CoGames

Development of Adaptive Smartphone-Based Games for Remote Cognitive Assessment in Multiple Sclerosis

Silvan Pless, PhD

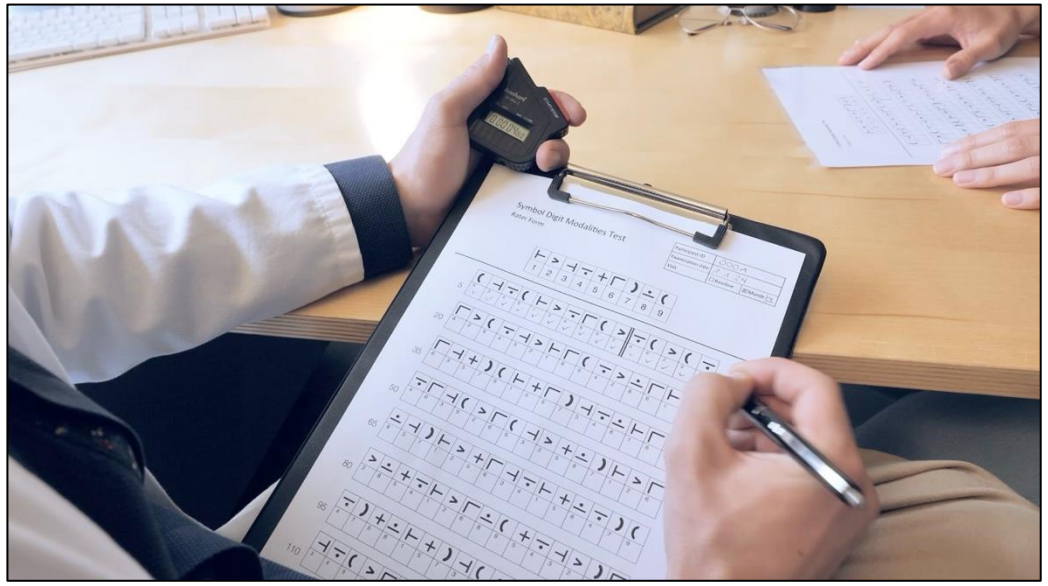
27.08.2026

Digital Endpoints, Basel Biometrics Society

Shortcomings of cognitive assessment in MS

Cognitive impairment in MS

- Prevalence up to 70%
- Strong negative impact on quality of life
- Assessment often neglected



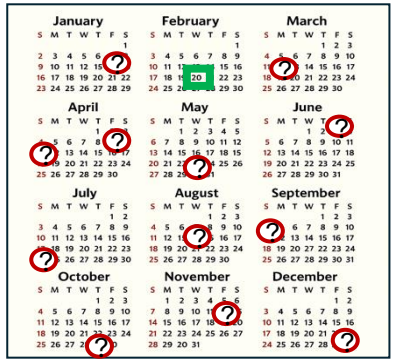
Time-consuming

Inconvenient

Requires experts
(neuropsychologist)

Costly

Only a snapshot

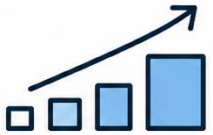


Improving cognitive assessment

Accessible

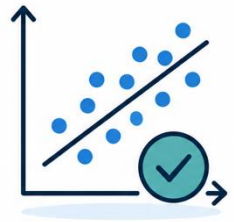
Motivating & enjoyable

Rigorous Psychometrics



Scalable

User friendly

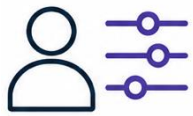


Validity

High frequency

Reliability

Sensitivity to change



Personalized / adaptive

Low cost

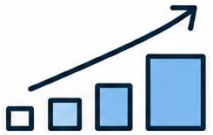
Convenient: *Time-efficient*

Convenient: *Remote*

Improving cognitive assessment

Accessible

Motivating & enjoyable



Scalable

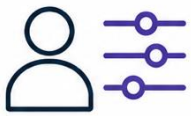
User friendly

Validity

High frequency

Reliability

Sensitivity to change



Personalized / adaptive

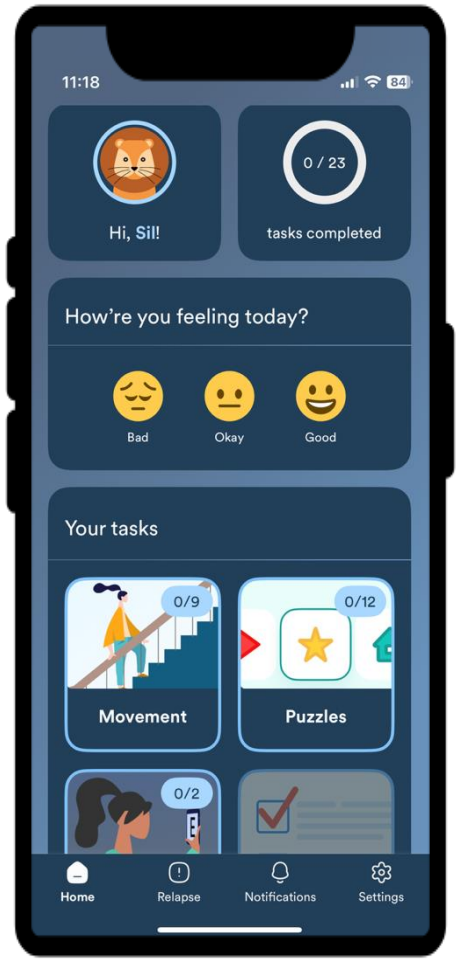
Low cost

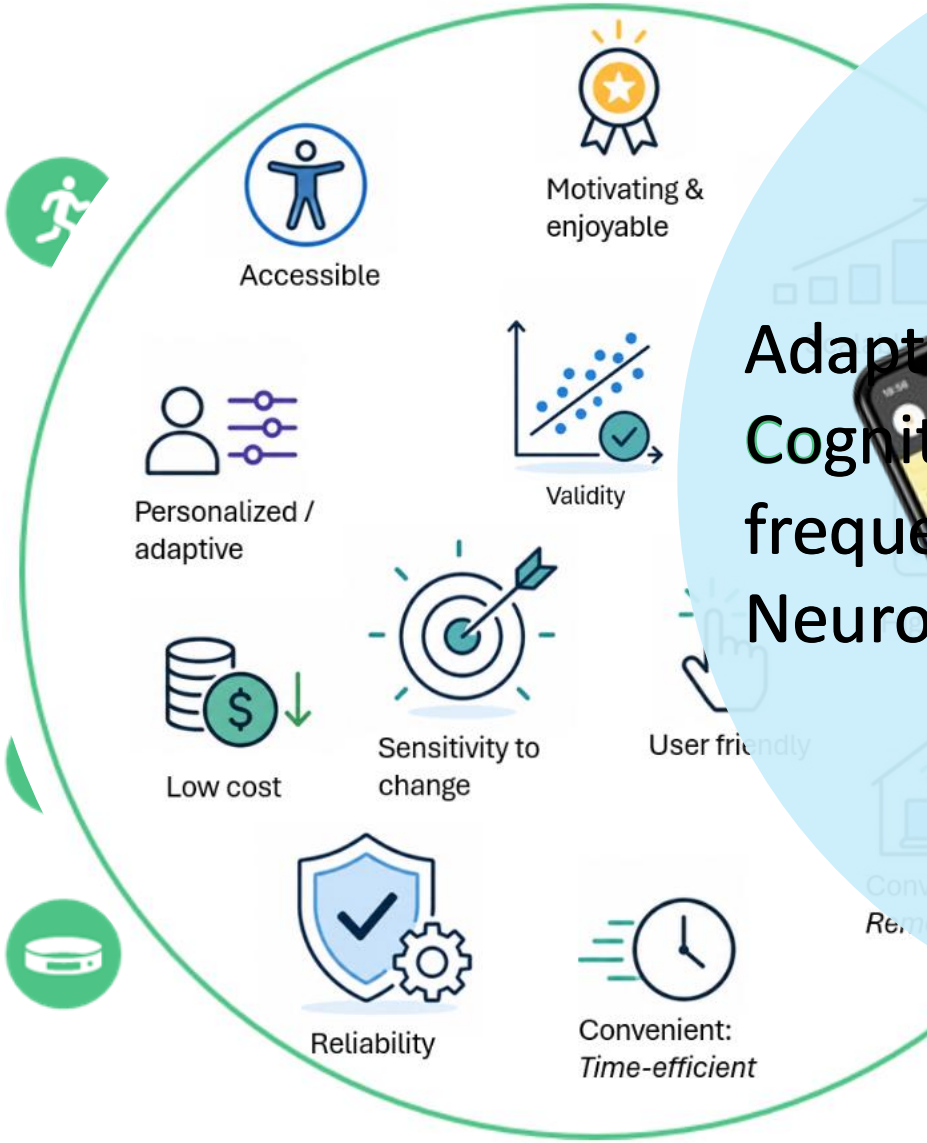
Convenient: *Time-efficient*

Convenient: *Remote*

DreaMS Cognitive assessment

DreaMS App





CoGames

Adaptive, Smartphone-based Cognitive Games for Remote, High-frequency, and Longitudinal Neuropsychological Assessment

