

Innovative Culture in the Age of AI

EXECUTIVE SUMMARY

May 7, 2026



Table of Contents

- Welcome and Opening Remarks | p. 3
- Executive Connection & Framing | p. 5
- Inside Pixar: A Case Discussion | p. 7
- Measuring Innovative Culture | p. 9
- Tactics and Practices to Create a Balanced Culture | p. 13
- The Paradoxes of Innovative Cultures | p. 17
- Q&A with Gary Pisano | p. 21
- Extending These Ideas | p. 23
- Closing Remarks | p. 25

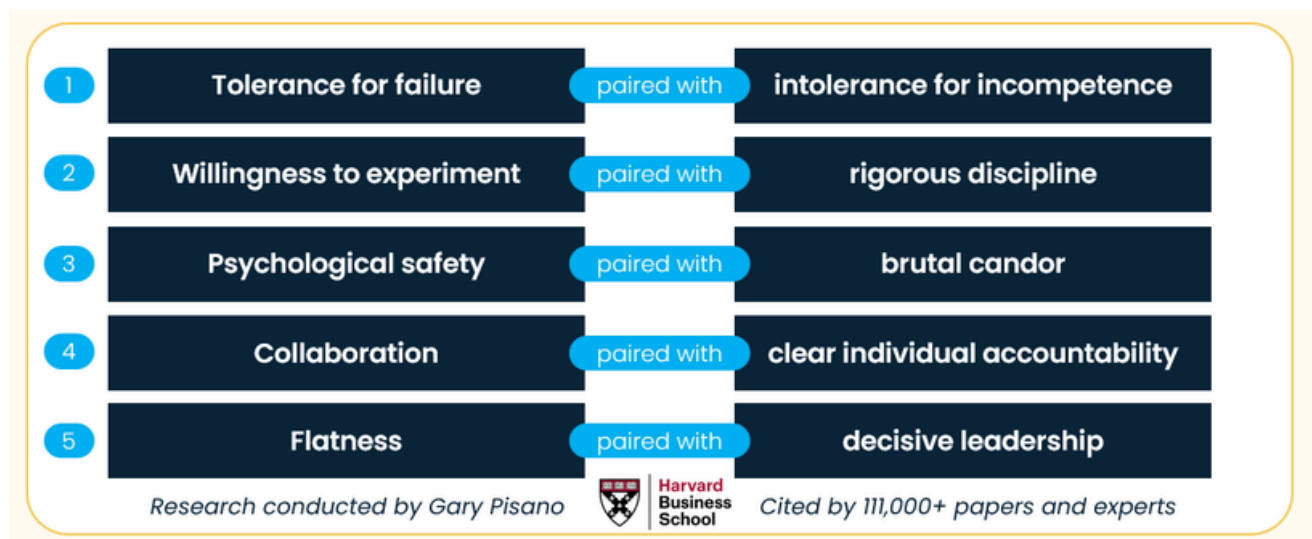
Welcome and Opening Remarks

Michael McCarroll, SVP & GM, Teamraderie

Michael framed the day around a paradox every leader is living: boards demanding speed and discipline, CEOs demanding experimentation and accountability, investors demanding transformation and predictability. The friction leaders feel inside their organizations—exciting pilots that never scale, AI initiatives stalled in cross-functional committees—is not a talent problem and not an investment problem. It is a problem of operating norms built for one era trying to function in another.

He anchored the urgency with three data points. MIT: ~50% of corporate AI budgets flow to sales and marketing use cases, but the strongest measurable returns are in back-office functions. McKinsey's gen AI paradox: 80% of companies invest in AI; almost none report material impact on earnings. Teamraderie research: ~40% of stage-one stalled AI initiatives are stalled because no single person owns the next step.

He introduced the day's structure—see it, measure it, practice it, scale it, start it—and one ground rule: the unit of focus is the team, not the individual. Everything of significance in an organization is done by teams.



Welcome and Opening Remarks

Michael McCarroll, SVP & GM, Teamraderie

Key Takeaways

1. The Mandate Is Paradoxical

Adopt AI faster than competitors while preserving the standards and judgment that make the company great. Both halves are required.

