Gerko Vink

November 2022

G.Vink@uu.nl — Web — Linkedin — GitHub

UTRECHT UNIVERSITY

Department of Department of Methodology and Statistics
Padualaan 14, room B.110
Utrecht University
3584 CH Utrecht

Education

2015 **PhD**, *Utrecht University - Department of Methodology and Statistics*, Utrecht, Defended on 13 March 2015

Title of dissertation: Restrictive imputation of incomplete survey data. pdf

- 2009 MSc Methodology and Statistics in Psychology (with honours), Leiden University, Leiden
- 2008 BSc in Psychology, Leiden University, Leiden
- 2005 2008 Photography, Royal Academy of Arts, The Hague

Professional experience

- 2020 now Associate Professor / UHD, Department of Methodology and Statistics, Utrecht University

 June 2020 present
- 2015 2020 Assistant Professor / UD, Department of Methodology and Statistics, Utrecht University
 September 2015 May 2020
- 2016 2017 **Adjunct Associate Research Scholar**, *Applied Statistics Center*, Columbia University, New York
 February 2016 January 2017
- 2014 2015 **Postdoctoral Fellow**, *Department of Methodology and Statistics*, Utrecht University September 2014 July 2015
- 2012 2014 **Lecturer [.5 fte]**, *Department of Methodology and Statistics*, Utrecht University September 2012 August 2014
- 2012 2014 **AIO (doctoral candidate) [.5 fte]**, Department of Methodology and Statistics, Utrecht University, and Department of Methodology and Quality, Statistics Netherlands (CBS)

 September 2012 August 2014
- 2009 2012 **AIO (doctoral candidate)**, Department of Methodology and Statistics, Utrecht University, and Department of Methodology and Quality, Statistics Netherlands (CBS)

September 2009 - August 2012

Management positions

- 2020 now Member Programmaraad ADS master, Applied Data Science, Utrecht University
- 2020 now **Coordinator ADS education**, Department of Methodology and Statistics, School of Social Sciences, Utrecht University
- 2017 2021 **PhD counselor (mentor)**, Department of Methodology and Statistics, Faculty of Social Sciences, Utrecht University
- 2012 2019 **Project Coordinator**, *Interactive digital education*, Research master program *Methodology and Statistics for the Behavioural, Biomedical and Social Sciences*, Faculty of Social Sciences, Utrecht University
- 2015 2018 **Educational Advisor**, *Programme Advisory Committee*, Research master program *Methodology and Statistics for the Behavioural, Biomedical and Social Sciences*, Faculty of Social Sciences, Utrecht University
- 2013 2016 **Coordinator**, *Minor Methodology and Statistics in the Social and Behavioural Sciences*, Faculty of Social Sciences, Utrecht University
- 2012 2016 **Student counselor (tutor)**, Research master program *Methodology and Statistics* for the Behavioural, Biomedical and Social Sciences, Faculty of Social Sciences, Utrecht University
- 2012 2015 **Educational Representative**, Educational Advisory Committee, Research master program Methodology and Statistics for the Behavioural, Biomedical and Social Sciences, Faculty of Social Sciences, Utrecht University

Received funding ≥ €5.000

- 2022 **Project lead**, *Dessert: An R-package for standardized model evaluation*, For €15,000, Sept 2022 August 2023
- 2022 **Project lead**, *KOE: Kenniscentrum Open Educatie. FUSO project*, For €24,000, September 2022 December 2023
- 2022 **Consortium member**, *Education with professionals. USO project*, September 2022 December 2023
- 2020 **Project lead**, *Shinymice: model building and evaluation for multiple imputation*, For €25,000, June 2020 December 2020
- 2020 **Investigator**, *Machine Learning for Eye-tracking data*, For €15,000, DoY, Utrecht, the Netherlands
 February 2020 December 2020
- 2019 **Investigator**, *Multiple imputation of oil production rate time series*, For €5,000, TNO, the Netherlands

 July 2019 December 2019
- 2014 Investigator, Multiple imputation of Parent Management Training Oregon (PMTO) data, with Stef van Buuren (UU/TNO) for €12,000 (my share: €6,000), Maastricht University, the Netherlands