Development of CoGames

Step 1

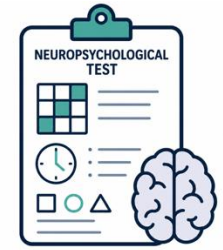
DreaMS
Feasibility Study

Step 2

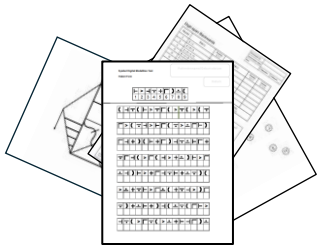
Development
CoGames 1.0

1. Scientific Basis: Domains and their Assessments

Cognitive Domains



Neuropsychological Tests



Measures

2. Gamification: UX / UI

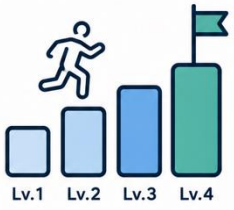
Feedback

Rewards

Clear and appealing Design

3. Dynamic Difficulty Adjustment (DDA)

Progression



Challenge

Adaptiveness

Development - *Scientific basis*

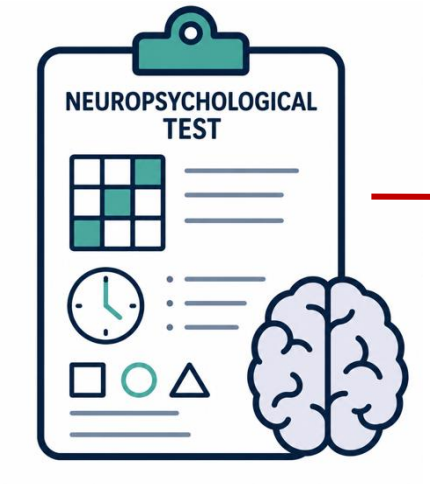
Step 1

DreaMS
Feasibility Study

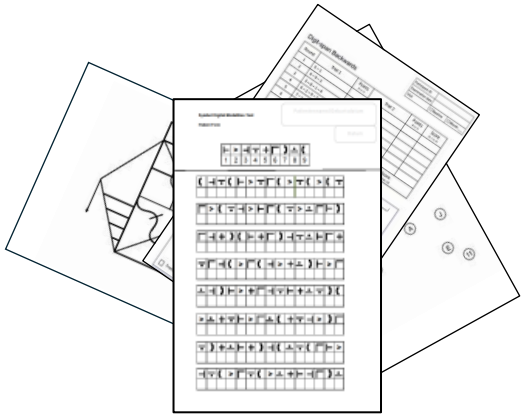
Step 2

Development
CoGames 1.0

Cognitive Domains

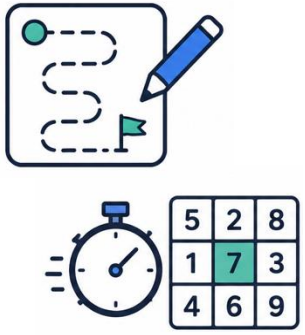


Neuropsychological Tests

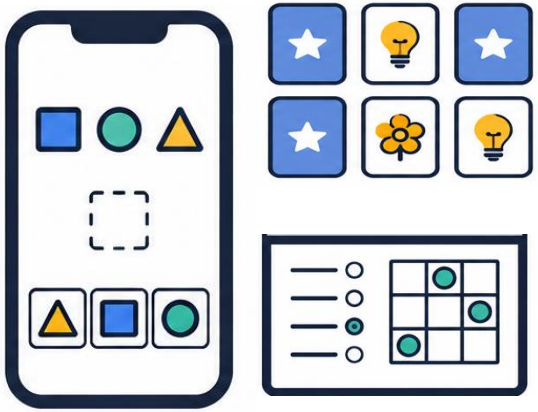


Measures

Test Design

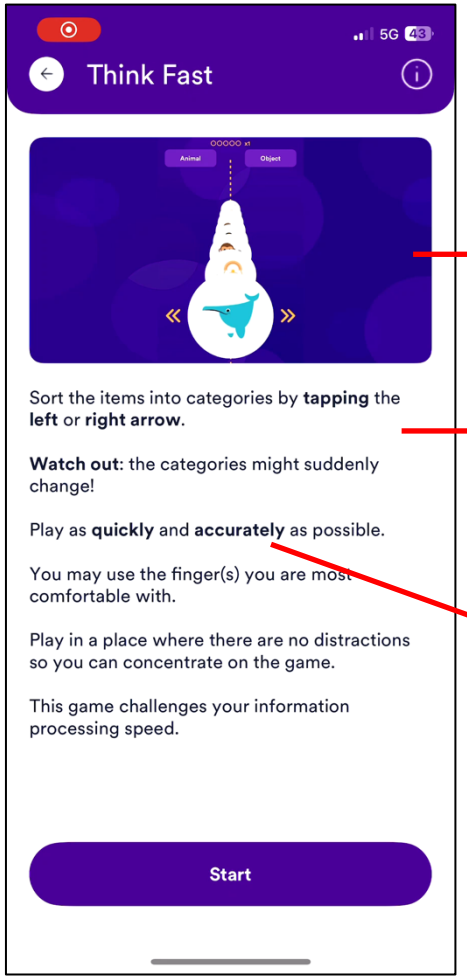


Gameplay



Development - *UI / UX*

Instruction Screen

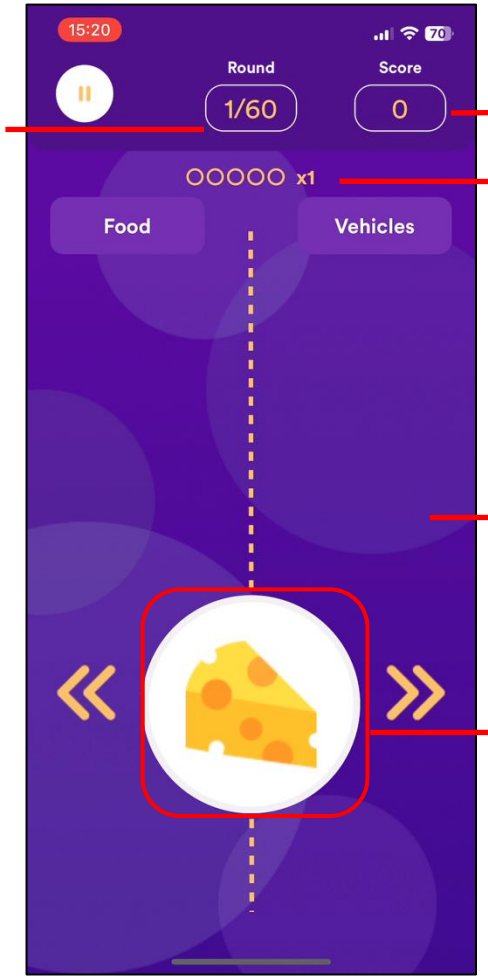


Animated instruction

Short written instruction

Important information in bold

Gameplay



Progress indicator

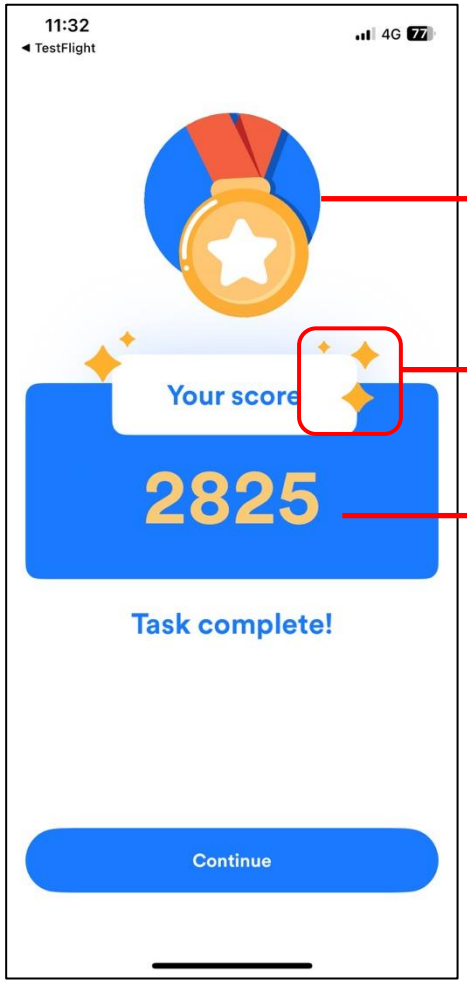
Points / Score

Multiplier

Colorful but non-distracting background

Appealing Graphic Design

Victory Screen

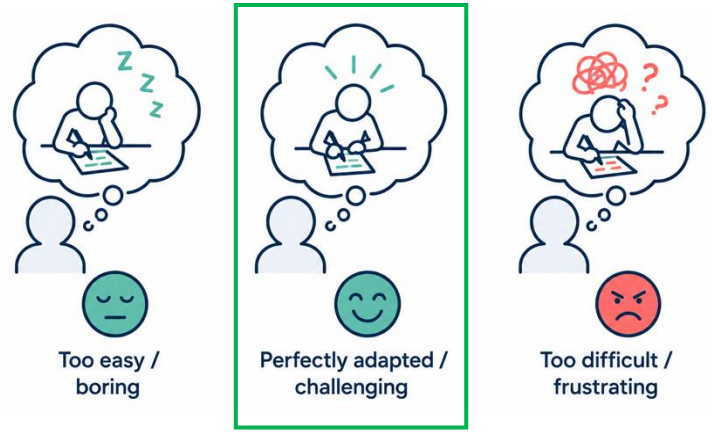


«Trophies»

«Victory animations» (e.g. fireworks, stars)

Total Score

Development - *Dynamic Difficulty Adjustment*



- Engagment & Enjoyment
- Challenge & Concentration

- Motivation & Adherence
- Maximum performance
- Prevents Ceiling- and floor effects

