*Gyöngyvér Molnár*
*University of Szeged, Institute of Education*
*http://www.staff.u-szeged.hu/~gymolnar*

edia.hu
**eDia**
**SZTE OK**

# DEVELOPMENT OF AN ONLINE DIAGNOSTIC ASSESSMENT SYSTEM FOR THE SUPPORT OF PERSONALIZED TEACHING AND LEARNING

"If you can not measure it, you can not improve it."
Kelvin

*ECER 2021*
*September 9. 2021*

**edia.hu**     **Center for Research on Learning and Instruction**
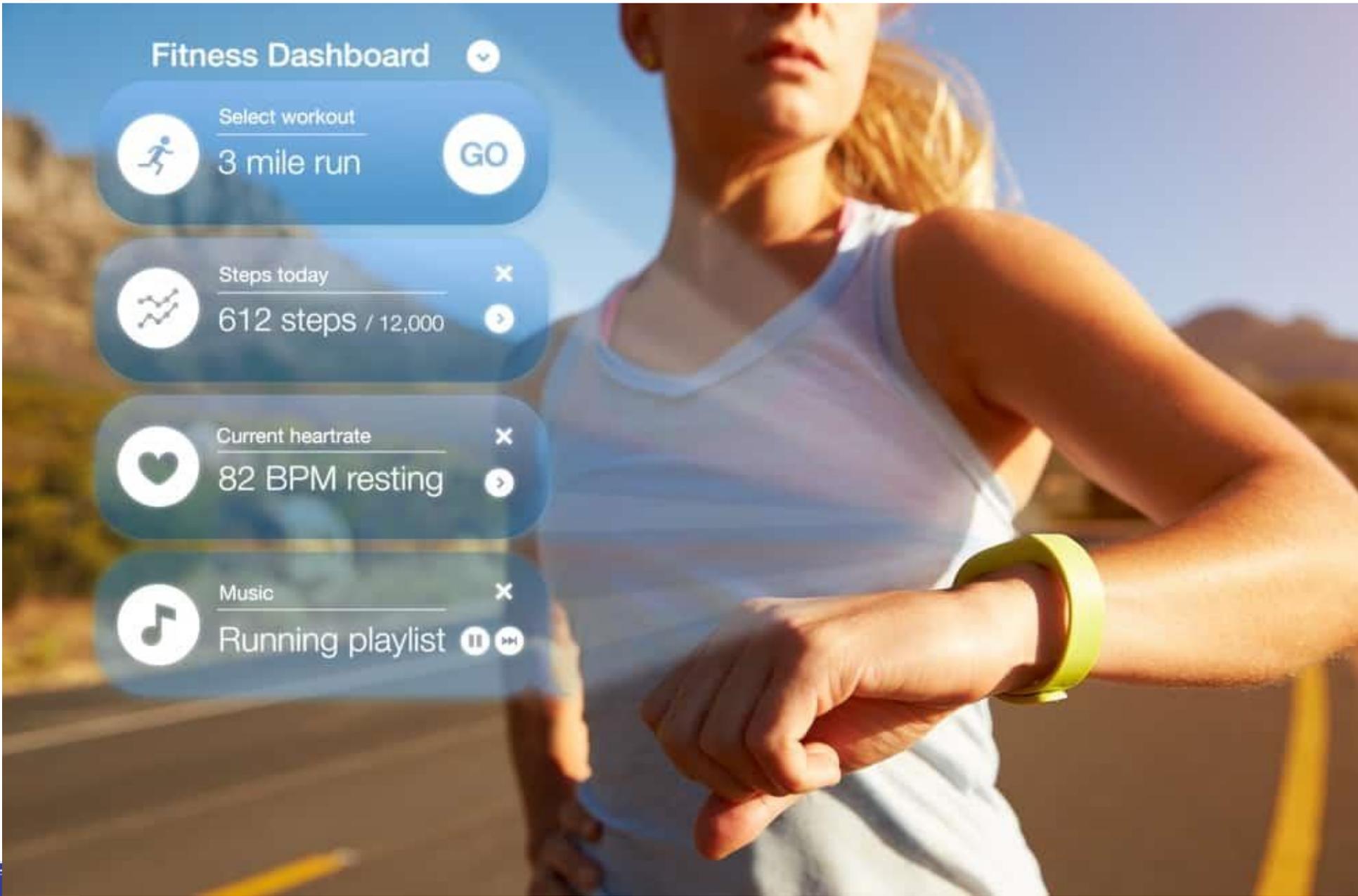
# Why is the "one size fits all" approach inefficient?

# How can we best use technology to help students learn?

# Children are very diverse and in many different ways diverse.
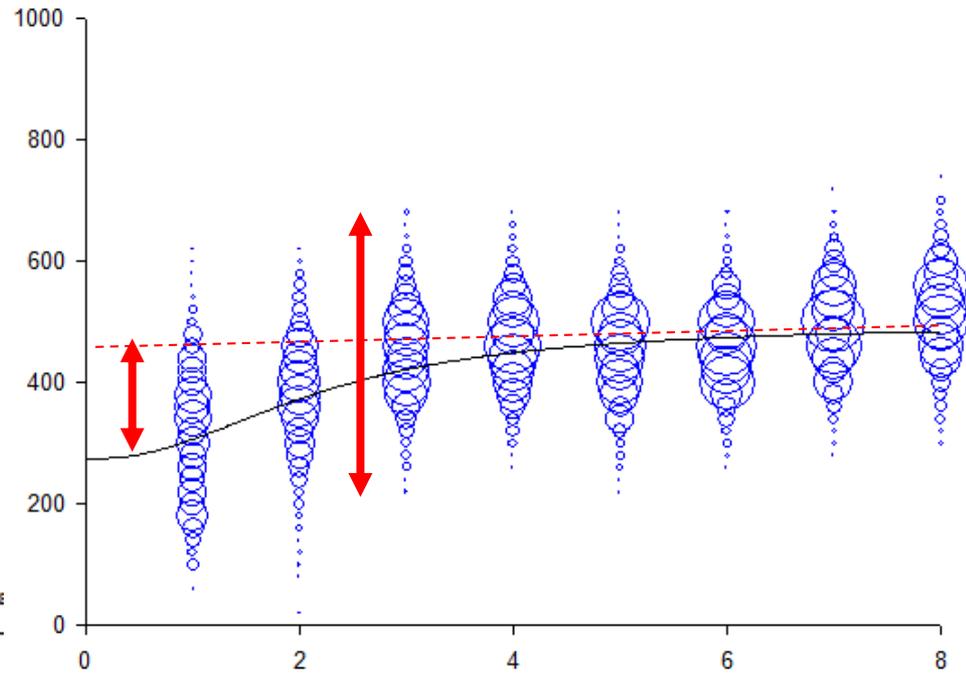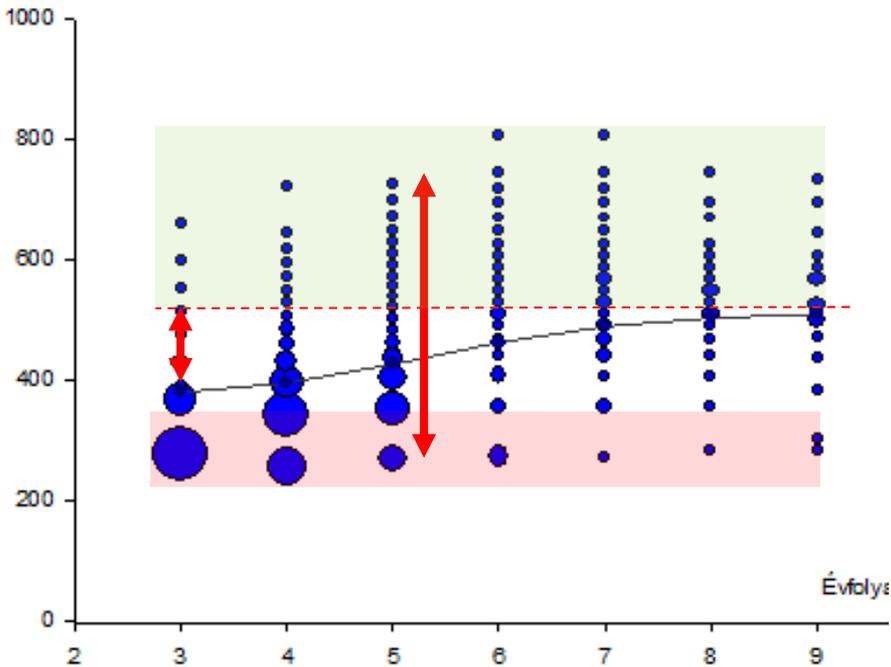