 November 2014 December 2014

2013 Investigator, Multilevel Bayesian estimation of multiply imputed social-educational advantage scores over time, with Stef van Buuren (UU/TNO) for €42,477 (my share: €21,238.50), Australian Curriculum, Assessment and Reporting Authority, Sydney, Australia

March 2013 - August 2013

2012 **Investigator**, *Multiple imputation of large register data with multilevel structures*, with Stef van Buuren (UU/TNO) for €53,503 + €5,000 bonus (my share: €29,251.50), Australian Curriculum, Assessment and Reporting Authority, Sydney, Australia

August 2012 - December 2012

Relevant professionalisation

- 2020 Corporate Entrepeneurship, Columbia Business School, New York, NY, USA
- 2016 Big Data and Social Analytics, Massachusetts Institute of Technology, Cambridge, MA, USA

Teaching experience

Undergraduate courses

	Undergraduate courses
2019 – 2021	Fundamental Techniques in Data Science, Utrecht University, Course coordinator
2015 - 2021	Missing Data Theory and Causal Effects, Utrecht University, Course coordinator
2018	Analysis of Behavioural Data, University College Utrecht
2016 – 2018	Methods & Statistics 2, Utrecht University
2012 – 2018	Conducting a Survey, Utrecht University
2011 – 2015	Applied Multivariate Statistics, University College Utrecht
2013	Methods & Statistics I, Utrecht University
2013	Ondersteuning Leeronderzoek, Utrecht University
2009 – 2011	Methods & Statistics I, University College Utrecht
2008 - 2009	Inleiding in de Methoden en Technieken, Leiden University
2008 - 2009	Psychometrie, Leiden University
2008 - 2009	Experimenteel en Correlationeel Onderzoek, Leiden University
2008 - 2009	Toetsende statistiek, Leiden University
2008 - 2009	Multivariate Data Analyse, Leiden University
	Graduate courses
*2021 – now	Supervised Learning and Visualization, Utrecht University, Course coordinator
*2017 – now	Markup Languages and Reproducible Programming in Statistics, <i>Utrecht University</i> , Course coordinator
2013 - now	Master Thesis Supervision, Utrecht University
2012 – 2016	Research Seminar I, Utrecht University
2014	Research Seminar II, Utrecht University

2012 – 2015 Statistical Programming in R elective course, Utrecht University

2009 – 2013	Imputation Techniques, Utrecht University
2012	Statistical Programming in R, Utrecht University
2009	Applied Multivariate Data Analysis, Leiden University
	Post-graduate courses
2016	Multiple Imputation in Practice, Utrecht University
2013	Applied Multiple Imputation using SPSS and R, Utrecht University
	Summerschool courses
*2016 – now	Data Science: Multiple Imputation in Practice , <i>Utrecht University</i> , Course coordinator, rated 4.65/5
2015 – 2021	Data Science: Statistical Programming in R , <i>Utrecht University</i> , Course coordinator, rated 4.7/5
2018	Advanced Multilevel Analysis, Utrecht University
2014 - 2016	Advanced Issues in Survey Design and Analysis
2013	Survey Research: Statistical Analysis and Estimation
	Supervised students
2021	Thom Volker. Internship Thesis
	Anaïs Fopma. Graduate Thesis
2020	Jihane Cherif. Undergraduate Thesis
2020	Paulina von Stackelberg. Graduate Thesis
2019	Hanne Oberman. Graduate Thesis
2019	Lonneke Theeuwes. Undergraduate Thesis
2018	Arjan Huizing. Graduate thesis
2016	Rianne Schouten. Graduate thesis
2016	Xynthia Kavelaars. Graduate thesis
2016	Arnold Lau (Columbia University). Graduate thesis
2015	Lisanne Rijnveld. Research Master paper
2015	Wouter Smink. Graduate thesis
2015	Tom Bergkamp. Undergraduate thesis
2015	Henk Jan Stevens. Graduate thesis
2014	Bo Kuiper. Graduate thesis
2014	Laura Boeschoten. Graduate thesis
	Phd students
2021 – now	Hanne Oberman [ongoing project]
2018 - 2022	Mingyang Cai
2016 - 2020	Duco Veen
2017 - 2019	Rianne Schouten

2022 – now Mingyang Cai [ongoing project]

