

Rapport onderzoekscommissie Wetenschappelijke integriteit

Datum: 1 juni 2012

Betreft: Onderzoek naar mogelijke schending van de wetenschappelijke integriteit

Inhoudsopgave

1	Commissiesamenstelling:	2
2	Inleiding	2
3	Inhoud melding	2
4	De onderzoekscommissie	2
5	Werkwijze commissie.....	2
5.1	Ad 1: Methode [REDACTED]	3
5.2	Ad 2: Informatie uit gesprekken	3
5.3	Ad 3: Analyse data	5
6	Conclusies	8
6.1	Conclusies algemeen	8
6.2	Aanbevelingen:	9
6.3	Algemene aanbevelingen:	9
7	Bijlage 1: Mails van [REDACTED] aan [REDACTED]	10
8	Bijlage 2: Mail Smeesters aan Editor [REDACTED]	11
9	Bijlage 3: Bijlagen m.b.t. analyse papers	13
9.1	De methode [REDACTED] toegepast op vergelijkbare artikelen	13
9.2	De methode toegepast op de artikelen van Smeesters	15
10	Bijlage 4: Analyse van samengestelde bestanden	27
11	Bijlage 5: Tijdslijn gebeurtenissen en procedure onderzoekscommissie wetenschappelijke integriteit aangaande Smeesters.....	30
12	Bijlage 6: Verslagen gesprekken Onderzoekscommissie met Smeesters.....	33

1 Commissiesamenstelling:

Prof.dr. [REDACTED] (voorzitter)

Prof.dr. [REDACTED]

Prof.dr. [REDACTED]

Drs. [REDACTED] (secretaris)

2 Inleiding

De onderzoekscommissie is ingesteld door het College van Bestuur op 12 januari 2012 op verzoek van de decaan van de [REDACTED], naar aanleiding van een melding van dr. [REDACTED] (Universiteit [REDACTED]) aan prof.dr. [REDACTED] ([REDACTED]) aangaande een vermoeden van een wetenschappelijke fout of fraude door prof.dr. D.H.R.V. Smeesters (RSM). (zie bijlage 1).

Professor Smeesters heeft zelf op 30 november 2011 een gesprek met de vertrouwenspersoon Wetenschappelijke integriteit aangevraagd. Deze afspraak was ingepland op 7 februari 2012, maar is in overleg met Smeesters afgezegd en in plaats daarvan is een gesprek met de commissie ingepland.

De commissie heeft in haar eerste bijeenkomst op 18 januari 2012 geconcludeerd dat op basis van de gegevens van dr. [REDACTED], die ontvangen zijn via prof. [REDACTED], en de mail van prof. Smeesters zelf, nader onderzoek gewenst is en de melding ontvankelijk is.

3 Inhoud melding

De [REDACTED] onderzoeker Dr. [REDACTED] heeft op basis van analyses van een gepubliceerd papier van Smeesters geconcludeerd dat de data van Smeesters "too good to be true" waren en daarover met Smeesters gecorrespondeerd. [REDACTED] heeft daarop van Smeesters samengestelde data ontvangen en trok hieruit dezelfde conclusie; namelijk dat de data te mooi waren om waar te zijn. Het betrof artikel [REDACTED] ([REDACTED]).

4 De onderzoekscommissie

De onderzoekscommissie heeft zich ten doel gesteld om aan de hand van de methode van dr. [REDACTED] de wetenschappelijke betrouwbaarheid van de artikelen van prof. Smeesters te onderzoeken.

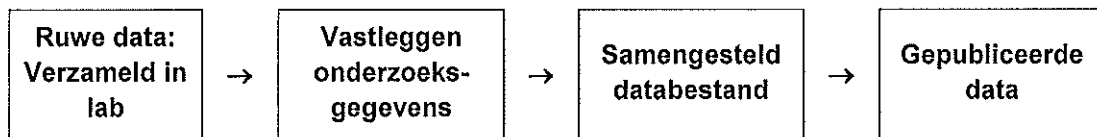
5 Werkwijze commissie

1. De commissie heeft allereerst dr. [REDACTED] via de mail benaderd om nadere informatie en uitleg te krijgen over zijn methode om fraude te achterhalen. Vervolgens hebben dr. [REDACTED] ([REDACTED]) en dr.