2. AI Surfaces Existing Problems

AI does not create new cultural problems; it makes existing ones more urgent and more expensive.

3. The Numbers Show the Stall

McKinsey: 80% of companies invest in AI; <10% see bottom-line impact.
~40% of stalled AI initiatives stall because no one owns the next step.

4. The Team Is the Unit

Everything of significance in an organization is done by teams. The day's focus is on changing team behavior, not individual mindset.



“The mandate is two things at once. The friction you feel is not a talent or investment problem—it’s operating norms built for one era trying to function in another.”

Executive Connection & Framing

Anja Svetina Nabergoj, Stanford University

Anja set the tone with a deliberately disorienting warm-up. After a quick table exchange, participants paired with someone they didn't know and counted 1-2-3 alternately. Then 1 became a clap. Then 1 and 2 became a clap and snap. Each round, more people stumbled—and each round, the instruction was to keep going rather than apologize, freeze, or laugh it off.

The exercise made the day's first tension viscerally real: before a team can experiment with discipline, it has to be willing to experiment at all. Willingness gets blocked by fear of failure, fear of embarrassment, and the evolutionarily wired stress response that treats a missed snap like physical danger.

Changing a team's relationship with failure is a prerequisite, not a nice-to-have. The session set up the entire day's work: experiential learning sticks; the team has to feel the tension before they will change behavior.

The Science of Failure

Stress system evolved to help us survive physical danger.

In humans it can also activate around social meaning: embarrassment, judgment, exclusion, and loss of status.

Failure can feel threatening because our brain treats social risk as real risk.

Robert Sapolsky, Stanford University

**Stanford
d.school**

[Anja Svetina Nabergoj] Hasso Plattner Institute of Design at Stanford. Shared with Creative Commons 4.0 license.

Executive Connection & Framing

Anja Svetina Nabergoj, Stanford University

Key Takeaways

1. Willingness Precedes Discipline

Before a team can experiment with discipline, it has to be willing to experiment at all.

2. Failure Is Evolutionarily Wired

Stress responses to embarrassment are hardwired—even a low-stakes counting game triggers them.

3. Change the Response, Not the Risk

Pre-train the recovery: keep going when you mess up. Strategy is to absorb failure, not avoid it.

4. Experiential Beats Conceptual

Teams change their relationship with failure faster through felt experience than through frameworks.



“Notice what you do when you mess up—and keep going anyway. Willingness to experiment is the prerequisite that has to be trained, not assumed.”

Inside Pixar: A Case Discussion

Francesca Gino, Behavioral Scientist

Francesca led the room through a case discussion built around five video clips of Ed Catmull, Pixar's co-founder, reflecting on how Pixar built and sustained an innovative culture. The discussion was structured to surface the paradoxes of the framework rather than describe them—participants experienced candor, plus-ing, and the deliberate flattening of authority that the Brain Trust depends on.

Key moments from the clips: a hidden second class of production managers on Toy Story that Catmull had to identify and address; an American Dog story trust meeting where the chief creative officer's positive opening shut down all candor; the Brain Trust as peer-to-peer, not boss-to-peer; the principle that the meeting must be safe for the most junior person to speak; and Up's original pitch being unrecognizable from the final film, illustrating tolerance for failure paired with intolerance for incompetence.

The discussion also surfaced what makes this hard to copy. Underlying beliefs have to be made explicit. Candor is about the problem, not the person. "Sham participation"—asking for input on a decision already made—is the fastest way to manufacture cynicism. Pixar Academy trained managers explicitly on listening, plus-ing ideas, and substantive judgment with curiosity. The capability was institutionalized, not assumed.

Inside Pixar: A Case Discussion

Francesca Gino, Behavioral Scientist

Key Takeaways

1. A Hidden Second Class

On Toy Story, Pixar managers quietly felt second-class. The leader's job is to identify what's been missed, including their own role in missing it.

2. Candor Is the System

The American Dog meeting failed because no one would speak after the chief creative officer led with positive feedback. Candor is not optional; it is the operating system.

3. The Brain Trust

Peer-to-peer not boss-to-peer. Authority stays quiet for the first 10–15 minutes. No decisions made in the room. The director still owns the film.

