

# Assessment, measurement and evaluation within public mental health

Citation for published version (APA):

Hoofs, H. M. (2018). *Assessment, measurement and evaluation within public mental health: methodological and psychometric challenges*. [Doctoral Thesis, Maastricht University]. Ipskamp Printing BV. <https://doi.org/10.26481/dis.20180507hh>

## Document status and date:

Published: 01/01/2018

## DOI:

[10.26481/dis.20180507hh](https://doi.org/10.26481/dis.20180507hh)

## Document Version:

Publisher's PDF, also known as Version of record

## Please check the document version of this publication:

- A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
- The final author version and the galley proof are versions of the publication after peer review.
- The final published version features the final layout of the paper including the volume, issue and page numbers.

[Link to publication](#)

## General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the "Taverne" license above, please follow below link for the End User Agreement:

[www.umlib.nl/taverne-license](http://www.umlib.nl/taverne-license)

## Take down policy

If you believe that this document breaches copyright please contact us at:

[repository@maastrichtuniversity.nl](mailto:repository@maastrichtuniversity.nl)

providing details and we will investigate your claim.

# Valorisation addendum

## Introduction

As the findings have various implications for future public mental health practice and policy, valorisation of these results is important. In this chapter an outline is given regarding the valorisation potential of the research findings presented in the thesis. Valorisation during the project (e.g. sharing of results) will also be addressed.

The goal of this thesis was to identify potential latent defects of measurement instruments and evaluate their impact within the public mental health. Therefore, some specific potential latent defects of measurement instruments were addressed using  $\beta$ -psychometrics.

As addressed in the discussion identification of latent defects and increasing the awareness in public mental health professionals is of the utmost important to increase and maintain a high level of quality in research and practice. This thesis, furthermore, shows that using the right preventive strategies measurement of public mental health can achieve this high level of quality. With adequate preventive measures the measurement, assessment, and evaluation of public mental health can be increased.

This high quality measurement of public mental health is important as conclusions, both for individuals but also for policies on different levels, are often (partly) based on this measurement. The challenges of measuring public mental health may appear small and unimportant at first glance but, they may have far-reaching consequences. That is, if latent defects prevail they will be lurking throughout the whole process hindering valid judgment and adequate decision making. As such, hampered measurement can impede the care for vulnerable children within the preventive youth health care, but also nation-wide policy that addresses preventive programs (e.g. to decrease physical and psychological job demands of employees within specific economic sectors).

## Valorisation of findings during the PhD trajectory

Results of the current thesis were shared with both researchers and professionals within the public mental health. Especially the results of a pilot study (Hoofs, Jansen, Mohren, Reijs, et al., 2015) were shared at multiple occasions. The subject of this pilot study was the usability of the electronic health records of the preventive youth health care for research purposes. It was within this same preventive youth health care that the results were shared with doctors, managers, and IT-specialists. All these groups were involved with the electronic health records and were, as such, in a position to increase their quality. With interactive presentations the results of this pilot study provided insight in the data quality of the health records. As this quality was highly diverse these results also illustrated the importance of guidelines to facilitate and assure uniform reporting. Shortcomings in the electronic health records were, furthermore, not only hampering research but also the application within the individual setting.<sup>8</sup> To show the benefits of a higher data-quality potential use-cases were also explored within these presentations. Education is, as indicated in the discussion of this thesis, crucial for the understanding of the importance of the quality of the measurement within public mental health. These results were, therefore, not only shared with current professionals

---

<sup>8</sup>It should be furthermore noted that some variables within the electronic health records are relevant and important for the Dutch Health Care Inspectorate

within the preventive youth health care but also with doctors in training. The importance of a critical assessment of measurement implications within public mental was therefore addressed by showing the implications of the pilot study. This will not only result in a higher data quality regarding the electronic health records, it can also facilitate their evidence based practice with respect to measurement (instruments).

Valorisation of the results took also place with respect to the findings of chapter 2 investigating the Strengths and Difficulties Questionnaire (SDQ). One of the findings of this chapter was that cut-off points of the SDQ cannot be used directly between different settings. Future application of these cut-off points should therefore be based on thresholds that are based on the setting at hand. If cut-off points from a setting are, nevertheless, applied within a different setting these cut-off points should first be re-evaluated within this different setting. To increase the likelihood that these findings are actually implemented within the public mental health and increase the valorisation of this thesis as such, two specific methods were used. First, the findings were not only published within an international (English) journal but also within a national journal specifically directed at (health) professionals (i.e. TSG). Large parts of the manuscripts were therefore rewritten in order to increase accessibility and readability for a wider public. This was further strengthened by the second method, illustrating the effect of cross-use of cut-off points on actual data. Using the cut-off points based on data of one setting (i.e. collective) on the data from a different setting (i.e. individual) showed that the observed percentage of children that scored above the cut-off point was much lower than expected, further underscoring the relevance of these findings.

