LumiKin

Methodology

Version 1.1 · Published May 1, 2026

lumikin.org/methodology

Overview

LumiKin measures two things about every game and reports them independently. The first is what a child stands to gain from playing — measured as a Benefits Development Score (BDS) across cognitive, social-emotional, and motor dimensions. The second is the design pressure the game exerts on a developing brain — measured as a Risk Influence Score (RIS) across dopamine manipulation, monetization pressure, and social risk. From these two scores, LumiKin derives two outputs: a LumiScore (0–100), which is a harmonic mean of benefit and safety that penalises games that are high-risk or low-benefit on either axis; and a daily time recommendation in five tiers, derived from the RIS with adjustments for unusually high or low benefit scores. Many games score meaningfully on both axes simultaneously, and LumiKin is designed to make that visible rather than collapse it into a single judgement.

LumiKin does not measure how much a child enjoys a game, how skilled a player is, or whether a particular title is "good" or "bad" in any general sense. It measures structural design choices — what the game asks of a player and what mechanisms it uses to keep them playing. A beloved family favourite can carry meaningful design risks; a forgettable puzzle game can quietly deliver real cognitive benefit. The framework is deliberately separated from taste, popularity, and critical reception, all of which are well-served by other rating systems.

The framework is designed to be auditable. Every dimension has a fixed scoring scale, every category has a documented weight, and every score carries the methodology version that produced it. When the rubric changes, a new version is published and prior versions remain queryable. Reviewers — whether human or automated — score against

the same rubric, with reference titles used to calibrate against drift. Limitations and edge cases (UGC platforms, live-service content, automated scoring) are documented openly in the Limitations section below rather than buried.

The Scoring Model

Every LumiKin game profile produces two independent composite scores and a time recommendation derived from them.

Benefits Development Score (BDS)

The BDS combines three benefit categories using fixed weights:

CATEGORY	WEIGHT
B1 — Cognitive development	50%
B2 — Social-emotional development	30%
B3 — Physical and motor development	20%

Each category score is normalised to 0–1 before weighting.

Formula: `BDS = (B1_norm × 0.50) + (B2_norm × 0.30) + (B3_norm × 0.20)`

Risk Influence Score (RIS)

The RIS combines three risk categories. Content risk (R4) is reported separately and does not feed into the time recommendation — this preserves compatibility with ESRB and PEGI ratings.

CATEGORY	WEIGHT
R1 — Dopamine manipulation design	45%
R2 — Monetization pressure	30%
R3 — Social and emotional risk	25%

Formula: `RIS = (R1_norm × 0.45) + (R2_norm × 0.30) + (R3_norm × 0.25)`

Time recommendation tiers

The primary output is a recommended daily time limit derived from the RIS, with adjustments for high-benefit and low-benefit games.

RIS RANGE	BASE RECOMMENDATION
0.00 – 0.15	120 min/day
0.16 – 0.30	90 min/day
0.31 – 0.50	60 min/day
0.51 – 0.70	30 min/day
0.71 +	15 min/day (not recommended for daily play)

Tier adjustments:

- If `BDS ≥ 0.60` and `RIS ≤ 0.70`: recommendation extends one tier (e.g. 60 → 90 min)
- If `BDS < 0.20` and `RIS > 0.30`: recommendation drops one tier (e.g. 60 → 30 min)

These adjustments do not stack.

B1 Cognitive

Rationale: Cognitive development is the largest single weight in the BDS (50%) because the age range LumiKin targets — roughly six to sixteen — overlaps with significant maturation of executive function, working memory, spatial reasoning, and metacognition. Games occupy a meaningful share of discretionary time in this window. The mechanics a child engages with repeatedly during this period are not neutral: they constitute a form of practice. Some games offer rich, varied cognitive challenges; others offer narrow, repetitive ones. The B1 dimensions attempt to characterise that landscape so parents and partners can distinguish between superficially similar titles that differ substantially in what they actually train.