4. Safety for the Least Powerful

Every meeting must be safe enough for the most junior person to speak—and interns are explicitly invited to do so.

5. Tolerance for Failure, Intolerance for Incompetence

Up's original pitch was unrecognizable from the final film. Directors get replaced when needed. Failure is celebrated, sloppy work is not.



“It has to be safe for the least powerful person in the room to talk. When you get that right, it leads to a healthy culture.”

— Ed Catmull, Pixar

Measuring Innovative Culture

Francesca Gino, Behavioral Scientist

The cohort's pre-event scan results were used to demonstrate how a team would actually use the assessment. The framework measures pairs of cultural attributes—one from each side of each tension—and reports the level of agreement on each dimension. The most useful question is not “which bar is lowest?” but “where is the gap between the two halves of a tension widest?”

Working benchmark: scores below 20% warrant attention; scores above 40% indicate strength. Across the broader Teamraderie dataset, three patterns surface consistently. Psychological safety often outpaces candor—teams interpret “safety” as “be nice,” so candor only happens in the meeting-after-the-meeting. Strong leadership often reads low because of unclear ownership; when teams don't know if a meeting is brainstorm or decision, they read it as weak leadership. Willingness to experiment without discipline produces motion without learning.

The recommended approach: take the scan, debrief with Francesca to identify the one or two tensions where the gap is widest and the business cost is highest, then run targeted interventions on those specific behaviors. Innovation looks different in different teams—shop-floor cleanup, customer-form redesign, AI adoption—but every team needs experimentation, decision-making, and candor.

Measuring Innovative Culture

Francesca Gino, Behavioral Scientist

Key Takeaways

1. Working Benchmarks

Scores below 20% warrant attention. Above 40% indicate strength. The gap between paired dimensions is the diagnostic.

2. Psychological Safety Often Outpaces Candor

Teams interpret “safety” as “be nice,” producing meetings where candor only happens after the meeting.

3. Strong Leadership Reads Low When Ownership Is Unclear

When teams don't know if a meeting is brainstorm or decision, they read it as weak leadership.

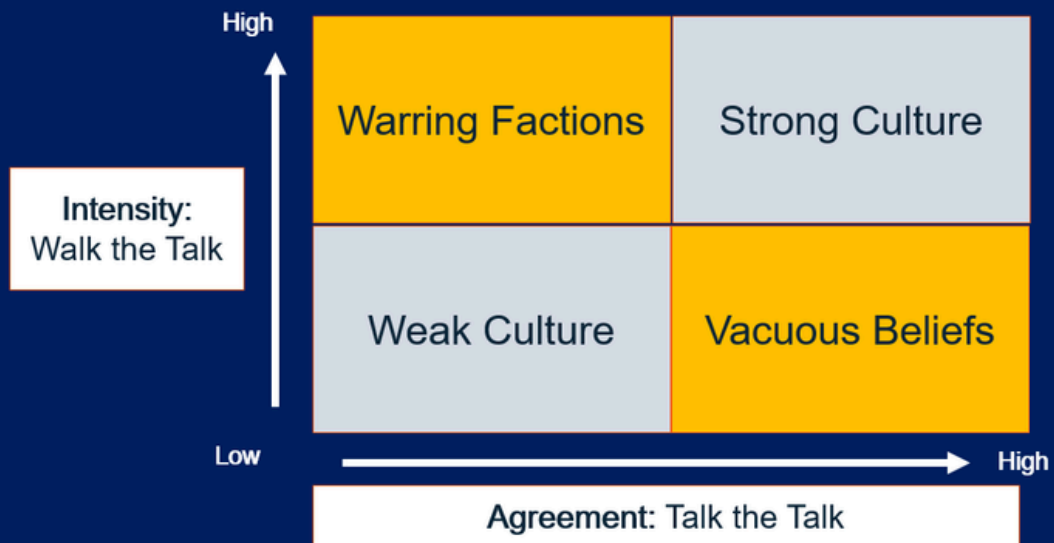
4. Discipline Is the Tractable Gap

In the cohort data, willingness to experiment was decent; discipline was weak. The gap is solvable.

