Borrowed theories: how frontier-technology concepts explain an investment bank's daily work

借来的理论：前沿技术的概念如何解释一家投行的日常

庚辛研究院 / Glacier Institute*†

From The Wittgenstein Institute

Glacier Institute（庚辛研究院） | Working Paper No. GI-WP-2026-P11-EN | Date: 2026-09-22 | Language: English, with Chinese title and abstract

Cite as:

Glacier Institute (2026). Borrowed theories: how frontier-technology concepts explain an investment bank's daily work. Glacier Institute Working Paper GI-WP-2026-P11-EN. https://glacier.mba/research/GI-WP-2026-P11-EN.html

庚辛研究院（2026）.《借来的理论：前沿技术的概念如何解释一家投行的日常》. 庚辛研究院通讯论文 GI-WP-2026-P11. https://glacier.mba/research/GI-WP-2026-P11.html

Key Takeaways

1. **That a client's theory explains us is rarely a coincidence. We can only hear the part we already know.**
2. **"We understand this industry" has one testable meaning: we can borrow its concepts, and we borrowed them correctly.**
3. **Four in ten of the borrowed concepts are sixty years old, simply put to use again.**
4. **The better an analogy reads, the more it feels proven — when all that increased is the resemblance.**

Abstract

A financial intermediary that spends years working with frontier-technology companies keeps noticing the same thing: the concepts its clients use to explain their own technology also explain the intermediary's own organization and way of working. This paper treats that observation as a phenomenon worth studying rather than a figure of speech.

The source text is public. The firm lists twelve pairs on its website under the heading "their theory, our phenomena": on the left a concept from a frontier-technology field, on the right one of its own daily practices. All twelve are reproduced verbatim, with line numbers in the public plain text.

The mechanism section draws on five bodies of work: analogical reasoning and conceptual transfer (structure-mapping theory and multiconstraint theory), absorptive capacity (an organization can only take in what it already has related prior knowledge for), organizational routines and the cumulative growth of capability, knowledge conversion and tacit knowledge in professional service firms, and a set of boundary references — the representativeness heuristic that treats similarity as evidence, the role and danger of metaphor in organization studies, and misreadings of exploration and exploitation.

The contribution lies in the third paragraph of each pairwise analysis. The website states what the theory says and what it is mapped onto; it does not state the conditions under which the mapping holds and the conditions under which it fails. This paper supplies those for all twelve. Three of the twelve do not hold under the premises of the source theory; the paper says why, and what part remains usable.

From this we build one tool: a nine-row analogy audit, each row written as criterion, what passing looks like, what failing looks like. The twelve pairs are then scored against it, producing a table computed here rather than quoted. The

central claim is stated as a falsifiable proposition together with the evidence that would refute it. Five boundary cases close the paper, including two that must be stated: an analogy that reads well is not the same as a decision that is right, and the pull of after-the-fact explanation.

No client company is named, no description that could identify one is used, and none of the firm's own financial figures appear.

Keywords: analogical reasoning; structure mapping; conceptual transfer; absorptive capacity; organizational routines; professional service firms; primary market

JEL Classification: D83（搜寻、学习与信息）, O33（技术变迁：选择、后果与扩散）, L84（专业与商业服务）, G24（投资银行与创业投资）

摘要 (Chinese abstract)

一家长期服务前沿科技公司的中介机构，会反复发现一件事：客户用来解释自己技术的概念，恰好也解释了这家机构自己的组织与作业方式。本文把这件事当成一个可研究的现象，而不是一句修辞。底本是官网第 18 章的十二对「他们的理论 / 我们的现象」，逐字抄录并标行号。机制部分用五组文献：类比推理与概念迁移、吸收能力、组织惯例、知识转化与默会知识，以及一组边界文献。本文的学术贡献在逐对分析的第三段：官网写了理论是什么、迁移到什么上，没有写这个迁移在什么条件下成立、什么条件下不成立；十二对逐一补上，其中三对在原理论的前提下并不成立。工具是一张九行的类比自查表，并用它给十二对打分。主张写成可以被推翻的命题，文末给出五类边界。本文不出现任何客户公司的名字，不出现这家机构自身的任何经营数字。

Keywords (Chinese): 类比推理；结构映射；概念迁移；吸收能力；组织惯例；专业服务机构；一级市场

1. The Problem: Why a Client's Theory Explains Us

A boutique investment bank spends eight years on hard technology and nothing else. For eight years, almost everything it reads, discusses and hears is someone else's technology. It may not master the technical terms, but it hears the concepts used to *explain* that technology thousands of times: training and inference are different things; errors accumulate across the seams of a segmented pipeline; what caps a model is not the volume of data but its quality; generalization needs a world model as a prior.