Invited courses / workshops

- 2022 Creating synthetic data with mice in R [with Thom Volker], UMCU, 4 November 2022, Utrecht, the Netherlands.
- 2022 Inferential and predictive statistics and machine learning, Rabobank, 4-day course, Utrecht, the Netherlands.
- 2021 Data science and predictive machine learning, Erasmus Academy, 4-day course, Rotterdam, the Netherlands.
- 2020 Imputing health care data with mice in R [with Stef van Buuren], UMCU, 30 October 2020, Utrecht, the Netherlands.
- 2019 2022 Multiple imputation in practice [with Stef van Buuren], Data Science Rijkstrainees, The Hague, the Netherlands.
- 2019 2022 Statistical Programming with R, Data Science Rijkstrainees, The Hague, the Nether-
 - 2019 Statistical computation with R, Wageningen Economic Research, September 17; 24&25, The Hague, the Netherlands.
 - 2019 Statistical programming with R, Wageningen Economic Research, September 3; 10&11, The Hague, the Netherlands.
 - 2019 Advanced Multilevel Imputation [with Stef van Buuren], 12th International Multilevel Conference, April 8, Utrecht
 - 2018 Introduction to Statistics with R, September 3-5, Utrecht
 - 2018 Introduction to Statistics with R, July 2-4, Utrecht
 - 2017 Introduction to R: How to get started, August 28, Utrecht
 - 2017 Multiple imputation in R. RECSM / Pompeu Fabra, June 18-29, Barcelona, Spain.
 - 2017 Statistical programming in R. NorthWest University, March 14-17, Vanderbijlpark, South Africa.
 - 2017 Multiple imputation in practice [with Stef van Buuren]. ETH Zürich, February 23-24, Zürich, Switzerland.
 - 2017 Missing value imputation [with Stef van Buuren]. University of Zürich, February 20-21, Zürich, Switzerland.
 - 2016 Multiple Imputation in Practice. TU Berlin/Berlin HECOR [with Stef van Buuren], September 12-14, Berlin, Germany.
 - 2015 Statistical programming with R. Radboud University, December 1, 2015, Nijmegen, the Netherlands.
 - 2015 Applied multiple imputation in R [with Stef van Buuren]. Pre-conference course at the 2015 conference of the International Society for Clinical Biostatistics, August 23, 2015, Utrecht, the Netherlands.

- 2015 Handling missing data in R with MICE [with Stef van Buuren]. Tutorial course at the MissData conference. June 17, 2015, Rennes, France
- 2015 Applied multiple imputation with SPSS. Course within the Boerhaave Clinical Trials educational series. April 14, 2015, Leiden, the Netherlands
- 2015 Working with R: a full day crash-course. Pre-conference course for the 10th International Multilevel Conference, April 7, 2015, Utrecht, the Netherlands
- 2013 Incomplete data workshop: Flexible Imputation of Missing Data [with Shahab Jolani], 1st Southern European Conference on Survey Methodology, December 2013, Barcelona, Spain
- 2013 Applied Multiple Imputation in R [with Stef van Buuren]. CE Course, Joint Statistical Meetings, Aug 4, 2013, Montreal, Canada
- 2013 Multiple imputation for multilevel data using MICE [with Stef van Buuren]. Preconference course. March 26, 2013, Utrecht
- 2011 Pre-conference workshop "How to deal with missing data". Mplus user conference, October 2011, Utrecht (with invited contributions by Stef van Buuren and Craig Enders [UCLA])

Invited presentations and keynotes

- 2019 CMstatistics, Advances in complex data modeling, December 16, London, United Kingdom
- 2019 University of Padua, Advances in Multilevel and Longitudinal Modelling, January 28, Padua, Italy
- 2018 Leiden University, Methods & Statistics Alumni Event, November 22, Leiden, the Netherlands
- 2017 Tilburg University, Statistical colloquium, May 9, Tilburg, the Netherlands
- 2017 ETH Zürich, Statistical symposium, February 22, Zürich, Switzerland
- 2016 Joint Statistical Meetings, August 2016, Chicago, IL, USA, topic contributed session
- 2015 Social Sciences M&S Power Snacks, November 2015, Utrecht, the Netherlands
- 2014 Annual symposium of the Dutch-speaking platform for survey research, 19 May 2014, Amsterdam, the Netherlands
- 2013 Meeting on multiple imputation of multilevel data, June 2013, London, UK
- 2011 Second meeting of the Dutch Mplus users group, May 2011, Utrecht, the Netherlands
- 2011 ISED research dagen 2011, November 2011, Utrecht, the Netherlands

