University of Groningen

Passive digital phenotyping

Jongs, Niels

DOI:
[10.33612/diss.171368248](https://doi.org/10.33612/diss.171368248)

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:
2021

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

Citation for published version (APA):
Jongs, N. (2021). *Passive digital phenotyping: Objective quantification of human behaviour through smartphones*. [Thesis fully internal (DIV), University of Groningen]. University of Groningen.
<https://doi.org/10.33612/diss.171368248>

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).

The publication may also be distributed here under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license. More information can be found on the University of Groningen website: <https://www.rug.nl/library/open-access/self-archiving-pure/taverne-amendment>.

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.

Passive Digital Phenotyping

*Objective quantification of human behaviour
through smartphones*

Niels Jongs

This work was financially supported by the PRISM project (www.prism-project.eu) which has received funding from Innovative Medicines Initiative 2 Joint Undertaking under grant agreement No 115916. This Joint Undertaking receives support from the European Union's Horizon 2020 research and innovation programme and EFPIA. This publication reflects only the authors' views neither IMI JU nor EFPIA nor the European Commission are liable for any use that may be made of the information contained therein.

The printing of this thesis was financially supported by the University of Groningen and the Graduate School of Science and Engineering (GSSE).

**rijksuniversiteit
 groningen**

Passive Digital Phenotyping

Objective quantification of human behaviour through
smartphones

Proefschrift

ter verkrijging van de graad van doctor
aan de Rijksuniversiteit Groningen
op gezag van de
rector magnificus prof. dr. C. Wijmenga
en volgens besluit van het College voor Promoties.

De openbare verdediging zal plaatsvinden op

vrijdag 18 juni 2021 om 12.45 uur

door

Niels Jongs

geboren op 28 augustus 1989
te Delfzijl

Promotores

Prof. dr. M.J.H. Kas

Prof. dr. ir. M.J.C. Eijkemans

Copromotor

Dr. J.A.S. Vorstman

Beoordelingscommissie

Prof. dr. R.A. Schoevers

Prof. dr. W.G. Staal

Prof. dr. R. van de Schoot

A Klee painting named *Angelus Novus* shows an angel looking as though he is about to move away from something he is fixedly contemplating. His eyes are staring, his mouth is open, his wings are spread. This is how one pictures the angel of history. His face is turned toward the past. Where we perceive a chain of events, he sees one single catastrophe which keeps piling wreckage upon wreckage and hurls it in front of his feet. The angel would like to stay, awaken the dead, and make whole what has been smashed. But a storm is blowing from Paradise; it has got caught in his wings with such violence that the angel can no longer close them. The storm irresistibly propels him into the future to which his back is turned, while the pile of debris before him grows skyward. This storm is what we call progress.

Walter Benjamin
[Theses on the Philosophy of History](#)

Paul Klee – The Israel Museum, Jerusalem

TABLE OF CONTENTS

General Introduction: The smartphone as a tool to quantify human behaviour	9
Chapter 1: Effect of disease related biases on the subjective assessment of social functioning in neuropsychiatric patients	21
Chapter 2: A framework for assessing neuropsychiatric phenotypes by using smartphone-based location data	39
Chapter 3: Towards unbiased and data-driven readouts of daily social functioning derived from longitudinal smartphone data and deep learning	65
Chapter 4: Digital behavioural signatures reveal trans-diagnostic clusters of Schizophrenia and Alzheimer's Disease patients	99
General Discussion: Methodological and statistical perspectives on smartphone-based digital phenotyping	119
Appendix:	
Nederlandse samenvatting	139
List of publications	145
Dankwoord	147

General Introduction

The smartphone as a tool to quantify human behaviour

INTRODUCTION

The quantification of human behaviour is considered a fundamental aspect of various research disciplines: some examples include psychiatry, sociology, economy and anthropology. These disciplines have in common that they predominantly utilize subjective research methods such as in-person interviews, questionnaires and self- or proxy-rated measures to study various aspects of behaviour. While these methods undeniably have led to numerous important insights, they also have their limitations that preclude their ability to objectively quantify behaviour. Most notably is that these methods rely on the subject's own (or the subject's proxy) account of behaviour, and are often obtained post-hoc. Due to these two limitations and several others, traditional methods for quantifying human behaviour are susceptible to a wide variety of method and response biases¹. These biases have the potential to induce systematic and random measurement errors in data² and therefore impede the validity and interpretation of findings³. For example, various types of biases as well as certain symptoms, such as cognitive dysfunction or lack of disease insight (which are often seen symptoms in several neuropsychiatric disorders) affect the ability to subjectively report about various aspects of behaviour. Due to these measurement errors the comparison between different sample populations is severely hampered.

