

## University of Groningen

### The non-existent average individual

Blaauw, Frank Johan

**IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.**

*Document Version*

Publisher's PDF, also known as Version of record

*Publication date:*

2018

[Link to publication in University of Groningen/UMCG research database](#)

*Citation for published version (APA):*

Blaauw, F. J. (2018). *The non-existent average individual: Automated personalization in psychopathology research by leveraging the capabilities of data science*. University of Groningen.

**Copyright**

Other than for strictly personal use, it is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license (like Creative Commons).

**Take-down policy**

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

*Downloaded from the University of Groningen/UMCG research database (Pure): <http://www.rug.nl/research/portal>. For technical reasons the number of authors shown on this cover page is limited to 10 maximum.*

# **The non-existent average individual**

Automated personalization in psychopathology research  
by leveraging the capabilities of data science

Frank Johan Blaauw

*The non-existent average individual: Automated personalization in psychopathology research by leveraging the capabilities of data science*

© Copyright 2018, F. J. Blaauw, the Netherlands

All rights reserved. No parts of this dissertation may be reproduced or transmitted in any form or by any means, without the written permission from the author or, when appropriate, from the publishers of the publications.

Published by *Ridderprint* – [www.ridderprint.nl](http://www.ridderprint.nl) – Ridderkerk. The illustration on the cover was provided by *Patrick Léger* – [www.patrick-leger.com](http://www.patrick-leger.com) – who generously permitted usage of this illustration for the cover. The color palette used throughout this dissertation is based on the *Nord* color palette – [www.git.io/nord](http://www.git.io/nord).

This dissertation was realized in collaboration with the Espria Academy. Espria is a health care group in the Netherlands consisting of multiple companies targeted mainly at the elderly population.

ISBN: 978-94-034-0405-9 (printed)  
ISBN: 978-94-034-0404-2 (electronic)

The logo for Espria, featuring the word "espria" in a lowercase, sans-serif font. The "e" is red, and the remaining letters "spria" are blue.



university of  
 groningen

# The non-existent average individual

Automated personalization in psychopathology research  
 by leveraging the capabilities of data science

**PhD thesis**

to obtain the degree of PhD at the  
 University of Groningen  
 on the authority of the  
 Rector Magnificus Prof. E. Sterken  
 and in accordance with  
 the decision by the College of Deans.

This thesis will be defended in public on  
 Monday 12 February 2018 at 11.00 hours

by

**Frank Johan Blaauw**

born on 14 June 1989  
 in Groningen

**Supervisors**

Prof. P. de Jonge

Prof. M. Aiello

**Co-supervisor**

Dr. J.A.J. van der Krieke

**Assessment Committee**

Prof. J.W. Romeijn

Prof. N. Petkov

Prof. S. Dustdar

---

# Contents

<b>Acknowledgments</b>	<b>xi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 A Classification System . . . . .	3
1.2 Group and Individual Data to Improve Well-being . . . . .	8
1.3 Back to Dimensionality, and its Curse . . . . .	10
1.4 Scope and Contribution of this Dissertation . . . . .	11
1.5 Outline . . . . .	12
<b>2 E-mental Health and Personalized Psychiatry</b>	<b>15</b>
2.1 Precision Medicine . . . . .	16
2.2 Psychopathology as a Time Series . . . . .	19
2.3 Predicting and Explaining Psychopathology . . . . .	22
2.3.1 Time Series Analysis . . . . .	22
2.3.2 Machine Learning Perspective . . . . .	26
<b>I Monitoring and Measuring Psychopathology Online</b>	<b>31</b>
<b>3 An Online Platform for Personalized Well-being</b>	<b>33</b>
3.1 HowNutsAreTheDutch and Leefplezier . . . . .	34
3.2 Crowdsourcing Procedure . . . . .	35
3.3 Shifting Perspectives: From the Population to the Individual . . . . .	36
3.3.1 Cross-sectional Study . . . . .	37
3.3.2 Ecological Momentary Assessments . . . . .	38
3.4 Discussion and Concluding Remarks . . . . .	43

