EUROPEAN CONGRESS OF PSYCHOLOGY

2023

INTERNATIONAL TEST COMMISSION

EAPA
European Association of Psychological Assessment

BFP Testcommissie

efpa
European Federation of Psychologists' Associations

Mark Schittekatte PhD

Research Support Office / Assessment Lab

Ghent University Belgium

Belgian delegate of **EFPA's** *Board of Assessment*

& President of the **Belgian Test Commission**

Nigel Evans C.Psychol

NEC Director

Advisor to **BPS Psychological Test Centre**

UK delegate and **Chair** of **EFPA's Board of Assessment**

(1320 *Paper 1*)

Updating the EFPA BoA *Test Review Model*: a *necessary titanic* work with many *angles* and supported by even more *shoulders*
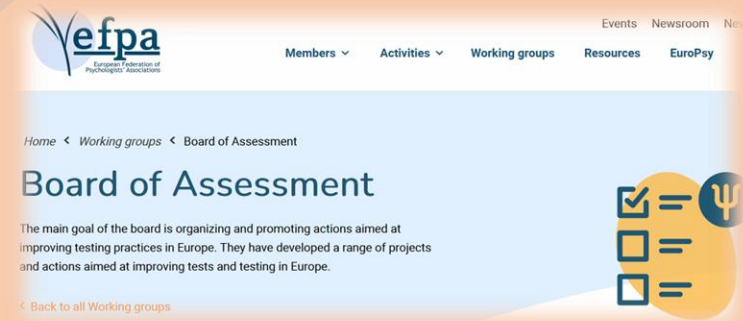
Symposium: *Updates on the Work of the EFPA Board of Assessment* (1136)
Weds 5 July 16:30 – 18:00   Chair: Nigel Evans

efpa
European Federation of Psychologists' Associations

Events Newsroom Ne

Members ⌄ Activities ⌄ Working groups Resources EuroPsy

Home ‹ *Working groups* ‹ Board of Assessment

Board of Assessment

The main goal of the board is organizing and promoting actions aimed at improving testing practices in Europe. They have developed a range of projects and actions aimed at improving tests and testing in Europe.

‹ Back to all Working groups

https://www.efpa.eu/working-groups/assessment

Updating the EFPA BoA Test Review Model: a necessary titanic work with many angles and supported by even more shoulders

The main goal of the EFPA Test Review Model (TRM) is to provide a description and a detailed and rigorous assessment of the tests, scales and questionnaires mainly used in the field of psychological and educational assessment. In other words, offering a tool for assessing the quality of tests.

<= What?

This information is made available to test users and professionals, in order to improve tests and testing, and help them to make the right assessment decisions. The EFPA TRM is part of the information strategy of the EFPA, which aims to provide all necessary mainly technical information about the tests in order to enhance its use.

<= For who?

This EFPA test review model aims further to support and encourage the process of harmonising the quality standards and the reviewing of tests across Europe.

<= Why?

However the latest version of this Test Review Model dates from 2013, so a revision is very neccessary!

The EFPA Board of Assessment (BoA) is in the middle of updating the TRM in this period, and how this is handled, with who (different stakeholders) and in what time frame, are few of the issues to be discussed.

<= How?

Further attention is given to: what topics are the most urgent to update (e.g. online testing, gamification, AI algorithms) and how can this TRM be implemented by local test commissions. Also the hurdles in this process in different European countries will be highlighted (e.g. reviewed by who, whether or not an overall final score for each reviewed instrument, need for financial resources, and considering the impact of negative reviews).

<= When?

This paper presentation asks explicitly inspiring input from the audience.

Schittekatte M , Evans N

In the early (BoA) days.. =>

"BoA" Madrid 1996

EFPA Board of Assessment Document 110c

**EFPA REVIEW MODEL FOR
THE DESCRIPTION AND EVALUATION OF
PSYCHOLOGICAL AND EDUCATIONAL TESTS**

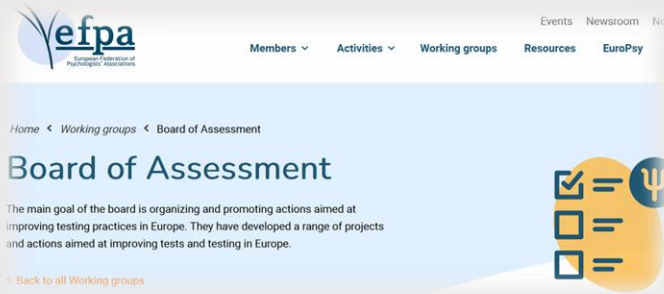*TEST REVIEW FORM AND NOTES FOR REVIEWERS*

VERSION 4.2.6

Version 4.2.6 is a major revision of Version 3-42
(2008) by a task force of the Board of Assessment
of EFPA consisting of:

Arne Evers (chair, the Netherlands)
Carmen Hagemeister (Germany)
Andreas Høstmælingen (Norway)
Patricia Lindley (UK)
José Muñiz (Spain)
Anders Sjöberg (Sweden)

Approved by the EFPA General Assembly, 13-07-2013

Dave Bartram   José Muñiz   Arne Evers   Dušica Boben   Krunoslav Matešić

The shoulders..