Problem-solving skills   Mathematical knowledge

# Age does not determine skills and abilities.

# Technology can help to address this issue and personalise education.

# "Children of today have been surrounded by digital technology" (Livari et al., 2020)

But

- Used not for learning

- => Much larger digital divide

- Need for proactive technology usage: critically considering how it could and should be used

# Are they prepared to use technology-based educational programs – school infrastructure?

**Pop all the bubbles** before you run out of time.

→

Switch on the radio.

**Click on the button on the radio.**

# 45 min. training - mean achievement: 93% (SD=10%)
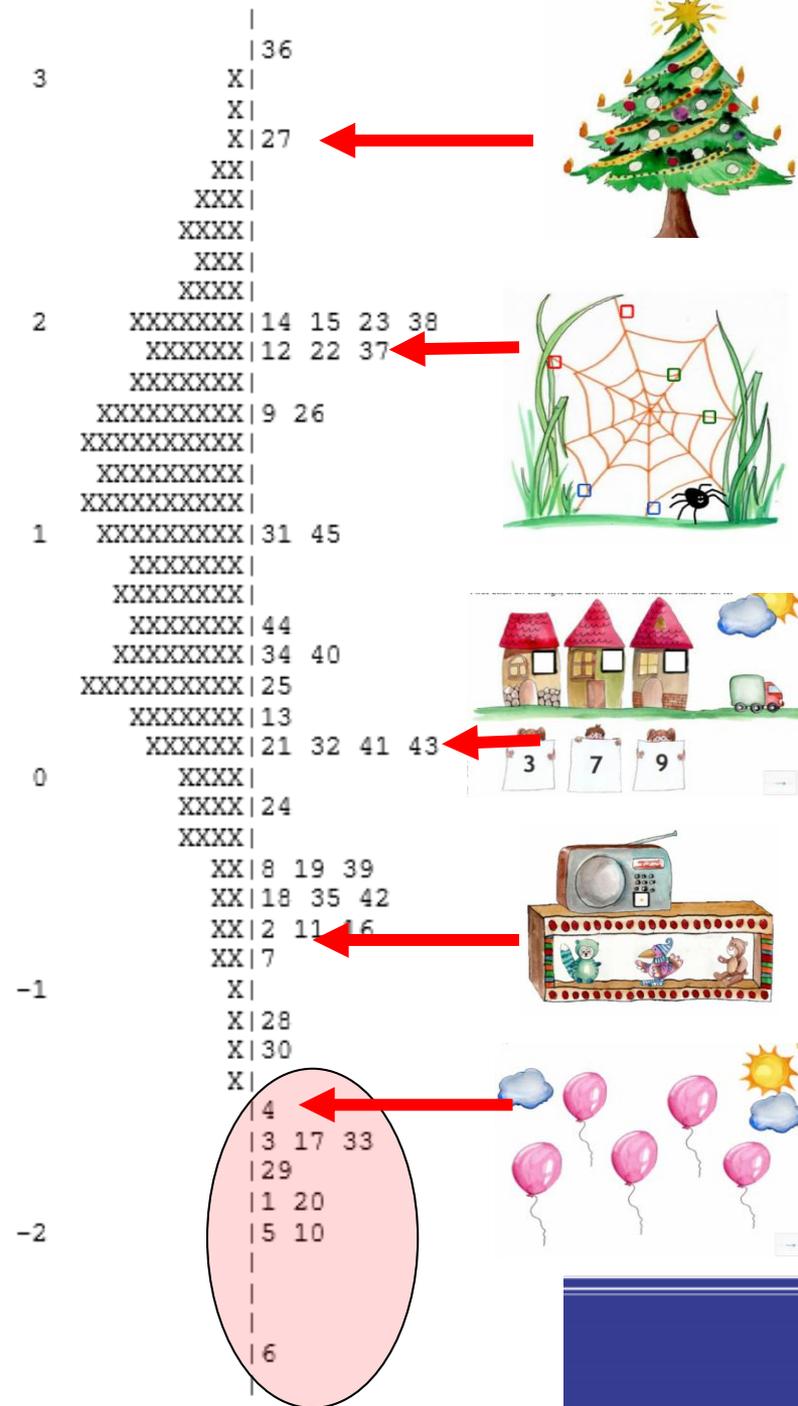
# What do we need to know?

- How skilled our students are?
- What do they know in the most important domains of education?
- The profile of assessment from a summative approach to a <span style="color:red">diagnostic, more learning-centred view</span> - use <span style="color:red">assessment to facilitate learning</span>.

# Possibilities of technology based assessment

- among the most dynamically developing areas
- huge improvement of data transfer technology and data analysis methods
- qualitative change of assessment needed

# Computer-based assessment

- extraordinary opportunities
- more realistic, application-oriented, engaging and authentic context
- innovative item development opportunities, producing dynamic, interactive multimedia items
- more valid assessments
- provide instant, objective, standardised feedback
- adaptive test algorithms

# Instant feedback+adaptivity

- make video games so popular
- technology-based assessment and game-based learning are converging

# What do we know in 2021?

- we can develop complex, real-world, authentic, high-quality tests
- the 'one-size-fits-all' approach is not effective
- summative test results do not meet the individual needs of students
- tests needed: improving the learning process -> re-think the essence of assessment
- more a learning-centred, low-stakes approach
- Diagnostic assessment to facilitate learning.