<b>4</b>	<b>Architecture and Infrastructure of HowNutsAreTheDutch and Leefplezier</b>	<b>45</b>
4.1	Service-Oriented Architectures in E-mental Health . . . . .	46
4.2	Two Case Studies . . . . .	47
4.2.1	The Architecture of HowNutsAreTheDutch . . . . .	47
4.2.2	The Architecture of Leefplezier . . . . .	49
4.2.3	Technical Overview . . . . .	50
4.3	Comparison . . . . .	50
4.3.1	Data Security . . . . .	51
4.3.2	Conducting Questionnaires . . . . .	52
4.3.3	Feedback Generation . . . . .	53
4.3.4	Feedback Visualization . . . . .	54
4.3.5	Content Management . . . . .	54
4.4	Requirements of E-mental Health Applications . . . . .	55
4.4.1	Data Security and Patient Privacy . . . . .	55
4.4.2	Maintainability of the E-mental Health Platform . . . . .	55
4.4.3	Availability and Reliability for Data Collection . . . . .	56
4.5	Proposed Architecture . . . . .	56
4.6	Discussion and Concluding Remarks . . . . .	59
<b>5</b>	<b>HowNutsAreTheDutch Descriptives and Results</b>	<b>61</b>
5.1	Cross-Sectional Results . . . . .	61
5.1.1	Sample Characteristics . . . . .	62
5.1.2	Key Results . . . . .	62
5.1.3	Evaluation . . . . .	67
5.2	Diary-study Results . . . . .	67
5.2.1	Sample Characteristics . . . . .	67
5.2.2	Adherence and Completion Rates . . . . .	68
5.2.3	Automatically Generated Feedback and Evaluation . . . . .	70
5.2.4	Between-Persons and Within-person Associations . . . . .	71
5.3	Discussion and Concluding Remarks . . . . .	71
<b>II</b>	<b>Automatically Personalizing Psychopathology Research</b>	<b>75</b>
<b>6</b>	<b>Personalized improvement of well-being: Automated Impulse Response Analysis</b>	<b>77</b>
6.1	Automated Diary Study Data Analysis . . . . .	78
6.2	Impulse Response Function Analysis and Ecological Momentary Assessment Advice . . . . .	80
6.3	From Variable Selection to Advice Generation . . . . .	81

## Contents

---

6.3.1	Initialization . . . . .	81
6.3.2	Simulation . . . . .	84
6.3.3	Variable Selection . . . . .	86
6.3.4	Advice Generation . . . . .	88
6.4	Algorithms . . . . .	90
6.4.1	Selecting Variables and Determining Advice . . . . .	90
6.4.2	Time Complexity . . . . .	93
6.5	Experimental Results . . . . .	95
6.5.1	Most Influential Node . . . . .	95
6.5.2	Length of the Effect . . . . .	96
6.5.3	Percentage Effect . . . . .	97
6.6	Real-world Application of Automated Impulse Response Analysis . . . . .	98
6.6.1	Aims of the Study . . . . .	100
6.6.2	Methods . . . . .	100
6.6.3	Study Results . . . . .	103
6.6.4	Discussion . . . . .	108
6.7	Discussion and Concluding Remarks . . . . .	111
<b>7</b>	<b>Machine Learning for Precision Medicine in Psychopathology Research</b>	<b>113</b>
7.1	Methods . . . . .	115
7.1.1	The Machine Learning Procedure . . . . .	118
7.1.2	Random Hyperparameter Search Procedure . . . . .	121
7.1.3	Synthetic Minority Over-sampling . . . . .	122
7.1.4	Performance Measures . . . . .	123
7.1.5	Application and Implementation Details . . . . .	124
7.2	Results . . . . .	125
7.3	Discussion and Concluding Remarks . . . . .	127
<b>8</b>	<b>Exploring the causal effects of activity on well-being using online targeted learning</b>	<b>129</b>
8.1	Quick Historical Overview of the Targeted Learning Methodology . . . . .	130
8.2	HowNutsAreTheDutch . . . . .	131
8.2.1	The Study Protocol and Data set . . . . .	132
8.2.2	Formalization . . . . .	133
8.3	Causal and Probabilistic Perspectives . . . . .	135
8.3.1	Probabilistic Framework . . . . .	135
8.3.2	Unrealistic Ideal Experiments . . . . .	136
8.3.3	Causal Model, Counterfactuals, and Quantity . . . . .	137
8.4	Statistical Model . . . . .	139