Contributed presentations

- 2015 10th International Multilevel Conference, April 2015, Utrecht, the Netherlands
- 2014 IOPS Graduate School Winter Conference, Amsterdam, the Netherlands
- 2014 Joint Statistical Meetings, August 2014, Boston, MA, USA
- 2014 European Congress of Methodology, July 2014, Utrecht, the Netherlands
- 2014 18th ISA world Congress of Sociology, July 2014, Yokohama, Japan

- 2013 1st Southern European Conference on Survey Methodology, December 2013, Barcelona, Spain
- 2012 Fifth European Congress of Methodology, July 2012, Santiago de Compostela, Spain
- 2010 CBS Methodologiedagen, Heerlen, the Netherlands

Review activities

2010 - now Recurring article and book reviewer for at least:

R Journal, Springer Nature, PLOS one, Psychometrika, Sociological Methods and Research, Biometrics, JRSS Series C, MDPI Data, Biometrical Journal, European Journal on Criminal Policy and Research, Journal of Educational and Behavioral Statistics, CRC Press / Taylor & Francis, Springer, Psychological Test and Assessment Modeling, Survey Methodology, Statistica Neerlandica, The Spanish Journal of Psychology, International Journal of Geo-Information

Editorships

2016 - now Recurring article editor for Sage Open

Development

2022 - now dessert in R. Creator

2018 - now SHINYmice: Model building and evaluation for Multiple Imputation. Author.

2017 - now BAMI: Benchmarking and Assessment for Multiple Imputation. Creator.

2016 - now parlMICE in R. Creator.

2016 - now ampute in R. Author.

2010 - now mice in R. Core Developer and GitHub Collaborator

2013 - 2015 correlate in R. Author.

other Programming and Scripting experience: Python, R, HTML, C++, LATEX, markdown.

Publications

- [1] Cai, M., van Buuren, S., and **Vink, G.** Generalizing univariate predictive mean matching to impute multiple variables simultaneously. In Arai, K., editor, *Intelligent Computing*, pages 75–91, Cham, (2022). Springer International Publishing. ISBN 978-3-031-10461-9.
- [2] Muijzer, M., Hoven, C., Frank, L., **Vink, G.**, and Wisse, R. (2022). A machine learning approach to explore predictors of graft detachment following posterior lamellar keratoplasty: a nationwide registry study. *Scientific reports*, 12(1):1–9.
- [3] Claessens, J. L., Godefrooij, D. A., **Vink, G.**, Frank, L. E., and Wisse, R. P. (2022). Nationwide epidemiological approach to identify associations between keratoconus and immune-mediated diseases. *British Journal of Ophthalmology*, 106(10):1350–1354.
- [4] Cai, M., van Buuren, S., and **Vink, G.** (2022). How to relate potential outcomes: Estimating individual treatment effects under a given specified partial correlation. *arXiv* preprint *arXiv*:2208.12931.
- [5] Cai, M., van Buuren, S., and Vink, G. (2022). Joint distribution properties of fully conditional