BoA TRM Workgroup

EC Liaison

BoA Convenor

BoA & COTAN member

Main questions raised..

- What is the *purpose* of the Model?
- To be used by..?
- How to use, how *not* to use..?

- Diversity and inclusivity (D&I) / Fairness
- IRT developments
- Adaptative testing
- Hidden algorithms, gamification,
  and continuous norming.
- Privacy issues with implications for norm groups
- Routine Outcome Monitoring (ROM)
- Unproctored data gathering
- ...

EFPA Board of Assessment Document 110c

CONTENTS

'Chapter by chapter'

In association with the COTAN

Nederlands Instituut van Psychologen NIP

-    We are also going to need other experts and reviewers to give outside opinions; especially on issues that are new and/or highly specialized.
     o   Example: how do we write items in this new age of gender and diversity?
     o   Example: complicated psychometrics.
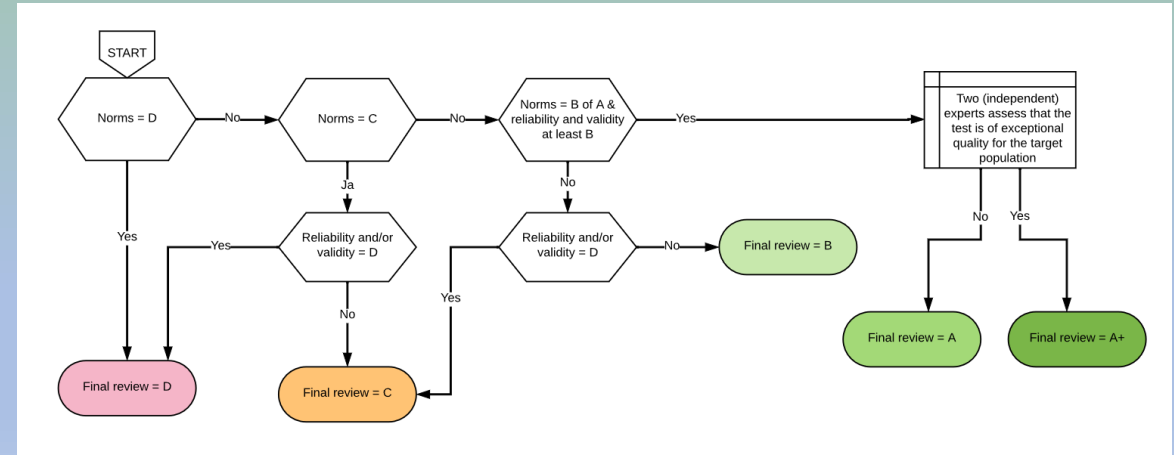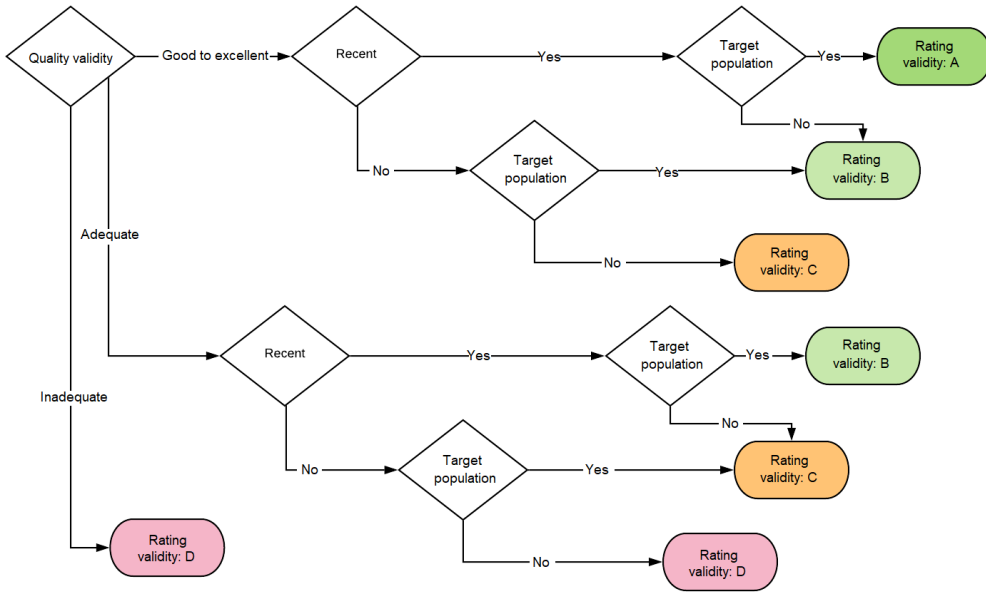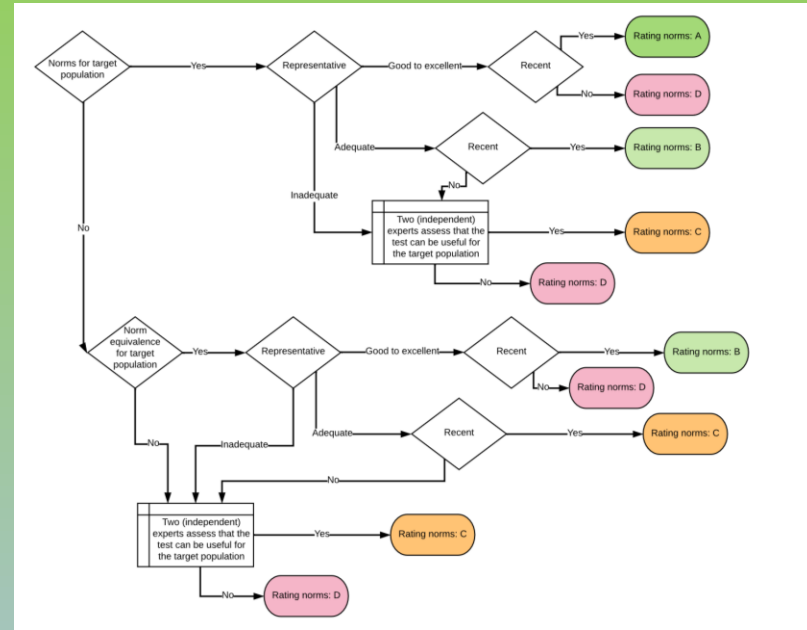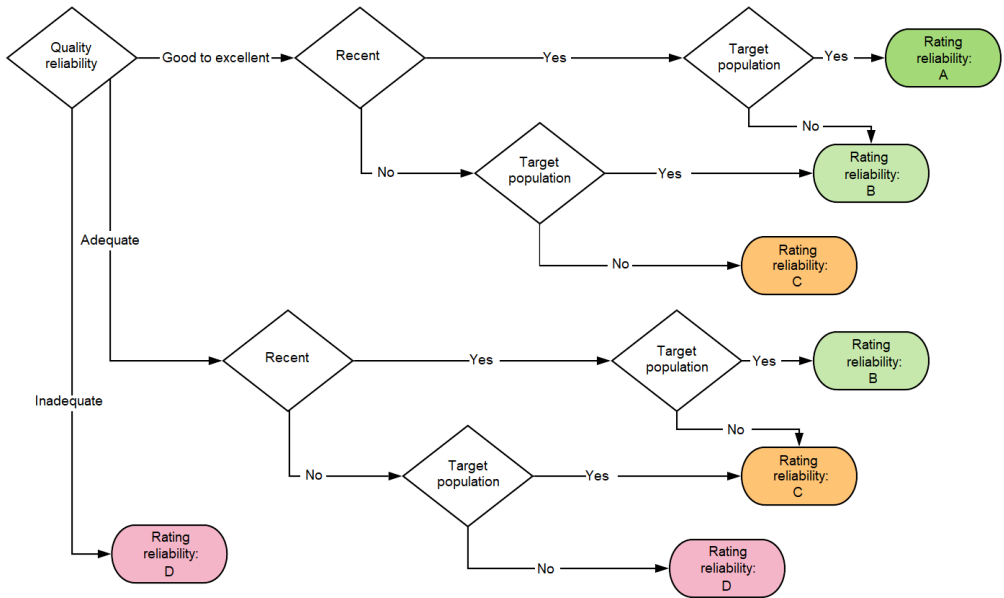
Not more *severe*, but more *explicit*

