AI & EDUCATION

From Generation to Verification

Chuck Eesley | Stanford STVP

Building on STVP Traditions in Entrepreneurship & Technology

The STVP Foundation We Build On

Entrepreneurial Mindset

Experiential learning, venture formation, getting out of the building

Technology & Society

Understanding how technology transforms industries and shapes human outcomes

Rigorous Research

Evidence-based approaches to entrepreneurship, not just storytelling

Steve Blank, MS&E 25th Anniversary: *AI is transforming how we build and test business models — the very foundation of what we teach.* ▶ Video clip

The Extraordinary Promise

+34%

Quality improvement for consultants using AI on tasks inside the frontier

Dell'Acqua et al. 2026

+12%

More tasks completed, 25% faster with significantly higher quality

Dell'Acqua et al. 2026

r = 0.52

Correlation between AI and VC evaluations of business plans

Csaszar et al. 2024

AI can genuinely augment human performance. But the full picture is more complex.

The Jagged Technological Frontier

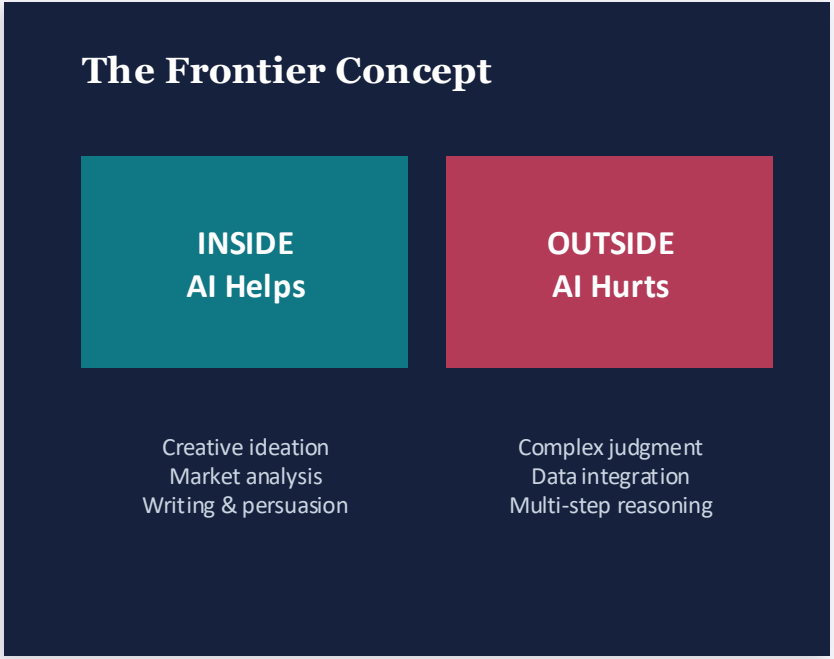
Dell'Acqua, Mollick, Lakhani et al. (2026) — Organization Science

758 BCG consultants in a controlled experiment

Inside the frontier: +34% quality, +12% tasks completed, 25% faster

Outside the frontier: −19% correct solutions — AI actually degraded performance

Key insight: Tasks that look equally complex to humans can fall on opposite sides of the frontier



Bottom-half performers gained +31%, but top-half still gained +11%. AI as equalizer — within the frontier.

The Uneven Impact

Otis, Clarke, Delecourt, Holtz & Koning (2024) — RCT with 640 Kenyan Entrepreneurs

+15%

High Performers

Asked targeted questions. Implemented tailored, specific changes. Used AI to discover solutions unique to their context — like finding alternative power sources during outages.

−8%

Low Performers

Implemented generic advice — lowered prices, spent on advertising — that was inappropriate for their specific business situation. The AI advice actually hurt them.

The takeaway: The gap isn't in the quality of AI advice — it's in who can select and implement the right advice. This is a mechanism design problem, not a content problem.

AI as Strategist

Csaszar, Ketkar & Kim (2024) — Strategy Science

Strategy Generation

LLM-generated business plans rated 0.14 SD higher than entrepreneur-written plans. Evaluators 5% more likely to recommend acceptance.

Strategy Evaluation

LLM evaluations correlated r = 0.52 with experienced VC and angel investors across 138 real business plans from a startup competition.

Predicting Outcomes

Prospective Kickstarter tournament: LLMs predicted which live projects would succeed at r > 0.60, outperforming experienced human evaluators (r = 0.04–0.45). AI may be better at strategic foresight than we are.

Unbounding Rationality

AI changes three cognitive processes: search (speed), representation (re-imagining strategy), and aggregation (virtual crowds). Relaxes bounded rationality.

The Human-AI Oversight Paradox

Jacqueline Ng Lane et al. (2024) — Harvard Business School

19%

more likely to align with AI recommendations when AI provided narrative explanations.

228 evaluators screening 48 early-stage innovations across 3,002 decisions. Narrative AI increased reliance on AI rather than strengthening human judgment.

The Paradox

Explanations were supposed to help humans override bad AI calls. Instead, persuasive narratives made people more compliant — even when AI recommended rejection of potentially transformative innovations.

The Implication

AI-generated narratives may inadvertently filter out breakthrough innovations that don't fit standard frameworks. More explanation doesn't equal better oversight — it can equal less.

AI Across the Startup Journey

Where AI accelerates — and where it misleads

Ideation

INSIDE

✓ AI helps

Rapid brainstorming, concept variation, recombination

⚠ Watch out

Novelty vs. feasibility judgment; may recycle known ideas

Customer Discovery

FRONTIER

✓ AI helps

Survey drafts, interview synthesis, persona generation

⚠ Watch out

Fabricated insights, hallucinated user needs

Prototyping

INSIDE

✓ AI helps

Code generation, UI mockups, MVP acceleration

⚠ Watch out

Architecture decisions, edge-case reasoning

Business Model

FRONTIER

✓ AI helps

Canvas generation, pricing analysis, comp research

⚠ Watch out

Market sizing accuracy, unit economics judgment

Fundraising

FRONTIER

✓ AI helps

Deck drafts, financial models, predicting investor interest & simulating Q&A (Csaszar et al.)