# 3 axioms of the developments:

- the role of teachers remains central in the teaching and learning processes
- large differences between pupils
- regular feedback is essential for learning

# Diagnostic assessment – what do we need?

- Theory – frameworks

- Assessment platform

- Tasks (based on theory) -> Itembank (running in the assessment platform) -> Scaled itembank

# Diagnostic assessment – what do we need? - Recipe

- **Theory – frameworks**
- Assessment platform
- Tasks (based on theory) -> Itembank (running in the assessment platform) -> Scaled itembank

PISA
PISA 2018 Assessment and Analytical Framework

TIMSS
TIMSS Advanced 2015
Assessment Frameworks

Ina V.S. Mullis
Michael O. Martin, Editors

IEA TIMSS & PIRLS
International Study Center

PROGRESS IN INTERNATIONAL READING LITERACY STUDY
PIRLS
PIRLS 2021
Assessment Frameworks

Ina V.S. Mullis and Michael O. Martin, Editors

TIMSS & PIRLS
International Study Center
Lynch School of Education
BOSTON COLLEGE
IEA
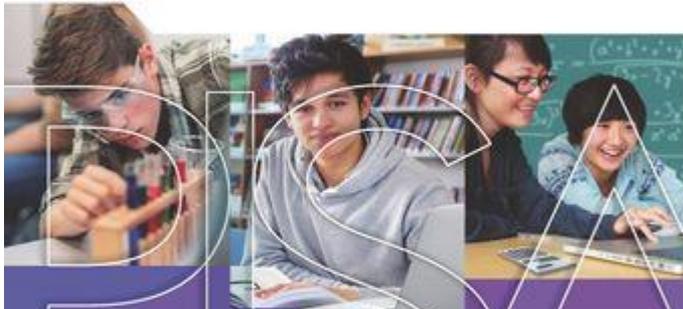
READING FRAMEWORK
FOR THE 2019 NATIONAL ASSESSMENT OF EDUCATIONAL PROGRESS

NATIONAL ASSESSMENT GOVERNING BOARD
U.S. Department of Education
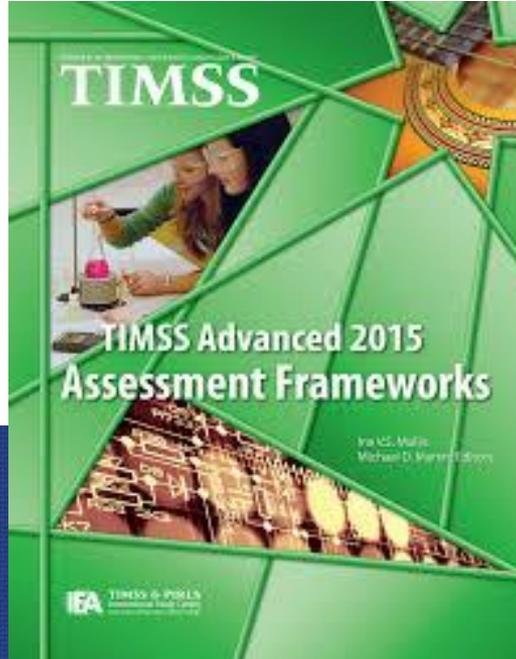
Mathematics framework for the 1996 National Assessment of Educational Progress

NAEP Mathematics Consensus Project

United States. National Assessment Governing Board.

Reprint from the collections of the University of California Libraries
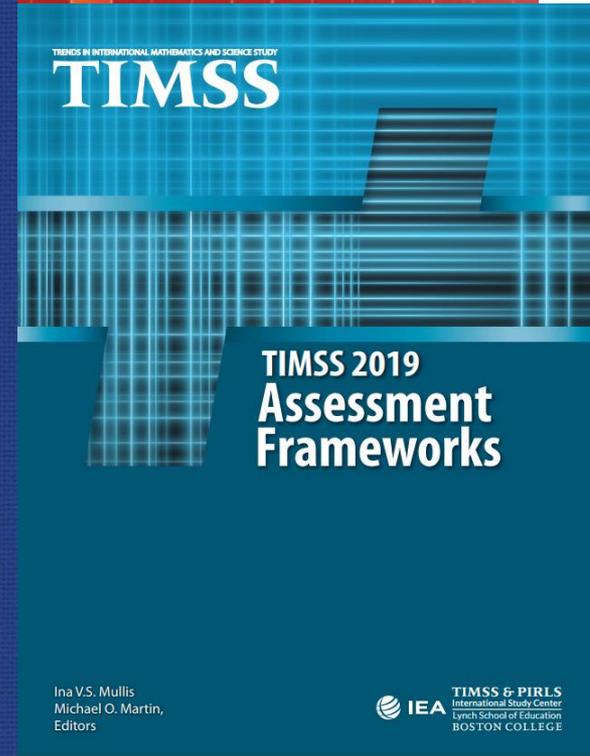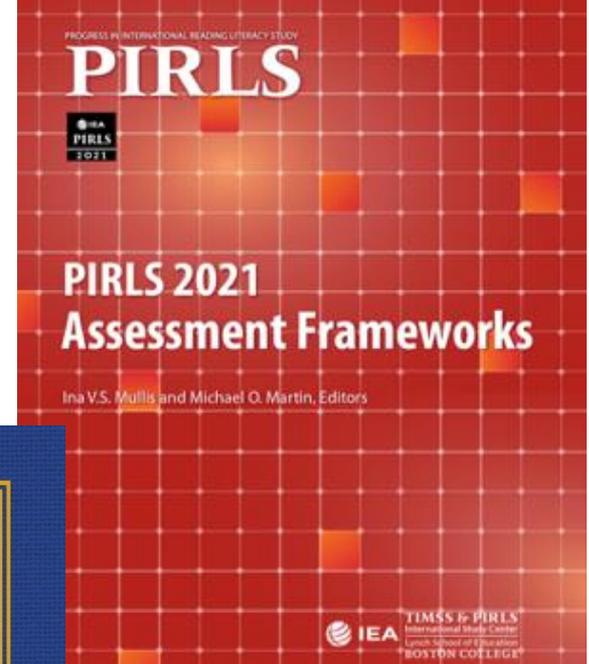
TRENDS IN INTERNATIONAL MATHEMATICS AND SCIENCE STUDY
TIMSS
TIMSS 2019
Assessment Frameworks

Ina V.S. Mullis
Michael O. Martin,
Editors

IEA TIMSS & PIRLS
International Study Center
Lynch School of Education
BOSTON COLLEGE

# Diagnostic assessment – what do we need? - Recipe

- Theory – frameworks

- **Assessment platform**

- Tasks (based on theory) -> Itembank (running in the assessment platform) -> Scaled itembank
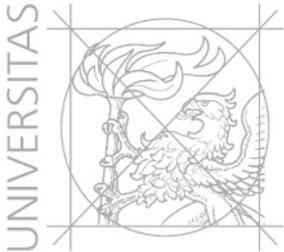
# Diagnostic assessment – assessment platform

- learning centred
- easy-to-use, but innovative possibilities
- Item builder module for first-, second- and third-generation tasks
- administer both fix and adaptive tests
- any device + even low-speed internet
- prompt or quick scoring
- good feedback module – embedded visualization

# Diagnostic assessment – what do we need?

- Theory – frameworks

- Assessment platform

- Several thousand tasks based on the theory -> Itembank (running in the assessment platform) -> Scaled itembank

# Diagnostic assessment – what do we have?

- Theory – frameworks

- Assessment platform

- Several thousand tasks based on the theory -> Itembank (running in the assessment platform) -> Scaled itembank

# eDia online diagnostic assessment system for personalised learning

- Development started in 2009
- Theory: three-dimensional model of knowledge

# Diagnostic assessment - theory

- identifying all important knowledge elements

- three main goals of schooling:
  - to cultivate students' minds, general cognitive abilities and thinking skills
  - to develop usable, applicable knowledge
  - to transmit content knowledge