- specification under the normal linear model with normal inverse-gamma priors. arXiv preprint arXiv:2208.12930.
- [6] Fopma, A., Cai, M., van Buuren, S., and **Vink, G.** (2022). A blended distance to define" people-like-me". arXiv preprint arXiv:2207.04650.
- [7] Gopalakrishna, G., Wicherts, J. M., **Vink, G.**, Stoop, I., van den Akker, O. R., ter Riet, G., and Bouter, L. M. (2022). Prevalence of responsible research practices among academics in the netherlands. *F1000Research*, 11(471):471.
- [8] Vink, G. (2022). Roderick J. Little and Donald B. Rubin: Statistical Analysis with Missing Data. *Psychometrika*. ISSN 1860-0980. doi: 10.1007/s11336-022-09856-8. URL https://doi.org/10.1007/s11336-022-09856-8.
- [9] Gopalakrishna, G., ter Riet, G., **Vink, G.**, Stoop, I., Wicherts, J. M., and Bouter, L. M. (2022). Prevalence of questionable research practices, research misconduct and their potential explanatory factors: A survey among academic researchers in the netherlands. *PLOS ONE*, 17 (2):1–16. doi: 10.1371/journal.pone.0263023.
- [10] Cai, M. and **Vink, G.** (2022). A note on imputing squares via polynomial combination approach. *Computational Statistics*. doi: 10.1007/s00180-022-01194-8.
- [11] Volker, T. B. and **Vink, G.** (2021). Anonymiced shareable data: Using mice to create and analyze multiply imputed synthetic datasets. *Psych*, 3(4):703–716.
- [12] Gopalakrishna, G., Wicherts, J. M., **Vink, G.**, Stoop, I., Van den Akker, O., Riet, G. t., and Bouter, L. Prevalence of responsible research practices and their potential explanatory factors: a survey among academic researchers in the netherlands, (2021). URL osf.io/preprints/metaarxiv/xsn94.
- [13] Gopalakrishna, G., Riet, G. t., **Vink, G.**, Stoop, I., Wicherts, J. M., and Bouter, L. Prevalence of questionable research practices, research misconduct and their potential explanatory factors: a survey among academic researchers in the netherlands, (2021). URL osf.io/preprints/metaarxiv/vk9yt.
- [14] Wisse, R., Muijzer, M., Hoven, C., van Luijk, C., Vink, G., and Frank, L. A machine learning approach to identify predictors of adverse events in posterior lamellar keratoplasty; a nationwide registry study. In ACTA OPHTHALMOLOGICA, volume 99, pages 6–6. WILEY 111 RIVER ST, HOBOKEN 07030-5774, NJ USA, (2021).
- [15] Claessens, J. L., Godefrooij, D. A., **Vink, G.**, Frank, L. E., and Wisse, R. P. (2021). Nationwide epidemiological approach to identify associations between keratoconus and immune-mediated diseases. *British Journal of Ophthalmology*.
- [16] Oberman, H. I., van Buuren, S., and **Vink, G.** (2021). Missing the point: Non-convergence in iterative imputation algorithms. *arXiv preprint arXiv:2110.11951*.
- [17] **Vink, G.** (2020). mice: Combining inferences after imputation with mice in r. $https://www.gerkovink.com/MiceVignettes/Combining_inferences.html.$
- [18] **Vink, G.** and Van Buuren, S. (2019). The best prediction can be the worst imputation. https://www.gerkovink.com/need_for_uncertainty/text1.html.

