere beoordelingschaal voor gladheid/ruwheid van het spiegeloppervlak te ontwikkelen. We komen hier in hoofdstuk 5 op terug.
- c. Eerder (paragraaf 3.2.3) zagen we al, dat er een tendens was spiegel B als ruwer te beoordelen dan spiegel A, hetgeen juist is.

4.3.2 De kwaliteit van de spiegelrand

In tabel 4 hebben we de kwaliteit van de rand van spiegel A beoordeeld als *'lichte TDE'* en de rand van spiegel B als *'matige TDE'*. De meeste testers (tabel 19) beoordeelden de rand als hebbende *'iets TDE'*: 58% van de testers van spiegel A en 72% van de testers van spiegel B kwamen tot dit oordeel. Omdat de antwoordschaal in het scoringsformulier geen onderscheid maakte tussen *'lichte'* en *'matige'* TDE accepteren we – hoewel eigenlijk onjuist voor spiegel B – *'iets TDE'* voor beide spiegels als het juiste antwoord. Geen van beide spiegels had een TUE, dus degenen die dat hadden ingevuld maakten een beoordelingsfout.

Dit geldt ook voor testers die geen TDE waarnamen en de rand als 'goed' beoordeelden en eveneens voor de testers die meenden 'veel TDE' te zien. De rand van spiegel B werd (paragraaf 3.2.4) als slechter beoordeeld dan de rand van spiegel A, wat klopt met de werkelijkheid. Ervaren testers gaven niet vaker het juiste antwoord dan minder ervaren testers.

De conclusie inzake de beoordeling van de rand van de spiegels op basis van de hierboven genoemde percentages luidt als volgt: in het algemeen werd de rand van de spiegels niet juist beoordeeld, hoewel dit sterker gold voor spiegel A dan voor spiegel B. Ook hier verdient het aanbeveling de beoordelingsschaal te verbeteren. We komen hier eveneens in hoofdstuk 5 op terug.

4.4 Het eindoordeel van de testers over de spiegels

In tabel 21 is het eindoordeel van de testers inzake de beide spiegels samengevat. In welke mate komt dat overeen met onze eigen oordeel zoals geformuleerd in tabel 4? Wij beoordeelden spiegel A als 'vrij goed' en spiegel B als 'slecht'. Dit laatste met name vanwege de grote hoeveelheid astigmatisme.

Spiegel A werd, zoals we zien in tabel 21, door 6 testers - 55% - als 'vrij goed' beoordeeld. Eén persoon vond de spiegel 'goed' en 4 personen vonden hem 'slecht' of 'matig'. Door de meeste testers die buiten de 55% 'vrij goed' vielen werd de kwaliteit van de spiegel als totaal dus onderschat.

Spiegel B werd slechts door 1 tester als 'slecht' beoordeeld, daarmee als enige overeenstemmend met ons eigen oordeel. Dit heeft vooral te maken met het feit dat bijna niemand de hoeveelheid astigmatisme juist inschatte. Wel gaven nog 10 testers aan dat ze de spiegel 'matig' vonden, wat in dit geval een te positieve kwalificatie is.

Het eindoordeel werd, zo hebben we gezien in tabel 28 en 30, vooral ingegeven door de PV-wavefrontfout die men vond, door het oordeel over de ruwheid van het oppervlak en het oordeel van de testers over de kwaliteit van de rand van de spiegel. Meer kennis van het waarnemen van astigmatisme en van de *lage gevoeligheid* voor astigmatisme van de doorsnee Foucault- en Ronchitest, zou de juistheid van het eindoordeel zeker kunnen bevorderen.

