WHITEPAPER · NOVEMBER 2026

NEXT CHAPTER

Game-Based Assessment in the AI Era

What the evidence says about its benefits for external and internal employee selection. A systematic literature review for HR leaders.



Junialdi Dwijaputra
Head of Level Up, powered by Agate

Level Up
powered by Agaté

EXECUTIVE SUMMARY

Where game-based assessment actually pays off

We reviewed 35 sources published up to November 2026, plus 34 articles synthesized in an earlier review, to show HR leaders where game-based assessment (GBA) pays off: in external hiring, in internal selection, and in an era when AI can sit a traditional online test for a candidate.

35+34
sources reviewed directly, plus articles synthesized through a prior review

r ≈ .50
convergence between cognitive GBAs and conventional ability tests

58
Net Promoter Score from 4,778 real applicants for one cognitive GBA

95th
percentile GPT-4 reached on a text-based verbal ability test

Six findings at a glance

Proven

Employer attractiveness is the most consistent win

Enjoying the game makes the employer look innovative and competent.

Georgiou & Lievens (2022)

Comparable

Validity is on par, not superior

Cognitive GBAs track ability tests (r ≈ .50) and predict performance.

Landers et al. (2022); Leutner et al. (2023)

Partly proven

Less faking, not zero faking

Distortion on a game-based honesty measure was significantly smaller.

Ohlms, Melchers, Kanning, & Barends (2025)

Monitor

Fair, if monitored

No more adverse impact than tests, but gaming habits can bias scores.

Ohlms, Hohner, & Melchers (2025)

Evidence gap

Internal selection is untested

No study has tested GBA in real promotion decisions yet.

This review

Plausible

AI raises the stakes

LLMs ace text tests; interactive, process-scored games are harder to outsource.

Hickman et al. (2024)

BOTTOM LINE FOR HR

Use GBA as one part of a multi-method process: validate it locally, monitor it for fairness, and stress-test it against AI before go-live.

ABOUT THIS WHITEPAPER

Contents

01	Why game-based assessment, and why now	4
02	How we reviewed the evidence	7
03	Benefits for external selection	10
04	Can GBA resist AI-assisted cheating?	14
05	Benefits for internal selection	15
06	Myths, risks and contradictions	17
07	A practical playbook for HR	19
08	Limitations, research agenda and conclusion	21
09	How Next Chapter puts the evidence to work	22
	References	23
	Appendix: studies included in the review	25

Who this is for

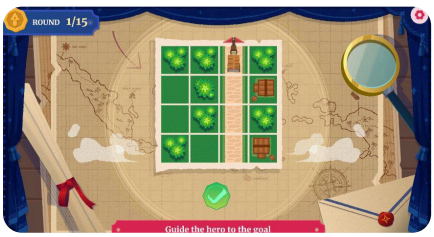
Talent acquisition teams, HR business partners and assessment specialists who run external hiring and internal mobility.

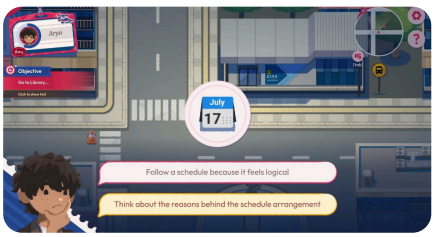
About the author

Junialdi Dwijaputra is Head of Level Up, powered by Agate.

About Level Up

A PT Agate International business line building gamification for marketing, learning, training and assessment, including **Next Chapter**.







Every claim is traceable to a source in the references. Where evidence is thin or indirect, we say so. Screens: Next Chapter. Learn more at agatelevelup.com/game-based-assessment.

01 · CONTEXT

Why game-based assessment, and why now

GBA is already widely used in hiring, but the science only caught up after 2022.

BEFORE 2022

Practice ran ahead of science

Vendors spread GBA faster than independent research could test it.

2022

First selection-focused review

GBA can be used, but its advantages are "fewer than imagined".

Ramos-Villagrasa et al. (2022)

2023–2026

A new wave of evidence

Real-applicant field data, faking experiments, fairness studies and a meta-analysis.

25 of 28 new studies in this review

From test to game: four kinds of game-related assessment

CLOSER TO A TRADITIONAL TEST — CLOSER TO A GAME PLAYED FOR FUN

Gamified assessment

An existing test with game elements added: narrative, points, progress bars.

Gamified card-sorting test (Hommel et al., 2022)

Gamefully designed

A new assessment designed with game elements from the start.

Owiwi SJT with characters and story (Georgiou et al., 2019)

Game-based assessment

A full game built to measure job-related constructs.

Virus Slayer, US Air Force cyber roles (Wiernik et al., 2022)

Playful game

A commercial game repurposed to collect assessment data.

VR *Job Simulator* and intelligence (Simons et al., 2023)

Typology: Landers & Sanchez (2022); Ramos-Villagrasa et al. (2022). In this whitepaper "GBA" covers all four.

Two ways to build the scoring

Theory-driven

The game is designed around a psychological construct, such as general cognitive ability, and scored against it.

Landers et al. (2022)

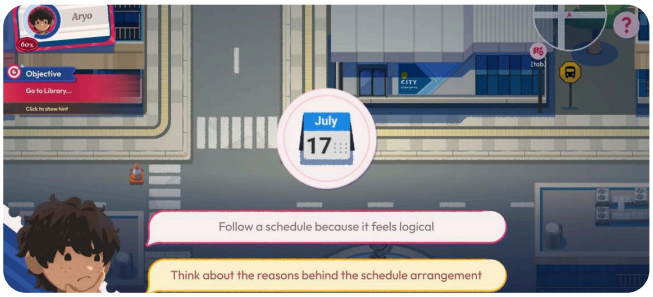
Data-driven

Play behaviour is linked directly to job outcomes, often with machine learning on trace data.

Auer et al. (2022)



Cognitive module: a route-planning puzzle.



Personality module: a narrative decision scene.

External and internal selection are different problems

External selection

Choosing among applicants from outside

- **Attract talent** in a competitive market
- **Screen at volume** while keeping the candidate experience good
- **Limit faking** and outside help on unproctored tests

VS

Internal selection

Promotion, rotation, placement and talent programmes

- **Unsuccessful candidates stay**, so their reactions affect morale and retention
- **Fairness drives advocacy**: in police promotion exams, fairness predicted recommending the exam
- **Internal movers** initially outperform external hires and leave less often

Ford et al. (2009); McCarthy et al. (2009), N = 498 and 182; Bidwell (2011)

The research gap: games for training, not for selection

Research focus of 41 HR gamification studies, 2020–2025



↑ Recruitment

85.4%
of studies targeted existing employees

7.3%
addressed recruitment specifically

0
reviews separating GBA's benefits for external and internal selection, until this one

Prianto & Sulaksono (2025)

Questions this whitepaper answers

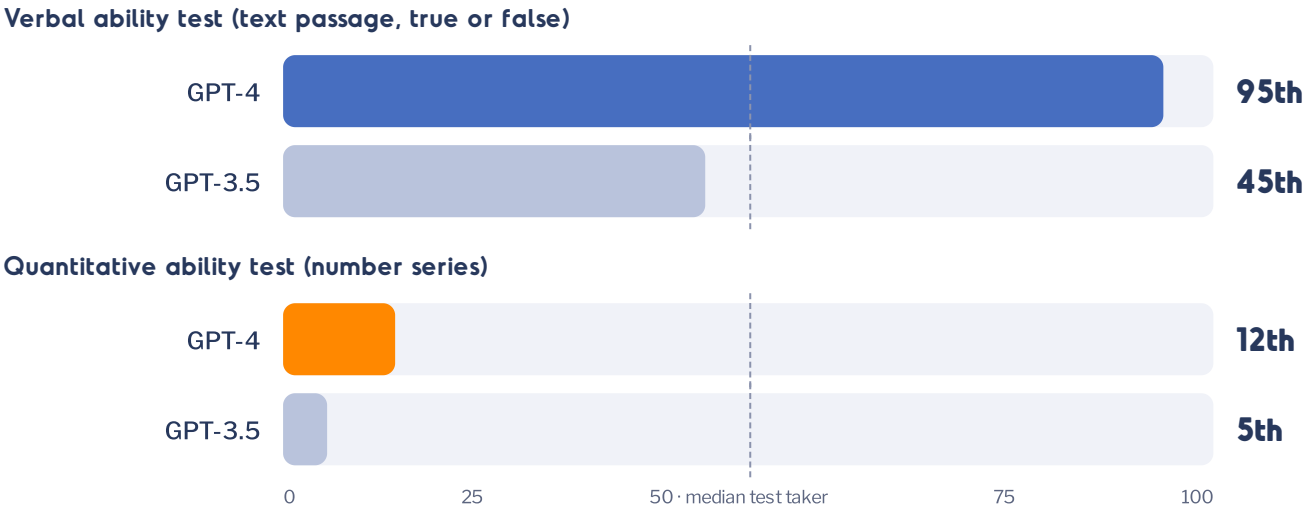
- What does GBA offer **external** selection?
- What does it offer **internal** selection?
- Which **conditions** strengthen or weaken it?
- How does it hold up against **AI**?
- What should **HR do** about it?