'Open process' with *all* stakeholders

With 3 doc's
     - 'TRM share' 70 pages with (a lot of) comments ☺
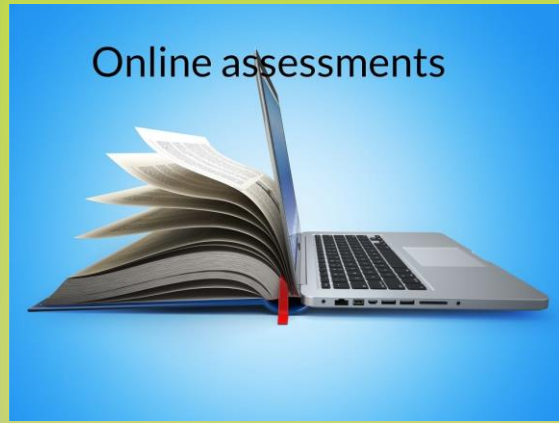   - List of Experts
     - List of Literature

WOW! AMAZING

(Belgian) Q-Scan Decision trees: *reliability, validity & norms and an END SCORE*

Reliability decision tree (top-left):

- Quality reliability
  - Good to excellent → Recent
    - Yes → Target population
      - Yes → Rating reliability: A
      - No → Rating reliability: B
    - No → Target population
      - Yes → Rating reliability: B
      - No → Rating reliability: C
  - Adequate → Recent
    - Yes → Target population
      - Yes → Rating reliability: B
      - No → Rating reliability: C
    - No → Target population
      - Yes → Rating reliability: C
      - No → Rating reliability: D
  - Inadequate → Rating reliability: D

Norms decision tree (top-right):

- Norms for target population
  - Yes → Representative
    - Good to excellent → Recent
      - Yes → Rating norms: A
      - No → Rating norms: D
    - Adequate → Recent
      - Yes → Rating norms: B
      - No → Two (independent) experts assess that the test can be useful for the target population
        - Yes → Rating norms: C
        - No → Rating norms: D
    - Inadequate → Two (independent) experts assess that the test can be useful for the target population
  - No → Norm equivalence for target population
    - Yes → Representative
      - Good to excellent → Recent
        - Yes → Rating norms: B
        - No → Rating norms: D
      - Adequate → Recent
        - Yes → Rating norms: C
        - No → Two (independent) experts assess that the test can be useful for the target population
          - Yes → Rating norms: C
          - No → Rating norms: D
      - Inadequate
    - No

Validity decision tree (bottom-left):

- Quality validity
  - Good to excellent → Recent
    - Yes → Target population
      - Yes → Rating validity: A
      - No → Rating validity: B
    - No → Target population
      - Yes → Rating validity: B
      - No → Rating validity: C
  - Adequate → Recent
    - Yes → Target population
      - Yes → Rating validity: B
      - No → Rating validity: C
    - No → Target population
      - Yes → Rating validity: C
      - No → Rating validity: D
  - Inadequate → Rating validity: D

End score decision tree (bottom-right):

- START
  - Norms = D
    - Yes → Final review = D
    - No → Norms = C
      - Ja → Reliability and/or validity = D
        - No → Final review = C
        - Yes
      - No → Norms = B of A & reliability and validity at least B
        - No → Reliability and/or validity = D
          - No → Final review = B
          - Yes → Final review = C
        - Yes → Two (independent) experts assess that the test is of exceptional quality for the target population
          - No → Final review = A
          - Yes → Final review = A+

N(ormen)	B(etrouwbaarheid)	V(aliditeit)	R(elevantie)	T(estmateriaal)	TQS eindlabel
A	B	C	A+	A	B(a)

Test qualification table for IQ tests

Online assessments

Digitale diagnostiek: een goed idee!?

DOOR VEERLE BRIERS & MARK SCHITTEKATTE

ℹ Kwalificatie cognitieve vaardigheidstests

Beoordeling	Meetpretentie & testoverzicht		
	(a) Algemene intelligentie	(s) Specifieke cognitieve vaardigheden	(c) Cognitief ontwikkelingsniveau
A+ (kwalificatie "uitstekend")	CoVaT-CHC Basisversie WISC-V-NL		
A (kwalificatie "goed")	WPPSI-IV-NL (+4j.) WAIS-IV-NL WPPSI-III-NL (+4j.)[5] RAKIT-II	WPPSI-IV-NL (-4j.)	
B (kwalificatie "voldoende")		WPPSI-III-NL (-4j.)[5]	
C (kwalificatie "met reserve") *		RAVEN's 2 PM SON-R (2-8) SON-R (6-40) WNV	Bayley-III
D (kwalificatie "onvoldoende")	GIT-2 IST KAIT LEM RAKIT-R WAIS WAIS-R WAIS-III WISC-III WISC-R WPPSI-R	SON-R (2½-7.11) SON-R (5½-17)	BSID-II GOS 2½-4½ MOS MSEL DOS
Geen beoordeling	IDS-2		