What's measured: Ten dimensions, each scored 0–5 by a trained reviewer:

DIMENSION	WHAT IS ASSESSED
Problem solving	Novel problems requiring reasoning, experimentation, or logic
Spatial awareness	Mental rotation, 3D navigation, map reading, spatial planning
Strategic thinking	Planning ahead, resource management, evaluating trade-offs
Critical thinking	Evaluating information, questioning assumptions, evidence-based decisions
Memory and attention	Working memory, sustained attention, pattern recognition
Creativity and expression	Open-ended tools for building, designing, composing, or storytelling
Reading and language	Vocabulary, reading comprehension, narrative understanding
Math and systems thinking	Numerical reasoning, economic systems, statistical thinking
Learning transfer	Knowledge or skills that apply outside the game
Adaptive challenge	Difficulty that scales with player skill (flow state design)

Scoring: Sum of 10 dimensions (0–5 each). Max 50. Normalised to 0–1 for weighting.

Weighting: B1 contributes 50% of the BDS — the largest single weight, reflecting the centrality of cognitive development in the age ranges LumiKin targets.

Research basis:

Bavelier, D. & Green, C.S. (2019). Enhancing attentional control: Lessons from action video games. *Neuron*, 104(1), 147–163.

Granic, I., Lobel, A., & Engels, R.C.M.E. (2014). The benefits of playing video games. *American Psychologist*, 69(1), 66–78.

Sala, G., Tatlidil, K.S., & Gobet, F. (2018). Video game training does not enhance cognitive ability: A comprehensive meta-analytic investigation. *Psychological Bulletin*, 144(2), 111–139. (Included deliberately as a critical counterweight — LumiKin does not claim broad transfer effects, and the B1 framework distinguishes specific trained skills from general cognitive enhancement.)

B2 Social-Emotional

Rationale: Cooperative play, narrative-driven games, and well-moderated multiplayer environments can develop genuine social and emotional skills — perspective-taking, emotional regulation under pressure, ethical reasoning in low-stakes contexts. The same structural feature that makes games effective at training these skills (sustained social interaction with consequence) also makes them vulnerable to exploitation, which is why the social-emotional dimensions are scored separately from social-emotional risks (R3). A game can simultaneously develop authentic teamwork (high B2) and expose a child to identity pressure or unmoderated strangers (high R3); these are independent axes and the framework treats them that way.

What's measured: Six dimensions, each scored 0–5:

DIMENSION	WHAT IS ASSESSED
Teamwork and cooperation	Genuine collaboration where players depend on each other
Communication skills	Meaningful communication between players
Empathy and perspective-taking	Understanding other viewpoints, cultures, or emotional experiences
Emotional regulation	Persistence, managing frustration, coping with loss — through design, not punishment
Ethical reasoning	Moral dilemmas, consequences for choices, opportunities to consider fairness
Positive social interaction	Moderated multiplayer, prosocial incentives, constructive community design

Scoring: Sum of 6 dimensions (0–5 each). Max 30. Normalised to 0–1 for weighting.

Weighting: B2 contributes 30% of the BDS.

Research basis:

> Greitemeyer, T. & Mügge, D.O. (2014). Video games do affect social outcomes: A meta-analytic review of the effects of violent and prosocial video game play.

Personality and Social Psychology Bulletin, 40(5), 578–589.

Harrington, B. & O'Connell, M. (2016). Video games as virtual teachers: Prosocial video game use by children and adolescents from different socioeconomic groups is associated with increased empathy and prosocial behaviour. *Computers in Human Behavior*, 63, 650–658.

B3 Motor

Rationale: Motor development is the smallest weight in the BDS (20%) because the dominant concern for screen time in this age range is sedentary behaviour rather than fine-motor deficit. The dimension is retained because some games — particularly motion-controlled, VR, and AR titles — meaningfully extend into physical activity, and because hand-eye coordination training has documented co-benefits in attention and reaction tasks. Games that contribute negligibly to motor development are scored low here without penalty to the overall BDS provided they perform well on B1 and B2.

What's measured: Four dimensions, each scored 0–5:

DIMENSION	WHAT IS ASSESSED
Hand-eye coordination	Precise timing, aiming, coordination between visual input and motor response
Fine motor skills	Precise small-muscle movements, dexterity, touch precision
Reaction time	Quick reflexes and rapid decision-making
Physical activity	Whole-body movement (VR, motion controls, AR)

Scoring: Sum of 4 dimensions (0–5 each). Max 20. Normalised to 0–1 for weighting.

Weighting: B3 contributes 20% of the BDS.