HOW TO READ THIS WHITEPAPER

Chapters 3 and 4 cover external hiring, Chapter 5 covers internal selection, and Chapter 7 turns both into an adoption playbook.

AI can now sit a text-based test for the candidate

Average percentile reached by large language models on two ability tests used in real high-stakes selection



Hickman, Dunlop, & Wolf (2024), Table 1: averages across prompts and temperatures; quantitative results are for the open-ended version.

What makes an assessment harder to outsource to AI

MOST EXPOSED

Untimed, text-based items

Copy, paste, answer. Ghost-writer-level help is now available to everyone.

Sindre (2023)

LESS EXPOSED

Timed, non-text items

Images resist pasting; time limits restrict outside help.

Hickman et al. (2024)

LEAST EXPOSED (PLAUSIBLY)

Interactive, process-scored games

Real-time play, varied scenarios and scoring from the play process itself.

Auer et al. (2022); Torkestani & Mansouri (2025)

74–92%

of students have tried generative AI for academic work.

Torkestani & Mansouri (2025)

Design for visibility

Process visibility, personalization, multimodal and real-time interaction resist AI.

Torkestani & Mansouri (2025)

AI inside the game

Machine-learning scoring can be optimized for validity and fairness, and must be governed.

Leutner et al. (2023)

WHAT IT MEANS FOR HR

If your screening relies on unproctored, text-based tests, assume every candidate has access to AI help. Tests once considered hard to fake may now be weaker signals.

Implication drawn from Hickman et al. (2024).

02 · METHOD

How we reviewed the evidence

A systematic review following PRISMA 2020 (Page et al., 2021), built from four search routes.

ROUTE 1
7 core documents
Prior reviews and studies supplied at the start

+

ROUTE 2
Database search
OpenAlex, three Boolean queries, 24 Nov 2026, no date limit

+

ROUTE 3
Snowballing
264 citations in the core documents checked against Crossref

+

ROUTE 4
Targeted search
LLMs and selection-test integrity

=

RESULT
35 sources
plus 34 articles via the 2022 review

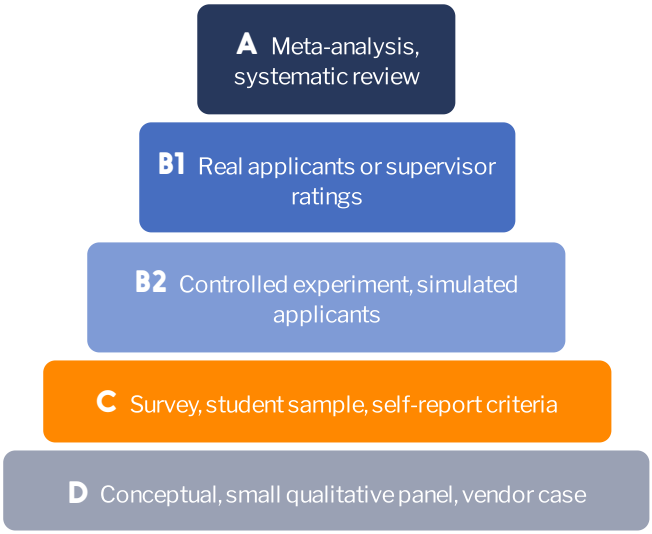
Included

- Any of the four GBA types
- Selection, recruitment, promotion, placement or talent identification, incl. simulated applicants and employee samples
- At least one benefit outcome: validity, reliability, reactions, attractiveness, faking, fairness, efficiency, learning
- Peer-reviewed articles, scholarly chapters, full proceedings, reviews, meta-analyses

Excluded

- Education-only, clinical, health or marketing contexts
- Dissertations, preprints, conference abstracts, vendor reports
- Records without abstract or full text

Grading the evidence



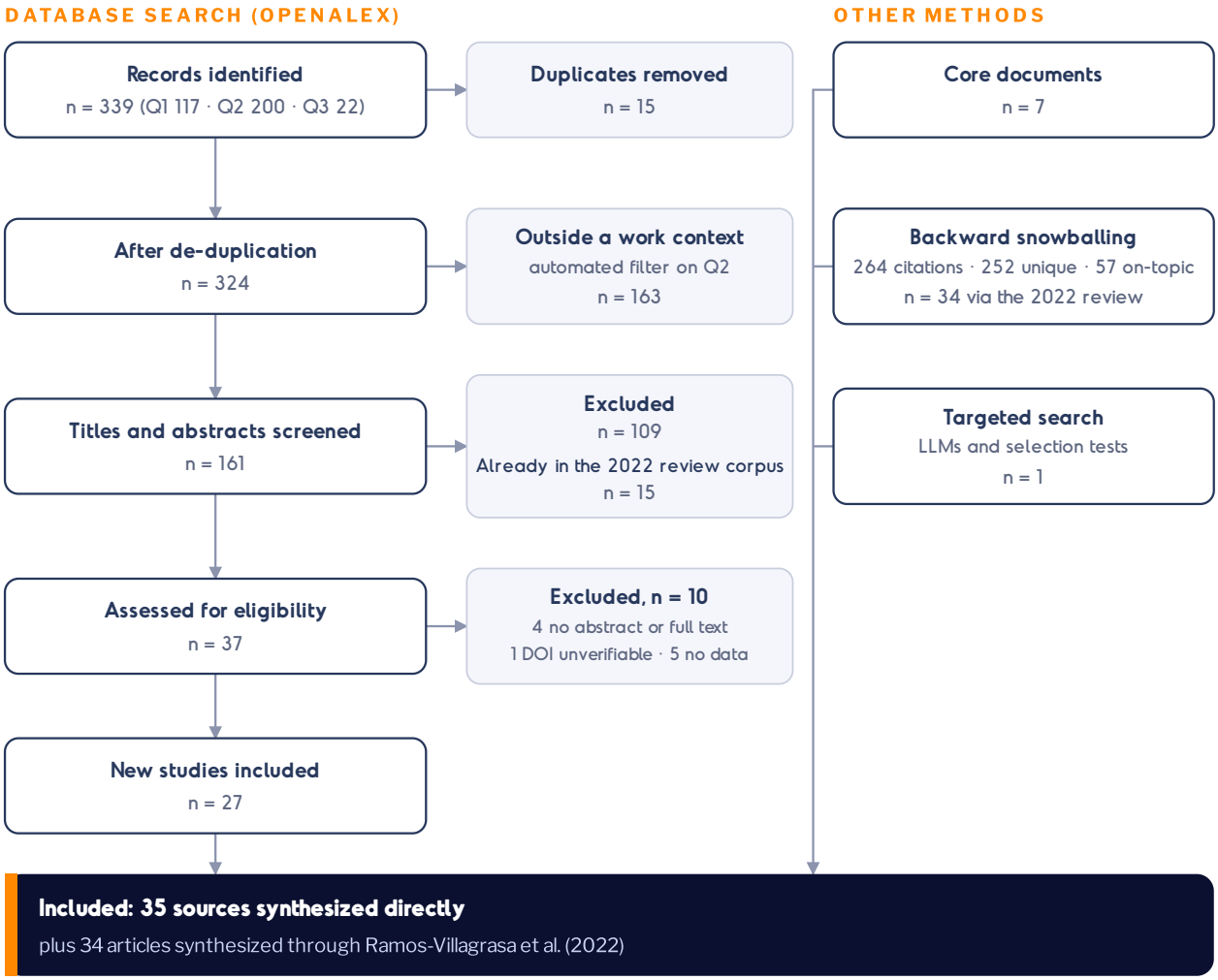
Findings were synthesized narratively by benefit, recording direction and evidence level. No new meta-analysis: designs and outcomes are too heterogeneous. Single reviewer, no independent double coding.