4.5 Het effect van ervaring

Het leek ons plausibel, dat de testresultaten van ervaren testers beter in overeenstemming zouden zijn met meer objectieve, interferometrische tests en dat er ook meer overeenkomst zou zijn tussen hun antwoorden op kwalitatieve vragen (astigmatisme, rand en ruwheid), vergeleken met de resultaten van minder ervaren testers. Dit bleek in zijn algemeenheid niet op te gaan. Alleen in het geval van de beoordeling van de kwaliteit van de rand van de spiegel speelde ervaring een rol (correlatie =0.44, zie ook tabel 26): meer ervaren testers beoordeelden de rand van de spiegel als beter dan minder ervaren testers. Kijken we naar de cijfers, dat zien we dat meer ervaren dan onervaren testers de rand als 'goed' beoordeelden, terwijl hij dat niet was, bij geen van de beide spiegels. Kennelijk valt er ook voor ervaren

testers op dit punt nog bij te leren, want die gaven wel andere antwoorden, maar hadden het *niet vaker* bij het rechte eind dan minder ervaren testers. Ditzelfde geldt voor het beoordelen van de ruwheid van het oppervlak.

Hoe komt het dat, tegen de verwachting in, ervaren testers geen kwantitatieve testresultaten hebben die beter overeenkomen met interferometrische analyses dan onervaren of minder ervaren testers? Het onderzoek geeft hier geen uitsluitsel over. Mogelijk worden ervaren testers nonchalanter omdat ze denken dat het wel goed zit met hun meetvaardigheden, terwijl minder ervaren testers meer hun best doen om zo nauwkeurig mogelijk te meten. Waar het gaat over de *kwalitatief* waarneembare defecten, speelt waarschijnlijk een rol, dat hierover weinig eenduidige informatie voorhanden is. Voorbeelden op internet stellen vaak extreem afgevalen randen, zeer ruwe oppervlakken en zeer grote hoeveelheden astigmatisme voor omdat men dan het defect duidelijk kan laten zien. Mogelijk denken zowel ervaren als minder ervaren testers daarom, dat het bij de spiegel die zij testen wel meevalt, waardoor gemakkelijk onderschatting plaatsvindt. Ook hier komen we in hoofdstuk 5 op terug.

5. Adviezen

Op grond van de uitkomsten van het onderzoek is het mogelijk om adviezen te formuleren voor het verbeteren van iemands testkwaliteiten. Deze adviezen hebben betrekking op alle typen metingen die in dit onderzoek zijn verricht. We richten ons eerst op de kwalitatieve oordelen.

5.1 Astigmatisme

De Foucaulttest en de Ronchitest zoals door amateurs gebruikt is niet erg gevoelig voor astigmatisme. Gewoonlijk wordt aangenomen dat, als astigmatisme *'met enige moeite'* te zien is, het ook om *slechts weinig* astigmatisme gaat. Het lijkt erop dat de redenering men de redenering volgt: *'je ziet weinig astigmatisme dus er is weinig astigmatisme'* Die redenering is helaas onjuist. Veel testers kennen de voorbeelden van astigmatisme in de Foucaulttest en de Ronchitest die op internet te vinden zijn. Daarin is astigmatisme zeer duidelijk te herkennen, maar het is dan ook zeer veel. Om $1/4$ lambda PV-wavefront astigmatisme te zien, moet je goed je best doen en beter dan $1/6$ lambda - als dat al lukt – is niet meer waarneembaar. Dat wil zeggen dat, als je bij nauwkeurig observeren astigmatisme in Foucault- en/of Ronchitest ziet, het minimaal $1/6$ lambda is en zo goed als altijd meer. Anders gezegd: *als je het ziet is het teveel om nog op een goede spiegel uit te komen*.

Ons advies luidt dan ook: zie je astigmatisme in de Foucault en/of Ronchitest, neem er geen genoegen mee. Laat de spiegel eventueel interferometrisch testen of doe een stertest onder goede omstandigheden. Wees ook daarbij *zeer kritisch*, want ook hiermee wordt astigmatisme vaak onderschat.

5.2 Ruwheid van het oppervlak

Voor het beoordelen van ruwheid bestaat eigenlijk geen goed referentiekader. Als classificatie voor het beoordelen van de mate van ruwheid, zou de schaal die in dit onderzoek werd gebruikt bijvoorbeeld uitgebreid²⁷ kunnen worden tot de vier categorieën in tabel 34 hieronder.

