

# D-score for international comparisons

*Stef van Buuren*

*2018-04-04*



# Contents

<b>Preface</b>	<b>5</b>
<b>1 Introduction</b>	<b>7</b>
1.1 SDG 4.2.1 indicator . . . . .	7
1.2 Quick scan of instruments . . . . .	7
1.3 Care-giver vs direct assessment . . . . .	7
1.4 Individual-programmatic-population settings . . . . .	7
<b>2 Comparability</b>	<b>9</b>
2.1 Challenge of comparability . . . . .	9
2.2 Bridging instruments by mapping items . . . . .	9
2.3 Overview of promising item mappings . . . . .	9
<b>3 Equate groups</b>	<b>11</b>
3.1 Statistical framework for equate groups . . . . .	11
3.2 Common latent scale . . . . .	11
3.3 Differential item functioning . . . . .	11
3.4 Quantifying equate fit . . . . .	11
<b>4 Modeling equates</b>	<b>13</b>
4.1 GCDG data: design and description . . . . .	13
4.2 Empirical and fitted item response curves . . . . .	13
4.3 Active and non-active equate groups . . . . .	13
4.4 Splitting and combining equate groups . . . . .	13
4.5 modeling strategies . . . . .	13
<b>5 Comparing ability</b>	<b>15</b>
5.1 D-score distribution by study . . . . .	15
5.2 Impact of measurement error . . . . .	15
<b>6 Validity</b>	<b>17</b>
6.1 Discriminatory validity . . . . .	17
6.2 Concurrent validity . . . . .	17
6.3 Predictive validity . . . . .	17
<b>7 SDG 4.2.1 Indicator</b>	<b>19</b>
7.1 Application I: Estimating the SDG 4.2.1 indicator from existing data . . . . .	19
7.2 Data sets . . . . .	19
7.3 Equate groups . . . . .	19
7.4 Definition <i>developmentally on track</i> . . . . .	19
7.5 Country-level estimates . . . . .	19
7.6 Relation to other estimates . . . . .	19

<b>8</b>	<b>Who is on-track?</b>	<b>21</b>
8.1	Application II: What determines who is developmentally on-track? . . . . .	21
8.2	Types of explanatory factors . . . . .	21
8.3	Relative importance . . . . .	21
8.4	Opportunities for intervention . . . . .	21
<b>9</b>	<b>Discussion</b>	<b>23</b>
9.1	Potential of existing data for SDG 4.2.1 . . . . .	23
9.2	Options for improving health policy . . . . .	23
9.3	Suitability of D-score metric . . . . .	23
9.4	Suggestions for better measurement . . . . .	23

# Preface

This is an introductory booklet on the measurement of child development by means of the D-score. The D-score is a one-number summary that quantifies generic neurocognitive development for children with ages 0-4 years.

This is the *second* in a series of three booklets. The series consists of the following titles:

1. D-score for measuring development of children 0-4 years
2. D-score for international comparisons
3. D-score for creating better instruments

The development of this series is kindly supported by the Bill & Melinda Gates Foundation.



# Chapter 1

## Introduction

1.1 SDG 4.2.1 indicator

1.2 Quick scan of instruments

1.3 Care-giver vs direct assessment

1.4 Individual-programmatic-population settings





## Chapter 2

# Comparability

2.1 Challenge of comparability

2.2 Bridging instruments by mapping items

2.3 Overview of promising item mappings



## Chapter 3

# Equate groups

3.1 Statistical framework for equate groups

3.2 Common latent scale

3.3 Differential item functioning

3.4 Quantifying equate fit



## Chapter 4

# Modeling equates

- 4.1 GCDG data: design and description
- 4.2 Empirical and fitted item response curves
- 4.3 Active and non-active equate groups
- 4.4 Splitting and combining equate groups
- 4.5 modeling strategies



## Chapter 5

# Comparing ability

5.1 D-score distribution by study

5.2 Impact of measurement error





## Chapter 6

# Validity

6.1 Discriminatory validity

6.2 Concurrent validity

6.3 Predictive validity



## Chapter 7

# SDG 4.2.1 Indicator

- 7.1 Application I: Estimating the SDG 4.2.1 indicator from existing data
- 7.2 Data sets
- 7.3 Equate groups
- 7.4 Definition *developmentally on track*
- 7.5 Country-level estimates
- 7.6 Relation to other estimates



## Chapter 8

# Who is on-track?

- 8.1 Application II: What determines who is developmentally on-track?
- 8.2 Types of explanatory factors
- 8.3 Relative importance
- 8.4 Opportunities for intervention



## Chapter 9

# Discussion

9.1 Potential of existing data for SDG 4.2.1

9.2 Options for improving health policy

9.3 Suitability of D-score metric

9.4 Suggestions for better measurement