Search strings

```
Q1 selection  ("game-based assessment" OR "game-based assessments" OR "gamified assessment" OR "gamified assessments" OR "game-related assessment"
              OR "gamified selection" OR "gamified test") AND ("personnel selection" OR "employee selection" OR recruitment OR hiring OR applicants
              OR "applicant reactions")
Q2 internal   ("game-based assessment" OR "game-based assessments" OR "gamified assessment" OR "serious game" OR "serious games") AND (promotion
              OR "internal selection" OR "talent identification" OR "high potential" OR "internal mobility" OR "talent management" OR "assessment center")
Q3 serious    ("serious game" OR "serious games" OR "video game" OR "simulation game") AND ("personnel selection" OR "employee selection"
              OR "job applicants" OR "applicant reactions")
```

Contextual evidence: Capatina et al. (2024), Chen et al. (2025) and Prianto & Sulaksono (2025) inform internal uses; Sindre (2023) and Torkestani & Mansouri (2025) frame the AI threat and are not evidence about GBA itself.

What the search returned



PRISMA 2020 flow. Search run on 24 November 2026; Q2 retrieved the 200 top-ranked of 385 records.

The corpus at a glance

25/28
new studies published 2023–2026: a young, fast-moving field

6
studies used real applicants or supervisor ratings (level B1)

0
studies tested GBA in an actual promotion decision

9
studies from one German research group: replication needed

PART TWO

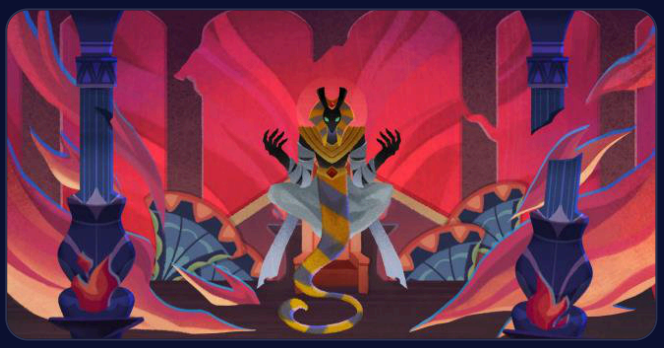
What the evidence says

Thirty-five sources, eight benefit dimensions, two selection contexts, and one new threat from generative AI.









> "GBA's most consistent benefit is how it makes the organization look to candidates. Its validity is comparable to conventional tests, not superior."

03 · EXTERNAL SELECTION

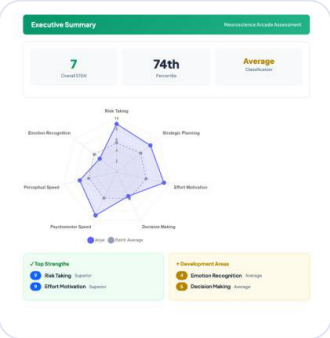
Benefits for external selection

Eight benefits, scored by the direction of the evidence and the strength of the best study behind it.

BENEFIT	EVIDENCE	VERDICT	KEY RESULT
Employer attractiveness	●●●●○	Supported	Enjoyment signals an innovative, competent employer Georgiou & Lievens (2022); Gkorezis et al. (2021)
Candidate experience	●●●●○	Design-dependent	NPS 58 from 4,778 applicants, but some GBAs rate worse than paper tests Leutner et al. (2023); Ohlms et al. (2024a, 2025a)
Validity	●●●●●	Comparable	r ≈ .50 with ability tests; predicts performance; not superior Fadillah et al. (2025); Landers et al. (2022)
Resistance to faking	●●●○○	Partly	Faking significantly smaller, not eliminated Ohlms, Melchers, Kanning, & Barends (2025)
Resistance to AI cheating	●○○○○	Untested	Plausible from indirect evidence and practice Hickman et al. (2024); Torkestani & Mansouri (2025)
Fairness and inclusion	●●●●○	No worse	Parity across 3,107 applicants; gaming-experience bias Leutner et al. (2023); Ohlms, Hohner, & Melchers (2025)
Realistic job preview	●●●○○	Supported	Job-related storylines clarify what the work involves Ohlms, Voigtländer, et al. (2024); Ohlms et al. (2025)
Efficiency and data richness	●●○○○	Plausible	Trace data predict ability; shorter sessions; costlier to build Auer et al. (2022); Leutner et al. (2021)

Evidence strength of the best study:

A	B1	B2	C	D
●●●●●	●●●●○	●●●○○	●●○○○	●○○○○



From play to profile

Benefits only count if decision-makers can read the results. A Next Chapter report turns play data into an executive summary, job-fit percentage, competency and sub-competency profile, and a final recommendation.

3.1 How a game makes an employer more attractive

STEP 1
Candidate plays
Enjoyment and flow during the assessment

STEP 2
Organization sends a signal
"This employer is innovative and competent"

STEP 3
Attractiveness rises
Higher organizational attractiveness

STEP 4
Candidates advocate
Stronger intentions to recommend the employer

Georgiou & Lievens (2022); Gkorezis et al. (2021)

Moderator: gaming experience

A gamified SJT raised attractiveness only for candidates with video-game experience.

Gkorezis et al. (2021)

Moderator: framing

Simply calling a test a "game" made the organization look more innovative.

McChesney et al. (2022)

72% vs 28%
AHP preference weight for employers using GBA vs traditional tests (35 students). A relative weight, not a "72% advantage".

Gürbüz et al. (2022)

158
recruiters named organizational image and applicant reactions as top reasons to adopt GBA.

Ohlms, Melchers, & Kanning (2024b)

179
test takers rated GBAs higher on attractiveness, intentions, enjoyment and modernity.

Ohlms, Melchers, & Lievens (2026)

3.2 Candidate experience: positive, but design decides

NPS 58

from 4,778 real applicants for one cognitive GBA; some questioned how abstract games relate to the job.

Leutner et al. (2023)

But not automatically

Another cognitive GBA was rated worse than a paper test on every reaction measure (N = 183).

Ohlms, Melchers, & Kanning (2024a)

Six games, six verdicts

All six GBAs were enjoyable, but fairness and attractiveness varied widely by design, more than by age or gaming experience.

Ohlms & Melchers (2025a), N = 147

Labs underestimate it

Real applicants rated a gamified test fairer and more motivating than simulated applicants did (N = 210).

Ohlms & Melchers (2025b)

3.3 Validity: on par with good tests, not above them

Convergent validity: how strongly GBA scores correlate with established measures (correlation r)



Realistic range r = .30–.50. Diamonds = meta-analyses. Ortigoza et al. (2025) appears in a journal we could not yet verify.

Predictive evidence: GBAs forecast real outcomes

ρ .97
latent convergence of a general-ability GBA with g (N = 633)
Landers et al. (2022)

r .29
with supervisor ratings for 49 employees
Landers et al. (2022)

5
work-performance dimensions predicted for 115 employees
van Lill et al. (2023)

ΔR² .107
adaptive performance predicted beyond the Big Five (N = 182)
Ramos-Villagrasa & Fernández-del-Río (2023)

Assessment-centre link

Simulation-game scores of high-stakes applicants correlated with assessment-centre performance.
Melchers & Basch (2022)

Not superior

Adding a conventional test improved prediction beyond one cognitive GBA, not the reverse. Conscientiousness GBAs mostly measured cognitive ability.
Landers et al. (2022); Wu et al. (2022)

3.4 Faking: reduced, not removed

EXPERIMENT · N = 171

Honest vs applicant condition

Same people took a GBA and a traditional honesty–humility test

RESULT 1

Both could be faked

Scores shifted on the GBA and on the traditional test

RESULT 2

Smaller shift on the GBA

Faking effect was significantly smaller on the game

Ohlms, Melchers, Kanning, & Barends (2025). A story-based personality test also resisted faking better than the original (Landers & Collmus, 2022).