█ (█) als deskundigen op het gebied van statistische analyse de methode van █ bestudeerd.

2. De commissie heeft gesprekken gevoerd met Smeesters (op 1 en 16 februari en op 23 mei 2012). Het derde gesprek is drie keer kort tevoren afgezegd om medische redenen en heeft uiteindelijk plaats gevonden op 25 mei te █ ten huize van Smeesters. Hiernaast zijn gesprekken gevoerd met █ (█), █ (vroegere PhD student van Smeesters, nu █ aan █), █ en █ (PhD studenten van Smeesters).
3. Naast de gesprekken heeft de commissie haar onderzoek via drie sporen uitgevoerd, te weten het onderzoeken van (zie Figuur 1) :
 - Ruwe data, zoals ingevulde vragenlijsten of bestanden op labcomputers.
 - Door de onderzoeker samengestelde bestanden waarop statistische analyses kunnen worden uitgevoerd.
 - De rapportage van beschrijvende en verklarende statistiek in een (al dan niet gepubliceerd) manuscript.

Hiervoor is gebruik gemaakt van de data die door Smeesters en █ zijn aangeleverd, zijn de EUR netwerkschijven van Smeesters bekeken en zijn papers van Smeesters onderzocht en van een aantal anderen in het vakgebied (random gekozen) geanalyseerd met de methode van █.



Figuur 1: Onderscheid in type data ten behoeve van onderzoek.

5.1 Ad 1: Methode █

De statistische experts hebben aangegeven dat de techniek die █ gebruikt valide is. De commissie heeft de methode vervolgens, met toestemming van █ en onder embargo, toegepast op al het gepubliceerde werk van Smeesters.

De analyse van █ is gebaseerd op de aanname dat er, █
█
█
█. █ beargumenteert dat als █
█ dit een aanwijzing is dat er problemen zijn met data.

5.2 Ad 2: Informatie uit gesprekken

Gesprekken met professor Smeesters:

Smeesters heeft de commissie verzekerd dat hij heel coöperatief wil zijn, maar dat hij ook veel aan zijn hoofd heeft ten gevolge van gezondheidsproblemen en problemen in de privésfeer.

Smeesters vertelde de commissie dat hij schrok van de mail van [REDACTED] en direct [REDACTED] heeft aangeboden een replicatiestudie te doen. De patronen die Smeesters aangaf keren nu ook terug in de eerste resultaten van de replicatiestudie. Hij wil bekijken of mensen die een moeilijke vraag goed hadden ook altijd een andere moeilijke vraag goed hadden.

Smeesters vertelde verder dat hij misschien wel een fout gemaakt heeft. Hij heeft aan de editor van [REDACTED] laten weten dat hij mogelijk een fout gemaakt heeft (zie bijlage 2). De betreffende editor heeft niets van zich laten horen, aldus Smeesters. Smeesters gaf aan dat hij bij 2 vragen vermoedelijk een coderingsfout heeft gemaakt. Het betreft: [REDACTED]

De commissie heeft in het eerste gesprek met Smeesters laten weten dat zij het zorgelijk vindt dat Smeesters de replicatie zelf doet, omdat hij dan zijn eigen controleur is. Smeesters weet niet of de methode van [REDACTED] klopt.² Hij ziet het probleem dat [REDACTED] aankaart veel voorkomen. Smeesters stelt dat zijn data er zo nu eenmaal uitzien, dat hij achter zijn data en dataverwerking staat en dat hij ook benieuwd is naar de uitkomst van het onderzoek.

Wat betreft de ruwe data verklaarde Smeesters dat deze data op zijn eigen computer thuis stonden. De harde schijf hiervan is gecrasht in september 2011. Volgens Smeesters heeft zijn zwager aan hem verteld dat de schade onherstelbaar was omdat het een fysieke crash was van de harde schijf. De papier-en-potlood data zijn ook verdwenen na een verhuizing van Smeesters bij [REDACTED].³

[REDACTED] heeft over zijn werkwijze verteld dat hij vaak alleen of samen met onderzoeksassistenten de data verzamelde. In samenwerkingsverbanden met collega's uit het buitenland (bijvoorbeeld [REDACTED] van [REDACTED] en [REDACTED] van de [REDACTED]) was het zo dat gezamenlijk experimenten werden ontworpen, Smeesters de data verzamelde en analyseerde en de resultatensecties van experimenten schreef en aan de manuscripten toevoegde.