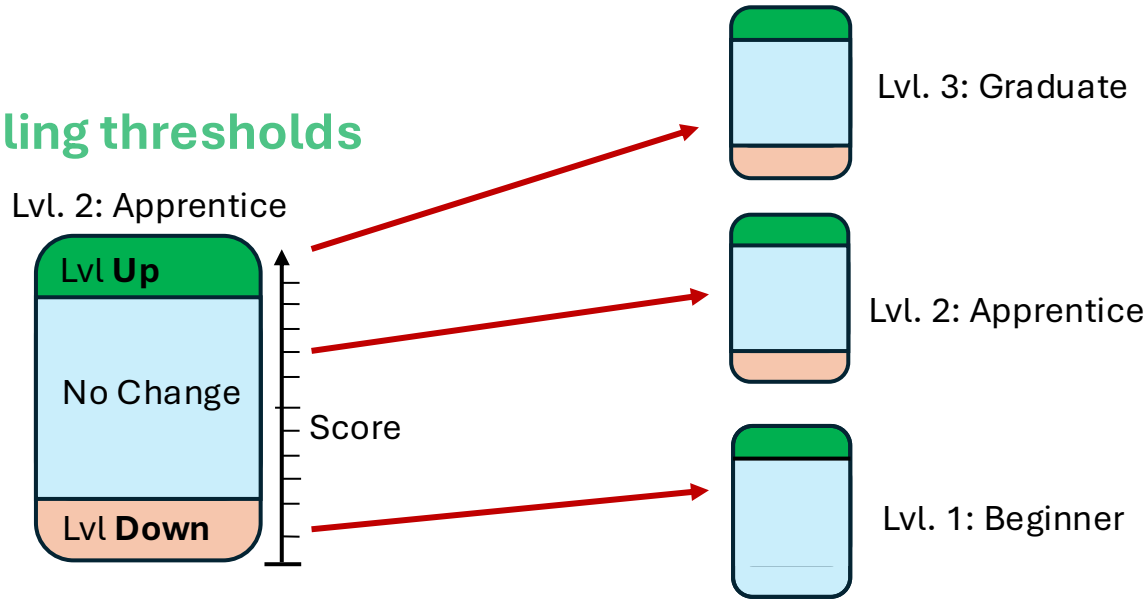
CoGames Levels

Difficulty Parameters

- No. of items to memorize
- Time until recall
- Distractions
- …

Lvl. 1 Beginner	Lvl. 2 Apprentice	Lvl. 3 Graduate	Lvl. 4 Expert
5	6	7	8
2’’	5’’	8’’	10’’
no	no	yes	yes

Leveling thresholds

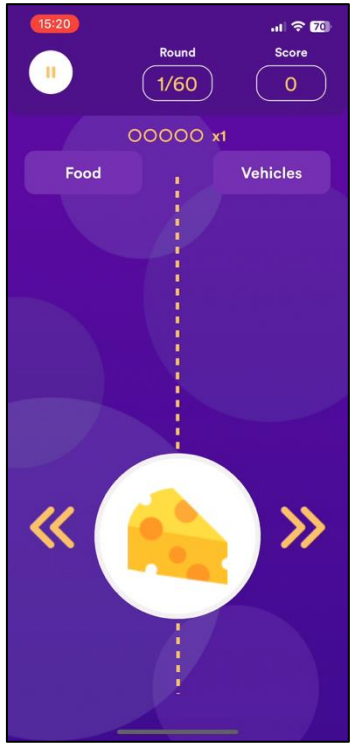


CoGames 1.0

RC2NB | Research Center for Clinical Neuroimmunology and Neuroscience Basel

Processing speed

Think Fast



No. of correct responses in 60 sec.

Numbers

No. of correct responses in 60 sec.

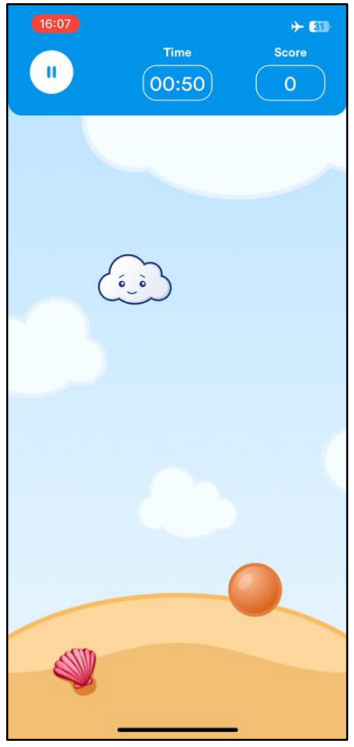
Short- term memory

Treasure Hunt

% of correctly connected squares

Psycho- motor speed

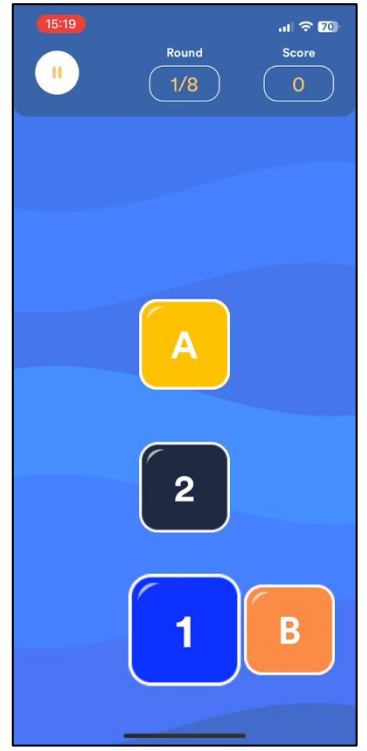
Catch a Cloud



No. of correct responses in 50 sec.

Mental flexibility

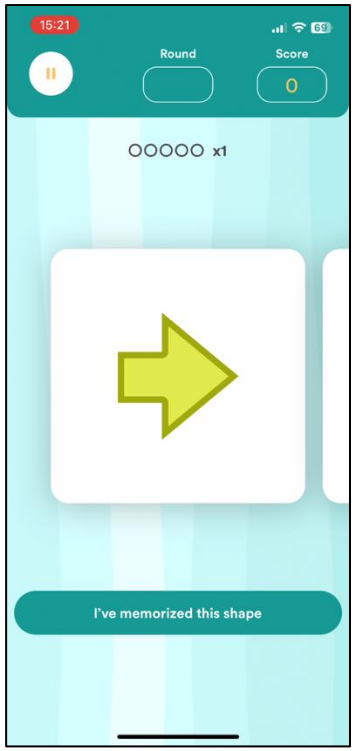
Mixer



No. of correct responses in 60 sec.

Working memory

Match Maker



No. of correct responses in 60 sec.

DreaMS Validation Study

Step 1
DreaMS
Feasibility Study

Step 2
Development
CoGames 1.0

Step 3
dreaMS
Validation Study

DreaMS Validation Study (VS)

- Ongoing (start: April 2022, end Dec. 2026)
- Swiss Multicenter study
- 2 years per participant
- n = 275 people with MS; 50 healthy volunteers
- CoGames 1x/month with intensive (weekly) phases

DreaMS Validation Study - *Data-look*

Step 1 — DreaMS Feasibility Study

Step 2 — Development CoGames 1.0

Step 3 — dreaMS Validation Study

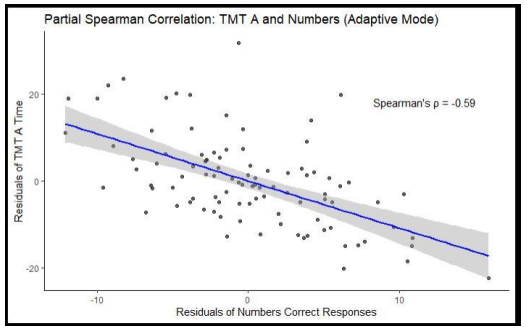
DreaMS VS data-look

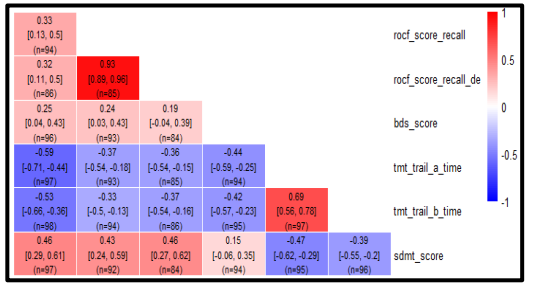
- *Data: **first 6 weeks** of the VS*
- ***n = 100** MS patients*
- *CoGames 1x/Week*

1. **Test-retest reliability**
   - Good-excellent correlations (range: r **= 0.75 to 0.9**)
2. **Correlation with reference tests**
   - 5/6 Moderate-strong correlation (range: r_s **= 0.4 to 0.66)**
3. **Adherence rates**
   - High adherence rates (range: **86.6% - 92%**)

Monitoring cognition in multiple sclerosis via adaptive smartphone games—first insights from a validation study
Pless, S. et al. *Front. Digit. Health* 2026

Test-retest reliability		
Game	Pearson Correlation [95% CI]	n
Match Maker	0.91 [0.87, 0.94]	98
Think Fast	0.9 [0.86, 0.93]	98
Numbers	0.79 [0.7, 0.86]	98
Treasure Hunt	rho: 0.41 [0.23, 0.58] rho: 0.75 [0.64, 0.83] rho: 0.71 [0.59, 0.81]	97
Mixer	0.87 [0.82, 0.91]	99
Catch a Cloud	rho: 0.86 [0.77, 0.92]	99