3.5 Fairness: no worse than tests, with one GBA-specific risk

Evidence of parity

- Outcome parity across **3,107 applicants** with fairness-optimized scoring (Leutner et al., 2023)
- **263 autistic** job seekers performed broadly on par with 323 graduates (Willis et al., 2021)
- No score gaps by gender, race or education on two GBAs (Hilliard et al., 2022; Brown et al., 2022)

Watch-outs

- **Gaming habits** raised GBA scores but not academic performance: test bias (Ohlms, Hohner, & Melchers, 2025; Burt et al., 2018)
- Small advantage for **men and younger** applicants on one simulation game (Melchers & Basch, 2022)
- Cognitive GBAs show race gaps **similar to cognitive tests** (Landers et al., 2022)

3.6 Realistic stories beat fantasy

Realistic workday storyline

✓ Job-relatedness ✓ Procedural fairness ✓ Organizational attractiveness ✓ Clarity of work

Fantasy storyline

Equally enjoyable and modern, but lower on all four outcomes on the left.

Ohlms, Melchers, & Lievens (2025), N = 195. Job-related storylines also raised clarity about what to expect (Ohlms, Voigtländer, et al., 2024, N = 212).

3.7 Efficiency: promising, costly to build

Every interaction becomes trace data that can predict ability.

Auer et al. (2022)

Shorter assessment sessions and a more language-agnostic experience.

Leutner et al. (2021); Landers et al. (2023)

Higher development cost, so benefits do not automatically outweigh costs.

Gürbüz et al. (2022)

04 · THE AI ERA

Can GBA resist AI-assisted cheating?

GBA very likely raises the cost of cheating compared with text-based tests, but no independent study has yet tested it directly against generative AI.

Faking

Presenting yourself more favourably on a self-report measure. **Evidence: reduced on GBA (B2).**

Cheating

Getting outside help: a proxy test taker, a leaked answer key, or an AI model. **Evidence: indirect only (D).**

PRACTITIONER OBSERVATION

Preparation material for well-known game-based screens such as McKinsey's Solve is easy to find, from video walkthroughs to paid courses. Yet, in our observation as practitioners, comparatively few candidates still get through. Solve pass rates are not publicly reported, so this cannot yet be verified empirically.

Four mechanisms that support the observation

Text is the weak spot

LLMs excel on text-based verbal tests but struggle with number series.

Hickman et al. (2024)

Time and images slow AI down

Image items resist pasting; time limits restrict outside help.

Hickman et al. (2024)

The process is scored

Every interaction becomes trace data, matching design elements that target AI's limits.

Auer et al. (2022); Torkestani & Mansouri (2025)

Less distortion

Faking effects are smaller on GBAs and story-based tests.

Ohlms et al. (2025); Landers & Collmus (2022)

Three caveats

Low pass rates may reflect strict cut-offs and selectivity, not cheat-resistance.

Practice effects can still raise scores; compare prepared and unprepared candidates.

AI moves fast: multimodal models may do better soon.

Hickman et al. (2024); Sindre (2023)

VERDICT

A plausible advantage, not yet a proven one. Treat it as a hypothesis to monitor, and stress-test every assessment against current AI before go-live.

05 · INTERNAL SELECTION

Benefits for internal selection

No study has yet tested GBA in a real promotion decision. Internal benefits are extrapolated from employee samples, promotion research and training studies.

Extrapolated

Pre-screening for managerial promotion

A problem-solving simulation game whose scores track assessment-centre performance, before the costly centre itself.

Melchers & Basch (2022), external applicants

Extrapolated

Placement, rotation, internal mobility

Cognitive GBAs predict employees' supervisor ratings and five performance dimensions.

Landers et al. (2022); van Lill et al. (2023)

Extrapolated

High-potential identification

A personality serious game predicts adaptive performance beyond the Big Five, a core trait for new roles.

Ramos-Villagrasa & Fernández-del-Río (2023)

Contextual

Certifying safety and SOP competence

Game-based step-sequencing tasks assess procedures and teach them at the same time; a safety GBA routes people to training.

Chen et al. (2025); Burt et al. (2018)

Contextual

Assessment linked to development

Gamified training raised knowledge retention and job performance in 110 employees across seven countries.

Capatina et al. (2024), self-report

Hypothesis

Protecting morale of those not promoted

Positive reactions and an innovative image may soften rejection for internal candidates who stay.

Ford et al. (2009); Georgiou & Lievens (2022)

73.5 vs 58.7

paper-test scores after a gamified safety assessment vs control (N = 151 students)

Chen et al. (2025)

r .29

GBA with supervisor ratings in an employee sample (N = 49)

Landers et al. (2022)

d −0.39

work experience shifted one game score: check incumbent effects

Ramos-Villagrasa & Fernández-del-Río (2023)

A recommended internal selection pathway

our synthesis of the evidence



Talent pool
Nomination or internal posting

Online GBA
Cognitive, personality and simulation

Supervised check
Brief in-person session on play strategy

Assessment centre
Finalists only

Decision
GBA plus performance data

Feedback for all
A development plan for every candidate

Evidence behind the steps: 2 Landers et al. (2022); 3 Sindre (2023); 4 Melchers & Basch (2022); 6 Ford et al. (2009). Step 5 is a design principle: never decide on GBA alone.

Why candidate reactions matter more inside the organization

They stay

Unsuccessful internal candidates remain employees: their reactions shape satisfaction, morale, retention and how they work with whoever was promoted.

Ford et al. (2009)

Fairness drives advocacy

In police promotion exams, motivation predicted scores and fairness predicted recommending the exam.

McCarthy et al. (2009), N = 498 and 182

GBA may help

GBAs feel more enjoyable and modern, and on one cognitive GBA anxiety did not predict scores. Still a hypothesis for promotion.

Ohlms et al. (2026); Leutner et al. (2023)

What is different about internal populations

Older, more experienced candidates

Younger applicants had a slight edge on one simulation game, and job experience shifted one game score. Check age and tenure effects.

Melchers & Basch (2022); Ramos-Villagrasa & Fernández-del-Río (2023)

Job insiders judge relevance

Employees who know the job may question abstract games; occupational background shapes perceptions of GBA.

Georgiou (2021); Ohlms, Melchers, & Lievens (2026)

AI applies inside too

Unproctored online promotion exams and certifications are as exposed to AI as external tests.

Hickman et al. (2024)

Verification is cheap

Candidates are already on site, so a supervised follow-up costs little. Face-to-face exams carry the lowest cheating risk.

Sindre (2023)

06 · RISKS

Five myths, checked against the evidence

Each popular claim is contradicted by part of the evidence. Here is what holds up.

MYTH

"Candidates always like GBA"

Usually, not always

EVIDENCE

Field data are positive (NPS 58, 4,778 applicants), but one GBA rated worse than a paper test, Polish applicants preferred traditional methods, and workers favour work samples and interviews. Design decides.

Leutner et al. (2023); Ohlms et al. (2024a, 2025a); Balcerak & Woźniak (2021); Zibarras et al. (2025)

MYTH

"GBA is bias-free"

No worse, not bias-free

EVIDENCE

Parity in several studies, yet small sex and age effects, bias from gaming habits, and cognitive GBAs inherit the race gaps of cognitive tests.

Leutner et al. (2023); Melchers & Basch (2022); Ohlms, Hohner, & Melchers (2025); Landers et al. (2022)

MYTH

"GBA measures personality well"

Moderately

EVIDENCE

Two meta-analyses disagree ($r = .516$ vs about .30 for personality games) and both are highly heterogeneous. Some "personality" games mostly measure cognitive ability.

Fadillah et al. (2025); Ortigoza et al. (2025); Wu et al. (2022)

MYTH

"GBA is more valid than tests"

Not supported

EVIDENCE

Cited support reports convergent validity and reliability, not higher prediction. A GBA added nothing beyond a conventional cognitive test. Validity is comparable.