## **Future applications of the generated knowledge**

As this thesis shows that latent defects can have a large impact on the quality of measurement instruments within the public mental health, the awareness for latent defects should be increased. Target groups for which this awareness should be raised are considered in a separate section of this valorisation chapter. Nevertheless, this thesis clearly illustrates that for future application of measurement instruments their quality should be critically assessed (both for newly developed and existing instruments). This thesis furthermore shows that this assessment is best served by using statistical methods within the  $\beta$ -psychometrics. As argued in the discussion this assessment of the impact of latent defects is also possible on a post-hoc basis. While a-priori testing should be favoured,  $\beta$ -psychometrics makes it possible to successfully identify latent defects if the measurement instrument is already used. This makes it possible to continuously test (critical) quality issues of a measurement instrument. This is furthermore advantageous as latent defects can strongly differ between assessments.

While the identification of latent defects within existing measurement instruments is crucial, preventing latent defects within measurement instruments is of course preferable. Developers of new measurement instruments should therefore critically assess if latent defects could be (or become) present within their measurement instrument. This does not only increase the quality of measurement instruments, it often also increases the understanding of an instrument during the development phase as  $\beta$ -psychometrics requires an active modelling approach. This approach is only possible if a conceptual framework of the instrument is present.

A specific (future) application of the generated knowledge within this thesis is the newly proposed and validated fit index within Bayesian confirmatory factor analysis (CFA), the Bayesian

Root Mean Square Error of Approximation (BRMSEA). The BRMSEA will increase the accessibility and validity of advanced statistical tools within Bayesian CFA. As the BRMSEA makes it possible to assess model fit within large samples Bayesian CFA, public mental health may be more easily served by the tools this methodology includes. That is, research within the public mental health is often characterised by its large sample size. While analysing large datasets has many benefits, assessing the fit of the statistical model can be challenging. The introduction of the BRMSEA is motivated by this challenge and chapter 6 shows that the BRMSEA is up to this task. Again, chapter 6 also includes a section with a clear (empirical) illustration to increase its valorisation. As such the BRMSEA is not only validated within a simulation study, but its added value is also shown within an illustration based on empirical data. Another future application of the BRMSEA is its potential implementation within Mplus (Muthen & Muthen, 1998-2012), one of the most used software package when it comes to (Bayesian) CFA, and Blavaan (Merkle & Rosseel, 2016), an open source software library for R. This will greatly increase the ease with which researchers can use the BRMSEA for the benefits of their research. With the adaptation of the BRMSEA within a software package that is used by so many researchers within the field of public mental health the impact of the generated knowledge within this thesis is greatly enhanced. Especially as researchers do no longer have to 'actively' search for the results of this study as these findings are now 'brought' to them through this software package.

As already indicated at the start of this addendum and in the conclusion of this thesis, the main goal of this thesis is to promote identification of latent defects and to increase the awareness within public mental health professionals for these defects. This thesis has all the ingredients to facilitate and achieve this increased awareness for the measurement of public mental health. More precisely, to increase the awareness for the valid assessment of measurement of public mental health. That is, as already indicated in the introduction of this thesis, the *validation* of measurement instruments within the public mental health is a difficult challenge. This thesis shows, however, that with this increased awareness, the right tools, modelling strategy, and conceptual framework it is possible to successfully identify potential latent defects. With  $\beta$ -psychometrics it is, to some extent, even possible to adjust for these defects without the need to completely re-design and -implement measurement instruments. If these three steps will set foot in the public mental health, it will increase the quality of results and conclusions based on measurement instruments. This will positively influence the public mental health on all different levels and dimensions (World Health Organization, 2001).

## **Novelty in the topic of the PhD thesis and novelty in the findings**

When it comes to measurement in the public mental health it sometimes looks a bit sidelined. Measurement has to be shorter, faster, and more appealing but at the same time also more reliable, valid, and dynamic. While it is always a good approach to strive for improvement regarding measurement, its quality sometimes looks somewhat neglected. This can be costly, as it is difficult to change an instrument during research while a population could be potentially more easily increased (if money permits). The main novelty of this PhD is, therefore, that it allows for a moment of contemplation regarding the issues that could potentially violate the quality of measurement within public mental health. Not only does this thesis offers such an

insight with advanced techniques it also shows and suggest potential measures to overcome or at least limit the effects of these issues.

