Heterogeneous Effects of Theory-Based Entrepreneurship Training

A Systematic Framework

Helena Montoya Calero · April 2026

The question. Theory-based entrepreneurship training works, on average. The literature has documented this across multiple RCTs and contexts. What remains open is a systematic account of *for whom* it works and *why*. Existing papers study one moderator at a time, without a framework that justifies why that moderator and not others. This paper proposes such a framework.

The gap. The heterogeneity literature on entrepreneurship training has two limitations. First, moderators have been studied in isolation: startup development stage (Novelli & Spina 2024), country setting (Agarwal et al. 2025), environmental uncertainty (Gambardella & Messinese 2025). No paper organizes these and other candidates into a coherent taxonomy. Second, individual-level moderators, the cognitive and experiential traits that entrepreneurs bring to training, have not been examined at all. McKenzie, Woodruff et al. (*VoxDevLit*, 2025) identify this as the primary open question in the field: "There is a lot of heterogeneity in both samples and results, so there is less guidance on which groups benefit most."

The framework. I propose organizing heterogeneous effects into three categories:

Category	Question	Key variables
Person-level	Who is the entrepreneur?	SDM capability, ambiguity aversion, prior experience, education, learning orientation
Venture-level	What is the startup?	Business model development stage, sector, team composition, idea scope
Dosage	How much did they absorb?	Attendance, sessions completed, mentor engagement

These categories are not arbitrary. Person-level heterogeneity speaks to whether individual cognitive traits shape learning from structured training. Venture-level heterogeneity captures whether the startup's stage and structure make theory-based reasoning more or less applicable. Dosage speaks to the role of treatment intensity, rarely studied in this literature despite its policy relevance.

Method. For each category, I pre-specify the key moderators and estimate heterogeneous treatment effects via:

$$Y_{it} = \alpha + \beta_1 T_i + \beta_2 M_i + \beta_3 (T_i \times M_i) + \mathbf{X}_i'\gamma + \delta_r + \varepsilon_{it}$$

where T_i is the randomized treatment, M_i is the moderator, $\mathbf{X}_i$ are controls, and δ_r are RCT/country/cohort fixed effects. Standard errors clustered at the cohort level. A data-driven complement using recent ML-based HTE methods (Chernozhukov, Demirer, Duflo & Fernandez-Val, *Econometrica* 2025) is also feasible given the pooled sample size.

Contributions. This paper reframes the heterogeneity question in entrepreneurship training from "does X moderate the effect?" to "what categories of moderators matter, and why?" The taxonomy is replicable: future RCTs can build it into their design from the start. For the external literature, it provides a systematic answer to the gap McKenzie et al. document. For the lab, it lays the groundwork for designing the next generation of experiments with heterogeneity in mind.