5. Targeted Beats Broad

Pick the one or two tensions with the widest gap and highest business cost —run targeted interventions on those behaviors.

What makes organizational culture strong?



Lesson 1

Culture won't help unless it's strategically relevant

Lesson 2

Everyone has to uphold the culture... even when:

- It's hard
- It is not absolutely necessary
- No one is looking

Survey: Assessing Specific Cultural Attributes*

- Tolerance for Failure
 - Willingness to Experiment
 - Collaborative
 - Psychological Safety
 - Flat/Empowered
- Intolerance for Incompetence
 - Disciplined
 - Individual Accountability
 - Candor
 - Strong Leadership

* Method: Multiple Questions Per Attribute

Summary of Method

Step 1: Normalize responses (re-order so that 1 corresponds to low performance and 7 corresponds to high performance)

Step 2: For each question, sum % of responses in category 6 and 7 to create a question level score.

Question level score (e.g. question 1) = % of responses in cat 6 + % of responses in cat 7

Step 3: Average question level scores across all question corresponding to a particular cultural attribute (e.g. tolerance for failure)

Cultural Attribute Score (e.g. Tolerance for Failure) = Average (Q1 score, Q2 score, Q3 score)

Step 4: Analyze data (if available, also analyze by group)

Tactics and Practices to Create a Balanced Culture

Anja Svetina Nabergoj, Stanford University

Anja ran abbreviated versions of two of the team experiences participants will deploy with their own teams: Failing Smart (experimentation discipline) and a Decision Ownership demo. Both rest on the principle that experiential learning sticks—the team has to feel the tension before they will change behavior.

Failing Smart: tables shared a recent experiment—product, process, or way-of-working—and stress-tested it against four questions: Was it clear what you were trying to learn? Could you have learned it any other way? What told you whether it worked? Who else needs to know what you learned? Most experiments break on questions 1 or 2. The reframe: experiments are not “successful or failed”—they are productive or unproductive. A failure that produces clear learning is productive.

Decision Ownership: tables ran the same hiring decision twice. Round 1 had no decision owner and required consensus in four minutes. Round 2 had one decision owner (the person closest to the door); everyone else advised. The difference was visible. In Round 1, people leaned back and ran out of time. In Round 2, posture shifted, the owner felt the weight of the choice and asked sharper questions, and devil’s-advocate inputs landed differently because the role was clear. The takeaway tool: name the owner, confirm authority, actively pull in collaboration, set a no-later-than date.

Tactics and Practices to Create a Balanced Culture

Anja Svetina Nabergoj, Stanford University

Key Takeaways

1. Productive vs. Unproductive Experiments

The reframe: not “successful or failed”—productive or unproductive. Failure that produces clear learning is productive.

2. Most Experiments Break Early

Tested against four questions, most break on “What were we trying to learn?” or “Could we have learned it another way?”

3. Decision Reality

Only ~20% of orgs rate themselves good at decisions. Respondents spend ~40% of time deciding. >50% rate that time ineffective.

4. Posture Reveals Ownership

In the hiring exercise, Round 1 (consensus) leaned back. Round 2 (one decision owner) leaned in.

5. Ownership Diagnostic

Is the owner named? Do they have authority? Have they pulled in collaboration? Is there a no-later-than date?



“The team has to feel the tension before they will change behavior. The framework comes after the experience, not before.”

Productive vs Nonproductive Experiment

1. Was it clear what we were trying to learn? ✓
2. Could we have learned it another way? ✓
3. Was the cost worth the answer? ✓
4. Who else needs to know what we learned? ✓

Adapted from Thomke (2003), Pisano (2019), Edmondson (1999), and Ries (2011).

[Anjo Svetina Nobergo] Hasso Plattner Institute of Design at Stanford. Shared with Creative Commons 4.0 license.

Stanford
d.school

Design

→ Run →

Debrief

Before Experiment

1. What do we want to learn?
2. How small can we keep it?
3. What would tell us it worked (or didn't)?

Run the experiment

~ 2 weeks

The team runs the experiment.

After Experiment

1. What did we actually learn?
2. Move forward, tweak it, or drop it?
3. Who else needs to know what we have learned?

[Anjo Svetina Nobergo] Hasso Plattner Institute of Design at Stanford. Shared with Creative Commons 4.0 license.