Waargenomen	ruwheid	oppervlak is
Direct in het oog springende ruwheid van het oppervlak	zeer ruw	slecht
Ruwheid goed te zien bij nauwkeurig kijken	matig ruw	matig
Ruwheid met moeite te zien bij nauwkeurig kijken	licht ruw	vrij goed
Ruwheid niet of slechts zeer moeilijk te zien ²⁸	glad	goed

Tabel 34: classificatie mate van ruwheid spiegeloppervlak in Foucaulttester

²⁷ Ook hier laten wij de microruwheid die alleen met een Lyottest of professionele apparatuur is waar te nemen buiten beschouwing

²⁸ De alleen met een Lyottester of geavanceerde professionele apparatuur zichtbaar te maken microruwheid laten we hier buiten beschouwing.

5.3 De kwaliteit van de rand

Ook voor het beoordelen van de rand bestaat geen duidelijk referentiekader. Net als voor ruwheid adviseren wij om ook voor de kwaliteit van de rand het referentiekader te objectiveren. We zouden daarbij tabel 35 kunnen gebruiken als leidraad.

Waargenomen	TDE/TUE	Rand is
Aan een kant heldere diffractierand, andere kant geheel donker	sterk	slecht
Duidelijk verschil in helderheid diffractierand beide kanten	matig	matig
Net zichtbaar verschil in helderheid diffractierand	licht	vrij goed
Geen zichtbaar verschil in diffractierand beide kanten	geen	goed

Tabel 35: classificatie kwaliteit rand in Foucaulttester

5.4 Het kwantitatieve meetresultaat

Een van de conclusies inzake de gebrekkige overeenkomst tussen het Foucaulttestresultaat en het interferometrische resultaat voor sferische aberratie was, dat de Foucaulttest vaak een te gunstige waarde oplevert (paragraaf 4.2.1), vooral wat de Strehlratio betreft. Op grond daarvan adviseren we om, bij het *maken* van een spiegel, niet te gauw tevreden te zijn. Als je een spiegel wilt met een PV-waarde van minimaal $1/8$ lambda, stop dan niet met paraboliseren bij $1/8$ lambda, maar ga door tot $1/15$ lambda of nog beter. Dat lijkt overdreven, maar het verschil tussen $1/8$ en $1/15$ lambda is hetzelfde verschil als tussen $1/8$ en $1/5.5$ lambda, dus zeker relevant. Het is namelijk het verschil tussen een 'zeer goede' en 'vrij goede' spiegel (zie tabel 1). Ga je door tot $1/15$ lambda, dan kun je je veroorloven dat de werkelijke kwaliteit iets minder goed uitpakt, omdat je dan nog steeds een zeer goede spiegel kunt hebben. Oriënteer je je liever op de Strehlratio, ga dan door tot bijvoorbeeld 0.99 en stop niet bij 0.95. Houd er verder bij het *testen* van een bestaande spiegel met de Foucaulttest rekening mee, dat het testresultaat dat je vindt vaak *geflatteerd* is, vooral als je de Strehlratio als referentie gebruikt voor de spiegelkwaliteit, maar ook in minstens de helft van de gevallen als je uitgaat van de PV-wavefrontfout.

Het onderzoek – paragraaf 3.3.1 – leverde een aanwijzing op dat het aantal assen waarover men test een nauwkeuriger resultaat kan geven. Het testen over meer assen gebeurt vaak al, maar toch zouden we dit ook als *dringend advies* willen meegeven. Ten slotte zagen we in de laatste alinea van paragraaf 3.2.1, dat de gemiddeld door de testers gevonden PV-waarde van de beide spiegels $1/6$ lambda was en de Strehlratio's *0.92 en 0.94* voor respectievelijk spiegel A en B. Zeer dicht in de buurt van de interferometrisch gevonden resultaten dus. Met meer mensen dezelfde spiegel testen kan dus een beter resultaat geven dan het alleen testen van een spiegel met de Foucaulttest.