# Three-dimensional model of knowledge

**REASONING DIMENSION**
**Psychological**

**SOCIAL AND CULTURAL DIMENSION**
**Application**

**DISCIPLINARY DIMENSION**
**Content**

# Three-dimensional model of mathematical knowledge

**REASONING DIMENSION**
**Psychological**

the interaction between students' cognitive development and learning mathematics at school

content knowledge described in the national core curriculum

**SOCIAL AND CULTURAL DIMENSION**
**Application**

**DISCIPLINARY DIMENSION**
**Content**

the type of skills that make mathematics useful in areas other than the immediate school context
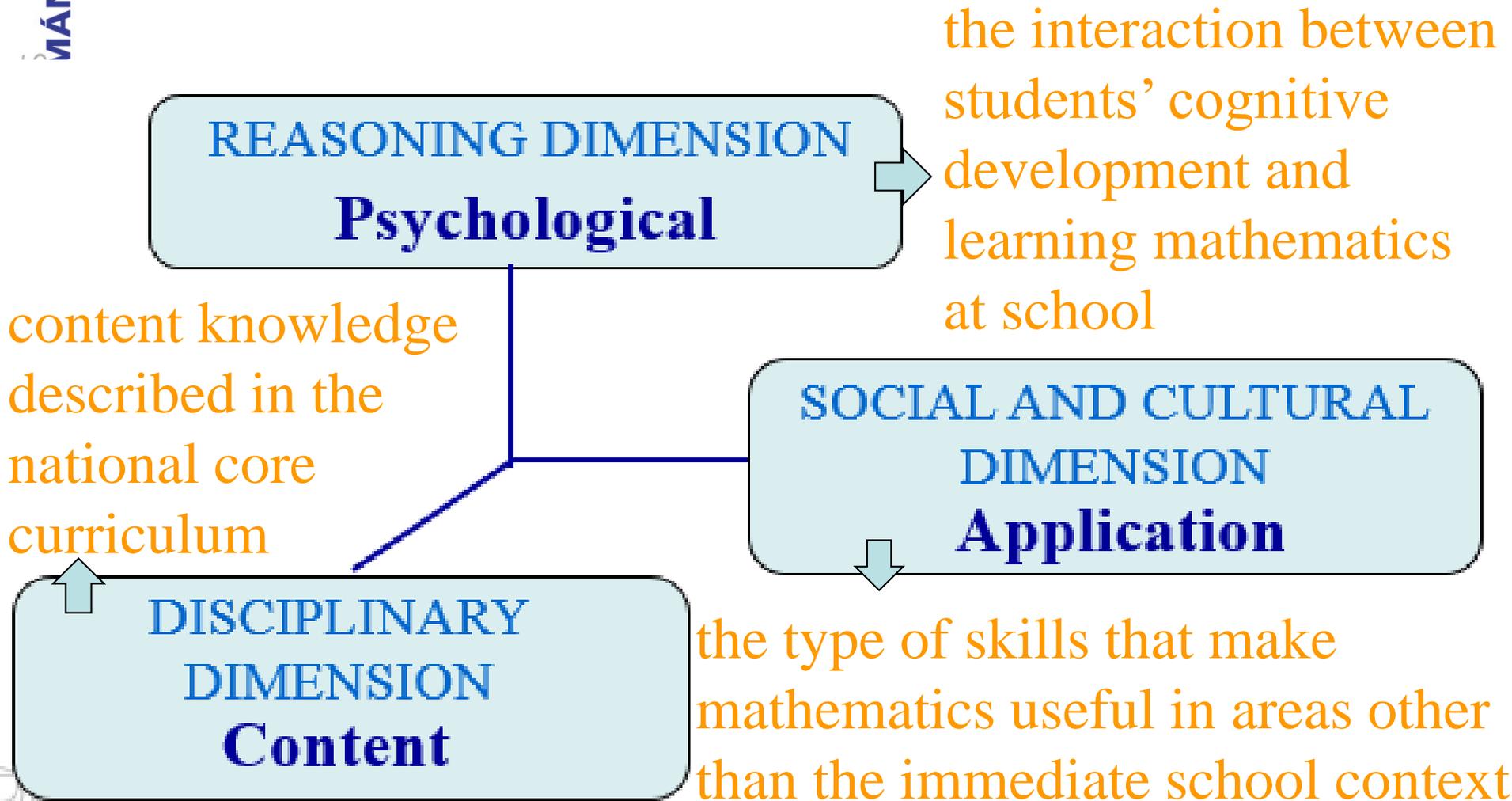
FRAMEWORK FOR
DIAGNOSTIC ASSESSMENT
OF READING

Edited by
Benő Csapó · Valéria Csépe

NEMZETI TANKÖNYVKIADÓ

FRAMEWORK FOR
DIAGNOSTIC ASSESSMENT
OF MATHEMATICS

Edited by
Benő Csapó · Mária Szendrei

NEMZETI TANKÖNYVKIADÓ

FRAMEWORK FOR
DIAGNOSTIC ASSESSMENT
OF SCIENCE

Edited by
Benő Csapó · Gábor Szabó

NEMZETI TANKÖNYVKIADÓ

# Diagnostic assessment – what do we have?

- Theory – frameworks

- Assessment platform

- Tasks (based on theory) -> Itembank (running in the assessment platform) -> Scaled itembank

# eDia-platform

Rose Frog is jumping from one lily pad to the next with the numbers on the lily pads going up in each row.

Click on the lily pads with the numbers on them going up in each row.
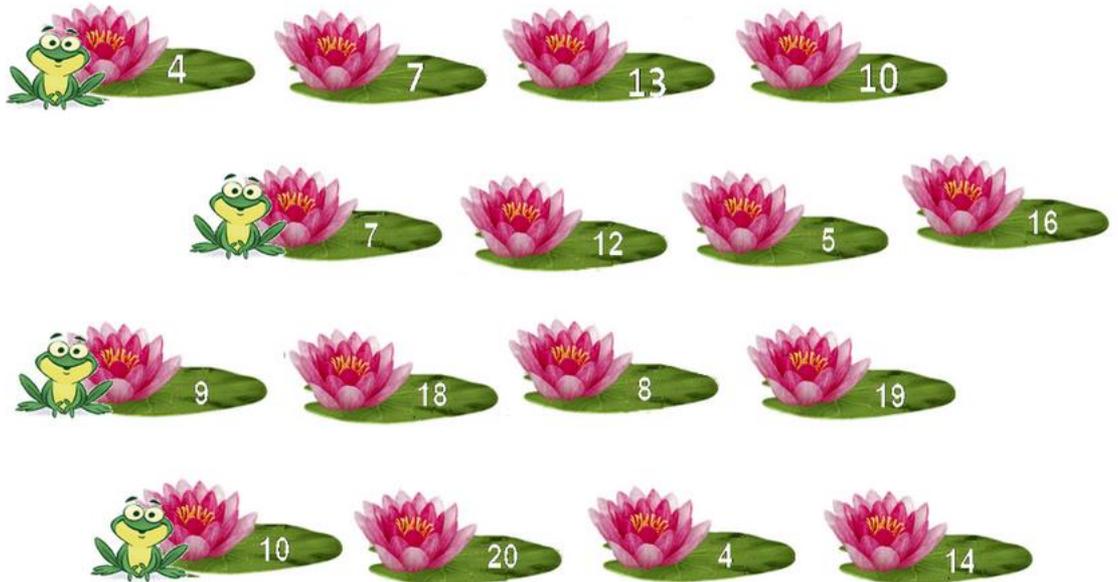
Which animal sound do you hear? Click on the right picture.