Smeesters heeft aangegeven dat de coauteurs, [REDACTED] en [REDACTED], geen databestanden hebben gekregen en hier ook niet om gevraagd hebben. De coauteurs werden vaak op congressen benaderd, waarbij Smeesters te kennen gaf toegang te hebben tot een uitstekend lab met een subject pool, zodat hij de dataverzameling gemakkelijk voor zijn rekening kon nemen.

Over de samenwerking met [REDACTED] ([REDACTED]) heeft Smeesters verteld dat [REDACTED] programma's schreef met het softwarepakket Authorware, waarmee de data door Smeesters en/of zijn onderzoeksassistenten verzameld werden. [REDACTED] zette de ruwe data vervolgens op een ftp site van de [REDACTED]. [REDACTED] kreeg zo toegang tot de data, die [REDACTED] samenvoegde in één SPSS bestand en daarna onderwierp aan statistische analyses. In overige gevallen had Smeesters geen controle over de data omdat deze werden verzameld en geanalyseerd door anderen, bijvoorbeeld aio's zoals [REDACTED] en [REDACTED].

Smeesters heeft op verzoek van de commissie aangegeven wie er per paper verantwoordelijk was voor de dataverzameling en -analyse (zie bijlage 3).

Het laatste gesprek vond plaats in [REDACTED] ten huize van Smeesters. Smeesters vermeldt in dit gesprek dat hij data heeft "gemasseerd" in 3 artikelen (artikel 1, 7 en 27, zie bijlage 3), om significantie te bereiken en dat hij de ruwe data niet adequaat heeft opgeslagen waardoor ze verloren zijn gegaan. De data massage gebeurde via de blue dot techniek waarbij proefpersonen

¹ Het artikel is nog online beschikbaar.

² In het tweede gesprek met [REDACTED] zegt hij dat er op het SPSP-forum veel kritiek is op de methode van [REDACTED]. Dit is niet correct; de kritiek betreft een andere methode.

³ Na herhaald verzoek heeft [REDACTED] in een laat stadium de naam en het telefoonnummer van zijn zwager aan de commissie doorgegeven; ondanks enkele pogingen is het de commissie echter niet gelukt om deze zwager te spreken te krijgen zodat niet nagegaan kan worden wat er is geprobeerd om de computerbestanden te herstellen.

geïdentificeerd worden die instructies niet goed gelezen hebben. De massage vindt plaats door eerst de analyse met alle proefpersonen te doen en als het effect niet sterk genoeg is de analyse te herhalen zonder de proefpersonen die de instructie niet goed gelezen hebben. Smeesters geeft aan dat "data massage d.m.v. de Blue dot techniek frequent gebruikt wordt en dat hij dus onderdeel van een cultuur is. Hij is het eens met de commissie dat deze cultuur discutabel is en dat dit onverlet laat dat hij deze techniek heeft toegepast zonder dit in de papers te vermelden.

Gesprek met prof. ██████████

Prof. ██████████ geeft aan dat hij in eerste instantie bedenkingen had bij de methode van ██████████. Hij had de indruk dat ██████████ en Smeesters samen bezig waren om uit te vinden wat er aan de hand was. Later ontving ██████████ van ██████████ documenten met meerdere analyses (23 december 2011, (zie bijlage 1) waarbij ██████████ toch ging vermoeden dat er wat aan de hand is met de artikelen

██████████ geeft desgevraagd aan dat hij het eens is met de conclusie dat de gemiddelden wel erg dicht bij elkaar liggen. Op de vraag of het gebruikelijk is in dit soort studies om bijvoorbeeld 20 experimenten te doen en er maar 1 te rapporteren geeft ██████████ aan dat marketing journals minimaal 3 tot 4, maar meestal 7 tot 8 experimenten vragen. ██████████ denkt dat in het vakgebied van Smeesters de helft van de studies uitkomen en dat studies die experimenten die geen significant verschil geven doorgaans niet in het artikel worden opgenomen. Het is wel belangrijk om het core effect een aantal keren te repliceren.