All together, these limitations restrain our ability to further expand our understanding about variations in human behaviour and their underlying biological mechanisms. To further expand this understanding, we need behavioural assessment methods that have the ability to more objectively quantify human behaviour in natural setting and in a real-time longitudinal manner.

To adhere to these needs and to tackle those limitations that are inherent to traditional assessments of behaviour, researchers have started to explore and utilize the smartphone as a more objective tool to quantify human behaviour. By monitoring the embedded sensors (e.g., global Positioning System (GPS), Wi-Fi, and/or microphone) and usage patterns of smartphones in a passive and longitudinal manner a high-resolution trace of behavioural data can be generated. Subsequently, this trace of behavioural data could be used to derive relevant behavioural phenotypes. This passive behavioural monitoring approach is a sub strategy encompassed within the broader field of Digital Phenotyping⁴ which was defined by Jukka-Pekka Onnela in 2015 and was first introduced in Nature Biotechnology by Sachin H. Jain⁵. The term Digital Phenotyping includes both passive and active data collection and incorporates all devices that are able to collect behavioural data, not limited to the smartphone. However, the focus of this thesis is on passive and longitudinal monitoring of behaviour through smartphones; active data collected through smartphones (e.g., questionnaires via smartphones) or passive

data collected through devices other than smartphones, are not considered in the analyses reported here. Over the recent years, this digital approach to quantify behaviour through smartphones rapidly gained traction as a research tool (Figure 1). In 2009 about ten studies utilized a smartphone-based approach to quantify behaviour. This is in contrast to 2019 where 204 studies employed such an approach, a noteworthy increase of 1940% over ten years. This increasing trend in smartphone-based monitoring is primarily due to the ubiquity of smartphones in day-to-day life and to the increasing quality of data that is collected by smartphone sensors.

Figure 1 | Number of published studies that utilized a smartphone-based approach to quantify behaviour. Number of publications was obtained by searching PubMed and extracting the number the publications that concern smartphone-based or digital phenotyping of behaviour (search term PubMed: smartphone-based OR digital phenotyping AND behaviour).

Important to note is that the potential of using the smartphone as a tool to quantify behaviour is further increased by the relative low-cost of this approach combined with the fact that nowadays the majority of people worldwide own a smartphone.

Studies that employed such a smartphone-based approach are already starting to reveal the clinical potential of this approach, in particular within the context of neuropsychiatric research^{6–11}. For example, a study with 22 bipolar patients and 14 healthy controls showed that the variation in mobility patterns are indicative for a depressive state in patients diagnosed with bipolar disorder¹⁰. There is also emerging evidence suggesting that this approach can be employed to detect changes in behaviour related to (or possibly even preceding) a relapse of a psychotic episode^{9,12}. For example, fewer and shorter outgoing phone calls and fewer text messages were associated with relapses in schizophrenia¹³. For a complete and detailed overview of all the studies that employed a smartphone-based

approach we refer to one of several review articles that have been published over the recent years^{14–17}.

The rationale of a smartphone-based approach to quantify behaviour is that the observed behaviour is monitored in a more objective manner relative to traditional behavioural assessment methods (i.e. in-person interviews, questionnaires and self- or proxy-rated measures). Key features that contribute to this increase in objectivity are that data is collected in real-time, in the subject's natural environment, and without the need for any self- or proxy reporting. Inherent to using such a mobile personal device is that this source of data will reflect how an individual is functioning within his or her own real-life setting, where data is collected in real-time. For example, smartphone-based location data can be used to derive indicators of distance travelled, time spent at home and the number of places visited. By combining these indicators with the registration of outgoing and incoming calls they can for instance be used to study and/or monitor the withdrawal from social relations and mobility patterns over time. In addition, voice samples collected during phone calls could potentially be used to identify and derive various features that may be associated with changes in mood. For example, a recent study showed that by using voice analysis, manic and depressive states could be accurately identified in patients diagnosed with bipolar disorder⁷. Together, these prospects led to the development and introduction of several smartphone applications over the recent years, examples are the Beiwe App¹⁸, MONARCA system¹⁹ and Purple Robot app²⁰. These applications have in common that they all utilize smartphone sensors in order to quantify behaviour in a passive and unobtrusive manner. In sum, in contrast to the traditional methods to study human behaviour, the use of smartphones to collect and quantify human behaviour is 1) more objective, 2) unobtrusive, 3) can be continuously collected in “real-time”, and 4) within the natural “real-life” environment of the subject. Taken together, these features of this novel digital methodology allow a more objective and quantitative approach for assessing behaviour.