AQUALAND FUN BATH AND ADVENTURE PARK

23. How many pictures are demonstrating the experience options on the main page?

○ 1          ○ 5

○ 2          ○ 6

Next ○

# Deatiled feedback modul

- Assessment code
- Following students development
- GDPR proof

| Mérési azonosító | Mat | MA | MD | MG | Olv | OA | OD | OG | Term | TA | TD | TG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| A000-A000 | 476 | 544 | 350 | 524 | 545 | 529 | 635 | 517 | 522 | 486 | 527 | 548 |
|  | 557 | 533 | 592 | 577 | 437 | 499 | 424 | 416 | 531 | 610 | 470 | 521 |
|  | 694 | 694 | 667 | 737 | 533 | 468 | 531 | 559 |  |  |  |  |
|  | 534 | 560 | 524 | 470 | 450 | 424 | 505 | 453 | 474 | 601 | 423 | 440 |
|  | 616 | 663 | 470 | 657 | 563 | 591 | 609 | 532 | 511 | 610 | 517 | 463 |
|  | 584 | 657 | 568 | 567 | 522 | 517 | 505 | 540 | 571 | 584 | 548 | 568 |
|  | 484 | 497 | 493 | 444 | 412 | 377 | 346 | 477 |  |  |  |  |
|  | 520 | 524 | 475 | 657 | 410 | 346 | 436 | 406 | 559 | 672 | 484 | 610 |
|  | 471 | 470 | 470 | 473 | 450 | 438 | 375 | 487 | 429 | 568 | 385 | 392 |

Knowledge level (500/100 – Grade/Domain)

| | Mat | MA | MD | MG | Olv | OA | OD | OG | Term | TA | TD | TG |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Osztály eredménye: | 553 | 554 | 550 | 555 | 455 | 397 | 477 | 475 | 462 | 483 | 473 | 431 |
| Dél-Alföld régió eredménye: | 533 | 546 | 517 | 537 | 470 | 461 | 465 | 483 | 497 | 541 | 489 | 463 |
| Megyei jogú város településtípus eredménye: | 550 | 562 | 532 | 555 | 476 | 466 | 474 | 487 | 501 | 542 | 494 | 469 |
| Országos eredmény: | 538 | 556 | 515 | 543 | 485 | 477 | 482 | 495 | 501 | 542 | 493 | 470 |

# Dimension-level

| Domain/dimension | Grade 5 | Class | Score |
|---|---|---|---|
| Mathematics | 530 | 495 | 616 |
| Mathematics application | 569 | 536 | 607 |
| Mathematics disciplinary | 518 | 494 | 708 |
| Mathematics reasoning | 505 | 464 | 407 |

Light blue: classmates
Green: country level mean
Red: students' achievement
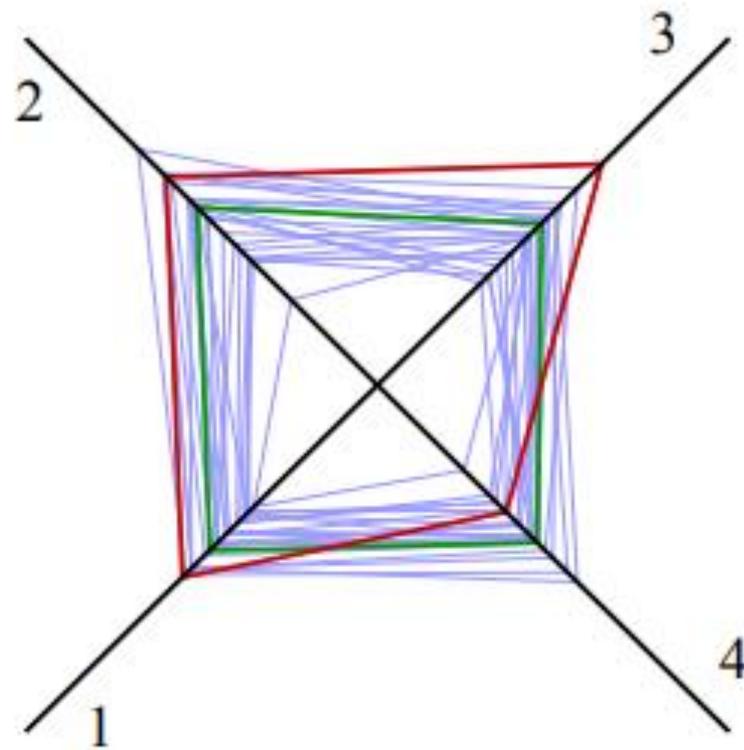
# Text-based evaluation

**2. Matematika alkalmazási kategória értékelése:**

A tanuló matematikai tudását hatékonyabban tudja alkalmazni, mint kortársai átlagosan teszik országos szinten. Képes arra, hogy korosztályában összetettnek számító, konkrét matematikai jellegű problémaszituációkban hatékonyan alkalmazzon modelleket, amelyek esetleg feltételhez kötöttek vagy feltételek megadását igénylik.

**3. Matematika diszciplináris kategória értékelése:**

A matematika diszciplináris tudása tekintetében a tanuló teljesítménye az évfolyama szerinti országosan legjobban teljesítő 2,5%-hoz tartozik. Magasan az átlag felett képes megoldani az iskolában, matematikaórán tanult feladatokat, a matematikaórán tanult ismereteket órai kontextusban használni. Képes egynél több megoldást találni egy tanórai kontextushoz igen hasonló szöveges feladatra. Az átlagosnál jóval magasabb szinten képes megoldani szöveges feladatokat korosztályának megfelelő (táblázatokban, kördiagramon, piktogramon, egynél nagyobb beosztású oszlopdiagramon) ábrán ábrázolt adatok értelmezésével. Igen hatékonyan képes különböző forrásból származó adatok összevetésére, majd azokból következtetések levonására (korosztályának megfelelő szinten).

**4. Matematika gondolkodási kategória értékelése:**

A tanuló megoldásaiból látszik, hogy alapvetően képes a feladatokat értelmezni és saját érveket alkotni. Ezen a típusú feladatokon nyújtott teljesítménye teljes mértékben megegyezik az országos átlagos teljesítménnyel (korosztálya vonatkozásában). Érti a mennyiségek és a számok közötti kapcsolatot, és érti az egész

A teszten elért eredményed annál magasabb, minél több lufit látsz Malacka fölött.

Az eredményed:

**100%-os**

Tantárgyi tudásod: 100%

Tudásod alkalmazása: 100%

Gondolkodtató feladatok: 100%

# Diagnostic assessment – what do we have?

- Theory – frameworks
- Assessment platform
- Tasks (based on theory) -> Itembank (running in the assessment platform) -> Scaled itembank

At about 20.000 tasks in the field of Mathematics, Reading and Science.

# Mathematical reasoning task: whole numbers + inductive reasoning

Dumpling Arthur got 20 bars of chocolate for his birthday. He ate a few pieces from each chocolate bar, and then he put the rest of the chocolate into groups according to a certain rule. There is an odd one out in each row. Which one is it? Click on it.