8.4.1	Nonparametric Statistical Model . . . . .	139
8.4.2	Counterfactual Nonparametric Statistical Model . . . . .	141
8.4.3	Target Statistical Parameter . . . . .	141
8.5	Online Targeted Learning . . . . .	142
8.5.1	Overview . . . . .	143
8.5.2	On Machine Learning of Infinite-dimensional Features . . . . .	144
8.5.3	Step one: infinite-dimensional features . . . . .	149
8.5.4	Step two: targeting the parameters of interest . . . . .	153
8.6	Simulation study . . . . .	156
8.6.1	Simulation scheme . . . . .	156
8.6.2	Implementation . . . . .	158
8.6.3	Simulation results . . . . .	159
8.7	Application to the HowNutsAreTheDutch data set . . . . .	160
8.8	Discussion and Concluding Remarks . . . . .	163
<b>9</b>	<b>Augmenting Ecological Momentary Assessments with Physiological Data</b>	<b>165</b>
9.1	Combining Sensor Technology With Ecological Momentary Assessments . . . . .	166
9.2	Background . . . . .	167
9.3	Physiqua . . . . .	169
9.3.1	Architecture . . . . .	170
9.3.2	Service Layer and Service Providers . . . . .	171
9.3.3	Aggregation and Processing Layer . . . . .	172
9.3.4	Imputation Layer . . . . .	175
9.3.5	Presentation Layer . . . . .	175
9.4	Case Study . . . . .	176
9.4.1	Ecological Momentary Assessments and Sensors . . . . .	176
9.4.2	Statistical Analyses . . . . .	177
9.5	Validation . . . . .	178
9.5.1	Effectiveness . . . . .	178
9.5.2	Accuracy . . . . .	180
9.6	Software Implementation . . . . .	180
9.7	Case Study Results . . . . .	181
9.8	Discussion and Concluding Remarks . . . . .	182
<b>10</b>	<b>Discussion: Personalization in Psychopathology Research</b>	<b>185</b>
10.1	HowNutsAreTheDutch and Leefplezier . . . . .	185
10.2	Automated Impulse Response Analysis . . . . .	187
10.3	Machine Learning for a More Precise Medicine . . . . .	190

## Contents

---

10.3.1 Interindividual Perspective . . . . .	190
10.3.2 Intraindividual Perspective . . . . .	193
10.4 Ecological Momentary Assessments and Wearables . . . . .	195
<b>11 Conclusion and Future Perspectives</b>	<b>197</b>
<b>A HowNutsAreTheDutch and Leefplezier — Supplement</b>	<b>199</b>
A.1 HowNutsAreTheDutch . . . . .	199
A.2 Leefplezier . . . . .	210
<b>B Automated Impulse Response Analysis — Supplement</b>	<b>217</b>
B.1 Impulse response calculation . . . . .	217
B.2 Time complexity . . . . .	218
<b>C Machine Learning for Precision Medicine — Supplement</b>	<b>221</b>
<b>D Online Super Learner — Supplement</b>	<b>229</b>
D.1 The questionnaire items . . . . .	229
D.2 Relevant source code . . . . .	229
D.2.1 The conditional density estimation algorithm . . . . .	230
D.2.2 Conditional density sampling . . . . .	230
D.2.3 Monte-Carlo sampling algorithm . . . . .	231
<b>Bibliography</b>	<b>239</b>
<b>Summary</b>	<b>289</b>
<b>Samenvatting</b>	<b>291</b>



---

## Acknowledgments

In pursuit of a Ph.D. many, but definitely not all people<sup>1</sup> experience various disorders listed in the Diagnostic and Statistical Manual of Mental Disorders (DSM). Disorders ranging from *major depressive disorders* via *adult attention deficit disorders* to *substance (ab)use disorders*. I count myself lucky that my pursuit towards this Ph.D. went without any major problems, and I was spared from any DSM classifications. I think a large part of this ‘luck’ was influenced by the people that supported me during the last years. And although thanking everybody who helped me achieve this goal would be impossible, I would like to mention a few people who have been (and still are) extremely important for me for achieving this goal.