Prianto & Sulaksono (2025); Leutner et al. (2023); Landers et al. (2022)

MYTH

"GBA employers are 72% more advantaged"

Misread statistic

EVIDENCE

The 72% is an AHP preference weight against 28%, from 35 students. The direction is reasonable; the size and generalizability are weak.

Gürbüz et al. (2022)

A risk map for GBA programmes

Design and measurement

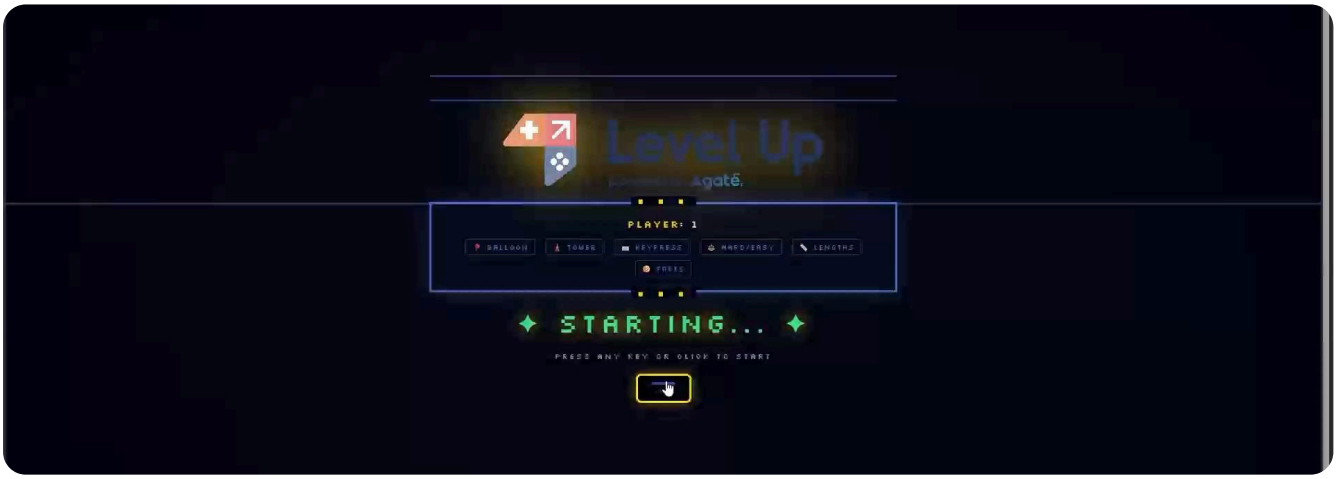
- **Deep gamification can hurt validity:** a game version of an SJT may have lost construct validity (Defresnes et al., 2026)
- **Use test–retest reliability**; one cognitive GBA reached r = .68 (Sanchez et al., 2022; Leutner et al., 2023)
- **Avoid commercial games:** very limited validity (Sanchez et al., 2022; Simons et al., 2023)
- **Check cross-national equivalence:** only partial invariance between two countries (van Lill et al., 2023)

Candidate experience

- **Gaming experience drives reactions** in real applicants (N = 376 and 284) (Naryniecki, 2026)
- **Bugs, not content,** trigger most negative app reviews (al-Qallawi & Raghavan, 2022)
- **Explain the purpose** beforehand to improve reactions (Georgiou, 2021)

Legal and governance

- **Privacy concerns are amplified** by behavioural data and GBA–AI confusion (Landers et al., 2023)
- **Bias audits and data protection** are the main legal exposure (Ortigoza et al., 2025)
- **Keep a human in the loop:** human mediation felt fairer, d = 0.68 (Santoriello et al., 2026)
- **Red-team with AI** before go-live (Torkestani & Mansouri, 2025)



Integrity starts before the first task: a proctored session start screen in Next Chapter.

d 0.68
fairness gain when selection is human-mediated rather than device-only

r .68
test–retest reliability of one commercial cognitive GBA

376 + 284
real applicants whose reactions tracked gaming experience

07 · PLAYBOOK

A practical playbook for HR

Use GBA as one component of a multi-method process, not as a replacement for validated tests.

External selection

- **Goal:** attract applicants and screen at volume
- **Role of GBA:** early screening that also builds the employer brand
- **Best-supported benefit:** attractiveness and candidate reactions
- **Main risks:** gaming-experience bias, faking on questionnaires, AI-assisted cheating
- **Validate by:** predictive validation on hired applicants

VS

Internal selection

- **Goal:** accurate, fair decisions that also develop people
- **Role of GBA:** complement to performance data and assessment centres
- **Best-supported benefit:** prediction of employee performance (extrapolated)
- **Main risks:** age and tenure effects, games seen as job-irrelevant, morale of those not selected
- **Validate by:** concurrent validation with current employees and supervisor ratings



Local settings make scenarios feel job-related.



Mini-games capture process data, not only answers.

FOR INTERNAL SELECTION

Validate on current employees, never promote on GBA alone, and give every candidate development feedback from their results.

Adoption roadmap: five phases, fifteen checks

PHASE 1 Define

- Set the purpose: volume hiring, promotion, placement or talent identification
- Match GBA type to construct; for personality, prefer gamified versions of validated tests
  Ramos-Villagrasa et al. (2022)
- Run a job analysis and derive constructs; favour theory-driven GBAs
  Landers et al. (2022)

PHASE 2 Select and design

- Use realistic, job-related storylines
  Ohlms, Melchers, & Lievens (2025)
- Plan accommodations for disabled and neurodivergent candidates
- Demand vendor evidence: validity, test–retest reliability, adverse impact
- Comply with Indonesia's Personal Data Protection Law (No. 27 of 2022)

PHASE 3 Pilot and validate

- Compare scores with current tests and employees' performance ratings
- Test language and cultural equivalence
  van Lill et al. (2023)
- Check score gaps by gender, age, gaming experience and occupation
- Stress-test with the latest LLMs and AI agents against human norms
  Hickman et al. (2024); Torkestani & Mansouri (2025)

PHASE 4 Launch

- Explain the purpose to candidates and keep a human touchpoint
  Georgiou (2021); Santoriello et al. (2026)
- Add a supervised verification step after the online GBA

PHASE 5 Monitor

- Track candidate reactions (e.g. NPS) and technical issues every hiring cycle
- Re-check adverse impact and predictive validity every hiring cycle

Before go-live, answer three questions

01 Does the GBA correlate with the tests and ratings you already trust?

02 Do scores differ by gender, age or gaming experience?

03 Can a current AI model or a proxy complete it convincingly?

If any answer is unclear, run a pilot before using GBA for high-stakes decisions.

08 · OUTLOOK

Limitations, research agenda and conclusion

What this review cannot tell you

Single reviewer
No independent double coding

One database
OpenAlex only; Q2 limited to the top 200 of 385

Abstract-based
Many studies assessed without full text

Narrow base
No Indonesian GBA validation study found

Concentration
Nine recent studies from one group; needs replication

Internal evidence
Extrapolated, not tested in promotion

AI evidence
Indirect; no LLM test on GBAs yet

Snowballing
57 on-topic references, not all synthesized

Research agenda

01
GBA in real promotion and talent-programme decisions, with post-placement outcomes

02
Reactions and retention of unsuccessful internal candidates

03
Fairness by age, tenure and occupation in employee populations

04
Validation and invariance with Indonesian and Southeast Asian samples

05
Utility and cost analyses at realistic hiring volumes

06
Effects of specific design elements on validity (Bedwell et al., 2012)

07
Direct tests of multimodal LLMs and AI agents on GBAs versus text-based tests, including effects on predictive validity

Conclusion

External
Proven: attractiveness, generally positive experience, less faking, comparable validity. Not proven: superior validity. Fairness depends on construct and gaming experience.

Internal
No direct evidence yet. Worth piloting as a complement to performance data and assessment centres, validated on current employees, never as the sole basis for promotion.

AI era
Interactive, timed, non-text and process-scored tasks are plausibly harder for AI than text tests. Monitor it; do not assume it.