Research basis:

Polechoński, J., Dębska, M., & Dębski, P.G. (2019). Exergaming can be a health-related aerobic physical activity. *BioMed Research International*, 2019, Article 1890527.

Li, R., Polat, U., Makous, W., & Bavelier, D. (2009). Enhancing the contrast sensitivity function through action video game training. *Nature Neuroscience*, 12(5), 549–551.

R1 Dopamine Design

Rationale: Dopamine manipulation design is weighted highest in the RIS (45%) because the mechanisms involved — variable ratio reward schedules, near-miss feedback, loss aversion triggers, time-limited urgency — are the most thoroughly documented in behavioural psychology and the most aggressively engineered into modern free-to-play games. The same techniques are used in regulated gambling products. The developmental concern is acute: the prefrontal regulation that helps adults resist these mechanisms is still maturing through adolescence. R1 attempts to characterise the intensity and variety of these mechanisms in a given title, on the principle that a game with one mild streak feature and a game built end-to-end around variable rewards present qualitatively different risks even if both technically contain "engagement design".

What's measured: Ten risk factors, each scored 0–3:

FACTOR	DESCRIPTION
Variable ratio rewards	Random reward schedules (loot boxes, gacha)
Streak mechanics	Daily login rewards with penalties for missing
Loss aversion triggers	Resources that decay or opponents that advance while absent
FOMO / time-limited events	Content that disappears, creating urgency
Artificial stopping barriers	Energy/lives systems that create frustration pressure
Notification and re-engagement	Notifications designed to create anxiety
Near-miss mechanics	Deliberate "almost won" feedback mimicking gambling
Infinite scroll / endless play	Auto-play, no natural endpoints
Escalating commitment	Sunk-cost design ("I've invested too much to stop")

FACTOR	DESCRIPTION
Variable reward frequency	Precisely calibrated reward timing (slot machine pattern)

Scoring: Sum of 10 factors (0–3 each). Max 30. Normalised to 0–1 for weighting.

Weighting: R1 contributes 45% of the RIS — the highest weight, reflecting the documented developmental harm of these mechanics in children's products.

Research basis:

Zendle, D. & Cairns, P. (2018). Video game loot boxes are linked to problem gambling: Results of a large-scale survey. *PLOS ONE*, 13(11), e0206767.

King, D.L. & Delfabbro, P.H. (2019). Predatory monetization schemes in video games (e.g., 'loot boxes') and Internet gaming disorder. *Addiction*, 114(11), 1967–1969.

UK Competition and Markets Authority (2022). Online choice architecture: How digital design can harm competition and consumers. Discussion paper.

The R1 dimension list draws on the SHARP-G framework (Saini, Hodgins et al., 2024 — *Journal of Behavioral Addictions*), adapted for children's contexts.

R2 Monetization

Rationale: Monetization pressure (R2, 30% of RIS) is scored separately from dopamine design because the mechanisms are distinct even when they co-occur. R1 measures psychological extraction — how strongly the game pulls a child into continuing to play. R2 measures financial extraction — how strongly the game pulls a child or parent into continuing to spend. A game can be high on one and low on the other: a free advertising-funded game may have minimal monetization pressure but heavy dopamine design; a flat-fee premium title may have low dopamine design but in-game purchase prompts targeting children. R2 is particularly load-bearing in regulatory contexts (UK Children's Code, FTC guidance on dark patterns in children's apps), and partner products typically need to filter on it independently of overall game quality.

What's measured: Eight risk factors, each scored 0–3:

FACTOR	DESCRIPTION
Spending ceiling	Whether unlimited spending is possible

FACTOR	DESCRIPTION
Pay-to-win mechanics	Gameplay advantage from spending
Currency obfuscation	Multiple virtual currencies obscuring real cost
Spending prompts	Prompts triggered by failure or frustration
Child-targeting design	Purchase UI explicitly targeting younger users
Ad pressure	Frequency and intrusiveness of advertising
Subscription pressure	Auto-renewal, locked features
Social spending pressure	Gifting mechanics, peer pressure to spend

Scoring: Sum of 8 factors (0–3 each). Max 24. Normalised to 0–1 for weighting.

Weighting: R2 contributes 30% of the RIS.

Research basis:

Federal Trade Commission (2022). Bringing dark patterns to light. FTC Staff Report.