Gesprekken met promovendi

Uit de gesprekken met de promovendi blijkt dat zij geen twijfels gehad hebben of hebben over de integriteit van Smeesters. Het vermoeden van ██████████ dat er wellicht iets mis is met de data komt voor hen volledig uit de lucht vallen. De promovendi verklaarden verder dat Smeesters veel in het lab aanwezig was.

5.3 Ad 3: Analyse data

Analyse papers / gepubliceerde data.

De analyse van ██████████ (persoonlijke communicatie en ongepubliceerd manuscript) is gebaseerd op de aanname dat er ██████████

██████████

██████████ beargumenteert dat als ██████████ onwaarschijnlijk is. Als

maat ██████████. Via een simulatie kan onderzocht worden hoe waarschijnlijk het is om de

██████████.

██████████

██████████

██████████

██████████

[REDACTED]
[REDACTED]

De onderzochte maten zijn:

[REDACTED]

[REDACTED] gebruikt zelf een derde maat

[REDACTED]

Probleem met deze maat is dat [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED] Daarom

wordt in dit rapport alleen maten a en b berekend en zal vooral gekeken worden naar maat b.

De opzet is om voor zoveel mogelijke artikelen van Smeesters waarvoor de gegevens beschikbaar zijn [REDACTED] Bij het doen van meerdere toetsen is dient een correctie toe te passen om te corrigeren voor meervoudig toetsen. In dit rapport maken wordt gebruik gemaakt van de False Discovery Rate (FDR).

De methode toegepast op vergelijkbare artikelen

Om te onderzoeken of de door [REDACTED] geconstateerde patronen mogelijk ook optreden bij andere willekeurige auteurs is de bovenstaande methode toegepast op vier externe artikelen voor de maten a en b. Deze artikelen zijn geselecteerd door een publicatie van Smeesters te nemen, dan het artikel te nemen dat zo dicht mogelijk voorafgaand aan de publicatie van Smeesters is verschenen, dat een vergelijkbare onderzoeksopzet heeft als het onderzoek van Smeesters, en dat de benodigde gegevens bevat. De geselecteerde artikelen zijn te vinden in bijlage 3.

Tabel 1 in bijlage 3 laat de resultaten zien van de analyse van [REDACTED] voor de vier externe artikelen. [REDACTED]

[REDACTED]
[REDACTED]
[REDACTED]
[REDACTED]

De methode toegepast op de artikelen van Smeesters

In eerste instantie zijn alle 29 artikelen (zowel gepubliceerd als nog onder review) met Smeesters als coauteur onderzocht. Vervolgens is Smeesters gevraagd om per artikel en experiment aan te geven wie de data-analist was, wie de experimentator, en is gevraagd de data aan te leveren. Dit overzicht staat in bijlage 3 Tabel 3. Vervolgens heeft de commissie die artikelen geselecteerd vanaf 2007 ([REDACTED] [REDACTED]) waarvoor Smeesters op enig moment de controle had over de data.

De bovenstaande kansen konden alleen berekend worden indien de gegevens (gemiddelde per groep, standaardafwijking per groep, grootte per groep) van een afhankelijke variabele vermeld werd in het artikel. Bovendien dient ook een verwachting uitgesproken te zijn in het artikel over het

verwachte effect zodat sets van groepen gemaakt kunnen worden. Indien er geen verwachting uitgesproken wordt in het artikel of als de set slechts een groep bevat, dan kan er geen kans berekend worden (omdat er ten minste twee groepen nodig zijn in een set om de standaardafwijking van groepsgemiddelden te kunnen bepalen). Verder is per artikel de kernvariabele meegenomen. In het totaal waren er voldoende gegevens om de methode toe te passen bij 22 afhankelijke variabelen in 22 verschillende experimenten.