Wat men ook kan doen is de Foucaulttest op een andere manier vorm geven, door de grijsheid van de maskerzones *niet visueel* te beoordelen op gelijkheid, maar gebruik te maken van een *webcam*. Een andere mogelijkheid is gebruik te maken van Robo-Foucault of vergelijkbare vormen van de Foucaulttest. Hiermee kunnen veel meer zones worden bekeken en van de Robo-Foucaulttest is het eerder vermelde Amerikaanse onderzoek aangetoond, dat de resultaten dicht in de buurt liggen van interferometrische testresultaten.

Bijlagen

Bijlage 1: het scoringsformulier

Bijlage 2: toelichting op het scoringsformulier

Bijlage 3: Scores van de deelnemers op de variabelen 'PV-verschil' en 'Strehl-verschil' . Toelichting in de bijlage zelf en in paragraaf 4.2.1 van het onderzoeksverslag.

Bijlage 1.

SCORINGSFORMULIER FOUCAULTTEST ONDERZOEK

Lees eerst de bijgaande instructie a.u.b.

VUL A.U.B. ALLE GEVENS IN DIE WORDEN GEVRAAGD, OOK AL ZOU EEN ANTWOORD AL UIT ANDERE ANTWOORDEN BLIJKEN

1. deelnemerscode
2. Land (kruisje achter land) Nederland
 België
3. Hoeveel ervaring heb je met Foucaulttesten?
Zet een kruisje in het juiste vakje
- | | | |
|--|---------------------|-------------------|
| | weinig ervaring | (zie toelichting) |
| | gemiddelde ervaring | |
| | veel ervaring | |

4. Gegevens van de gebruikte Foucaulttester (aankruisen)
- | | |
|--|--|
| <input type="checkbox"/> fixed source | (lampje beweegt <u>niet</u> mee met mes) |
| <input type="checkbox"/> moving source | (lampje beweegt <u>wel</u> mee met mes) |

5. Heb je ook van de Ronchitest gebruik gemaakt?
(kruis aan en licht toe waarvoor, indien 'ja'.

	nee
	ja

Toelichting indien 'ja' in vak hieronder

6. Hoe wordt de spiegel tijdens de test ondersteund
aan de rand onderaan? (aankruisen)

	sling (band die onder de spiegel doorloopt)
	2 pennen
	anders, namelijk (vul in):

7. Gebruikte software (aankruisen)

FXP
 FTA

	FigureXP
	Foucault Test Analysis

8. Vul de lichtkleur in van het lampje van de tester

	wit
	geel
	groen
	rood
	blauw
	anders, namelijk:

9. Waar heb je de test uitgevoerd (aankruisen)

thuis
 bij een sterrenwacht
 bij iemand anders thuis
 ergens anders

	namelijk:

10. In welke positie(s) heb je de spiegel getest om tot het eindresultaat te komen?

In de toelichting staat hoe dit te aan te geven in graden
vul zelf het aantal graden in

TESTRESULTATEN

11. Zelf gemeten Radius of Curvature

12. Best fit conic constant

13. Peak-to-Valley resultaat

14. Strehlratio resultaat

15. heb je astigmatisme aangetroffen in de spiegel
(kruis aan)

	geen astigmatisme
	weinig astigmatisme
	veel astigmatisme
	zeer veel astigmatisme

16. Als je astigmatisme hebt waargenomen zag je dat
dan in de Foucaulttest, Ronchitest of beide?
(aankruisen)

	Foucaulttest
	Ronchitest
	beide

17. Hoe beoordeel je de mate van oppervlakte gladheid of ruwheid (aankruisen)

zeer glad
 redelijk glad

ietsje ruw
 zeer ruw

18. Hoe beoordeel je de kwaliteit van de uiterste rand van de **spiegel** (aankruisen)

goed
 ietsje TDE (turned down edge, afgevalen rand)
 veel TDE
 ietsje TUE (turned up edge, omhoog gebogen rand)
 veel TUE

19. Vul **alle** zone readings (metingen) in die zijn meegenomen **in het eindresultaat**.

Reken de centrale zone (zone 1) terug naar 0.0 en zo verder tot zone 5.