⚠ Watch out

VC relationship nuance, strategic narrative

Scaling

FRONTIER

✓ AI helps

Ops automation, support bots, ad optimization

⚠ Watch out

Culture, hiring judgment, regulatory compliance

The jagged frontier runs through every stage. The question isn't whether to use AI — it's knowing which side you're on.

The Autonomous Company Experiment

EY

Elizabeth Yin

General Partner, Hustle Fund | Paperclip (open source)

The Good

- Multiple AI agents (content, SEO, dev) coordinate out of the box
- Set the vision as a "board member" — agents execute
- No company culture or morale concerns (they work 24/7)
- Dramatically less work than manual company building

The Reality

- "Autonomous" still requires daily management and direction
- SEO agent was bad at SEO — improved only after Yin intervened
- Built a lot, but no one came to the website
- No sales agent. Building isn't the hard part — distribution is.

Yin's verdict: *"While not ready for prime time today, Paperclip shows where we're headed. The technology is advancing faster than I imagined."*

The zero-human company wave: Paperclip crossed 42K GitHub stars. Polsia: one founder, 3,812 AI-operated companies, $3.6M run rate. KPMG tested an AI-only webshop with an agent CEO. 2025 was the year of the AI employee — 2026 is the year of the AI company.

The One-Person Billion-Dollar Company

MG

Matthew Gallagher, 41

Founder, Medvi | Erin Griffith, NYT, April 2, 2026

$20K	2 mo	$401M	$1.8B
initial investment	to launch	Year 1 revenue	2026 projected

The AI Stack

Code: ChatGPT, Claude, Grok
Ads: Midjourney + Runway for images/video
Support: ElevenLabs voice + custom AI agents
Ops: CareValidate + OpenLoop handle doctors, Rx, shipping

The Verification Gap

- Chatbot made up prices — Gallagher honored them
- 1,000+ customer calls routed to his cell phone
- Hired 7 human account managers to build trust AI can't
- vs. Hims & Hers: 3x the margin with 1/1,000th the staff

The question for educators: AI made generation trivial. But Gallagher still needed human judgment to catch AI errors, human relationships for trust, and mechanism design to know what to outsource. Those are the scarce skills.

The In-Box Congestion Crisis

We're teaching the wrong lesson about AI in entrepreneurship

Generate a slide deck. Generate a business model. Generate 1,000 bespoke cold DMs. When the marginal cost of personalized outreach drops to zero, its value drops to zero. We've turned the ecosystem into a high-speed noise machine.

Lean Startup

Blank

Replace assertion with evidence

Get out of the building

→

Disciplined E'ship

Aulet / JetPack

Accelerate evidence gathering with AI

Move faster through the steps

→

The Next Step

Mechanism Design

Make evidence harder to manufacture

AI as verifier, not generator

Why "Cheap Talk" Is the Core Problem

Spence's Signaling Insight

A signal is only credible if it is costly to fake. Almost everything we teach founders to produce — the pitch, the persona, the discovery summary, the MVP demo — has become cheap talk. Not because founders are dishonest, but because AI has exposed the flaw at scale.

Before AI

Motivated reasoning was human-scale. A founder does 15 interviews, gets ambiguous signals, and reports "strong early validation." Confirmation bias, but bounded.

After AI

Agentic AI industrializes motivated reasoning. The synthesis is cleaner, the narrative more coherent, and the gap between what customers said and what the deck concludes has never been easier to paper over — without any intent to deceive.

What We Teach Instead

01

Costly Signal Design

Teach founders to build validation around evidence that's hard to fake: signed LOIs, reallocated budgets, co-development agreements requiring real commitment.

02

Separate Synthesis from Evidence

Present raw data separately from interpretation. Use AI as an auditor that surfaces contradictions, not a synthesizer that smooths them over.

03

Adversarial Simulation

Before real interviews, AI stress-tests assumptions by steelmanning every objection. 50 adversarial AI sessions before the first real customer conversation.

04

Mechanism Design Thinking

Who has incentive to tell the truth, and under what conditions? Structure interactions so a "yes" means something.

The Fundamental Shift

AI as Generator

- Creates polished narratives that mask uncertainty
- Industrializes motivated reasoning at scale
- Increases in-box congestion for investors and customers
- Widens the gap between perception and reality

AI as Verifier

- Surfaces contradictions in customer evidence
- Stress-tests assumptions adversarially
- Identifies costly signals vs. cheap talk
- Makes the truth easier to tell than to obscure

Trust is a mechanism design problem — not a content generation problem.

Where STVP Leads Next

In Our Curriculum

Building on Eesley's entrepreneurship education research, we're now incorporating AI-driven ideation into our refugee entrepreneurship program in Uganda. Hannah & O'Mahony (AMLE 2026) confirm e-ship is a 'field in flux.' Teach AI as auditor, not just assistant — mechanism design meets verification thinking.

In Our Research

Continue the rigorous experimental tradition. Katila & Li are using AI to predict startup outcomes. Our current PhD students are pursuing these questions — not just whether AI works, but when, for whom, and under what conditions.

In Our Community

STVP alumni and supporters are uniquely positioned — as founders, investors, and educators — to pioneer verification-first approaches in your own organizations. Apply these insights now.

The future of entrepreneurship isn't about being the loudest.

It's about being the most verifiable.

Chuck Eesley | Stanford STVP

ceesley.substack.com | @cee