(*) "grijze zone": duidelijk positieve punten, maar (veel) tekorten m.a.w. voorlopig aanvaardbaar in uitzonderlijke gevallen en mits verantwoording van het gebruik

Future perspectives & actions

The *Board of Assessment* & *'TRM workgroup'* continue the next year updating the **EFPA BoA Test Review Model**

- With many shoulders
- Looking at different angles
- ....

And with an *'open perspective'*

Further need for (updated!) *European guidelines* and *spread the word*: there is still a lot of work at.. the *improvement of testing practices*

Thanks!

ANY QUESTIONS?

efpa
European Federation of Psychologists' Associations
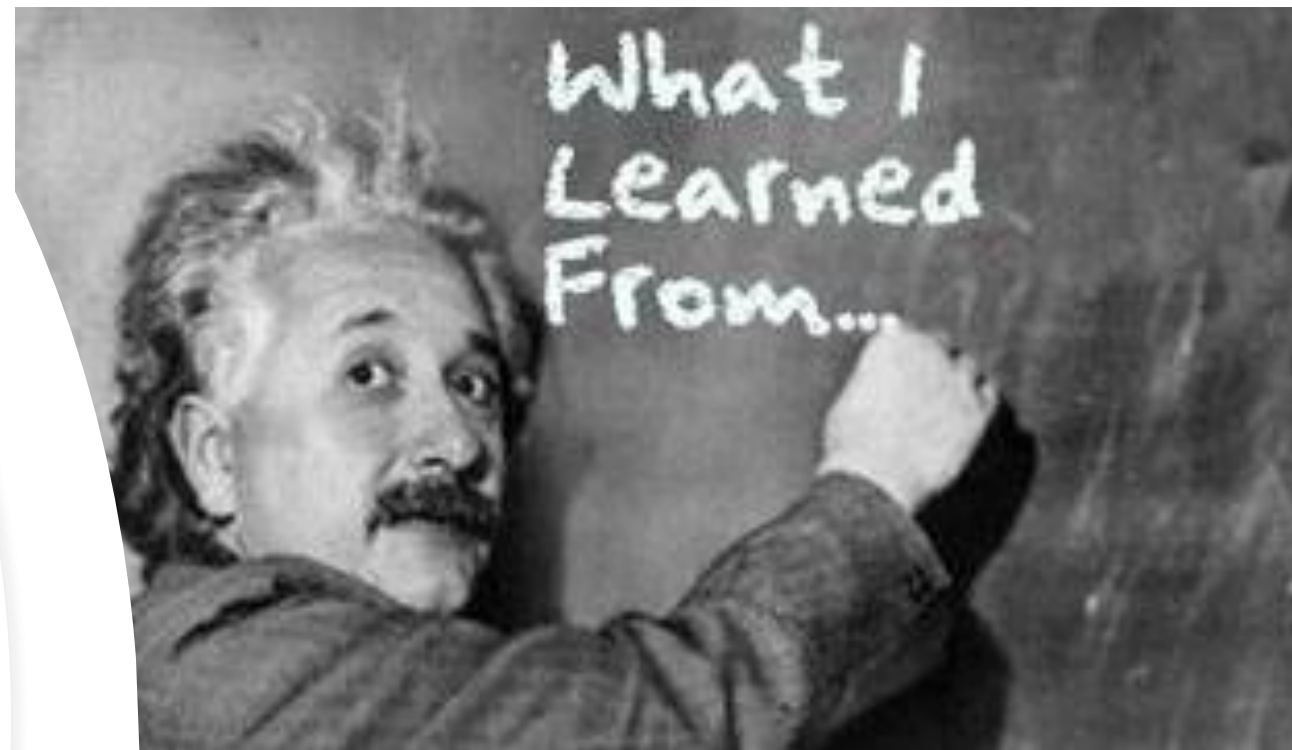
Events   Newsroom   News

Members ⌄   Activities ⌄   Working groups   Resources   EuroPsy

Home ‹ Working groups ‹ Board of Assessment

Board of Assessment

The main goal of the board is organizing and promoting actions aimed at improving testing practices in Europe. They have developed a range of projects and actions aimed at improving tests and testing in Europe.

‹ Back to all Working groups

mark.schittekatte@ugent.be
nigel@nigel-evans.com