De simulatie resultaten voor deze 22 geselecteerde afhankelijke variabelen staan in bijlage 3. Bij 22 toetsen (met een grenswaarde van 5%) mag verwacht worden dat er 1.1 een p-joint waarde kleiner dan 5% hebben. Tabel 2 (in de bijlage 3) laat zien dat p-joint waarden van 6 afhankelijke variabelen kleiner zijn dan 5%. Deze observatie duidt op een patroon van te weinig spreiding. Tegelijkertijd heeft deze observatie onvoldoende bewijskracht om een of meerdere individuele afhankelijke variabele aan te wijzen.

Met de FDR correctie voor meervoudig toetsen (p-FDR) is het wel mogelijk om individuele afhankelijke variabelen aan te wijzen. Er blijken drie variabelen met een p-FDR te hebben die kleiner is dan 5%:

- artikel 1, experiment 2, afhankelijke variabele [REDACTED], p-FDR = 3.85%
- artikel 7, experiment 1, afhankelijke variabele [REDACTED], p-FDR = 3.85%
- artikel 27, experiment 5, afhankelijke variabele [REDACTED], p-FDR = 4.97%

Voor deze drie afhankelijke variabelen is de sterkte van het bewijs tegen de hypothese dat personen willekeurig zijn toegewezen aan de condities is daarmee groot genoeg om deze hypothese te kunnen verwerpen. Op basis van deze gegevens is er voldoende bewijs om aan te nemen dat de groepsgemiddelden systematisch te dicht bij elkaar liggen.

Na confrontatie met deze resultaten heeft Smeesters op 25 mei verklaard dat bij deze drie papers hij proefpersonen weggelaten heeft die de manipulatiecheck niet doorstonden. Hij gaf aan dat het doel van het verwijderen van deze proefpersonen was om de sterkte van het effect te vergroten waarmee de p-waarden significant werden. Het zou gaan om niet meer dan 7 tot 10 proefpersonen per studie. Dit laatste kan de commissie niet verifiëren omdat de ruwe data ontbreken.

Deze verklaring strookt met de bovenstaande resultaten: de simulaties gaan uit van random toewijzing van proefpersonen aan condities. Bij de verzamelde data is dat niet meer het geval: Smeesters heeft een selectie achteraf gemaakt van respondenten zodanig dat het gewenste effect versterkt werd.

Analyse van samengestelde databestanden

De commissie beschikt over een aantal samengestelde databestanden. Deze zijn afkomstig uit drie bronnen: (1) direct aangeleverd door Smeesters, (2) aangeleverd door Smeesters' coauteur [REDACTED] en (3) aangetroffen op de netwerkschijf van [REDACTED]. Analyse van al deze bestanden vergt meer tijd dan de commissie tot haar beschikking heeft. De commissie heeft zich daarom beperkt tot bestanden uit twee artikelen waarin de gerapporteerde data onwaarschijnlijk leken op basis van de analyses beschreven in het vorige hoofdstuk. Deze bestanden liggen ten grondslag aan de volgende twee publicaties:

[REDACTED] & [REDACTED] (2011). [REDACTED]
[REDACTED]
[REDACTED]

De analyses van de betreffende databestanden zijn te vinden in bijlage 4.

Bij het bestuderen van deze samengestelde bestanden is de commissie op patronen gestuit die variëren van opmerkelijk tot uiterst onwaarschijnlijk. Bovendien is er een databestand aangetroffen op de netwerkschijf van Smeesters dat volgens wat hij heeft meegedeeld aan de commissie en wat zijn coauteur heeft bevestigd, niet in zijn bezit zou moeten zijn. Op basis van deze serie bevindingen heeft de commissie geen vertrouwen in de wetenschappelijke integriteit van de resultaten die in de twee genoemde artikelen gerapporteerd worden.

Ruwe data

Een systematisch search op de EBL computers leverde geen ruwe databestanden van Smeesters op. De praktijk van veel ██████████ is om alle data van de labcomputers te wissen. Er zijn ook geen ruwe data op de netwerkschijf van Smeesters aangetroffen.