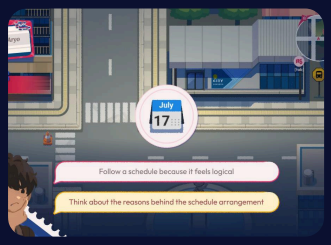
THE QUESTION TO ASK

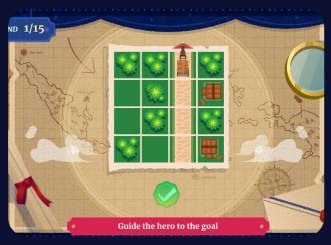
Not "is GBA better?" but "which GBA, for which construct, for which population, and has it been validated here?"

How Next Chapter puts the evidence to work

Level Up's game-based assessment for management trainees, supervisors and managers, combining simulation, personality and cognitive measurement in one enterprise-ready platform.

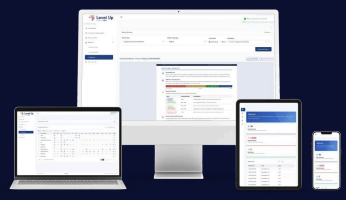
THE EVIDENCE RECOMMENDS	NEXT CHAPTER IS BUILT WITH
Theory-driven games, not repurposed commercial ones	A four-phase creation process (theory, items, validation, norming) guided by experts, with an expert blueprint and technical manual
Realistic, job-related storylines	Scenarios drawn from real situations, set in local Indonesian contexts; global themes for cognitive challenges
Multi-method assessment	Simulation, narrative Big Five personality, cognitive and neuroscience-based mini-games in one modular platform
Behaviour and process, not only self-report	Each personality facet operationalized through behavioural tendencies embedded in gameplay
Integrity in an era of AI and proxies	Built-in proctoring and anti-cheating logic with automatic scoring
Local validation and monitoring	Cut-off scores per competency, real-time status tracking, individual, batch and interactive reports











In line with Chapter 7, pilot and validate Next Chapter locally against your own performance data before high-stakes use.

Book a live demo

agatelevelup.com/game-based-assessment · enterprise@agate.id

SOURCES

References

al-Qallawi, S., & Raghavan, M. (2022). A review of online reactions to game-based assessment mobile applications. *International Journal of Selection and Assessment, 30*(1), 14–26. https://doi.org/10.1111/ijsa.12346

Aouam, H., Belmouffeq, B., & Mahil, A. (2023). Applying game-based approaches in personnel selection: A systematic literature review. *Management Dynamics in the Knowledge Economy, 11*(1), 16–29. https://doi.org/10.2478/mdke-2023-0002

Auer, E. M., Mersy, G., Marin, S., Blaik, J., & Landers, R. N. (2022). Using machine learning to model trace behavioral data from a game-based assessment. *International Journal of Selection and Assessment, 30*(1), 82–102. https://doi.org/10.1111/ijsa.12363

Balcerak, A., & Woźniak, J. (2021). Reactions to some ICT-based personnel selection tools. *Economics & Sociology, 14*(1), 214–231. https://doi.org/10.14254/2071-789X.2021/14-1/14

Bedwell, W. L., Pavlas, D., Heyne, K., Lazzara, E. H., & Salas, E. (2012). Toward a taxonomy linking game attributes to learning: An empirical study. *Simulation & Gaming, 43*(6), 729–760. https://doi.org/10.1177/1046878112439444

Bidwell, M. (2011). Paying more to get less: The effects of external hiring versus internal mobility. *Administrative Science Quarterly, 56*(3), 369–407. https://doi.org/10.1177/0001839211433562

Brown, M. I., Speer, A. B., Tenbrink, A. P., & Chabris, C. F. (2022). Using game-like animations of geometric shapes to simulate social interactions: An evaluation of group score differences. *International Journal of Selection and Assessment, 30*(1), 167–181. https://doi.org/10.1111/ijsa.12375

Burt, C. D. B., Crowe, L., & Thomas, K. (2018). Validation of a gamified measure of safety behavior: The SBT. In *Safety and reliability – Safe societies in a changing world* (pp. 263–270). CRC Press. https://doi.org/10.1201/9781351174664-34

Capatina, A., Juarez-Varon, D., Micu, A., & Micu, A. E. (2024). Leveling up in corporate training: Unveiling the power of gamification to enhance knowledge retention, knowledge sharing, and job performance. *Journal of Innovation & Knowledge, 9*(3), 100530. https://doi.org/10.1016/j.jik.2024.100530

Chen, K., Yang, S., Luh, D., Chen, Z., Ai, H., & An, Y. (2025). Gamification as an innovative approach for the assessment of procedural knowledge. *Electronics, 14*(8), 1573. https://doi.org/10.3390/electronics14081573

Defresnes, O., Huybens, W., Bertrand, F., Hansez, I., & Lo Bue, S. (2026). Gamifying a situational judgment test: Methods and challenges. *International Journal of Serious Games*, 137–167. https://doi.org/10.17083/ahsv5d92

Fadillah, Hidayat, R., & Santoso, A. (2025). Convergent validity of game-based assessment: A meta-analysis. *International Journal of Serious Games, 12*(4), 5–29. https://doi.org/10.17083/etxn7795

Ford, D. K., Truxillo, D. M., & Bauer, T. N. (2009). Rejected but still there: Shifting the focus in applicant reactions to the promotional context. *International Journal of Selection and Assessment, 17*(4), 402–416. https://doi.org/10.1111/j.1468-2389.2009.00482.x

Georgiou, K. (2021). Can explanations improve applicant reactions towards gamified assessment methods? *International Journal of Selection and Assessment, 29*(2), 253–268. https://doi.org/10.1111/ijsa.12329

Georgiou, K., Gouras, A., & Nikolaou, I. (2019). Gamification in employee selection: The development of a gamified assessment. *International Journal of Selection and Assessment, 27*(2), 91–103. https://doi.org/10.1111/ijsa.12240

Georgiou, K., & Lievens, F. (2022). Gamifying an assessment method: What signals are organizations sending to applicants? *Journal of Managerial Psychology, 37*(6), 559–574. https://doi.org/10.1108/JMP-12-2020-0653

Gkorezis, P., Georgiou, K., Nikolaou, I., & Kyriazati, A. (2021). Gamified or traditional situational judgement test? A moderated mediation model of recommendation intentions via organizational attractiveness. *European Journal of Work and Organizational Psychology, 30*(2), 240–250. https://doi.org/10.1080/1359432X.2020.1746827

Gürbüz, U., Afacan Fındıklı, M., & Özdemir, Ş. (2022). Digital HR: A study on how game-based assessments can attract talents. *Global Journal of Information Technology: Emerging Technologies, 12*(1), 22–33. https://doi.org/10.18844/gjit.v12i1.6941

Hickman, L., Dunlop, P. D., & Wolf, J. L. (2024). The performance of large language models on quantitative and verbal ability tests: Initial evidence and implications for unproctored high-stakes testing. *International Journal of Selection and Assessment, 32*(4), 499–511. https://doi.org/10.1111/ijsa.12479

Hilliard, A., Kazim, E., Bitsakis, T., & Leutner, F. (2022). Measuring personality through images: Validating a forced-choice image-based assessment of the Big Five personality traits. *Journal of Intelligence, 10*(1), 12. https://doi.org/10.3390/jintelligence10010012

Hommel, B. E., Ruppel, R., & Zacher, H. (2022). Assessment of cognitive flexibility in personnel selection: Validity and acceptance of a gamified version of the Wisconsin Card Sorting Test. *International Journal of Selection and Assessment, 30*(1), 126–144. https://doi.org/10.1111/ijsa.12362

Küpper, D. M., Klein, K., & Völckner, F. (2021). Gamifying employer branding: An integrating framework and research propositions for a new HRM approach in the digitized economy. *Human Resource Management Review, 31*(1), 100686. https://doi.org/10.1016/j.hrmr.2019.04.002

Landers, R. N., Armstrong, M. B., Collmus, A. B., Mujcic, S., & Blaik, J. (2022). Theory-driven game-based assessment of general cognitive ability: Design theory, measurement, prediction of performance, and test fairness. *Journal of Applied Psychology, 107*(10), 1655–1677. https://doi.org/10.1037/apl0000954

Landers, R. N., & Collmus, A. B. (2022). Gamifying a personality measure by converting it into a story: Convergence, incremental prediction, faking, and reactions. *International Journal of Selection and Assessment, 30*(1), 145–156. https://doi.org/10.1111/ijsa.12373

Landers, R. N., Lee, V., & Chen, C. (2023). Game-based assessment around the globe: International differences in privacy, legality, and applicant reaction concerns. In *Technology and measurement around the globe* (pp. 157–175). Cambridge University Press. https://doi.org/10.1017/9781009099813.008

Landers, R. N., & Sanchez, D. R. (2022). Game-based, gamified, and gamefully designed assessments for employee selection: Definitions, distinctions, design, and validation. *International Journal of Selection and Assessment, 30*(1), 1–13. https://doi.org/10.1111/ijsa.12376

Law of the Republic of Indonesia No. 27 of 2022 on Personal Data Protection.