Stanford
d.school

The Ownership Diagnostic

1. Can you name the one person who owns it?
2. Does that person have the authority to make it (or do they need someone's approval)?
3. Have they actively pulled in the input they need (or are they waiting for everyone to agree)?
4. Is there a date by which it will be made, no matter what?

Adapted from Rogers & Blenko, "Who Has the D?" (HBR 2006); Decide & Deliver (Rogers, Blenko & Mankins, 2010); McKinsey (2019).

[Anja Svetina Noberga] Hasso Plattner Institute of Design at Stanford: Shared with Creative Commons 4.0 license.

**Stanford
d.school**

Debrief

Round 1:

CONSENSUS

Everyone has to agree

4min

Round 1: How did it feel in your body?

Round 2: What was different?

What did the decision owner do that consensus group couldn't?

Round 2:

DECISION OWNER

One person decides.
Everyone else gives input.

4 min

[Anja Svetina Noberga] Hasso Plattner Institute of Design at Stanford: Shared with Creative Commons 4.0 license.

**Stanford
d.school**

The Paradoxes of Innovative Cultures

Gary Pisano, Harvard Business School

Gary opened with a question he asks senior leaders: when do you do culture? The answer is almost always “all the time.” Then he asked when they do budgeting—and got specific dates, processes, and deliverables. The point: culture is the only major business activity most organizations do not manage with a cadence. Until that changes, it stays squishy.

His working definition: culture is the expression of values through accepted and expected behavior. Values are hard to observe; behaviors can be measured. From running the five-tension scan across dozens of companies globally, Gary surfaced consistent patterns. Candor is the scarcest organizational resource—universally low. High psychological safety often masks low candor, with people “buying themselves safety” by not being candid. Tolerance for failure is constantly confused with tolerance for incompetence. Accountability gets read as threat when it should be read as ownership.

He walked through a healthcare client’s puzzle: a self-described risk-averse organization that kept having failed product launches. Tracing the data through the five tensions revealed that consensus decision-making syndicated risk, senior leaders ended up vetting and overriding everything, low candor meant real risks were never surfaced. Hence the paradox.

Gary’s template for change: identify behaviors to stop, identify behaviors to start, pick one or two, build mechanisms (e.g., project reviews dedicated only to risks). On scaling, his analogy was deliberate: the institutions best at propagating beliefs across millions of people for centuries are religions, and they all run the same playbook—prophets and disciples, stories and parables, designated days and rituals, coaching and training, institutionalized practices.

The Paradoxes of Innovative Cultures

Gary Pisano, Harvard Business School

Key Takeaways

1. Culture Is the Unmanaged Capability

Budgets, forecasts, and ops reviews have specific times and processes. Culture is treated as continuous and therefore drifts.

2. Candor Is the Scarcest Resource

Universally low across geographies. People avoid it because of rudeness, personal stakes, lack of accountability, or political cost.

3. The Risk-Averse Paradox

A risk-averse healthcare client kept failing launches: low candor buried real risks; consensus syndicated risk; senior leaders vetted everything.

4. Cultures Cascade Like Religions

Prophets, parables, designated days, rituals, coaching. You cannot decree culture from on top. Pick one or two behaviors to stop and start.



“You can’t get good at manufacturing quality without an operations team reviewing yields. You can’t get good at culture without the same cadence.”

Culture is the expression of values through accepted and expected behaviors

VALUES  **BEHAVIORS**

We tolerate failure in order to learn

People seek feedback early on **preliminary** ideas
Failed ideas morph into improved ideas
Experiment with multiple options early
We conduct systematic post-mortems after failures