	Meting nummer									
	1	2	3	4	5	6	7	8	9	10
Zone 1	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Zone 2										
Zone 3										
Zone 4										
Zone 5										

20. Sommige mensen meten met de Foucaulttest over 1 as, sommige over 2 en sommige over 3 assen.

Vul in over hoeveel assen je gemeten hebt

21. Als je readings gemiddeld hebt: geef dan hier per as waarover je gemeten hebt aan welke readings je voor die as gemiddeld hebt. Gebruik daarvoor de cijfers die boven de readings staan bij vraag 16.

As 1										
As 2										
As 3										

22. Welke zone(s) vond je het moeilijkst om uit te lezen in de Foucaulttester? Maximaal 2 invullen

zone nummer
 zonenummer

23. Geef je eigen eendoordeel over de spiegel (aankruisen)

slecht
 matig
 vrij goed
 goed
 zeer goed

24. Licht je antwoord op vraag 19 in dit kader toe (zie ook de toelichting).

25. Opmerkingen

Hier kun je alle verdere toelichtingen en opmerkingen kwijt die je wilt maken over de test

Bijlage 2

Toelichting scoringsformulier Foucaulttest onderzoek.

Lees a.u.b. alvorens te beginnen met testen eerst deze toelichting door.

Algemeen (deels overgenomen uit het eerder gepubliceerde reglement)

1. Test *uitsluitend individueel*. Overleg niet met iemand over de resultaten alvorens jezelf en je gesprekspartner (als die meedoet) de test helemaal heeft afgerond en *wijzig niets meer* op grond van een gesprek.
2. *Wijzig niets* aan het scoringsformulier door vragen toe te voegen of weg te halen.
3. Het is *niet de bedoeling* dat er nog discussie komt over het scoringsformulier. *Het wordt niet meer gewijzigd* en is voor iedereen hetzelfde.
4. Zet *niet je naam* op het formulier, alleen je deelnemerscode.
5. Is er ondanks deze toelichting iets niet duidelijk, neem dan contact op met Jan van Gastel (e-mailadres jhm.vangastel@ziggo.nl) of Jean-Pierre Grootaert: jeanpierre.grootaert@ugent.be (de 't' aan het eind van 'grootaert' is geen typefout).
6. Vul bij *elke vraag* een antwoord in, of kruis een antwoord aan waar dat gevraagd wordt. Soms zit een antwoord wellicht al impliciet verwerkt in het antwoord op een vorige vraag. *Vul in dat geval toch een antwoord in.*
7. Gebruik *alleen* het meegeleverde Foucaultmasker om te testen. Verschillen die zouden kunnen ontstaan door het gebruik van verschillende maskers, willen we juist uitsluiten.
8. Naast de Foucaulttest mag de Ronchitest –*als hulpmiddel*– worden gebruikt. Geef dan wel precies aan in het formulier waarvoor je die hebt gebruikt en hoe dat je conclusie uit de Foucaulttest heeft beïnvloed. Bijvoorbeeld: stel je denkt een zwakke afgevallen rand te zien in de Foucaulttester. Je kunt dan met de Ronchitest bekijken of die test dat wel of niet bevestigt.
9. Je mag de spiegel *maximaal een week* houden. Dan stuur je hem op naar de volgende persoon op de lijst. Neem contact op met Jan van Gastel of Jean-Pierre Grootaert als je ziet aankomen dat dat niet gaat lukken.
10. Print het formulier niet uit om het per post op te sturen, maar bewaar het op de computer ingevulde formulier en stuur dat als bijlage naar Jan van Gastel, email jhm.vangastel@ziggo.nl. Uiteraard mag men voor zichzelf wel een uitgeprint exemplaar bewaren.
11. Stuur het formulier in zo gauw je het geheel hebt ingevuld, *liefst zo snel mogelijk na het testen*. Stuur het ingevulde scoringsformulier als bijlage bij e-mail aan Jan van Gastel (e-mailadres bij de punten 5 en 10 hierboven). De *absolute deadline* voor het insturen van de resultaten is **25 februari 2015**. Later kan niet want er moet nog tijd zijn voor het maken van de rapportage.
12. De resultaten worden tijdens de *studiedag 'Optica' op 21 maart 2015, in Gent* bekendgemaakt. Na de studiedag worden de resultaten breder gepubliceerd en ontvangen alle deelnemers in elk geval ook een rapportage

Toelichting op enkele vragen

Vraag 1

Je deelnemerscode wordt je verstrekt door Jean-Pierre Grootaert of Jan van Gastel.