Then something happens. Those concepts start to explain the firm itself.

The firm has written this down. Chapter 18 of the Glacier Institute website lists twelve pairs of cards. On the left, "their theory". On the right, "our phenomena". The introduction is two sentences:

> 「庚辛长期与最前沿的公司一起工作，八年下来发现：他们用来理解世界的理论，恰好也解释了我们自己。理论来自这些公司与行业，现象是庚辛的日常。」
>
> *Glacier has worked alongside the most advanced companies for a long time. Eight years in, we found that the theories they use to understand the world happen to explain us as well. The theories come from those companies and industries; the phenomena are Glacier's ordinary days.*
>
> — Glacier Institute website, /institute.html [1]; the same text also appears on the home page /index.html [2]

The chapter closes with one line:

> 「向自己陪伴的公司学习，是这份工作的特权。」
>
> *Learning from the companies you accompany is the privilege of this work.*
>
> — Glacier Institute website, /institute.html [1]

This paper answers two questions that have actually been asked.

The first comes from clients and investors, usually in this form: when an intermediary says it "understands" a technical industry, does that claim have any testable content? What counts as understanding?

The second comes from inside, usually in this form: borrowing a client's theory to explain ourselves — is that insight, or is it just good phrasing? How do you tell them apart?

This paper treats them as one question. What concepts an organization can borrow depends on what it already knows (Section 3.2); whether a borrowed concept is insight or rhetoric depends on whether it maps structure or surface similarity, and on whether it can direct a specific action (Sections 3.1 and 5). Section 9 answers both directly.

The route: Section 2 lays out the source text, all twelve pairs, none omitted. Section 3 sets out the mechanism of conceptual transfer. Section 4 analyses each pair in three paragraphs, the third of which is this paper's addition. Section 5 gives the tool and scores the twelve pairs with it. Section 6 states the claim as a proposition that can be refuted. Sections 7 and 8 mark the boundaries.

表 1 / Exhibit 1 Pairs 1–6. The twelve "their theory / our phenomena" pairs of website Chapter 18, reproduced verbatim. Theory and phenomenon sentences are taken from the public plain text of https://glacier.mba/institute.html (retrieved 2026-09-22); line numbers refer to that plain text. English in italics is a translation supplied here for readers of this edition; the Chinese is the original.

No.	Their theory	Theory sentence (original)	Our phenomenon	Phenomenon sentence (original)	Lines
1	训推分离 Train–infer separation (from embodied-AI engineering practice)	「训练时靠多个辅助任务调优参数；到了推理，把所有的枝剪掉，只留一个输出。」 *In training, many auxiliary tasks tune the parameters; at inference, every branch is pruned and one output remains.*	这解释了「找钥匙」。*This explains "finding the key".*	「准备一场融资，我们穷尽技术、财务、结构的每一条辅助线；见投资人的那三十秒，剪掉全部枝蔓，只递一把钥匙。」 *Preparing a round, we exhaust every auxiliary line — technical, financial, structural. In the thirty seconds with an investor, we prune all branches and hand over one key.*	L3868–3873
2	专家模型融合 Merging expert models (from frontier foundation-model research)	「不同的专家模型融合成一个更强的模型；每一代不必推倒重训，智能连续增长。」 *Different expert models merge into a stronger one; each generation need not be retrained from scratch, and intelligence grows continuously.*	这解释了「理解的规模化」。*This explains "scaling understanding".*	「每个项目组都是一个专家模型，每周复盘把它们融合进同一个组织。八年，庚辛没有推倒重训过——能力是连续长出来的。」 *Every project team is an expert model; the weekly review merges them into one organization. In eight years Glacier has never retrained from scratch — capability grew continuously.*	L3874–3879
3	端到端 End-to-end (the tradition of autonomous driving and deep learning)	「从输入到输出，一个模型端到端负责；分成多段的流水线，误差会在段与段之间累积。」 *From input to output one model is responsible end to end; in a segmented pipeline, error accumulates between the segments.*	这解释了「全托管」。*This explains "full mandate".*	「从恢复事实到交割收口，同一个系统端到端负责，六个维度不换手——误差没有地方累积。」 *From restoring the facts to closing, one system is responsible end to end, six dimensions with no handoff — error has nowhere to accumulate.*	L3880–3885
4	数据质量与教师模型 Data quality and the teacher model (public engineering practice in autonomous driving)	「决定模型上限的不是数据量，是数据质量；要有教师模型持续评估：缺哪一类泛化数据，就补哪一类场景。」 *What caps a model is not data volume but data quality; a teacher model must keep evaluating: whichever class of generalization data is missing, go collect that class of scenes.*	这解释了我们的复盘制度。*This explains our review discipline.*	「复盘会就是那个教师模型——评估上一周的判断质量、找出缺口，指挥下一周去补缺的那类场景。」 *The review meeting is that teacher model — it grades last week's judgement, finds the gap, and directs next week to the class of scenes that is missing.*	L3886–3889, 3919
5	世界模型是先验 The world model is a prior (the consensus of world-model research)	「任何模型最终都需要一个世界模型做先验和约束，泛化能力才不靠运气。」 *Every model ultimately needs a world model as prior and constraint, so that generalization does not rest on luck.*	这解释了北坡、价格带与 4D。*This explains the North Slope, the price band and 4D.*	「它们是庚辛的资本世界模型：每一笔交易进入执行之前，先被这套先验约束过一遍。」 *They are Glacier's world model of capital: before any transaction enters execution, it is first constrained by this set of priors.*	L3920–3925
6	泛化与效率成反比 Generalization trades against efficiency (from embodied-AI scenario analysis)	「结构化场景要效率，非结构化场景要泛化；真正把模型训出来的，是非结构化场景的数据。」 *Structured scenes call for efficiency, unstructured ones for generalization; what actually trains a model is data from unstructured scenes.*	这解释了我们对复杂局面的偏好。*This explains our preference for complicated situations.*	「标准轮次是结构化场景，方法早已成熟；复杂局面才是高质量数据。我们据此选择走进非结构化的局面：每解开一个，组织就多一分泛化。」 *A standard round is a structured scene and the method is long settled; complicated situations are the high-quality data. On that basis we choose to walk into the unstructured ones: each one untangled adds a measure of generalization to the organization.*	L3926–3929, 3959