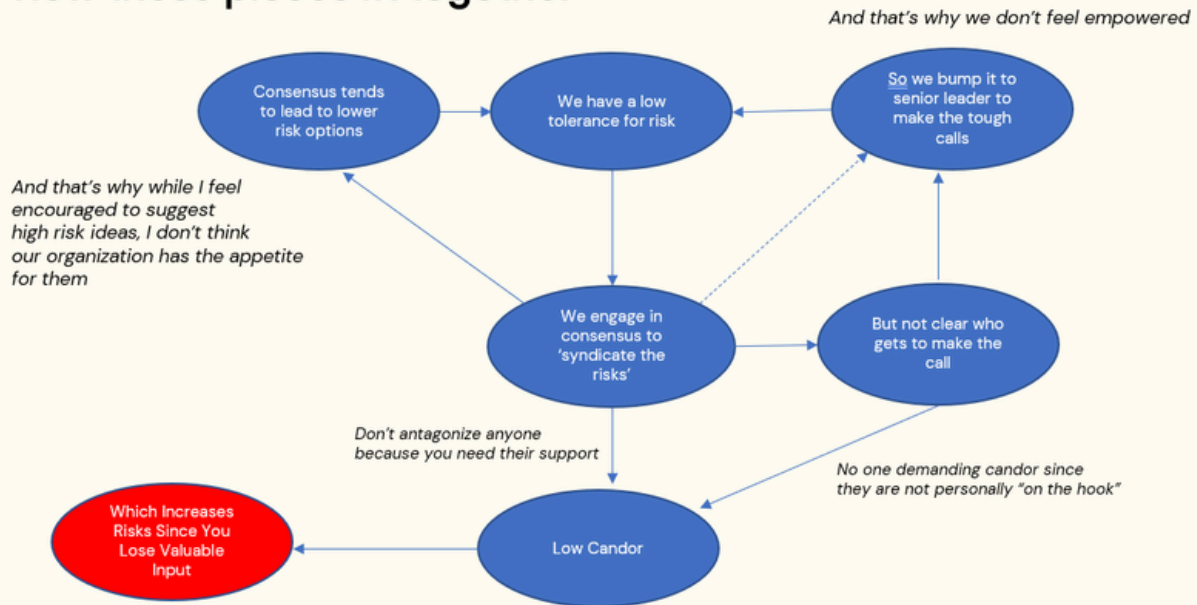


We can measure if these behaviors are present

Patterns Across Dozens of Companies, Across the World, and Across Industries

- **Candor vs. Psychological Safety**—Candor is the scarcest organizational resource!!!
- **Tolerance for Failure vs. Intolerance for Incompetence**—
hHow can we tell the difference?
- **Accountability**—High accountability getting confused with safety.

How these pieces fit together



Scaling Cultural Changes

1. Cascading—every level of the organization must be responsible for driving cultural changes
2. Telling the Stories—what does good look like?
3. Training and Coaching—how do I practice this as a leader?
4. Institutionalizing—culture management as a routine business process (measures, reviews, evaluations, corrective plans, etc.)

Q&A with Gary Pisano

Gary Pisano, Harvard Business School

An open Q&A on applying the 5 tensions in real organizations.

On candor across geographies: candor is expressed differently region to region. You need a cultural translator. Toyota is exceptionally candid about problems—the Toyota Production System is built around problem identification—it just looks different than candor in New York. On finding time: the same way leaders carve out three weeks of September for budgeting, they need to carve out specific time for culture. Treat it as an investment, not a continuous background activity. On deliverables for culture: the scan results are the deliverable. Set targets the same way operations sets targets for yield or service level.

On declaring you're working on culture: yes—but mean it. Performative culture work gets detected fast. People need to see culture treated as a strategic capability, not a buzzword. On culture in an AI-agent organization—the most provocative question of the day: agents will become carriers of culture; you can configure them to be more or less candid, more or less risk-seeking. As humans interact with agents as much as with each other, the human culture becomes the harder, more differentiating part.

Q&A with Gary Pisano

Gary Pisano, Harvard Business School

Key Takeaways

1. Candor Travels Differently

Candor is expressed differently across geographies. Toyota is exceptionally candid—it just looks different than candor in New York.

2. Make Time the Way You Make Budgets

The same way leaders carve three weeks of September for budgeting, carve specific time for culture.

3. Scan Results Are the Deliverable

Culture has metrics now. Set targets the same way operations sets targets for yield or service level.

4. Agents Will Become Cultural Carriers

Agent configuration becomes a culture decision. As humans interact with agents as much as with each other, the human culture becomes the more differentiating part.



“As people interact with agents as much as with each other, the human culture becomes the harder, more differentiating part.”