Vraag 3 (ervaring)

weinig ervaring is bijvoorbeeld het getest hebben van één spiegel

Gemiddelde ervaring is bijvoorbeeld het getest hebben van twee spiegels, of van een spiegel en

daarnaast bijvoorbeeld 'meegetest' met andere testers.

Veel ervaring is bijvoorbeeld het getest hebben van meer dan twee spiegels, of van twee spiegels en meegetest hebben met anderen, anderen leren Foucaulttesten, e.d.

Vraag 10.

Beide spiegels worden door ons vooraf aan de rand gemarkeerd. Start met testen met die markering boven. *Dat heet de 0-graden positie.* Als je in verschillende posities test, *geef dan in het scoringsformulier aan waar de markering zich bevond.*

Daarbij uitgaande van de volgende regels:

0° = markering boven

90° = markering aan de rechterkant, 90° gedraaid ten opzichte van 0°

180° = markering onder

270° = markering links, 270° gedraaid t.o.v. 0°

Er mogen ook andere, hiervan afgeleide posities worden gebruikt, zoals 45°, etc.

Vraag 19.

Het is niet verplicht om 10 metingen te verrichten. Doe zoveel metingen totdat je zelf vindt, dat je de test nauwkeurig hebt uitgevoerd.

Vraag 24.

Schrijf in deze toelichting hoe je aan het eindoordeel komt. Het kan bijvoorbeeld zijn dat je op grond van de Strehlratio en/of Peak-to-Valley waarde de spiegel als 'goed' zou beoordelen, maar dat je hem toch vanwege astigmatisme, ruwheid en/of een afgevalen rand toch als 'matig' of 'slecht'

Bijlage 3

Frequency Analysis for PV-verschil

FROM	UP TO	FREQ.	PCNT	CUM.FREQ.	CUM.PCNT.	%ILE RANK
-0.11	-0.09	2	0.07	2.00	0.00	0.03
-0.09	-0.08	3	0.10	5.00	0.17	11.67
-0.08	-0.07	1	0.03	6.00	0.20	18.33
-0.07	-0.05	0	0.00	6.00	0.20	20.00
-0.05	-0.04	5	0.17	11.00	0.37	28.33
-0.04	-0.02	0	0.00	11.00	0.37	36.67
-0.02	-0.01	5	0.17	16.00	0.53	45.00
-0.01	0.00	2	0.07	18.00	0.60	56.67
0.00	0.02	2	0.07	20.00	0.67	63.33
0.02	0.03	1	0.03	21.00	0.70	68.33
0.03	0.05	3	0.10	24.00	0.80	75.00
0.05	0.06	0	0.00	24.00	0.80	80.00
0.06	0.08	2	0.07	26.00	0.87	83.33
0.08	0.09	0	0.00	26.00	0.87	86.67
0.09	0.10	0	0.00	26.00	0.87	86.67
0.10	0.12	1	0.03	27.00	0.90	88.33
0.12	0.13	0	0.00	27.00	0.90	90.00
0.13	0.15	0	0.00	27.00	0.90	90.00
0.15	0.16	0	0.00	27.00	0.90	90.00
0.16	0.17	0	0.00	27.00	0.90	90.00
0.17	0.19	0	0.00	27.00	0.90	90.00
0.19	0.20	1	0.03	28.00	0.93	91.67
0.20	0.22	0	0.00	28.00	0.93	93.33
0.22	0.23	0	0.00	28.00	0.93	93.33
0.23	0.24	1	0.03	29.00	0.97	95.00
0.24	0.26	0	0.00	29.00	0.97	96.67
0.26	0.27	0	0.00	29.00	0.97	96.67
0.27	0.29	0	0.00	29.00	0.97	96.67
0.29	0.30	0	0.00	29.00	0.97	96.67

De cijfers in eerste twee kolommen vanaf links geven het verschil tussen interferometrisch gevonden PV-wavefrontfout en door de testers gevonden PV-wavefrontfout. Negatief betekent: testers vonden een lagere, dus *gunstiger* PV-fout dan de interferometrisch bepaalde PV-fout, positief betekent: testers vonden een hogere, dus *ongunstiger* PV-fout dan de interferometrisch bepaalde PV-fout.