Leutner, F., Codreanu, S.-C., Brink, S., & Bitsakis, T. (2023). Game based assessments of cognitive ability in recruitment: Validity, fairness and test-taking experience. *Frontiers in Psychology, 13*, 942662. https://doi.org/10.3389/fpsyg.2022.942662

Leutner, F., Codreanu, S.-C., Liff, J., & Mondragon, N. (2021). The potential of game- and video-based assessments for social attributes: Examples from practice. *Journal of Managerial Psychology, 36*(7), 533–547. https://doi.org/10.1108/JMP-01-2020-0023

McCarthy, J., Hrabluik, C., & Jelley, R. B. (2009). Progression through the ranks: Assessing employee reactions to high-stakes employment testing. *Personnel Psychology, 62*(4), 793–832. https://doi.org/10.1111/j.1744-6570.2009.01158.x

McChesney, J., Campbell, C., Wang, J., & Foster, L. (2022). What is in a name? Effects of game-framing on perceptions of hiring organizations. *International Journal of Selection and Assessment, 30*(1), 182–192. https://doi.org/10.1111/ijsa.12370

Melchers, K. G., & Basch, J. M. (2022). Fair play? Sex-, age-, and job-related correlates of performance in a computer-based simulation game. *International Journal of Selection and Assessment, 30*(1), 48–61. https://doi.org/10.1111/ijsa.12337

Naryniecki, T. J. (2026). Who most enjoys game-based assessment? Predicting applicant reactions by performance, gaming experience and demographic factors. *International Journal of Selection and Assessment, 34*(1), e70047. https://doi.org/10.1111/ijsa.70047

Ohlms, M. L., Hohner, A., & Melchers, K. G. (2025). Is test takers' video game usage a game changer? Criterion-related validity and potential bias of a cognitive game-based assessment. *Applied Psychology, 74*(5), e70038. https://doi.org/10.1111/apps.70038

Ohlms, M. L., & Melchers, K. G. (2025a). Are games always fun and fair? A comparison of reactions to different game-based assessments. *International Journal of Selection and Assessment, 33*(1), e12520. https://doi.org/10.1111/ijsa.12520

Ohlms, M. L., & Melchers, K. G. (2025b). Do applicant reactions to gamified cognitive ability tests differ between high- versus low-stakes settings? *Journal of Intelligence, 13*(3), 33. https://doi.org/10.3390/jintelligence13030033

Ohlms, M. L., Melchers, K. G., & Kanning, U. P. (2024a). Can we playfully measure cognitive ability? Construct-related validity and applicant reactions. *International Journal of Selection and Assessment, 32*(1), 91–107. https://doi.org/10.1111/ijsa.12450

Ohlms, M. L., Melchers, K. G., & Kanning, U. P. (2024b). Playful personnel selection: The use of traditional versus game-related personnel selection methods and their perception from the recruiters' and applicants' perspectives. *International Journal of Selection and Assessment, 32*(3), 381–398. https://doi.org/10.1111/ijsa.12466

Ohlms, M. L., Melchers, K. G., Kanning, U. P., & Barends, A. J. (2025). Game on, faking off? Are game-based assessments less susceptible to faking than traditional assessments? *Journal of Business and Psychology, 40*(6), 1323–1335. https://doi.org/10.1007/s10869-025-10019-6

Ohlms, M. L., Melchers, K. G., & Lievens, F. (2025). It's just a game! Effects of fantasy in a storified test on applicant reactions. *Applied Psychology, 74*(1), e12569. https://doi.org/10.1111/apps.12569

Ohlms, M. L., Melchers, K. G., & Lievens, F. (2026). Perceptions of game-based assessments: The role of test takers' occupational background. *International Journal of Selection and Assessment, 34*(1), e70040. https://doi.org/10.1111/ijsa.70040

Ohlms, M. L., Voigtländer, E., Melchers, K. G., & Kanning, U. P. (2024). Is gamification a suitable means to improve applicant reactions and convey information during an online test? *Journal of Personnel Psychology, 23*(4), 169–178. https://doi.org/10.1027/1866-5888/a000343

Ortigoza, A., Espinosa Aguilar, A., & Herrera Paredes, R. M. (2025). Neuroscience-branded and gamified hiring assessments: A systematic review of validity, adverse impact and legal exposure. *NeuroData, 2*, 124. https://doi.org/10.63688/neurodata2025124

Page, M. J., McKenzie, J. E., Bossuyt, P. M., Boutron, I., Hoffmann, T. C., Mulrow, C. D., … Moher, D. (2021). The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ, 372*, n71. https://doi.org/10.1136/bmj.n71

Prianto, C., & Sulaksono, A. N. R. (2025). Gamification in digital HR system: A systematic literature review of trends, game elements, and audiences. *Dinasti International Journal of Education Management and Social Science, 6*(6), 4856–4867. https://doi.org/10.38035/dijemss.v6i6

Ramos-Villagrasa, P. J., & Fernández-del-Río, E. (2023). Predictive validity, applicant reactions, and influence of personal characteristics of a gamefully designed assessment. *Journal of Work and Organizational Psychology, 39*(3), 169–178. https://doi.org/10.5093/jwop2023a18

Ramos-Villagrasa, P. J., Fernández-del-Río, E., & Castro, Á. (2022). Game-related assessments for personnel selection: A systematic review. *Frontiers in Psychology, 13*, 952002. https://doi.org/10.3389/fpsyg.2022.952002

Ramos-Villagrasa, P. J., Fernández-del-Río, E., Hermoso, R., & Cebrián, J. (2024). Are serious games an alternative to traditional personality questionnaires? Initial analysis of a gamified assessment. *PLOS ONE, 19*(5), e0302429. https://doi.org/10.1371/journal.pone.0302429

Sanchez, D. R., Weiner, E., & Van Zelderen, A. (2022). Virtual reality assessments (VRAs): Exploring the reliability and validity of evaluations in VR. *International Journal of Selection and Assessment, 30*(1), 103–125. https://doi.org/10.1111/ijsa.12369

Santoriello, M., Cortini, M., Talamo, A., Marocco, S., & Galanti, T. (2026). Human vs. device mediation in gamified selection: Fairness perceptions and technology acceptance. In *Proceedings of the 2026 European Conference on Cognitive Ergonomics* (pp. 386–390). ACM. https://doi.org/10.1145/3822301.3822353

Santoriello, M., Fantinelli, S., Cortini, M., Giffi, V., & Galanti, T. (2025). Does a positive attitude toward technology enhance job search self-efficacy? The role of gamification. *Informing Science, 28*, 034. https://doi.org/10.28945/5651

Simons, A., Wohlgenannt, I., Zelt, S., Weinmann, M., Schneider, J., & vom Brocke, J. (2023). Intelligence at play: Game-based assessment using a virtual-reality application. *Virtual Reality, 27*(3), 1827–1843. https://doi.org/10.1007/s10055-023-00752-9

Sindre, G. (2023). AI technology: Threats and opportunities for assessment integrity in introductory programming. *Norsk IKT-konferanse for forskning og utdanning (NIKT 2023)*.