- [19] **Vink, G.** (2019). prepr: A self-paced guided crashcourse in basic scripting with r. https://www.gerkovink.com/prepR.
- [20] Schouten, R. M. and **Vink, G.** (2018). The dance of the mechanisms: How observed information influences the validity of missingness assumptions. *Sociological Methods & Research*. URL https://doi.org/10.1177/0049124118799376.
- [21] Schouten, R. M., Lugtig, P., and **Vink, G.** (2018). Generating missing values for simulation purposes: a multivariate amputation procedure. *Journal of Statistical Computation and Simulation*, pages 1–22. URL https://doi.org/10.1080/00949655.2018.1491577.
- [22] Reijntjes, A., Vermande, M., Olthof, T., Goossens, F. A., Vink, G., Aleva, L., and van der Meulen, M. (2018). Differences between resource control types revisited: A short term longitudinal study. *Social Development*, 27(1):187–200.
- [23] **Vink, G.** and Van Buuren, S. (2017). mice: An approach to sensitivity analysis. https://www.gerkovink.com/MiceVignettes/Sensitivity_analysis.html.
- [24] **Vink, G.** and Van Buuren, S. (2017). mice: Imputing multi-level data. https://www.gerkovink.com/MiceVignettes/Multi_level_data.html.
- [25] **Vink, G.** and Van Buuren, S. (2017). mice: Passive imputation and post- processing. https://www.gerkovink.com/MiceVignettes/Passive_imputation_post_processing.html.
- [26] **Vink, G.** and Van Buuren, S. (2017). mice: The imputation and nonresponse models. https://www.gerkovink.com/MiceVignettes/Missingness_inspection.html.
- [27] **Vink, G.** and Van Buuren, S. (2017). mice: Algorithmic convergence and inference pooling. https://www.gerkovink.com/MiceVignettes/Convergence_and_pooling.html.
- [28] **Vink, G.** and Van Buuren, S. (2017). mice: Ad hoc methods and the mice algorithm. https://www.gerkovink.com/MiceVignettes/Ad_hoc_methods.html.
- [29] Boeschoten, L., **Vink, G.**, and Hox, J. (2017). How to obtain valid inference under unit non-response? *Journal of Official Statistics*, 33(4):963–978.
- [30] **Vink, G.** and van Buuren, S. (2017). micevignettes; flexible imputation with mice in r in a nutshell. https://gerkovink.github.io/miceVignettes/.
- [31] Schouten, R. and **Vink., G.** (2017). parlmice: faster, paralleler, micer. https://gerkovink.github.io/parlMICE/Vignette_parlMICE.html.
- [32] Schouten, R., Lugtig, P., Brand, J., and **Vink., G.** (2017). Generating missing values with ampute. https://github.com/RianneSchouten/mice/tree/master/vignettes.
- [33] Thijssen, J., **Vink, G.**, Muris, P., and de Ruiter, C. (2016). The effectiveness of parent management training—oregon model in clinically referred children with externalizing behavior problems in the netherlands. *Child Psychiatry & Human Development*, pages 1–15. doi: 10.1007/s10578-016-0660-5.
- [34] **Vink, G.**, Lazendic, G., and van Buuren, S. (2015). Partitioned predictive mean matching as a multilevel imputation technique. *Psychological Test and Assessment Modeling*, 57(4):577–594.

- [35] Peeters, M., Zondervan-Zwijnenburg, M., **Vink, G.**, and van de Schoot, R. (2015). How to handle missing data: A comparison of different approaches. *European Journal of Developmental Psychology*, 12(4):377–394.
- [36] Sorbi, M., Kleiboer, A., van Silfhout, M., **Vink, G.**, and Passchier, J. (2015). Medium-term effectiveness of online behavioral training in migraine self-management: A randomized trial controlled over 10 months. *Cephalalgia*, 35(7):608–618.
- [37] **Vink, G.** (2015). Restrictive Imputation of Incomplete Survey Data. PhD thesis, Utrecht University.
- [38] **Vink, G.** and van Buuren, S. (2015). Multiple imputation of parent management training oregon model data. Technical report, Maastricht University.
- [39] **Vink, G.** and van Buuren, S. (2014). Pooling multiple imputations when the sample happens to be the population. *arXiv* preprint *arXiv*:1409.8542.
- [40] **Vink, G.**, Frank, L. E., Pannekoek, J., and van Buuren, S. (2014). Predictive mean matching imputation of semicontinuous variables. *Statistica Neerlandica*, 68(1):61–90.
- [41] **Vink, G.** and Van Buuren, S. (2013). Multiple imputation of squared terms. *Sociological Methods & Research*, 42(4):598–607.
- [42] **Vink, G.** and Van Buuren, S. (2013). Multiple imputation of longitudinal naplan data. Technical report, Australian Curriculum, Assessment and Reporting Authority.
- [43] **Vink, G.**, Hox, J. J. C. M., and van Buuren, S. (2013). Report on bayesian estimation of social educational advantage. Technical report, Australian Curriculum, Assessment and Reporting Authority.
- [44] Vink, G., Hox, J. J. C. M., and van Buuren, S. (2013). Multiple imputation of the acara social educational advantage data. Technical report, Australian Curriculum, Assessment and Reporting Authority.
- [45] Banning, R. and Vink, G. (2010). Methoden en indicatoren voor het evalueren en beschrijven van controle- en correctieprocessen. CBS nota DMV-2010- 06-08-RBNG-GVNK, Centraal Bureau voor de Statistiek.
- [46] **Vink, G.** (2007). Meervoudige beoordelaarsovereenstemming met cohen's kappa: Dubbel-codering van strategiegebruik en fouttypen bij vermenigvuldigingsopgaven in onafhankelijke databeoordelaarscombinaties. Technical report, Leiden University.