First of all I would like to thank all people who supervised me in the past four (or more) years. Marco, you have been a tremendous support. Before I started my Ph.D. I did my master’s program and thesis with you. When you offered me a Ph.D. position it was clear that if it was a project in your group, it had to be cool and interesting. And it was. You convinced me to start this journey and I’m extremely grateful for that. I always thought you were a very inspiring person, and now I’m certain. From email conversations at 2:00 AM to drinking grappa together in Rome, it is always a pleasure to talk to you. Your opinions (or rants) related to science, bureaucracy, and, well, everything else, have been eye-opening and changed my naïve perspectives in a good way. Thanks!

Ten tweede, Peter, jij zorgde ervoor dat ik daadwerkelijk in de wetenschap ingebed werd. Je hebt me altijd met de juiste mensen weten te koppelen, waardoor de projecten en het onderzoek dat we deden sterk verbeterden. Daarnaast introduceerde je mij aan Mark, waardoor ik een fantastische ervaring in de Verenigde staten kon op doen. Het vertrouwen dat je me al deze jaren hebt gegeven voelde erg fijn.

---

<sup>1</sup>See for example this dissertation. . .

Ik kon altijd bij je aankloppen voor advies en daarvoor ben ik je erg dankbaar. Ik kijk ernaar uit om met je samen te blijven werken. Bedankt voor alles!

Ten derde, Lian, jij verdient een speciale plek. Ik ben ervan overtuigd dat ik zonder jou dit project niet had kunnen afronden. Op het moment dat het slecht ging trok jij aan de bel en stelde je me gerust. Je bent echt een fantastisch persoon en ik had onze samenwerking voor geen goud willen missen.

Besides my supervisors in the Netherlands, I also had several supervisors during my visit to UC, Berkeley. Mark, thank you for accepting me in your group and helping me with all my questions related to our project. It was a pleasure to work with you. Antoine, you too deserve a special thanks. I will never forget that I had only been in the U.S. for a couple of days after I terribly hurt my back and was unable to walk. Even though we had never met, your emails were amongst the kindest that I had ever received. You are a truly special individual with whom I really enjoy working. Our Skype sessions (and real-life sessions) are very inspiring and helped me get a much better understanding of the targeted learning methodology.

I would also like to thank the reading committee for assessing my dissertation. Prof. Dustdar, Prof. Petkov, and Prof. Romeijn, thank you for the time and effort you put in.

Een promotie traject is niet iets dat je alleen doet. En hoewel ik iedereen erg dankbaar ben voor alle samenwerkingen, verdienen mijn paranimfen een bijzondere plek. Ando en Maria, wat een fantastische tijd heb ik met jullie samen gehad. Ando, bij het ICPE en ontwikkelingspsychologie zijn wij toch een beetje de vreemde eenden in de bijt. Het was (en is) fijn om deze positie met jou te delen. Bedankt voor alle discussies, Gay pride uitjes, gesprekken, en lol die ik met jou heb gehad. Maria, ontzettend bedankt voor al je advies en opbeurende woorden. Vanaf mijn sollicitatiegesprek tot jouw vertrek uit het ICPE heb ik altijd met veel plezier met je samengewerkt en gepraat. De intense stress van de HoeGekIsNL downtime vijf minuten voor een radio interview zal ik nooit vergeten... Het is een eer dat jullie mijn paranimfen willen zijn. Behalve mijn paranimfen had ik nog meer kamergenoten met wie ik een erg fijne tijd heb gehad. Ricardo, Jan (van Bebbber) en Ella, ook jullie erg bedankt voor de leuke tijd samen.