Information Commissioner's Office, UK (2020). Age Appropriate Design: a code of practice for online services (Children's Code).

Fitton, D. & Read, J.C. (2019). Creating a framework to support the critical consideration of dark design aspects in free-to-play apps. *Proceedings of IDC '19*, 407–418.

R3 Social Risk

Rationale: Social risk is scored independently of the positive social dimensions in B2 because they describe different phenomena even when they share a surface — multiplayer interaction. A game with strong cooperative design and active community moderation can develop genuine teamwork (high B2) while remaining low-risk (low R3). A game with the same multiplayer surface but without moderation, with public ranking, and with unmoderated voice chat available to strangers can develop similar surface skills while exposing a child to qualitatively different risks. R3 attempts to characterise those risks distinctly: the chance of toxic competitive environments, the presence of unknown adults in voice or chat, social comparison pressure tied to spending or in-

game performance, and privacy practices. The most acute concern in this category — particularly for children under 13 — is unmoderated stranger contact, which is scored as its own factor and weighted heavily within R3.

What's measured: Six risk factors, each scored 0–3:

FACTOR	DESCRIPTION
Social obligation	Pressure to play from group expectations (guilds, team events)
Competitive toxicity	Rank anxiety, public shaming, limited moderation
Stranger interaction risk	Unmoderated communication with unknown adults
Social comparison	Prominent leaderboards, spending-visible cosmetics
Identity and self-worth	Self-worth linked to in-game performance or possessions
Privacy risk	Data collected on minors without transparent consent

Scoring: Sum of 6 factors (0–3 each). Max 18. Normalised to 0–1 for weighting.

Weighting: R3 contributes 25% of the RIS.

Research basis:

Kowert, R. (2020). Dark Participation in Games. *Frontiers in Psychology*, 11, 598947.

Internet Watch Foundation annual reports — relevant sections on grooming via gaming platforms.

5Rights Foundation (2021). Pathways: How digital design puts children at risk.

R4 Content Risk

Important: R4 is reported as a standalone content flag and **does not feed into the RIS or the time recommendation**. This is a deliberate design decision: content risk largely aligns with existing ESRB and PEGI ratings, and LumiKin's primary contribution is in areas those ratings don't cover (manipulation design, monetization, social risk). Including R4 in the time recommendation would create overlap with regulatory systems and reduce the distinctiveness of the LumiScore.

What's assessed: Five factors, each scored 0–3:

FACTOR	EXAMPLES
Violence level (R4.1)	None → mild cartoon → moderate realistic → graphic
Sexual content (R4.2)	None → mild suggestive → moderate themes → explicit
Language (R4.3)	None → mild → moderate profanity → extreme/frequent
Substance references (R4.4)	None → mild references → depicted without consequence → depicted positively
Fear and horror (R4.5)	None → mild tension → moderate scares → intense horror

Display: R4 scores are shown as a content advisory alongside the ESRB/PEGI rating on game pages, not as a component of the overall LumiScore.

Age-Floor Policy

Three of the five R4 dimensions produce a numeric `recommendedMinAge` for each game: violence (R4.1), sexual content (R4.2), and fear/horror (R4.5). Language (R4.3) and substance references (R4.4) inform the parent narrative and tips but do not set an age floor — they are contextual rather than developmentally threshold-linked.

DIMENSION	SCORE 0	SCORE 1	SCORE 2	SCORE 3
R4.1 Violence	0	7	13	17
R4.2 Sexual content	0	9	13	17
R4.5 Fear / horror	0	7	10	13

Formula: `recommendedMinAge = max(violenceFloor, sexualFloor, fearFloor)`, then context modifiers are applied.

Score-1 asymmetry. Violence and fear both step to 7 at score 1 because PEGI 7 explicitly permits "non-realistic violence in a child-friendly context" and "a game may be considered frightening". Mild romantic or suggestive content (sexual score 1) steps to 9 instead, because relational scaffolding for that material is not reliably present until approximately age 9. The asymmetry is developmental, not moral.

Fear caps at 13. Intense horror without graphic sexual or explicit violence content (R4.5 = 3 alone) places a game at the ESRB Teen / PEGI 16 boundary, not the adult boundary. The adult floor (17) is reserved for games where graphic violence or explicit sexual content is the primary driver. A pure horror atmosphere — however intense — does not independently warrant an adult floor.