THE POTENTIAL OF SMARTPHONE-BASED PHENOTYPING IN PSYCHIATRY

This concept of monitoring behaviour *in situ* through smartphones is particularly beneficial for psychiatry²¹. A measurement-based approach for the detection of medical events is at the basis of most medical disciplines. For example, blood lipids are used as early markers for cardiovascular disease and blood glucose is measured over the day to prevent hyper or hypoglycaemia in diabetes. However, in psychiatry, the overt manifestation of behavioural symptoms of the psychiatric

disorder marks in most cases the start of treatment. The development of preventative interventions at the population level is gaining interest but remains a formidable challenge^{22,23}. At present, psychiatry, lacks valid temporal and objective measures of behaviour that reflect day-to-day functioning. Such a measure could possibly detect subtle changes that precede the onset of a full-blown psychiatric disorder or episode. A more measurement-based approach would allow psychiatry to move towards the prevention and early detection of major psychiatric events by monitoring measures that indicate the onset of such an event. We know that several behavioural changes in day-to-day functioning are identified as prodromal symptoms for several neuropsychiatric disorders^{24,25}. Other cumulative risk factors that increase the vulnerability to mental disorder include, but are not limited to, prenatal environment (eg, poor nutrition, exposure to drugs, and maternal infections or stress), brain injuries, social factors (i.e., socioeconomic status and poverty) and stressful life events²³. For these populations at risk, such a smartphone-based approach could be used to identify the change from being at risk to the need for care and with that possibly prevent onset of a disorder.

In addition, this novel and digital approach has the ability to provide a better understanding of the variations in behaviour and their underlying biological pathways. Within the Psychiatric Ratings using Intermediate Stratified Markers (PRISM) project¹⁵, this ability is used to stratify neuropsychiatric patients according to their daily social behaviour. This stratification of patients is subsequently used to identify shared and/or different neurobiological mechanism(s) of social withdrawal in schizophrenia, Alzheimer's and depression. The investigators of this project argue that the use of quantitative biological parameters, such as real-world behavioural data collected through smartphones, might lead to the development of drugs that effectively target the neurobiological pathways involved in social withdrawal. While the PRISM project focuses on gaining a better understanding of social withdrawal in neuropsychiatric disorders, this approach has the potential to contribute to an overall better understanding of the neurobiological basis and classification of neuropsychiatric disorders^{4,16}. Given that this data is less sensitive to random and systematic measurement errors, it can be expected that this real-world behavioural data is more effectively linked to various biological parameters, such as genotypes, brain activity patterns or structural brain data to study the biological underpinnings of these disorders. A better understanding of the neurobiological basis neuropsychiatric disorders enhances the development of effective drugs that specifically targets the relevant pathways involved. Subsequently, the same approach could be used to evaluate the efficacy of newly development or existing drugs in terms of positive behavioural changes (i.e. more social engagement and/or increased mobility).

AIMS AND OUTLINE

Over the past decade several studies have appeared that utilized and explored a smartphone-based approach to quantify behaviour in a longitudinal and passive manner. These studies have demonstrated the potential of this digital approach on research and already provided unprecedented and unique insights into human behaviour^{5–9,23–25}. However, prospective studies concerning this approach should also address numerous technical and methodological considerations regarding the collection and the processing of the data generated by the embedded sensors in smartphones. One of the major challenges in the utilization of such a digital approach concerns the practice of deriving relevant and valid behavioural phenotypes from the raw multimodal and highly dimensional data generated by smartphones. So far, studies that employed such an approach to quantify behaviour utilized various statistical methods for deriving these phenotypes. From a methodological point of view, these statistical methods were often unvalidated, lack standardisation and not specifically developed for this type of data. The lack of a standardised and validated methodological approach for deriving these phenotypes 1) limits the ability to reproduce earlier findings, 2) impedes comparisons across studies and 3) negatively affects the validity of findings. In order to capitalize on this smartphone-based approach in research it is pivotal to develop and standardise statistical procedures for deriving behavioural phenotypes from this type of data. Therefore, in this thesis I aim to:

1. Emphasize the need for objective measures of behaviour in neuropsychiatric research by demonstrating that the severity of symptoms affects the ability of patients to assess their own behaviour (**Chapter 1**).
2. Introduce a standardised framework for deriving relevant behavioural phenotypes from smartphone-based location data (**Chapter 2**).
3. Develop and validate an unbiased and data-driven measure of daily social functioning by using a deep learning approach on longitudinal smartphone data (**Chapter 3**).
4. Apply these novel and digital measure of behaviour to study the shared behavioural and biological characteristics of schizophrenia and Alzheimer's disease (**Chapter 4**).

This thesis aims to be one of the first steps towards a more standardised approach of digital phenotyping through smartphones. It specifically addresses the issue of how relevant behavioural phenotypes are derived from complex smartphone sensor data by using statistical methods specifically developed for this type of data. These methods should be chosen or developed in such a manner so that they are able to handle longitudinal sensor data that is collected through various different

devices. For example, location data collected via various types of smartphones generates data of different quality and temporal resolution. To derive behavioural phenotypes from this type of data it is pivotal that the used methods take these characteristics of the data into account.

Figure 2 | Overview of the BEHAPP application. Until today the BEHAPP collects data by monitoring location data via the embedded GPS sensor in smartphones and the WiFi sensor. In addition, smartphone usage patterns and communication logs are monitored by registering events.

Given that this work is an integral part of the aforementioned PRISM project¹⁵, it mainly concerns the quantification of human social behaviour in neuropsychiatric patients (i.e. Alzheimer's disease, schizophrenia and depression) and healthy controls. All the smartphone data presented in this thesis were collected through several studies by using the BEHAPP application^{26,27} (www.behapp.org). BEHAPP is an Android application that passively and unobtrusively monitors behaviour in a longitudinal manner by utilizing the embedded sensors (i.e., GPS, Wi-Fi, communication logs, and smartphone usage) in smartphones (Figure 2). Given the sensitive matter of this data in terms of privacy, the collected data is stored and processed in such a manner that it is compliant with the General Data Protection Regulation (GDPR)²⁷. Important to emphasize here is that the data is securely stored in an encrypted data base and is collected with informed consent from all participants. We used this data to develop and validate statistical procedures for deriving relevant behavioural phenotypes that mainly relate to social behaviour and/or mobility. Together, the derived phenotypes were used to introduce a data-driven and unbiased construct of daily social functioning by using an unsupervised deep

learning approach. Most notable is that these approaches could be employed to derive behavioural constructs other than social behaviour.