# Mathematical literacy task: adding up to 10 in realistic application context

# Mathematical disciplinary task
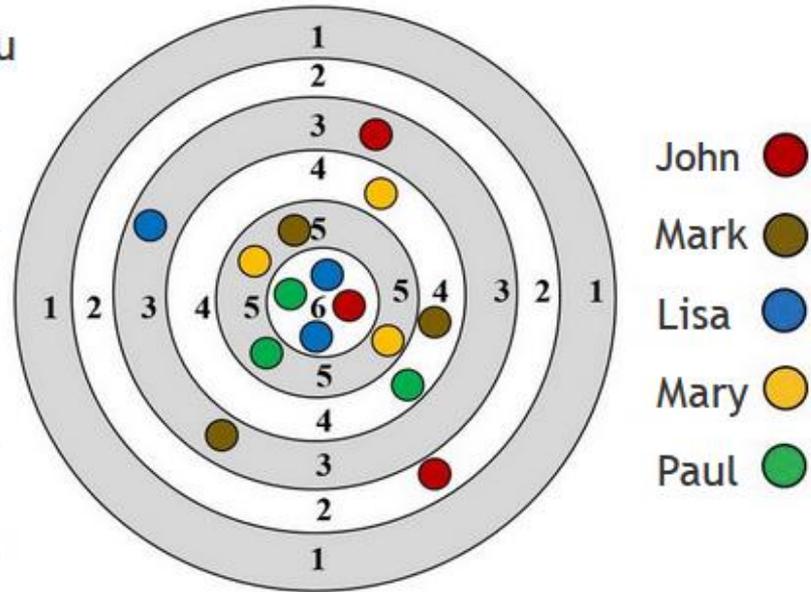
# Based on empirical data

- 10.418.422 task solution
- 461.446 solved tests
- 94.970 students
- 50.863 items
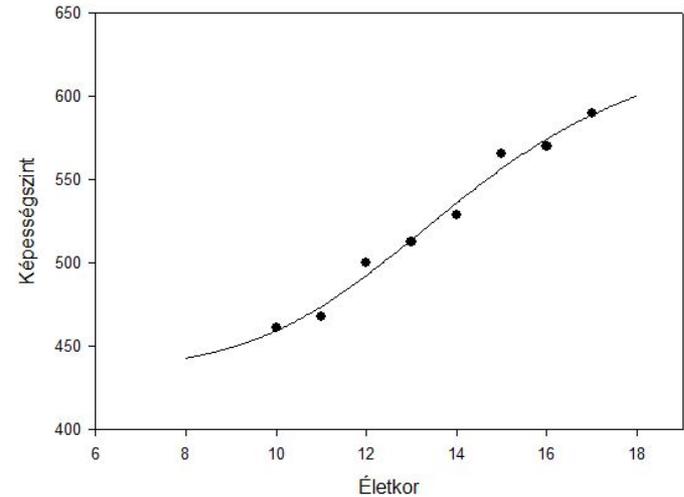
# eDia-system during the whole school year



Grade 1. mouse + keyborad + school readiness

Grade 1-8. Reasoning skills + ICT literacy

**Sept. Oct. Nov. Dec. Jan. Febr. March Apr. May Jun.**

Grade 2-7
Mathematics
Reading
Science

Grade 1-6
Mathematics
Reading
Science

Grade 1-6
Mathematics
Reading
Science

# Technology-Based Assessment is Applicable in an Educational Context
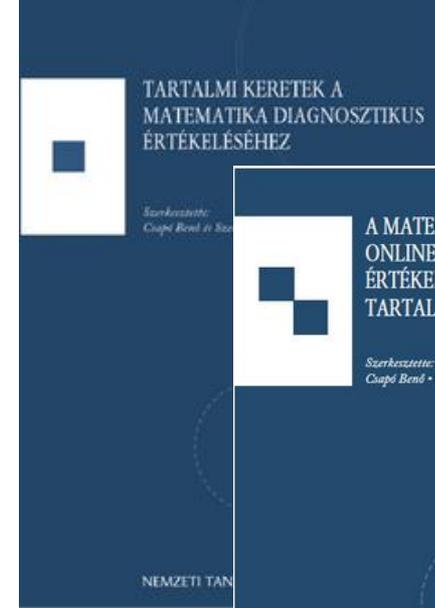
# Methods - sample

- 1-8 grade students (aged 7 to 14)
- 1500 to 2000 students in each cohort (N=14062)
- 656 classes from 134 schools
- The proportion of boys and girls was about the same

# Results



- Cronbach-$\alpha$=.942
- $\alpha$_MD=.89;
  $\alpha$_MR=.83;
  $\alpha$_ML=.89
- M=49.39% (SD=23.87)
- Math: only 18 out of the 5115 - not finishing the test
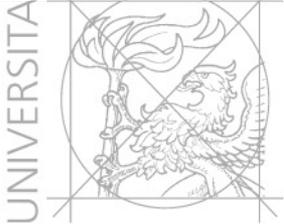
# Mathematics

90

88

248 | 540 | 542 | 567 | 569 | 570 | 89

237 | 390 | 419 | 541 | 579 | 580 | 83 | 85 | 86

188 | 289 | 380 | 389 | 427 | 428 | 511 | 512 | 84 | 87 | 98

176 | 196 | 197 | 199 | 200 | 212 | 214 | 215 | 217 | 297 | 325 | 328 | 364 | 387 | 405 | 420 | 425 | 426 | 460 | 461 | 492 | 513 | 559 | 560 | 568 | 573 | 576

127 | 128 | 175 | 180 | 210 | 219 | 256 | 347 | 350 | 379 | 449 | 462 | 539 | 562 | 575 | 581

114 | 151 | 187 | 198 | 213 | 236 | 257 | 280 | 281 | 296 | 298 | 34 | 349 | 448 | 459 | 494 | 561 | 95 | 96

111 | 112 | 185 | 186 | 211 | 216 | 255 | 295 | 327 | 337 | 361 | 403 | 410 | 413 | 450 | 490 | 491 | 508 | 525 | 528 | 582 | 591 | 592 | 593 | 594 | 97

137 | 194 | 195 | 20 | 201 | 220 | 276 | 277 | 278 | 279 | 317 | 346 | 348 | 362 | 388 | 404 | 447 | 473 | 489 | 556 | 558 | 574 | 587 | 598 | 600 | 602 | 73 | 7

113 | 122 | 163 | 193 | 206 | 235 | 245 | 246 | 247 | 27 | 28 | 313 | 315 | 359 | 415 | 417 | 475 | 504 | 507 | 51 | 510 | 571 | 572 | 588 | 597

08 | 150 | 152 | 162 | 168 | 190 | 208 | 209 | 21 | 221 | 226 | 227 | 23 | 234 | 24 | 265 | 273 | 283 | 305 | 310 | 311 | 316 | 35 | 355 | 356 | 363 | 383 | 402

02 | 04 | 05 | 06 | 102 | 135 | 136 | 157 | 160 | 165 | 167 | 169 | 207 | 22 | 228 | 229 | 230 | 231 | 232 | 233 | 243 | 25 | 26 | 263 | 274 | 275 | 284 | 285 | 28

01 | 03 | 07 | 100 | 101 | 120 | 149 | 158 | 159 | 161 | 166 | 170 | 171 | 172 | 174 | 189 | 191 | 192 | 203 | 244 | 252 | 261 | 264 | 268 | 270 | 271 | 272 | 28

106 | 117 | 119 | 121 | 133 | 134 | 146 | 148 | 17 | 173 | 19 | 204 | 205 | 218 | 254 | 293 | 302 | 307 | 308 | 309 | 312 | 321 | 322 | 323 | 326 | 334 | 336 | 3

10 | 103 | 107 | 109 | 110 | 115 | 116 | 12 | 126 | 13 | 14 | 141 | 147 | 15 | 16 | 178 | 184 | 202 | 242 | 251 | 253 | 262 | 267 | 269 | 292 | 294 | 299 | 300 |