Rood = aantallen/percentages die binnen het in de tekst van het onderzoeksverslag geformuleerde criterium vallen.

52% van degenen die niet binnen het criterium scoorden, scoorden gunstiger dan het criterium.

48% van degenen die niet binnen het criterium scoorden, scoorden ongunstiger dan het criterium

Door afronding kunnen cijfers iets verschillen van de tekst van het verslag.

Frequency Analysis for Strehl-verschil

FROM	UP TO	FREQ.	PCNT	CUM.FREQ.	CUM.PCNT.	%ILE RANK
-0.28	-0.26	1	0.03	1.00	0.00	0.02
-0.26	-0.25	1	0.03	2.00	0.07	5.00
-0.25	-0.24	0	0.00	2.00	0.07	6.67
-0.24	-0.23	0	0.00	2.00	0.07	6.67
-0.23	-0.22	0	0.00	2.00	0.07	6.67
-0.22	-0.21	0	0.00	2.00	0.07	6.67
-0.21	-0.20	0	0.00	2.00	0.07	6.67
-0.20	-0.18	0	0.00	2.00	0.07	6.67
-0.18	-0.17	0	0.00	2.00	0.07	6.67
-0.17	-0.16	0	0.00	2.00	0.07	6.67
-0.16	-0.15	0	0.00	2.00	0.07	6.67
-0.15	-0.14	0	0.00	2.00	0.07	6.67
-0.14	-0.13	1	0.03	3.00	0.10	8.33
-0.13	-0.12	0	0.00	3.00	0.10	10.00
-0.12	-0.10	0	0.00	3.00	0.10	10.00
-0.10	-0.09	0	0.00	3.00	0.10	10.00
-0.09	-0.08	0	0.00	3.00	0.10	10.00
-0.08	-0.07	0	0.00	3.00	0.10	10.00
-0.07	-0.06	0	0.00	3.00	0.10	10.00
-0.06	-0.05	0	0.00	3.00	0.10	10.00
-0.05	-0.03	4	0.13	7.00	0.23	16.67
-0.03	-0.02	1	0.03	8.00	0.27	25.00
-0.02	-0.01	2	0.07	10.00	0.33	30.00
-0.01	-0.00	1	0.03	11.00	0.37	35.00
-0.00	0.01	3	0.10	14.00	0.47	41.67
0.01	0.02	1	0.03	15.00	0.50	48.33
0.02	0.03	4	0.13	19.00	0.63	56.67
0.03	0.05	5	0.17	24.00	0.80	71.67
0.05	0.06	6	0.20	30.00	1.00	90.00

De cijfers in eerste twee kolommen vanaf links geven het verschil tussen interferometrisch gevonden Strehlratio en door de testers gevonden Strehlratio. Negatief betekent: testers vonden een lagere, dus *ongunstiger* Strehlratio dan de interferometrisch bepaalde Strehlratio, positief betekent: testers vonden een hogere, dus *gunstiger* Strehlratio dan de interferometrisch bepaalde Strehlratio.

Rood = aantallen/percentages die binnen het in de tekst van het onderzoeksverslag geformuleerde criterium vallen

Van de 4 personen binnen de categorie 0.02-0.03 scoorden er 2 binnen de 0.025 (dus binnen het criterium). In totaal dus 9 personen binnen het criterium.

62% van degenen die niet binnen het criterium scoorden, scoorden gunstiger dan het criterium.

38% van degenen die niet binnen het criterium scoorden, scoorden ongunstiger dan het criterium.

Door afronding kunnen cijfers iets verschillen van de tekst van het verslag.