Ik heb altijd met veel plezier op verschillende projecten gewerkt. Van het HoeGekIsNL project wil ik (los van de *usual suspects*) Stijn, Rob, Inge, Klaas, Hanneke, Marieke, Evelien en in het bijzonder Elske erg voor bedanken. Elske in het bijzonder omdat ook jij een cruciale rol hebt gespeeld in mijn promotie. Bescheiden als je bent zul je waarschijnlijk zeggen dat dit niet het geval is, maar onze meetings waren erg waardevol voor mij! Bij het Leefplezier project heb ik samengewerkt met ontzettend veel verschillende mensen, Ester, Chantal, Joris, en iedereen van Lifely, heel erg bedankt voor deze samenwerking. Bertus wil ik hierbij niet overslaan. Bertus,

jouw adviezen en ongeëvenaarde kennis waren altijd een feest. Heel erg bedankt voor alles! Hoewel al mijn collega's noemen teveel is, wil ik toch nog Fionneke, Jan (Houtveen), iedereen van RoQua (Erwin in het bijzonder), Esther, Margo, en Gerry bedanken.

Besides my time at the UMCG, I also very much enjoyed my time in the distributed systems group. Ang, Alexander, Andrea, Azkario, Brian, Esmee, Faris, Fatimah, Heerko, Laura, Talko, Tuan, and Viktoriya, although I only spent (at most) one day a week at the Bernoulliborg, I always felt like a full member of the group! Ilche, a special thanks to you. When I was in doubt whether to do a Ph.D. and asked you for advice, you replied with a great, and very helpful, honest message, which I would immediately cite if someone would ask me about doing a Ph.D., because indeed, it *"is quite pleasurable and arduous at the same time"* (Georgievski, 2013).

Another group I enjoyed being in was the Biostatistics group at UC, Berkeley. I would like to thank Alejandra, Caleb, Courtney, George, Johnathan, Kelly, Lina, Nima, Oleg, Yue, Sharon. A special thanks goes to Rachel and Robert, Sarah, and Tommy, with whom I spent hours and hours of struggling through math problems, and who treated me like family! Besides the people from the university I also would like to thank Steve, Monica, Aida, and of course Mallorie. Mallorie you are such an amazing person. The moment I arrived you had soup ready and you took care of me while I had my back problems. You introduced me to Dr. Squeeze and the Hugz and helped me with uncountably many things. Thank you so much!

Tenslotte zijn er naast mijn werk nog een aantal mensen die mij altijd erg geholpen hebben en mijn gezeur over mijn promotietraject wilden aanhoren. Allereerst wil ik mijn ouders en schoonouders bedanken. Willem, Eppy, Gert en Roelie, heel erg bedankt voor jullie opbeurende woorden als het even tegen zat en voor jullie interesse als het goed ging. Trijntje, Esther, Marten, Martijn, Heleen, Sjon en Gé, ook jullie bedankt!

Tijdens dit promotie traject heb ik helaas twee heel dierbare mensen moeten verliezen. Twee mensen die ontzettend trots zouden zijn als ze hierbij konden zijn. Lieve Oma's, wat waren jullie altijd geweldig. Ook al hadden jullie misschien geen flauw idee waar mijn onderzoek over ging, jullie waren altijd even geïnteresseerd. Ik mis jullie nog altijd, en zal jullie nooit vergeten.

Naast mijn familie stonden ook mijn vrienden altijd voor me klaar. Jarnick en Alexandra, Maarten, Alko en Juliët, Arjen en Josien (en Linn), Martijn en Kim, Geoffrey en Klaske, Simon en Nilanka, Mariëtte en Hans, Femke en Ine, hoewel mijn *substance use disorder* misschien lichtelijk is verhoogd door jullie, hielden jullie me wel met beide benen op de grond, en hielpen jullie me herinneren wat echt belangrijk is.

Dan rest mij nog één persoon om te bedanken, en zoals wel vaker gebruikelijk

is in de wetenschappelijke literatuur, is de laatste persoon de belangrijkste. Inge, wat ben je een geweldige vrouw. Ik heb je vaak geplaagd met hoe ik je in mijn dankwoord zou opnemen, maar eigenlijk maakt het niet uit wat ik hier zeg want woorden schieten toch te kort. Als er iemand is geweest die mij door dit avontuur heen heeft gesleept ben jij het wel. Dit promotie traject was voor jou af en toe net zo moeilijk (al dan niet moeilijker) als voor mij, maar jij was er altijd voor me. Ontzettend bedankt voor alles!

Frank Blaauw  
Groningen  
January 12, 2018