Adherence rates, n=100	
Game	Mean adherence
Match Maker	89.0%
Think Fast	89.2%
Numbers	89.2%
Treasure Hunt	86.6%
Mixer	89.6%
Catch a Cloud	92.0%

DreaMS Validation Study - *Data-look*

Step 1 — DreaMS Feasibility Study

Step 2 — Development CoGames 1.0

Step 3 — **dreaMS Validation Study**

DreaMS VS data-look

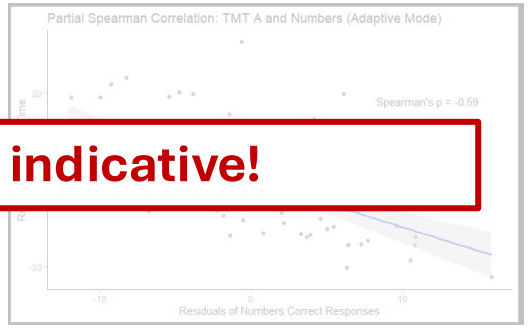
- Data: **first 6 weeks** of the VS
- **n = 100** MS patients
- CoGames 1x/Week

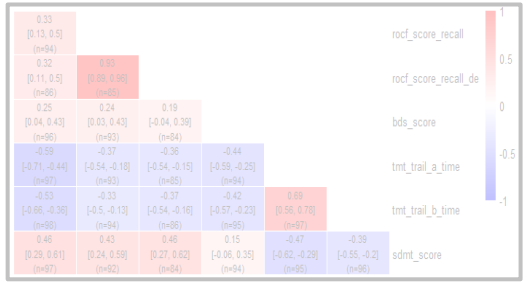
1. **Test-retest reliability**
   - Good-excellent correlations (range: r **= 0.75 to 0.9**)
2. **Correlation with reference tests**
   - 5/6 Moderate-strong correlation (range: r_s **= 0.4 to 0.66**)
3. **Adherence rates**
   - High adherence rates (range: **86.6% - 92%**)

Monitoring cognition in multiple sclerosis via adaptive smartphone games—first insights from a validation study
Pless, S. et al. *Front. Digit. Health* 2026

Data-look = Point estimates, only indicative!

Test-retest reliability		
Match Maker	0.91 [0.87, 0.94]	98
Think Fast	0.9 [0.86, 0.93]	98
Numbers	0.79 [0.7, 0.86]	98
Treasure Hunt	rho: 0.41 [0.23, 0.58] rho: 0.75 [0.64, 0.83] rho: 0.71 [0.59, 0.81]	97
Mixer	0.87 [0.82, 0.91]	99
Catch a Cloud	rho: 0.86 [0.77, 0.92]	99





Adherence rates, n=100	
Game	Mean adherence
Match Maker	89.0%
Think Fast	89.2%
Numbers	89.2%
Treasure Hunt	86.6%
Mixer	89.6%
Catch a Cloud	92.0%

Development – *Level Expansion*

Step 1 — DreaMS Feasibility Study

Step 2 — Development CoGames 1.0

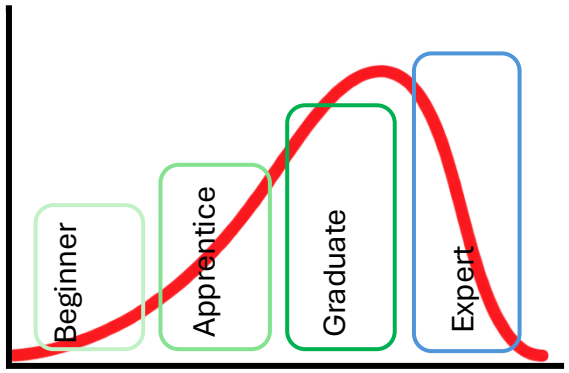
Step 3 — dreaMS Validation Study

Step 4 — **Development CoGames 2.0**

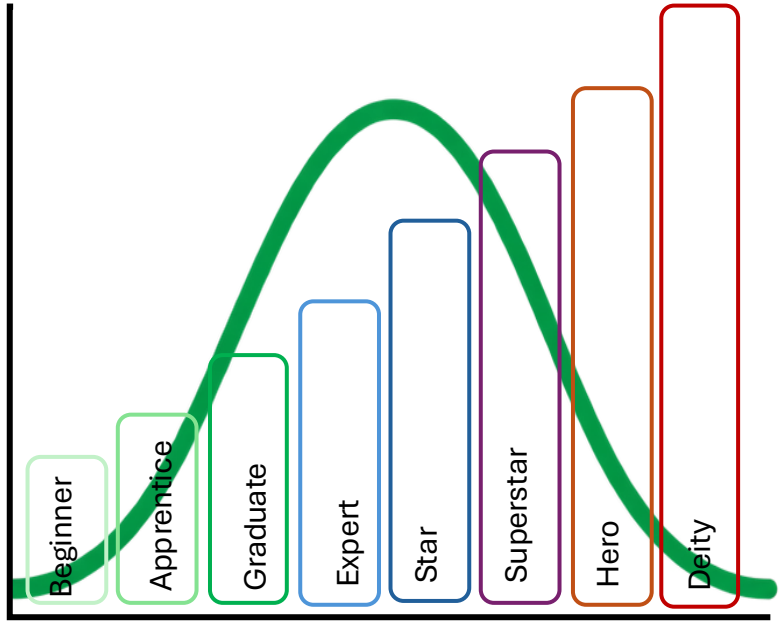
CoGames 2.0: additional levels

- More range = better distribution, more space for change
- Reduction of ceiling and floor effects (skewedness)

CoGames 1.0 (4 lvl)



CoGames 2.0 (8 lvl)



CoGames Study

Step 1 — DreaMS Feasibility Study

Step 2 — Development CoGames 1.0

Step 3 — dreaMS Validation Study

Step 4 — Development CoGames 2.0

Step 5 — **CoGames Study**

- *Study duration: 11 days, daily repetition*
- *n = 76 Healthy Volunteers*
- *CoGames 2.0 (8 expanded levels)*

1. **Test-retest reliability**
2. **Enjoyment ratings**
3. **Change over time (practice effects)**
4. **Differences between levels**

CoGames Study

Step 1 — DreaMS Feasibility Study

Step 2 — Development CoGames 1.0

Step 3 — dreaMS Validation Study

Step 4 — Development CoGames 2.0

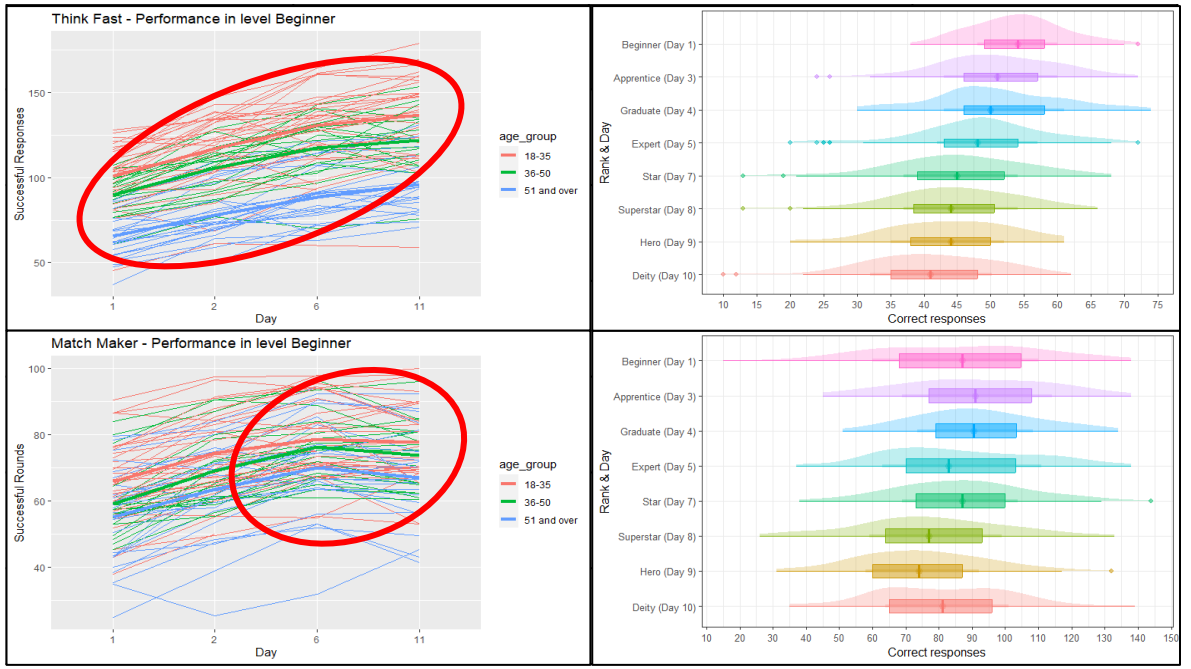
Step 5 — **CoGames Study**

Study Results

1. **Test-retest reliability**
   - 5/6 good-excellent (range: **r = 0.78 to 0.94**)
2. **Enjoyment ratings**
   - Good- very good (range: **3.13 to 4.28**)
3. **Change over time (practice effects)**
   - Observable increase in performance over time
4. **Differences between levels**
   - Quantifiable differences between level-difficulties

CoGames – Development of an adaptive smartphone-based and gamified monitoring tool for cognitive function in Multiple Sclerosis