Torkestani, M. S., & Mansouri, T. (2025). *Machine vs machine: Using AI to tackle generative AI threats in assessment* [Paper presentation]. Learning, Teaching & Student Experience 2025, Chartered Association of Business Schools, Nottingham, UK. https://doi.org/10.48550/arXiv.2506.02046

van Lill, X., McColl, L., & Neale, M. (2023). Cross-national applicability of a game-based cognitive assessment. *International Journal of Selection and Assessment, 31*(2), 302–320. https://doi.org/10.1111/ijsa.12425

Varghese, J., & Deepa, R. (2023). Gamification as an effective employer branding strategy for Gen Z. *NHRD Network Journal, 16*(3), 269–279. https://doi.org/10.1177/26314541231170438

Wiernik, B. M., Raghavan, M., Carretta, T. R., & Coovert, M. D. (2022). Developing and validating a serious game-based assessment for cyber occupations in the US Air Force. *International Journal of Selection and Assessment, 30*(1), 27–47. https://doi.org/10.1111/ijsa.12378

Willis, C., Powell-Rudy, T., Colley, K., & Prasad, J. (2021). Examining the use of game-based assessments for hiring autistic job seekers. *Journal of Intelligence, 9*(4), 53. https://doi.org/10.3390/jintelligence9040053

Wu, F. Y., Mulfinger, E., Alexander, L., Sinclair, A. L., McCloy, R. A., & Oswald, F. L. (2022). Individual differences at play: An investigation into measuring Big Five personality facets with game-based assessments. *International Journal of Selection and Assessment, 30*(1), 62–81. https://doi.org/10.1111/ijsa.12360

Zibarras, L. D., Castaño, G., & Cuppello, S. (2025). Applicant perceptions of selection methods: Replicating and extending previous research. *International Journal of Selection and Assessment, 33*(2), e70007. https://doi.org/10.1111/ijsa.70007

APPENDIX

Studies included in the review

35 sources synthesized directly, ordered by evidence level. E = external selection; I = internal selection; Ctx = contextual evidence. The 34 articles synthesized by Ramos-Villagrasa et al. (2022) are not listed individually.

Study	Context	Assessment	Sample	Design	Level	Key finding
Ramos-Villagrasa, Fernández-del-Río, & Castro (2022)	E	All GRA types	34 articles	SLR (PRISMA)	A	GRA usable; advantages smaller than claimed
Fadillah, Hidayat, & Santoso (2025)	E	Personality GBAs	18 effects, 13 articles	Meta-analysis	A	Convergence with self-report r = .516; heterogeneous
Ortigoza et al. (2025)	E, I	GBA, gamified, neuro	12 studies	SR + meta-analysis	A	Pooled r = .39; criterion validity still thin
Aouam, Belmouffeq, & Mahil (2023)	E	GBA, serious games	21 articles	SLR	A	Themes: candidate reactions, benefits and limits
Prianto & Sulaksono (2025)	E, I	HR gamification	41 studies	Automated SLR	A	85.4% target employees; 7.3% recruitment
Leutner, Codreanu, Brink, & Bitsakis (2023)	E	Cognitive GBA + ML	11,574 completions; 3,107 applicants	Field data	B1	r = .5; test–retest .68; outcome parity; NPS 58
Ohlms & Melchers (2025b)	E	Gamified cognitive test	210, real vs simulated applicants	Quasi-experiment	B1	Real applicants reacted more positively
Naryniecki (2026)	E	Two GBAs	376 and 284 real applicants	Field	B1	Scores and gaming experience predict reactions
Ohlms, Hohner, & Melchers (2025)	E, I	Cognitive GBA	156 workers + 92 potential applicants	Correlational	B1	Criterion validity; bias from gaming habits
van Lill, McColl, & Neale (2023)	E, I	Cognitive GBA	228 Australia, 239 South Africa	Invariance + predictive	B1	Partial invariance; predicts five performance dimensions
Willis et al. (2021)	E	Cognitive GBAs	263 autistic vs 323 graduates	Comparative	B1	Performance broadly on par
Ohlms, Melchers, Kanning, & Barends (2025)	E	Honesty–humility GBA	171	Experiment	B2	Faking smaller, not eliminated
Ohlms & Melchers (2025a)	E	Six GBAs	147	Experiment	B2	Reactions vary widely by design
Ohlms, Melchers, & Lievens (2025)	E	Storified tests	195	Experiment	B2	Realistic story beats fantasy
Ohlms, Melchers, & Lievens (2026)	E	GBA vs standard test	179	Experiment	B2	GBA higher on attractiveness; occupation matters
Ohlms, Voigtländer, et al. (2024)	E	Gamified cognitive test	212	Experiment	B2	Higher job clarity and opportunity to perform
Ohlms, Melchers, & Kanning (2024a)	E	Cognitive GBA	183	Lab	B2	Convergence r = .51; reactions worse than paper test
Santoriello, Cortini, et al. (2026)	E	Gamified selection	148 students	Scenario experiment	B2	Human mediation seen as fairer (d = 0.68)
Hickman, Dunlop, & Wolf (2024)	E (AI era)	Conventional ability tests	GPT-3.5 and GPT-4	LLM experiment	B2	GPT-4 95th percentile verbal; 12th on number series

Study	Context	Assessment	Sample	Design	Level	Key finding
Burt, Crowe, & Thomas (2018)	E, I	Gamified safety test (SBT)	Participants + independent informants	Criterion validation	C	Good criterion validity; affected by gaming experience
Leutner, Codreanu, Liff, & Mondragon (2021)	E	Game and video assessments	1,300	Validation	C	Convergence r = .21–.42; seen as more immersive
Ramos-Villagrasa & Fernández-del-Río (2023)	E	Personality serious game	182	Correlational	C	ΔR^2 = .107 for adaptive performance
Ramos-Villagrasa et al. (2024)	E	Gamified Big Five (VASSIP)	98 students	Correlational	C	Reliability similar to original; more attractive
Simons et al. (2023)	E	Commercial VR game	103	Lab correlational	C	Completion time correlates with intelligence
Ohlms, Melchers, & Kanning (2024b)	E	GRAs	158 recruiters, 144 potential applicants	Survey	C	Main reason: organizational image and reactions
Zibarras, Castaño, & Cuppello (2025)	E, I	17 methods incl. GBA	281 workers	Survey	C	GBA not among favourite methods
Santoriello, Fantinelli, et al. (2025)	E	Gamified selection	144 near-graduates	Survey	C	Openness to gamified selection lifts job-search self-efficacy
Varghese & Deepa (2023)	E	Gamified assessment	40 Gen Z students	Pilot	C	Employer-brand attributes of gamified assessment
Defresnes et al. (2026)	E (military)	Game-based SJT	Belgian Defence	Design + exploratory psychometrics	C	Risk of lower construct validity
Gürbüz, Afacan Fındıklı, & Özdemir (2022)	E	GBA	35 students	AHP	C	Preference weight 72% vs 28%
Capatina et al. (2024)	I (training)	Gamified training	110 employees	PLS-SEM survey + eye tracking	C	Higher knowledge retention and job performance
Chen et al. (2025)	Ctx (education)	Gamified assessment	151 students	Experiment	C	Paper-test score 73.5 vs 58.7 after GBA
Landers, Lee, & Chen (2023)	E (cross-cultural)	GBA	Vendors and stakeholders	Qualitative	D	Positive reactions across countries; privacy concerns amplified
Sindre (2023)	Ctx (AI)	Programming exams	CS1 tasks	ChatGPT trial + threat analysis	D	AI aggravates existing cheating threats
Torkestani & Mansouri (2025)	Ctx (AI)	Higher-education assessment	—	Theoretical framework	D	Eight design elements for AI-resistant assessment



Evaluate candidates through real gameplay.

Book a live demo and let your team play the assessment, review sample individual and batch reports, and plan a local validation pilot.

agatelevelup.com/game-based-assessment

enterprise@agate.id

Level Up
powered by Agaté.

Agate Headquarters · Kawasan Summarecon Bandung, Level Up Building, Jl. Sentra Raya Barat No. Kav. 2B, Cisaranten Kidul, Gedebage, Bandung, West Java 40295, Indonesia

© 2026 Level Up powered by Agate. All rights reserved.