

AI-supported startup evaluation and regional ecosystem context: sequential judgement revision in early-stage startup assessment

1. The question

Early-stage startup evaluation requires investors, accelerators, and funding organisations to allocate limited resources efficiently despite being prone to incomplete information and uncertain outcomes. Pre-seed and seed-stage ventures often lack established financial histories and verifiable performance benchmarks, which increases the reliance on qualitative signals and judgement. AI-supported screening tools may address these bottlenecks through structured summaries, while regional ecosystem information offers context. Existing research examines early-stage investment judgement, human interaction with AI-supported decision systems, and regional entrepreneurial ecosystems as separate literature streams. Little evidence exists on how professional evaluators revise their startup judgements when an AI-generated evaluation summary and regional ecosystem context are presented sequentially. This thesis asks: *To what extent do professional evaluators revise their judgements of early-stage Dutch startups after sequentially reviewing an AI-generated decision-support summary and standardised NUTS-2 regional ecosystem context?*

2. The approach

The study used a fixed-order, repeated-measures survey design. Evaluators assessed Philips Innovation Award 2023 cases across three stages: the PHIA Community Report, the

Kuanta.ai executive summary, and a standardised NUTS-2 ecosystem context page. After each stage, evaluators reported investment recommendation, perceived probability of follow-on funding within 12 months, and confidence. The sample contained 30 evaluators, 81 evaluator-startup dyads, and 28 startup cases, comprising 43 pre-seed and 38 seed-stage dyads. Confirmatory analyses used adjacent-stage change scores, two-tailed one-sample t-tests with Holm-adjusted p-values, and robustness analyses using Wilcoxon and mixed-effects models.

3. What the results show

Within the fixed sequence, all three outcomes changed after the AI-summary stage. Investment recommendation decreased by 0.37 scale points, 95% CI [-0.63, -0.11], Holm-adjusted $p = .011$, $d_z = -0.32$. Perceived follow-on funding probability decreased by 4.14 percentage points, 95% CI [-7.69, -0.58], Holm-adjusted $p = .023$, $d_z = -0.26$. This finding was less robust because the mixed-effects estimate was not statistically significant, $p = .074$. Confidence increased by 0.67 scale points, 95% CI [0.46, 0.88], Holm-adjusted $p < .001$, $d_z = 0.70$, and was the largest change. The recommendation and confidence findings were supported across the robustness analyses. By contrast, the NUTS-2 ecosystem stage produced no detectable average change across the outcomes; H1b, H2b, and H3b were not supported.

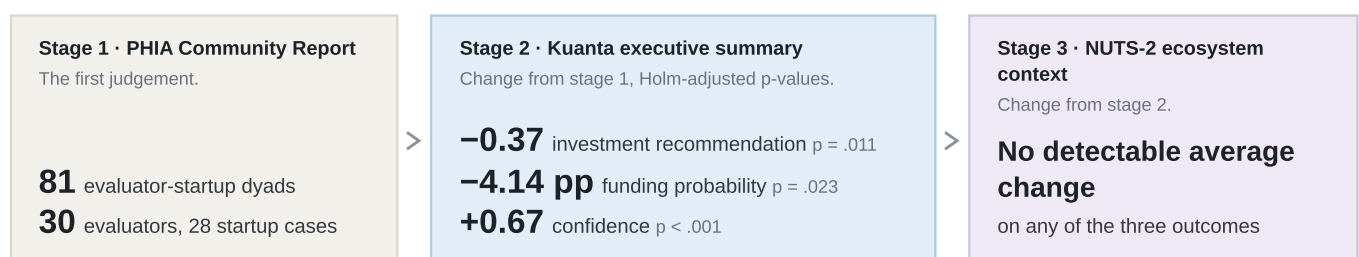


Figure 1. Judgements moved after the AI summary, not after the ecosystem context. Evaluators rated each case after every stage, in a fixed order. The funding-probability change was not significant in the mixed-effects model ($p = .074$).

Open-ended responses provided descriptive self-attributed context. Although ecosystem information was mentioned more often than the AI artefact overall, the AI artefact was deemed the primary influence in 28 dyads, compared to 12 for ecosystem information. Regional information was often acknowledged but less frequently considered decisive in judgement change. These self-attributions cannot directly establish information use.

The exploratory analysis examined alignment with qualifying public follow-on funding within 12 months. This was treated as limited near-term funding evidence, not as startup success or decision accuracy. The Kuantia.ai score and post-AI evaluator estimates showed positive but statistically inconclusive associations. Although the DeLong comparison was not statistically significant, $p = .728$, baseline discrimination was descriptively higher, $AUC = .732$, than post-AI discrimination, $AUC = .694$. This contrasted EH3's predicted direction. Post-ecosystem estimates showed no detectable improvement. These findings were imprecise because only 27 to 29 cases were available.

Overall the AI-supported information stage was associated with more conservative assessments and greater confidence. Greater confidence does not establish improved accuracy or decision quality. Similarly, the ecosystem-stage null does not imply the irrelevance of regional context. The information may have been too broad, varied insufficiently, or been constrained to the third-stage position. The fixed sequence prevents these explanations from being separated.

4. What it means in practice

The thesis contributes by measuring judgement change within professional evaluators and startup cases across sequential information stages. It treats the AI-generated

summary as an input into human judgement, separating this into investment recommendation, funding probability, and confidence. The divergence between evaluators' more conservative assessments and greater confidence suggests that structured AI support can affect judgement direction and subjective certainty differently, without demonstrating improved outcome alignment. The contrast between information stages furthermore implies that decision relevance may depend on the level of analysis at which information is presented. However, this cannot be isolated fully in this study.

Organisations using AI-supported screening tools should monitor how scores, visualisations, SWOT analyses, and structured summaries shape judgement and confidence. Regional information may need to be more specific to a startup's sector, market, or operating environment before becoming decision-relevant. These tools should be used as support rather than as replacements for professional judgement.

The central conclusion of this thesis is that decision support does not become valuable simply by adding more information. Its influence depends on how information is structured, where it enters the evaluation process, and how evaluators interpret it. Thus, the findings shift the question from whether AI can evaluate startups to how AI reshapes human judgement. The findings also suggest the need for startup-specific ecosystem context to become decision-relevant. Tools that alter decisions and confidence must therefore be designed as contestable, specific inputs to expertise instead of silent substitutes for it.

Originally submitted in 2026 as an MSc thesis in partial fulfilment of the MSc Business Information Management. This publicly shared version is the independent work of the author and is not an official publication of, nor endorsed by, any university. Limitations include the fixed order without counterbalancing or a control condition, the limited professional sample, and the restricted variation and broad geographical level of the ecosystem information. Caution with results also applies due to the small exploratory case sample, and the possibility that public sources omitted undisclosed financing.

Source: L. Rau (2026). *AI-supported startup evaluation and regional ecosystem context: sequential judgement revision in early-stage startup assessment*. Master's thesis, Rotterdam School of Management, Erasmus University.