Pless, S. et al. *J. Neurol*. 2025



Summary

CoGames features

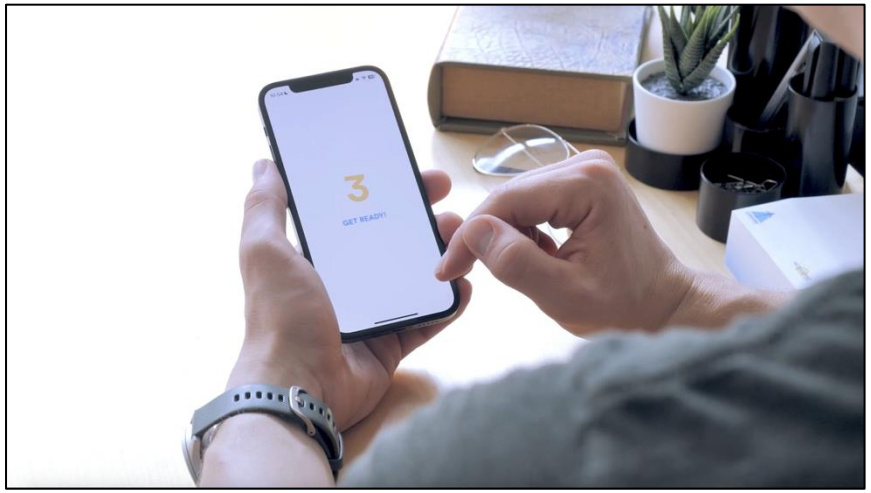
- Gamification
- Adaptiveness
- Remote self-administration
- Validated psychometrics (ongoing)

The road ahead...

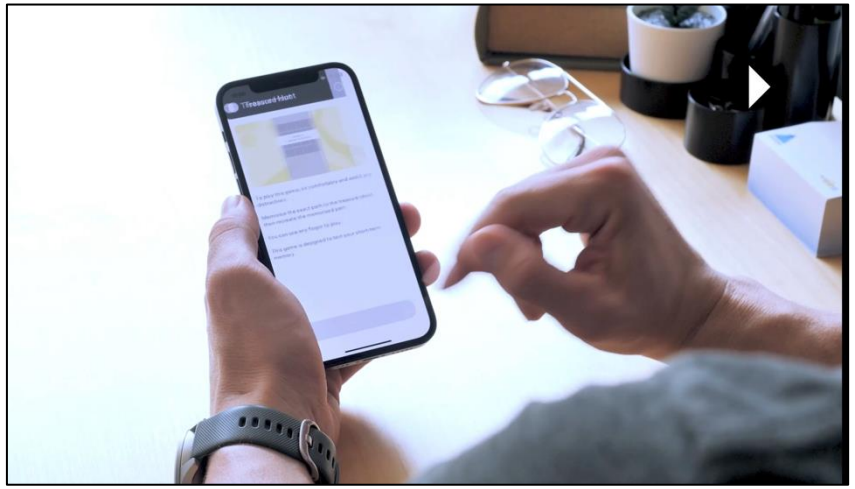
Next steps

- Longitudinal data (DreaMS VS)
- Managing practice effects
- CoGames for
  - Other diseases
  - Cognitive rehabilitation
- CoGames 3.0

Mixer



Treasure Hunt



Assessing Dynamic Difficulty Adjustment in Personalized Testing of Patients with Mild Cognitive Impairment:

The CoGames version 3

Emmanuel J. Bartholomé,

Indivi AG, Basel, Switzerland.
Basel Biometric Society, 27 August 2026, Basel, Switzerland.

Introduction

Problem statement

- The **responsiveness properties of clinical efficacy endpoints generally used to assess cognitive decline in early phases of drug development trials remain loosely characterized**. Few studies have properly assessed signal-to-noise ratio of the benchmark tools used in phase 2 PoC studies.

- A recent study of the responsiveness of **CDR-SB and ADAS-Cog-13 demonstrated at best a low to moderate sensitivity to change** within a 52-week interval[1].

- **A challenge of longitudinal quantification of cognitive decline in early development trials of neurodegenerative diseases such as AD/PD is the use of 'one-size-fits-all' instruments where the task difficulty is fixed and identical for all**. This measurement paradigm has inherent limitations that may hinder the detection of subtle changes in cognitive function:
  - **Practice effects** due to learning ability.
  - **Ceiling effects:** In higher-functioning individuals or those in early disease stages, a fixed difficulty of a cognitive task may be too simple for a sizeable proportion of study participants, causing scores to cluster at the maximum possible value. This "ceiling" effect prevents the detection of initial cognitive decline or improvements resulting from an intervention.
  - **Floor effects:** Conversely, if the fixed difficulty of a task is too challenging for impaired individuals, they may consistently score at the minimum. This "floor" effect makes it impossible to track further deterioration or identify small but clinically relevant therapeutic responses.

- To address these limitations, **CoGames** incorporates a **Dynamic Difficulty Adjustment (DDA)** system allowing to **adapt the task difficulty to individual performance** during a run-in phase (between screening and baseline visits) thus enabling prior to the onset of the trial to potentially saturate practice effects and alleviate floor/ceiling effects.

[1]McDougall F et al. J Prev Alzheimer's Dis. 2021, 8: 151-160.

AD, Alzheimer's disease; PD, Parkinson's disease; PoC, proof-of-concept; CDR-SB, Clinical Dementia Rating–Sum of Boxes; ADAS-Cog-13, Alzheimer's Disease Assessment Scale – Cognition Subscale 13 item version.

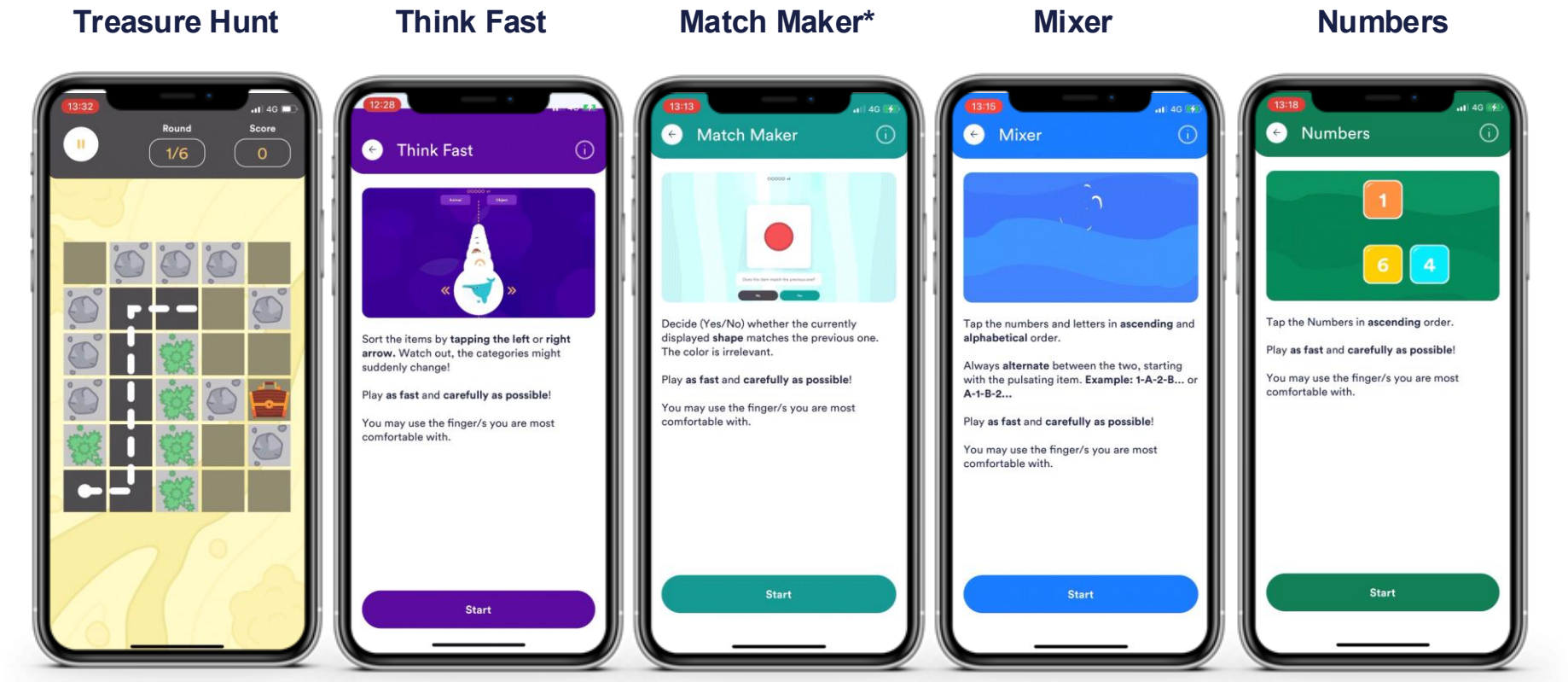
CoGames adaptation for neurodegenerative disorders

From CoGames v2.2 (Designed and for Multiple Sclerosis) to CoGames v3 (Adapted for MCI and other Neuro-Degenerative Disorders)

- **General principle:**
  - **excluding confounding factors** e.g. between numbers and letters (1 vs I),
  - **preventing mistakes not due to cognitive impairment** e.g. difficulty level based on font size,
  - **improving concept of interest focus** e.g. splitting Match Maker in 2 tests:
    - **Match One** (N-Back 1) and
    - **Match More** (N-Back 2 or 3),
  - improving **difficulty progression linearity**,
  - **aligning gamified scoring across tests**,
  - **removing in-level difficulty adaptability,** thus presenting systematically and consecutively all stages per level until the subject reaches 60 seconds,
  - Including the ranking logic inside the parent Indivi Go App, allowing **offline, daily rank adaptivity**

CoGames Proof-of-Concept User-Adaptive Cognitive Assessments

Likable, gamified and dynamically adaptive instruments for personalized measurement of longitudinal change in cognitive function



Visuo-spatial Memory
Executive Functions

Information Processing Speed
Semantic & Visual Recognition

Working Memory
Cognitive Flexibility (N-back)

Cognitive Flexibility
Divided Attention

Attention
Information Processing Speed

Pless S et al. J Neurol. 2025, 272: 119.
Rodríguez-Romero A. Front. Neurosci. 2025, 19:1653261.
Pless S et al. Front Digit Health. 2026, 8:1627226.
Zohaib M. Advances in Human-Computer Interaction. 2018, 1 5681652.