While most techniques applied are not a novelty as such, their application within the current framework of public mental health should be considered a novelty. As such this thesis brings together different paradigms and their corresponding techniques. The trait-state-occasion (TSO) model, for example, is often used in the field of developmental psychology (Cole, 2012). Applying it within the context of measurement within public mental health opened the possibility to analyse the stability of a measurement instrument over a longer period of time. The inclusion, furthermore, of the survival analysis within this TSO-model is, to the best of our knowledge, a first. As such this thesis shows the flexibility of structural equation modelling, and the new insights it can offer for a specific topic.

Another novel finding of this thesis is that the cross-use of cut-off points between different settings is not directly possible. As already addressed, this effect was also quantified by illustrating its effect on empirical data. Especially this quantification of this effect can be considered a novelty, as the effect of social desirability bias (SDB) is addressed in the literature (Krumpal, 2013). The majority of this research regarding SDB, however, focuses on its theoretical underpinning and mostly uses controlled settings to study its effect. As discussed in chapter 2, capturing the actual impact of the different setting within the public mental health (e.g. individual and collective) is almost impossible. Using real-world datasets from these settings makes it therefore possible to not only study its actual effect, but also to directly quantify this effect.

A technical novelty that was introduced in this thesis is the BRMSEA. The added value of the BRMSEA is already addressed in the previous section. It is, nevertheless, noteworthy that the implementation of this finding was met with some hefty criticism during the review process. While there is nothing wrong with strong critiques on novel findings, it is important that the scientific community is open to different approaches. The criticism was, furthermore, in stark contrast to the positive response that the manuscript has attracted. This seems to show that there is sometimes a difference between the gatekeepers of a (specific) methodology and its 'actual' users. While not necessarily invalid, it would sometimes be good if this former group would take this latter group into account when assessing the added value of a concept or methodology.

## Target groups

As already indicated in the discussion chapter of this thesis, many different target groups could be considered for the findings of this thesis. The two groups that were highlighted were, the users of the measurement instruments and their developers. The meetings with the PYHC professionals were a first step towards the goal of informing and creating awareness within this former group. For the latter group the findings in this thesis can be used to quantify specific assumptions that are violated, increasing awareness in researchers that develop measurement instruments. This thesis, as such, can create an increased awareness in both groups as it identified some potential latent defects and showed their impact. The thesis, furthermore, also showed (preventive) measures to these latent defects that these group can apply.

Through the means of education of public mental health professionals this increased awareness should become reality. In most educational programs the curriculum seem to jump from

reliability and simple validity issues directly towards inferential statistics (e.g. regression). As indicated in the discussion this introducing  $\beta$ -psychometrics at an early stage to public mental health professionals will not only be beneficial for their statistical knowledge, but will also enhance their general conceptual awareness regarding public mental health. As  $\beta$ -psychometrics is much more than only statistical analysis it will increase the critical appraisal of public mental health professionals regarding the measurement, assessment and evaluation of public mental health.

## Conclusion

The findings of this thesis have a wide range of valorisation targets. From professionals within the preventive youth health care to academics within the field of applied Bayesian statistics. Valorisation was furthermore actively pursued, most noteworthy by publishing in a national journal specifically directed at (health) professionals. I believe that this 'translation' of results of academic research is key towards a successful valorisation. It furthermore results in valuable input for further avenues of academic research. As such, the process of doing research becomes increasingly interactive. This will result in a more profound and faster iterative process from which the results will furthermore have wider support.

The findings of this thesis should have a continuing impact on the field of public mental health. It shows that many hidden and latent assumptions within this complex field of mental health *can* be made manifest. Doing so shows the strengths and weaknesses of currently used (measurement) instruments. While this thesis contains some specific examples, its rationale applies to the complete field of public mental health, and maybe even also within a broader context.

## References

- Cole, D. A. (2012). Latent trait-state models. In R. Hoyle (Ed.), *Handbook of structural equation modeling* (pp. 586–600). New York, NY: Guilford Press.
- Hoofs, H., Jansen, M. W. J., Mohren, D. C., Reijs, R., Jansen, N. W. H., & Kant, IJ. (2015). Is het DD-JGZ bruikbaar voor epidemiologisch onderzoek? [Is the DD-JGZ useful for epidemiological research?] *JA!-Plus*, 3, 1–7.
- Krumpal, I. (2013). Determinants of social desirability bias in sensitive surveys: A literature review. *Quality & Quantity*, 47(4), 2025–2047. doi:10.1007/s11135-011-9640-9
- Merkle, E. C. & Rosseel, Y. (2016). Blavaan: Bayesian structural equation models via parameter expansion. *arXiv 1511.05604*. Retrieved from <https://arxiv.org/abs/1511.05604>
- Muthen, L. K. & Muthen, B. O. (1998-2012). *Mplus user's guide: Seventh edition*.
- World Health Organization. (2001). *The world health report 2001: Mental health: New understanding, new hope*. World Health Organization.