09 | 108 | 11 | 118 | 123 | 124 | 125 | 132 | 145 | 164 | 183 | 258 | 260 | 266 | 29 | 319 | 329 | 330 | 332 | 341 | 342 | 352 | 365 | 366 | 396 | 411 | 412 | 43

104 | 105 | 130 | 131 | 138 | 143 | 179 | 18 | 181 | 182 | 224 | 225 | 240 | 259 | 33 | 331 | 395 | 406 | 44 | 45 | 46 | 93 | 94

129 | 139 | 140 | 142 | 144 | 153 | 154 | 155 | 156 | 177 | 222 | 223 | 238 | 391 | 394 | 421 | 422 | 424

239 | 241 | 344 | 345 | 392 | 393 | 423

249

250

333

WMM

# well structured item bank

# Technology-based assessment is applicable in the field of mathematics, reading and science from grade 1 to 8.

# Relations between MR, ML and MD (all coefficients are significant at p<.01)
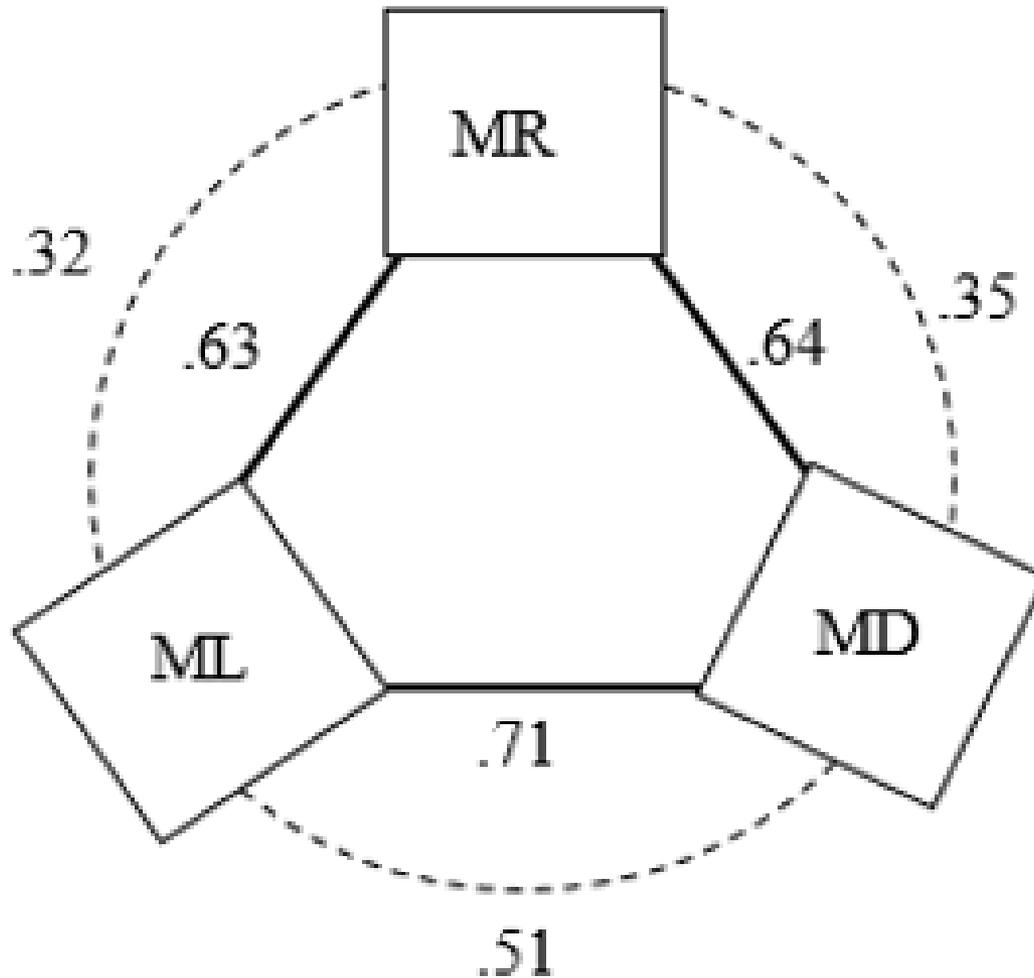


All bivariate relationships were influenced by the third construct

# Goodness of fit indices for testing dimensionality of mathematics

| Model | $\chi^2$ | $df$ | $p$ | CFI | TLI | RMSEA (90% CI) |
|---|---|---|---|---|---|---|
| 3-dimensional | 16955.213 | 1067 | .001 | .965 | .963 | .054 (.053−.055) |
| 1-dimensional | 31445.929 | 1073 | .001 | .931 | .928 | .075 (.075−.076) |

*Note: $df$* = degrees of freedom; CFI = Comparative Fit Index; TLI = Tucker-Lewis Index; RMSEA = Root Mean Square Error of Approximation; $\chi^2$ and $df$ are estimated by WLSMV.

At latent level:
$r_{MD\_MR}$=.685, $r_{MD\_ML}$=.749, $r_{ML\_MR}$=.634, p<.001

$\chi^2$=3389.111; df=6; p<.001
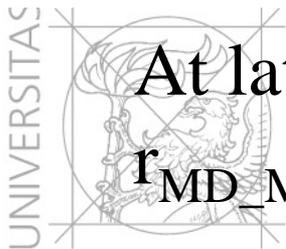
# Goodness of fit indices for testing dimensionality of mathematics

| Grade | Model | $\chi^2$ | df | p | $\Delta\chi^2$ | $\Delta$df | p | CFI | TLI | RMSEA |
|-------|-------|----------|-----|------|----------------|------------|------|------|------|-------|
| 1 | 3 dimension | 409.506 | 249 | .001 | 95.309 | 3 | .001 | .953 | .948 | .077 |
|   | 1 dimension | 586.328 | 252 | .001 |        |   |      | .902 | .893 | .110 |
| 2 | 3 dimension | 543.407 | 321 | .001 | 96.826 | 3 | .001 | .944 | .939 | .061 |
|   | 1 dimension | 734.133 | 324 | .001 |        |   |      | .897 | .889 | .083 |
| 3 | 3 dimension | 171.573 | 149 | .01  | 15.784 | 3 | .01  | .923 | .912 | .046 |
|   | 1 dimension | 194.581 | 152 | .01  |        |   |      | .855 | .837 | .063 |
| 4 | 3 dimension | 236.477 | 206 | .01  | 40.265 | 3 | .001 | .940 | .933 | .060 |
|   | 1 dimension | 268.352 | 209 | .01  |        |   |      | .883 | .871 | .083 |
| 5 | 3 dimension | 381.365 | 186 | .001 | 110.584 | 3 | .001 | .939 | .931 | .060 |
|   | 1 dimension | 675.939 | 189 | .001 |        |   |      | .847 | .830 | .095 |
| 6 | 3 dimension | 680.214 | 492 | .001 | 112.972 | 3 | .001 | .912 | .906 | .054 |
|   | 1 dimension | 966.684 | 495 | .001 |        |   |      | .780 | .765 | .085 |
| 7 | 3 dimension | 1182.063 | 816 | .001 | 205.034 | 3 | .001 | .968 | .966 | .047 |
|   | 1 dimension | 1882.948 | 819 | .001 |        |   |      | .908 | .903 | .079 |
| 8 | 3 dimension | 3021.062 | 557 | .001 | 165.118 | 3 | .001 | .876 | .867 | .124 |
|   | 1 dimension | 3412.642 | 560 | .001 |        |   |      | .856 | .847 | .133 |

Note. df: degrees of freedom; CFI: Comparative Fit Index; TLI: Tucker Lewis Index, RMSEA= Rood Mean Square Error of Approximation, $\chi^2$: and df are estimated by WLSMV. $\Delta\chi^2$ were estimated by Difference Test-procedure in MPlus (see Muthén & Muthén, 2012).