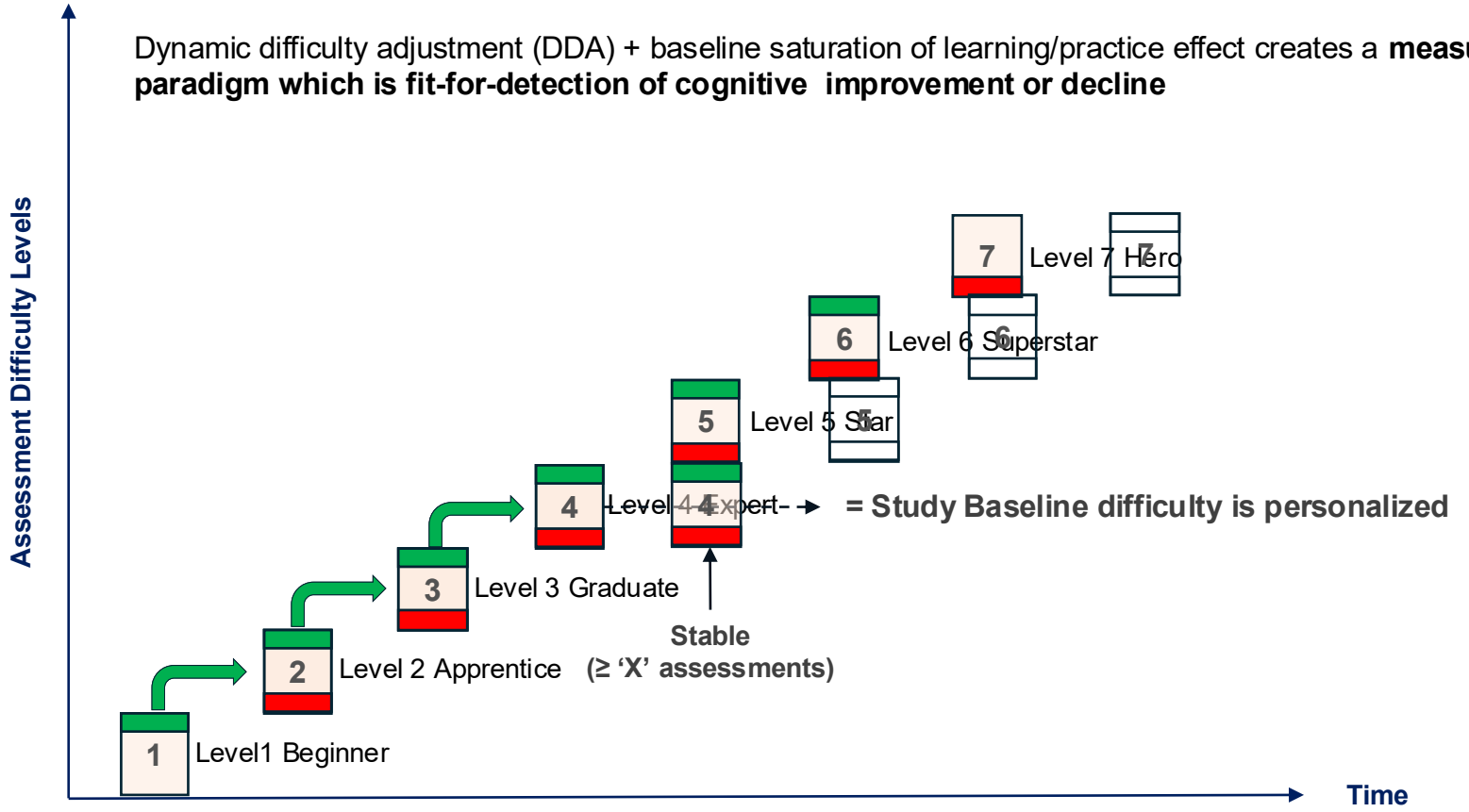
*Corresponds to two independent tests: **Match One** is a 1-back task (assessing working memory) and **Match More** is a 2/3-back task (assessing cognitive flexibility).

User-Adaptive Cognitive Assessments – What Does that Mean ?

Enabling to adapt the difficulty of the task to the subject's capacity and to saturate the learning/practice effect prior to study start

CoGames

Dynamic difficulty adjustment (DDA) + baseline saturation of learning/practice effect creates a **measurement paradigm which is fit-for-detection of cognitive improvement or decline**



p > XX% percentile performance in the average of last 'x' assessments triggers increase (+1) in Rank of Difficulty

p < YY% percentile performance in the average of last 'x' assessments triggers decrease (-1) in Rank of Difficulty

Range of score in a level of difficulty

Pless S et al. J Neurol. 2025, 272: 119.
Rodríguez-Romero A. Front. Neurosci. 2025, 19:1653261.
Pless S et al. Front Digit Health. 2026, 8:1627226.
Zohaib M. Advances in Human-Computer Interaction. 2018, 1 5681652.

Interim Analysis of the ADAPT-MCI-01 Study: Assessing Dynamic Difficulty Adjustment in Personalized Testing of MCI Patients

Antonio Rodríguez-Romero,[1] Alexa Burk,[2] Emmanuel J. Bartholomé,[1] Claudia Mazzà,[1] Carlos Luque,[1] Óscar Reyes,[1] Corrado Bernasconi,[1] Shibeshih Belachew,[1] Rhoda Au[2]

[1]Indivi AG, Basel, Switzerland; [2]Boston University, Boston, MA, USA.

Poster at the AAIC Conference, 12-15 July 2026, London, UK.

Conflict of Interest

- Antonio Rodríguez-Romero, Claudia Mazzà, Carlos Luque, Óscar Reyes, and Shibeshih Belachew, are employees of **Indivi AG**
- Corrado Bernasconi and Emmanuel J. Bartholomé are contractors of **Indivi AG**
- Rhoda Au is a scientific advisor to **Signant Health** and **NovoNordisk**

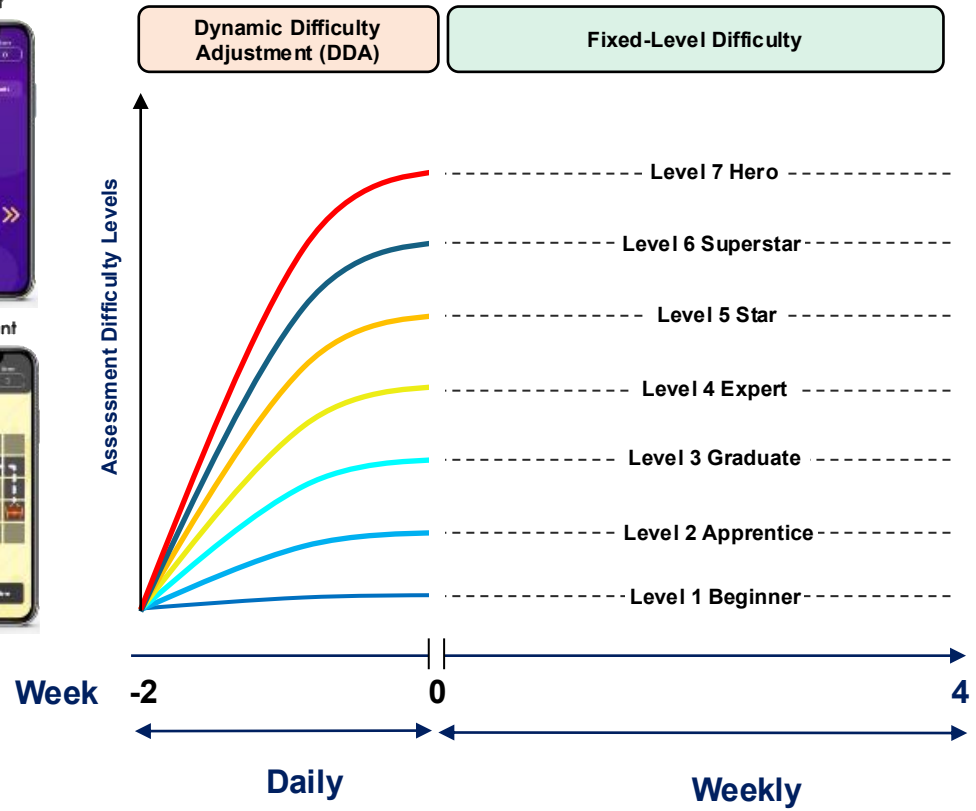
Acknowledgements

- This work is supported by Indivi AG and the NIH (Grant numbers: P30AG072978, R01-AG083735, U01NS086659, and U54NS115266).
- We acknowledge RC2NB and its collaborators for contributions to prior development and validation work on CoGames within the dreaMS program..