Op grond van bovenstaande is het voor de commissie onmogelijk om vast te stellen of

- de data waarover gepubliceerd is inderdaad verzameld zijn in het EBL. Er werden wel experimenten gedraaid door Smeesters, maar het is onmogelijk om te bepalen of deze corresponderen met gepubliceerde resultaten en
- er fraude is gepleegd, dan wel iets mis is gegaan tijdens het compileren van de individuele databestanden in samengestelde databestanden.

Overige Bevindingen

De commissie heeft navraag gedaan over mogelijkheid om data terug te vinden op harde schijf die gecrasht. ██████████ heeft aangegeven dat het zeer ongebruikelijk is dat er niets meer terug te vinden is op een harde schijf. In 99 van de 100 gevallen kan informatie nog gevonden worden.

6 Conclusies

6.1 Conclusies algemeen

De commissie is van mening dat de twijfels, gerezen door het onderzoek van dr. ██████████, ook na de huidige analyse van de gegevens blijven voortbestaan. Hierbij neemt de commissie naast de statische bevindingen, het ontbreken van ruwe data en het toepassen van selectie van data om significantie te bereiken in een aantal papers in haar conclusie en aanbevelingen mee.

Daarbij vallen een aantal zaken op:

1. Prof. Smeesters is een onderzoeker die vele data zelf verzamelt, zelf bewerkt en zelf analyseert, waarbij de controle van buitenaf zeer gering is. Binnen de afdeling is met hem weinig sprake van "peer review."
2. Ruwe data, zowel elektronisch als op papier van alle artikelen van dhr. Smeesters ontbreken. De commissie twijfelt aan de geloofwaardigheid van de redenen hiervoor.
3. De statistische analyse zoals uitgevoerd in het hoofdstuk gepubliceerde data/analyse papers toont een patroon van groepsgemiddelden aan dat te dicht bij elkaar licht over verschillende papers.

4. Bij drie artikelen (1, 7 en 27) is de commissie op patronen in de data gestuit die uiterst onwaarschijnlijk zijn. Deze patronen zijn vermoedelijk te wijten aan dataselectie door Smeesters. De commissie heeft geen vertrouwen in de wetenschappelijke integriteit van de resultaten in deze drie artikelen.
5. Smeesters heeft zelf (uiteindelijk) toegegeven bij de betreffende drie artikelen data te hebben verwijderd om significante effecten te krijgen.

6.2 Aanbevelingen:

- Prof. Smeesters op de hoogte stellen van de conclusies van de onderzoekscommissie wetenschappelijke integriteit alsmede van de verdere aanbevelingen.
- Prof Smeesters mededelen dat het verdwijnen van ruwe data en selectie van data in 3 papers hem persoonlijk wordt aangerekend
- Via professor [REDACTED] ([REDACTED]) de co-auteurs van deze artikelen hiervan op de hoogte brengen.
- Via professor [REDACTED] de betreffende tijdschriften benaderen met het verzoek de artikelen 1, 7 en 27 terug te trekken omdat de Erasmus Universiteit niet kan instaan voor de correctheid van de gebruikte data. Smeesters heeft immers toegegeven dat de gerapporteerde effecten alleen significant werden nadat er strategisch data waren verwijderd.

6.3 Algemene aanbevelingen:

- Aangezien Prof. Smeesters herhaaldelijk aangeeft dat de cultuur binnen zijn werkveld en zijn afdeling zodanig is dat hij zich als persoon niet schuldig voelt en ervan overtuigd is dat velen in publicaties op het gebied van de marketing en (in mindere mate) sociale psychologie gebruik maken van het bewust weglaten van data om significantie te bereiken, zonder dit te vermelden, dient er onderzoek / regelgeving op dit gebied te komen.
- Binnen de afdeling van Smeesters ([REDACTED]) maar mogelijk breder binnen de Erasmus Universiteit dient er een vast protocol te zijn voor dataverzameling en dataopslag waaraan iedere wetenschapper zich dient te houden. Het protocol moet er op gericht zijn om direct na dataverzameling de ruwe data te fixeren. Bovendien dient protocol te beschrijven hoe de data die bij uiteindelijke publicatie gebruikt zijn eenvoudig te herleiden zijn naar de ruwe data.