# Disciplinary, application and psychological dimensions of learning mathematics, reading and science can be empirically distinguished independent of the measured grade.
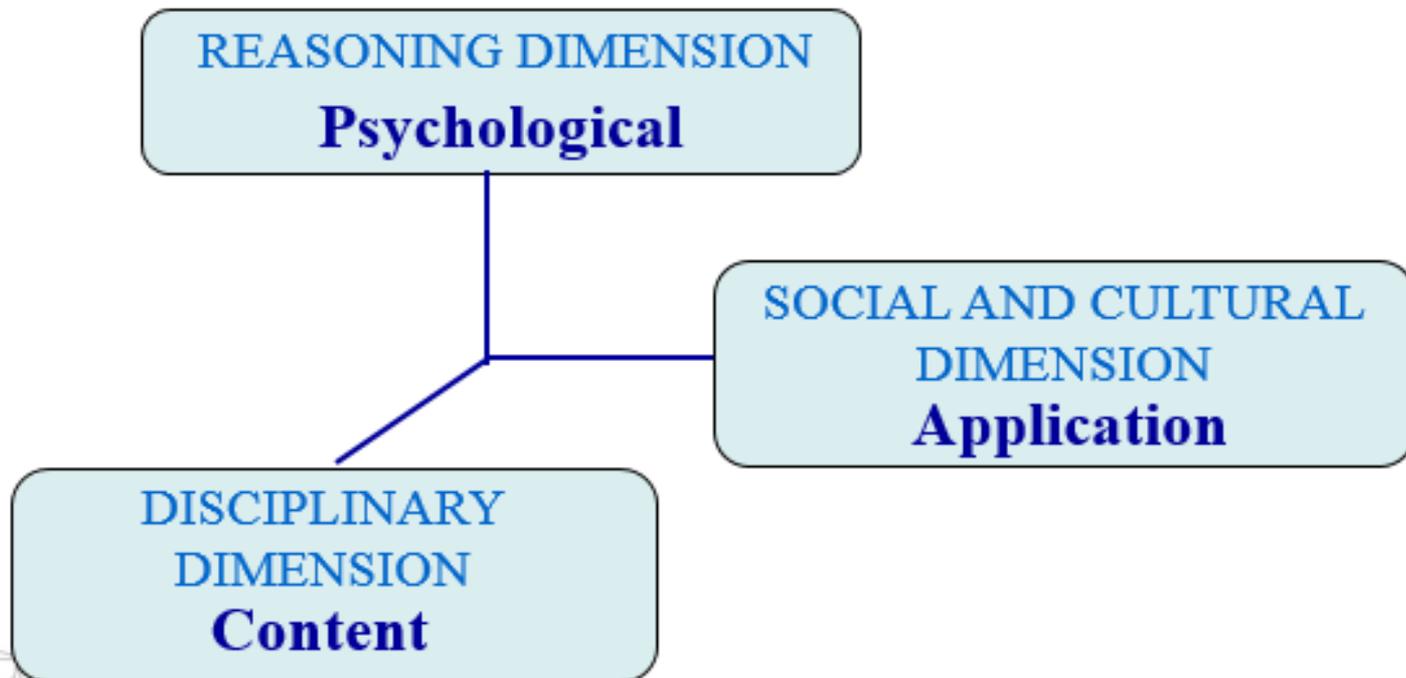
# The development of all three dimensions is important.
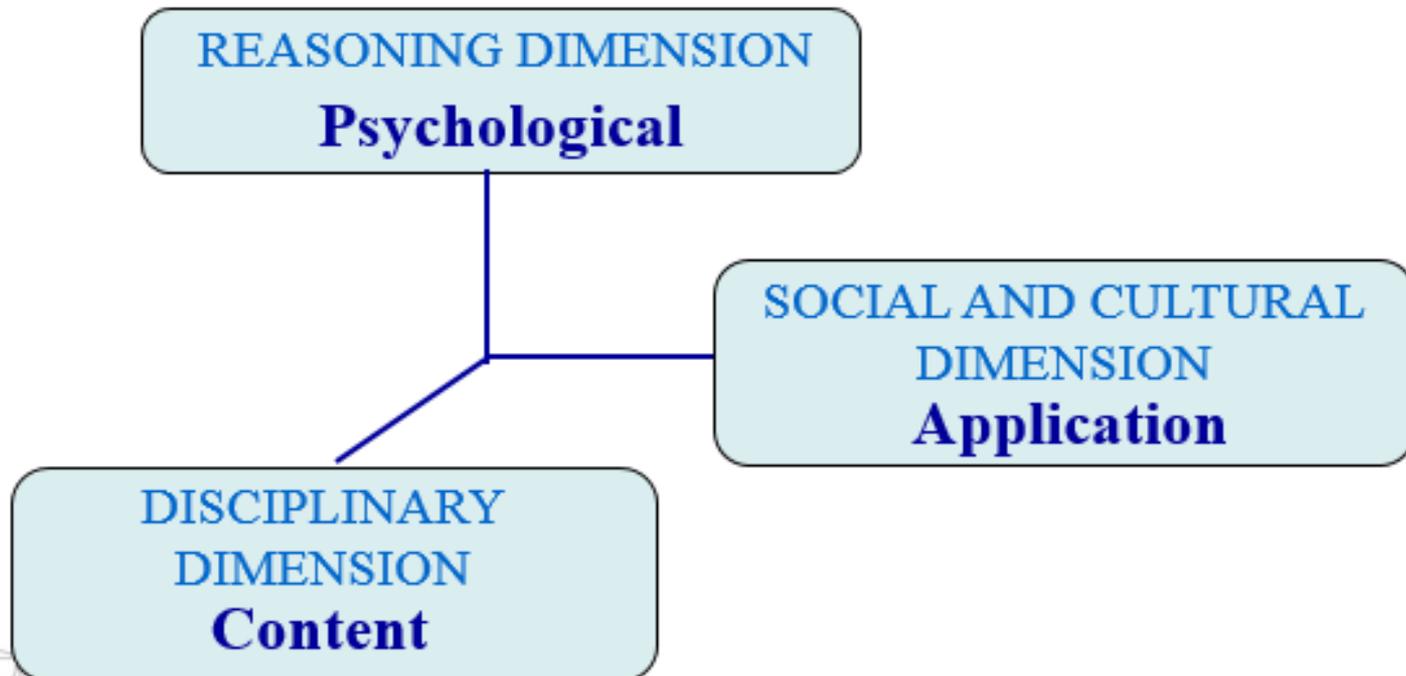
# Extra moduls

- eDia teacher test modul
  ([www.teszt.edia.hu](http://www.teszt.edia.hu))
  Almost 50.000 items (20.000 tasks)

- eDia kindergarten test modul
  ([www.ovi.edia.hu](http://www.ovi.edia.hu))
  2500 tasks
  Optimised for tablets

# Key Takeaways



- Leave the „fitting for all approach"

- Integrate knowledge from different fields

- Change the aim and type of assessment (what and how of teaching).

- Use the advantages of technology-based diagnostic assessment as a tool, supportive medium for <span style="color:red">personalize learning</span>