Assessing **D**ynamic Difficulty **A**djustment in **P**ersonalized **T**esting of Subjects with **M**ild **C**ognitive **I**mpairment – The ADAPT-MCI-01 Study

[NCT07174869] - In collaboration with **Prof. Rhoda Au** Boston University School of Medicine, **The Precision Brain Health Initiative**

CoGames v3



The number of levels may differ for each assessment.
HV: healthy volunteers; MCI: mild cognitive impairment, defined as a global clinical dementia rating (global CDR) of 0.5

Main Study Objectives:

In Persons with MCI (**PwMCI**) and Healthy Volunteers (**HV**):

- To investigate **adherence** to the schedule, **likability** and **usability** of the **CoGames** assessments
- To evaluate the **feasibility and performance of the DDA** system
- To establish the **construct validity** of each **CoGames** assessment

ADAPT-MCI-01 Study – Baseline Characteristics

Interim analysis

By 21 May 2026, 23 participants had completed the study; 2 subjects were lost to follow-up and excluded from the analysis. Baseline characteristics were well balanced between HV and PwMCI.

	Elderly HV (N=11)	MCI (N=12)	Overall (N=23)
Sex			
Female	7 (63,6%)	6 (50.0%)	13 (56.5%)
Male	4 (36.4%)	6 (50.0%)	10 (43.5%)
Age			
Median [Min, Max]	67.0 [51.0, 78.0]	69.5 [53.0, 85.0]	69.0 [51.0, 85.0]
Years of education			
Median [Min, Max]	16.0 [14.0, 18.0]	16.0 [13.0, 20.0]	16.0 [13.0, 20.0]

	Elderly HV (N=8)	MCI (N=8)	Overall (N=16)
Global CDR			
Median [Min, Max]	0 [0, 0]	0.50 [0.50, 0.50]	0.50 [0.00, 0.50]
CDR Sum of Boxes			
Median [Min, Max]	0 [0, 0]	0.50 [0.50, 1.00]	0.50 [0.00, 1.00]
CDR Memory Box			
Median [Min, Max]	0 [0, 0]	0.50 [0.50, 1.00]	0.50 [0.00, 1.00]
MoCA Total Score			
Median [Min, Max] Missing, n	29.5 [26.0, 30.0] 3 (27.3%)	26.0 [21.0, 30.0] 3 (25.0%)	28.0 [21.0, 30.0] 6 (26.1%)

HV: healthy volunteer; MCI: mild cognitive impairment; CDR: clinical dementia rating scale; MoCA: Montreal Cognitive Assessment.

ADAPT-MCI-01 Study – Adherence & Usability

CoGames v3 adherence remained above 75 % for all instruments in this pre-baseline run-in phase of daily assessment from Day -14 to end of the DDA (Day 0)

Adherence to Schedule of Assessment

Across all CoGames assessments, participants completed **87% and 78%** of scheduled assessments during the adaptive phase (daily) and the fixed-difficulty phase (weekly), respectively*.

Adherence was consistent among the assessments and groups (HV versus PwMCI).

Usability & Likeability at week 1 and/or week 4

	Elderly HV and PwMCI
System Usability Scale – Range [0-100]	
Median [Min, Max]	75.0 [67.5, 91.3]
Likeability score – Range [0-5]	
Median [Min, Max]	4.50 [3.00, 5.00]

HV: healthy volunteer; MCI: mild cognitive impairment; DDA: dynamic difficulty adjustment.
*Mean Adherence is defined as the percentage of scheduled tests completed by the subject.
System Usability Scale (SUS) from Brooke, J. (1996). SUS: A "quick and dirty" usability scale. In *Usability evaluation in industry*. CRC press, 1996.

ADAPT-MCI-01 Study – Dynamic Difficulty Adjustment & Measurement Ceiling

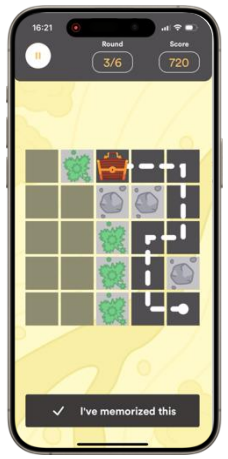
The performance-driven dynamic increase in assessment difficulty enables to remove the ceiling effect observed at lowest level of difficulty

During the 14-day adaptive phase, participants were able to increase the difficulty across a wide range of levels for all CoGames v3 assessments. On average, final difficulty levels were reached after **9 days** of daily assessments (range: 7–11 days)

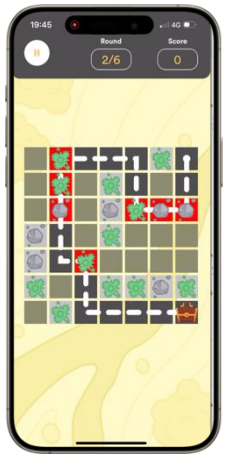
Treasure Hunt

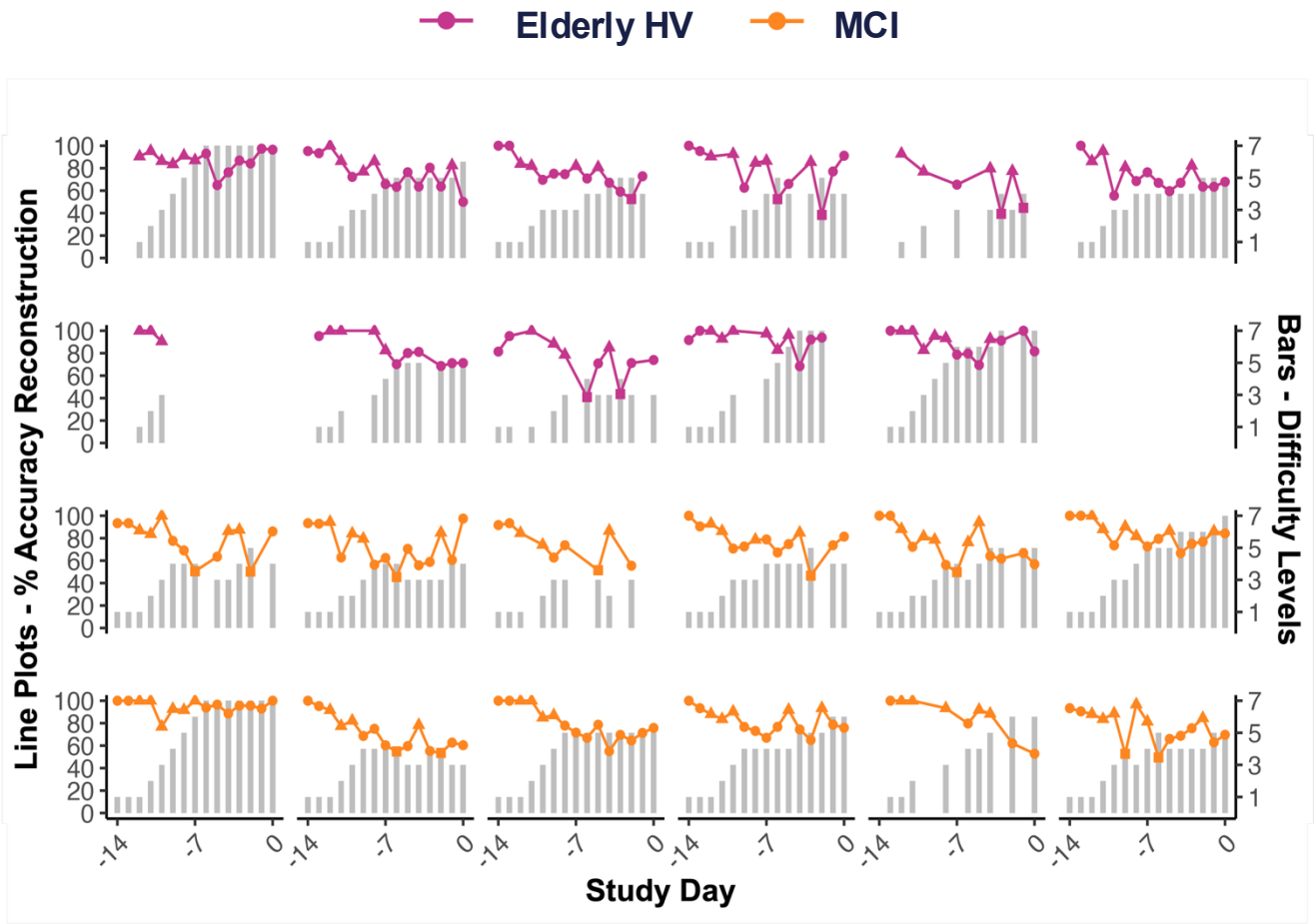
Visuo-spatial Memory & Executive Functions