Extending These Ideas

Mark Ozer and Mallory Lehrhoff, ExecOnline

ExecOnline (which acquired Teamraderie at the start of April) presented how the day's framework can be scaled to high-potential leaders across an organization. ExecOnline partners with eleven of the top business schools—Stanford, Wharton, Columbia, MIT, Berkeley among them—to deliver professor-led programs, on-campus immersives, on-demand short-form courses, and one-on-one leadership and executive coaching.

The combination of the two companies is the point: ExecOnline goes deep on a specific topic with the world's top faculty; Teamraderie shifts day-to-day operating behavior at the team level. A customer-suggested model that is now being rolled out: leaders go through an ExecOnline professor-led course (e.g., Columbia's strategy or team leadership programs), and their teams then go through matched Teamraderie experiences—so the leader's new framework is practiced by the people they lead, not just held in their own head.

Extending These Ideas

Mark Ozer and Mallory Lehrhoff, ExecOnline

Key Takeaways

1. Depth + Behavior Change

ExecOnline goes deep on a specific topic with the world's top faculty; Teamraderie shifts day-to-day operating behavior at the team level.

2. Eleven Top Business Schools

Stanford, Wharton, Columbia, MIT, Berkeley, and others power the curriculum.

3. Solution Stack

Professor-led virtual courses (3–6 weeks), on-campus immersives (3–5 days), on-demand short-form, leadership and executive coaching.

4. Pair the Leader and Their Team

Customer-suggested model: leaders take an ExecOnline course; their teams run matched Teamraderie experiences. The framework gets practiced by the people they lead.



“Depth from the top business schools, behavior change from Teamraderie—the leader’s framework gets practiced by the people they lead.”

Closing Remarks

Michael McCarroll, SVP & GM, Teamraderie

Michael closed by reaffirming the day's central premise: it is not too much to ask leaders to do two things at once given the consequence of their roles—and the leaders in this room are the ones their organizations need to make sure their teams understand both. He shared two participant quotes from prior cohorts. A biotech VP of R&D: "Other leadership programs made me sharper as an individual. The difference here was that my whole team went through it together." A CPG executive: "My team used to push back afterwards. Now they push back earlier." Candor moves earlier in the process, not louder.

The six-month program for each enrolled team totals about three and a half hours of team time, deliberately spaced. June: a 15-minute five-tensions assessment per team member. June: a 55-minute results debrief with Francesca to interpret results and pick the two tensions to work on. Treatment Session 1: a 55-minute virtual session targeting the chosen tensions. Treatment Session 2 (~60 days later): a second 55-minute session, intentionally spaced to let new behaviors absorb between sessions. October: re-run the assessment to measure change. Between sessions, leaders receive Evidence-Based Tips for Leaders—short, in-context behavior nudges grounded in research.

Cohort milestones: June 1 team selection plus a group "Failing Smart" experience together. Mid-July and mid-September check-ins with cross-company learnings. November conclusions and scaling. Walkaway: a framework, a cohort, and a team.

Closing Remarks

Michael McCarroll, SVP & GM, Teamraderie

Key Takeaways

1. The Whole Team Goes Through It Together

Shared language is what makes the change stick. Candor moves earlier in the process, not louder.

2. The 5-Step Plan (3.5 Hours Over 6 Months)

June diagnostic. June results debrief. Treatment Session 1. Treatment Session 2 (~60 days later). October post-measurement.

3. EBT4L + Cohort Milestones

Evidence-Based Tips for Leaders between sessions. June 1 team selection. Mid-July and mid-September check-ins. November conclusions.



“Other leadership programs made me sharper as an individual. The difference here was that my whole team went through it together.”

— VP R&D, biotech (prior cohort)

What Enrolled Teams Will Do

1. Diagnose	2. Decide	3. Session I	4. Session II	5. Measure
Five Tensions Assessment	Review with faculty. Commit to two tensions.	55-min virtual session (Tension #1)	55-min virtual session on (Tension #2)	Retake the assessment
~15 min per person	55 min, virtual	55 min, virtual	55 min, virtual	~15 min per person
June	June	July	September	October

What We Do Together May through November