REFERENCES

1. Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y. & Podsakoff, N. P. Common Method Biases in Behavioral Research: A Critical Review of the Literature and Recommended Remedies. *J. Appl. Psychol.* 88, 879–903 (2003).
2. Bagozzi, R. P. & Yi, Y. Multitrait-Multimethod Matrices in Consumer Research. *J. Consum. Res.* 17, 426 (1991).
3. Campbell, D. T. & Fiske, D. W. Convergent and discriminant validation by the multitrait-multimethod matrix. *Psychol. Bull.* 56, 81–105 (1959).
4. Onnela, J.-P. & Rauch, S. L. Harnessing Smartphone-Based Digital Phenotyping to Enhance Behavioral and Mental Health. *Neuropsychopharmacology* 41, 1–12 (2016).
5. Jain, S. H., Powers, B. W., Hawkins, J. B. & Brownstein, J. S. The digital phenotype. *Nature Biotechnology* (2015). doi:10.1038/nbt.3223
6. Faurholt-Jepsen, M. et al. Smartphone data as objective measures of bipolar disorder symptoms. *Psychiatry Res.* 217, 124–127 (2014).
7. Faurholt-Jepsen, M. et al. Voice analysis as an objective state marker in bipolar disorder. *Transl. Psychiatry* 6, e856 (2016).
8. Saeb, S. et al. Mobile phone sensor correlates of depressive symptom severity in daily-life behavior: An exploratory study. *J. Med. Internet Res.* 17, 1–11 (2015).
9. Barnett, I. et al. Relapse prediction in schizophrenia through digital phenotyping: A pilot study. *Neuropsychopharmacology* 43, 1660–1666 (2018).
10. Palmius, N. et al. Detecting Bipolar Depression from Geographic Location Data. 1–10 (2015).
11. Farhan, A. A. et al. Behavior vs. introspection: Refining prediction of clinical depression via smartphone sensing data. 2016 IEEE Wirel. Heal. WH 2016 30–37 (2016). doi:10.1109/WH.2016.7764553
12. Ben-Zeev, D. et al. CrossCheck: Integrating self-report, behavioral sensing, and smartphone use to identify digital indicators of psychotic relapse. *Psychiatr. Rehabil. J.* (2017). doi:10.1037/prj0000243
13. Buck, B. et al. Relationships between smartphone social behavior and relapse in schizophrenia: A preliminary report. *Schizophr. Res.* (2019). doi:10.1016/j.schres.2019.03.014
14. Reinertsen, E. & Clifford, G. D. A review of physiological and behavioral monitoring with digital sensors for neuropsychiatric illnesses. *Physiol. Meas.* 39, aabf64 (2018).
15. Huckvale, K., Venkatesh, S. & Christensen, H. Toward clinical digital phenotyping: a timely opportunity to consider purpose, quality, and safety. *npj Digit. Med.* 2, (2019).
16. Kas, M. J. et al. A quantitative approach to neuropsychiatry: The why and the how. *Neurosci. Biobehav. Rev.* 97, 3–9 (2019).
17. Insel, T. R. Digital phenotyping: a global tool for psychiatry. *World Psychiatry* 17, 276–277 (2018).
18. Torous, J., Kiang, M. V., Lorme, J. & Onnela, J.-P. New tools for new research in psychiatry: A scalable and customizable platform to empower data driven smartphone research. *JMIR Ment. Heal.* 3, (2016).
19. Frost, M., Doryab, A., Faurholt-Jepsen, M., Kessing, L. & Bardram, J. E. Supporting disease insight through data analysis: Refinements of the MONARCA self-assessment

- system. in UbiComp 2013 - Proceedings of the 2013 ACM International Joint Conference on Pervasive and Ubiquitous Computing (2013). doi:10.1145/2493432.2493507
20. Schueller, S. M., Begale, M., Penedo, F. J. & Mohr, D. C. Purple: A modular system for developing and deploying behavioral intervention technologies. *Journal of Medical Internet Research* (2014). doi:10.2196/jmir.3376
 21. Fortney, J. C. et al. A Tipping Point for Measurement-Based Care. *Psychiatr. Serv.* (2017). doi:10.1176/appi.ps.201500439
 22. Sommer, I. E. et al. Early interventions in risk groups for schizophrenia: What are we waiting for? *npj Schizophr.* (2016). doi:10.1038/npjjschz.2016.3
 23. Arango, C. et al. Preventive strategies for mental health. *The Lancet Psychiatry* (2018). doi:10.1016/S2215-0366(18)30057-9
 24. Fava, G. A. & Kellner, R. Prodromal symptoms in affective disorders. *Am. J. Psychiatry* (1991). doi:10.1176/ajp.148.7.823
 25. Stella, F. et al. Neuropsychiatric symptoms in the prodromal stages of dementia. *Current Opinion in Psychiatry* (2014). doi:10.1097/YCO.0000000000000050
 26. Alvarez-Lozano, J. et al. Tell me your apps and I will tell you your mood. *Proc. 7th Int. Conf. Pervasive Technol. Relat. to Assist. Environ. - PETRA '14* 1, 1–7 (2014).
 27. Torous, J. et al. Characterizing Smartphone Engagement for Schizophrenia: Results of a Naturalist Mobile Health Study. *Clin. Schizophr. Relat. Psychoses* (2017). doi:10.3371/csrp.jtps.071317
 28. Wang, R. et al. CrossCheck: Towards passive sensing and detection of mental health changes in people with schizophrenia. *Proc. Int. Conf. Ubiquitous Comput.* 1–12 (2016). doi:10.1145/2971648.2971740
 29. Eskes, P., Spruit, M., Brinkkemper, S., Vorstman, J. & Kas, M. J. The sociability score: App-based social profiling from a healthcare perspective. *Comput. Human Behav.* 59, 39–48 (2016).
 30. Mulder, T., Jagesar, R. R., Klingenberg, A. M., P. Mifsud Bonnici, J. & Kas, M. J. New European privacy regulation: Assessing the impact for digital medicine innovations. *European Psychiatry* (2018). doi:10.1016/j.eurpsy.2018.07.003