Level 2
25-box
6-8 step length
2 – 2.8 s intermission



Level 7
49-box size
18 step length
8.8-9.6 s intermission

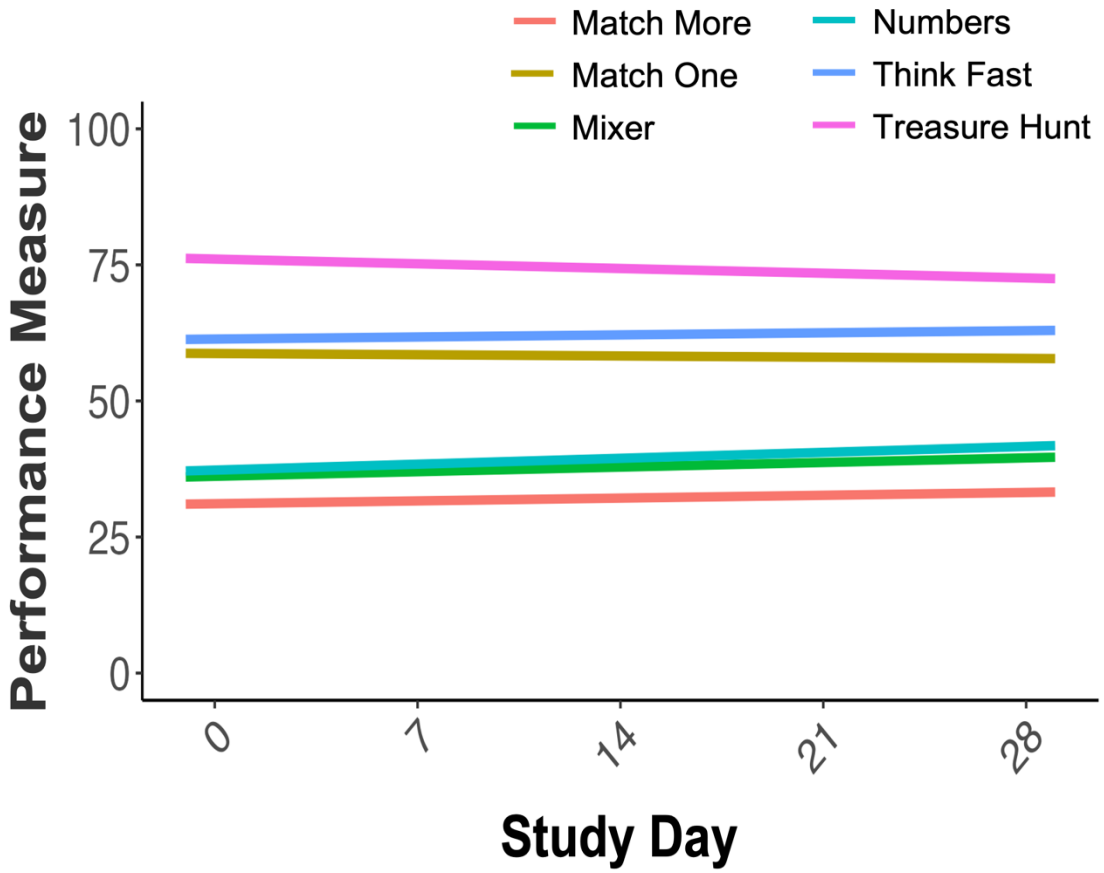




HV: healthy volunteer; MCI: mild cognitive impairment.

ADAPT-MCI-01 Study – Cognitive Performance Stability at Plateau of Learning Ability

For all assessments, performance measures remained stable in the fixed-difficulty phase



In the fixed-difficulty phase, the absolute median performance fluctuations around the mean of both groups over four weeks were lower than **10%** for the performance measures used for DDA. Linear random coefficients models (one for each assessment) did not identify any trends over time.

HV: healthy volunteer; MCI: mild cognitive impairment; DDA: dynamic difficulty adjustment.

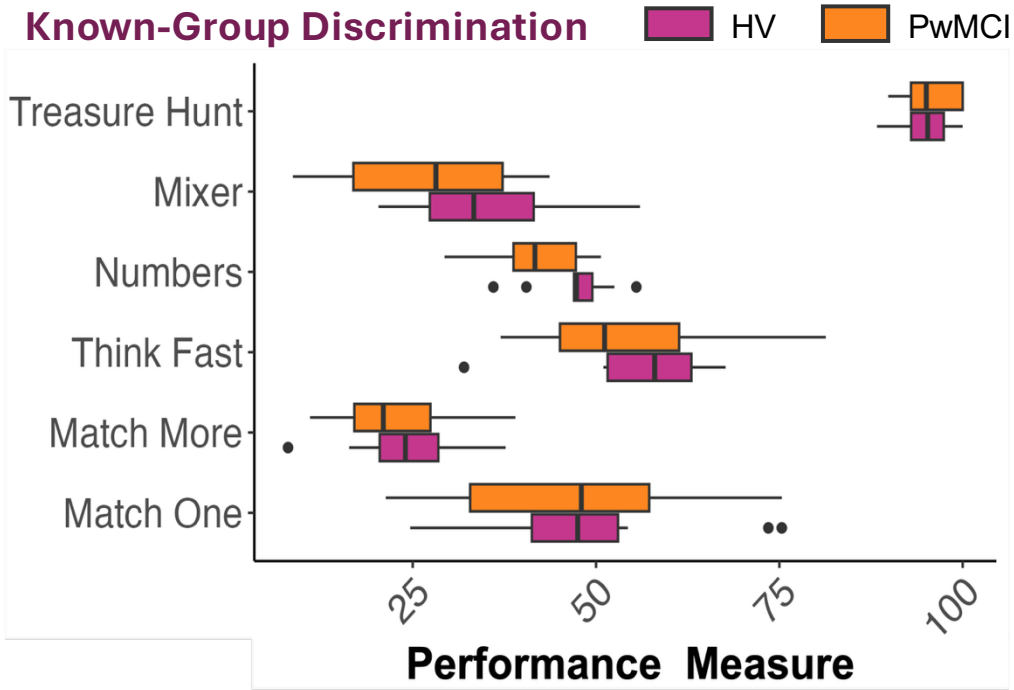
ADAPT-MCI-01 Study – Construct Validity

Performance measures showed known-group discrimination and convergent validity.

Performance measures during the first three days (all subjects at the lowest level of difficulty for each assessment) across the CoGames battery **were higher in the HV group compared to the PwMCI group** (Figure).

Convergent validity of CoGames v3 assessments was evaluated by the Spearman correlation of mean three first day performance measure with MoCA total score at last in-clinic visit (Table).

Known-Group Discrimination



Van Elteren's test stratified by assessment (**p =0.03**)

Convergent Validity vs MoCA Total Score

Assessment	Mixer	Numbers	Think Fast	Treasure Hunt	Match More	Match One
Spearman's rho	0.587*	0.558*	0.472*	0.411	0.321	0.309

HV: healthy volunteer; MCI: mild cognitive impairment.

ADAPT-MCI-01 Study – Summary

Interim analysis

- CoGames v3 user-adaptive battery and pre-baseline **performance-driven dynamic difficulty adjustment (DDA) of cognitive assessments appear to be feasible** in elderly HV and subjects with mild cognitive impairment.
- The user experience of MCI and HV subjects completing 6 weeks of high frequency assessments with CoGames v3 battery was associated with high scores of **usability and likeability**.
- CoGames v3 DDA system enabled a **distribution of MCI and elderly HV subjects at various steady-state levels of difficulty** for all assessments at the end of a 2-week pre-baseline run-in phase of the trial.
- A **plateau of learning ability** and difficulty increase was attained after about **7 to 11 days** of daily assessments.
- After 14 days of daily difficulty adaptive executions of the CoGames v3 tests, **performance measures remained stable** in the fixed-difficulty phase
- CoGames v3 performance measures showed preliminary evidence of **construct validity** in participants with mild cognitive impairment (MCI).

HV: healthy volunteer; MCI: mild cognitive impairment.

Thank you

Appendix Materials

AD/PD™ 2026
ADVANCES IN SCIENCE & THERAPY

Assessing Dynamic Difficulty Adjustment in Personalized Testing of Patients with Mild Cognitive Impairment – The ADAPT-MCI-01 Study

Antonio Rodríguez-Romero,[1] Alexa Burk,[2] Emmanuel J. Bartholomé,[1] Claudia Mazzà,[1] Carlos Luque,[1] Óscar Reyes,[1] Corrado Bernasconi,[1] Shibeshih Belachew,[1] Rhoda Au[2] . *[1]Indivi AG, Basel, Switzerland; [2]Boston University, Boston, MA, USA.*

AD/PD Conference Oral Presentation, Tuesday 17th March 2026, 8:30-10:30am, Copenhagen, Denmark.

Presentation Download

BU

ADAPT-MCI-01 Study – Inclusion and Exclusion Criteria

Group	Age Range	Key Inclusion Criteria
MCI/ mild AD	50 – 90 years old	• Part of BU-ADRC cohort • Diagnosed with MCI or mild AD dementia, as per BU-ADRC database label • Global CDR is 0.5 or 1.0 • CDR Memory Box score is 0.5 or 1.0 • MoCA score ≥24 or MMSE ≥22. • Able to use a smartphone.
Elderly HV	50 – 90 years old	• Part of BU-ADRC cohort • Labelled Cognitively Normal as per BU-ADRC database • Able to use a smartphone

Group	Exclusion Criteria
MCI and Elderly HV	• Any neurological condition that may be contributing to cognitive impairment (above and beyond that caused by the subject's MCI or AD). • A psychiatric diagnosis or symptoms, (e.g., hallucinations, major depression, or delusions) that could interfere with study procedures. • A history of stroke or seizures within the last 24 months. • A Geriatric Depression Scale (GDS) score ≥8. • A severe visual or hearing impairment that would interfere with capacity to perform cognitive assessments.

AD: Alzheimer's Disease; BU-ADRC: Boston University Alzheimer's Disease Research Center; CDR: Clinical Dementia Rating; HV: healthy volunteers; MCI: Mild Cognitive Disorder; MMSE: Mini-Mental State Exam; MoCA: Montreal Cognitive Assessment.

BU