R4 context modifiers

Context modifiers apply only to the violence (R4.1) and sexual (R4.2) dimensions and add **+2 to the affected dimension's floor, capped at 17**. Fear is assessed directly in R4.5 and has no modifiers. Modifiers are additive and stack within a dimension.

MODIFIER	APPLIES TO	CONDITION
`trivialized`	R4.1, R4.2	Violence or sexual content played for laughs or with no shown consequences
`defenceless_target`	R4.1 only	Violence against non-combatants, surrendered, or helpless characters
`mixed_sexual_violent`	R4.1 + R4.2	Sexual and violent content combined in the same scene; requires both R4.1 > 0 and R4.2 > 0

Update and Versioning Policy

Games are rescored under three triggers. First, automated rescore is initiated when the upstream metadata source (RAWG) reports a substantive update — DLC, platform expansion, or content patches that materially alter mechanics. Second, manual rescore is initiated for live-service titles where mechanics have changed in ways the automated pipeline did not detect. Third, ad-hoc rescore is initiated when LumiKin receives credible information about changes via community reports, regulatory action, or publisher disclosure.

The scoring weights, dimension definitions, scoring scales, and tier thresholds are locked within a methodology version. Any change to these — adding a dimension, adjusting a weight, redefining a tier boundary — requires publishing a new methodology version. Version numbers follow semantic versioning: minor version increments (1.0 → 1.1) for additive changes that do not alter existing scores, major

version increments (1.0 → 2.0) for changes that require all games to be rescored. Every score carries the methodology version that produced it. Historical scores remain queryable under their original version via the `?version=X` parameter on this page and via the `methodology_version` field in the API. Partners consuming the API receive `methodology_version` and `calculated_at` fields on every score response.

Limitations and Edge Cases

UGC platforms (Roblox, Fortnite Creative). LumiKin scores individual experiences on UGC platforms as well as the platform itself. The platform score reflects the host environment (moderation quality, monetization model, stranger interaction risk). Individual experience scores reflect the specific content. Consumers of the API should be aware that a "Roblox" platform score is not a score for Roblox experiences as a category — individual experience scores vary widely and should be queried separately.

Live-service games with evolving content. Games with weekly content updates, seasonal events, or ongoing battle passes are scored on their current state at review time plus any persistent structural mechanics. Time-limited events are scored under R1 (FOMO) but the content of those events may not be individually reviewed. Live-service titles are prioritised for periodic rescore when mechanics change materially.

Early access games. Early access titles are scored at the state available at review time. Scores may change substantially on full release. LumiKin does not guarantee rescore coverage for all early access titles on release.

Regional content differences. Some games ship with different content ratings by region (e.g., a game rated M in North America may be rated 16+ in Europe with modified content). LumiKin scores the international/default release where possible. Regional variants are noted where known but are not systematically tracked.

AI-generated scores. Some scores are produced by an automated review pipeline using the LumiKin rubric as a prompt. These are marked `review_tier: automated` in the API. Expert and community reviews are marked accordingly. Automated scores undergo periodic spot-check auditing but have not been individually verified by a human reviewer.

Changelog

VERSION	DATE	SUMMARY
1.1 (current)	2026-05-01	Age-Floor Policy formalised. R4.5 (fear/horror) added as a third age-floor dimension alongside R4.1 (violence) and R4.2 (sexual content). Fear floors: 0→0, 1→7, 2→10, 3→13 (capped at 13 — intense horror alone does not warrant an adult floor). Context modifiers (`trivialized` , `defenceless_target` , `mixed_sexual_violent`) introduced for the violence and sexual dimensions. Triggered by divergence analysis: psychological-horror titles such as *What Remains of Edith Finch* (R4.5 = 2) were systematically underaged under v1.0 because fear was scored but ignored in floor computation. Games with `fearHorror ≥ 1` may receive a higher `recommendedMinAge` than under v1.0; BDS, RIS, LumiScore, and time recommendations are unchanged.
1.0	2026-04-26	Initial published methodology. All weights, dimensions, and tier thresholds as described in this document.

LumiKin

Game ratings grounded in child development

BROWSE COMPARE METHODOLOGY FAQ PRESS PRIVACY TERMS RSS

GB EN