> "We understand this industry" has one testable meaning: we can borrow its concepts, and we borrowed them correctly.

2. The Source Text: Twelve Pairs of Theory and Phenomenon

2.1 How the source text was obtained

The source is a public web page and anyone can reproduce it. Method: fetch https://glacier.mba/institute.html and https://glacier.mba/index.html with curl; strip HTML

表 2 / Exhibit 2 Pairs 7–12. Columns and provenance as in Table 1a.

No.	Their theory	Theory sentence (original)	Our phenomenon	Phenomenon sentence (original)	Lines
7	范式的接力 The relay of paradigms (the public narrative of the foundation-model industry)	「一个范式到了效率下降的节点，增长不会停——接力棒交给下一个范式。」 *When a paradigm reaches the point of declining returns, growth does not stop — the baton passes to the next paradigm.*	这解释了「第八年」。*This explains "the eighth year".*	「八年，庚辛两次交棒：从单点服务到操作系统，从中国到全球——增长的曲线没有断过。」 *In eight years Glacier passed the baton twice: from point services to an operating system, from China to the world — the curve of growth never broke.*	L3960–3963, 3984
8	世界先于画面 The world precedes the picture (from the public debate between spatial-intelligence and world-model research programmes)	「屏幕上的一切呈现都只是底层世界的投影；智能应当建模投影之前的世界本身，而不是投影之后的像素。」 *Everything rendered on a screen is only a projection of an underlying world; intelligence should model the world before the projection, not the pixels after it.*	这解释了「恢复事实」。*This explains "restoring the facts".*	「BP、Memo、路演，都是投影；投影之前，先有那个真实的世界。所以每个项目的第一步，是像企业考古学家一样重建业务、技术、客户、交易历史与关键风险——先建世界，再渲染画面。材料只是这个世界的一次采样。」 *The deck, the memo, the roadshow are all projections; before the projection there is the real world. So the first step on every project is to rebuild, like a corporate archaeologist, the business, the technology, the customers, the transaction history and the key risks — build the world first, render the picture after. The materials are one sampling of that world.*	L3985–3990
9	解耦生长，统合交付 Grow decoupled, deliver combined (from public ideas in modular learning and complex-systems engineering)	「复杂能力不该由一个整体硬扛：按性质拆成层，让每一层独立生长，最后面向同一个目标重新耦合成一次交付。」 *A complex capability should not be carried by one monolith: split it into layers by nature, let each layer grow independently, then recouple them toward one goal into a single delivery.*	这解释了「二十七个子系统」。*This explains "twenty-seven subsystems".*	「4D、价格带、四层旋涡、GQW——每个子系统各居其层、各自成章；到了一单交易，再被重新编排成一次交付。拆开是为了各自变强，合拢是为了共同成事——名单会过期，系统不会。」 *4D, the price band, the four-layer vortex, GQW — each subsystem sits in its own layer and forms its own chapter; in a live transaction they are re-orchestrated into one delivery. Splitting is so each grows stronger; joining is so they accomplish one thing — a list expires, a system does not.*	L3991–3994, 4063
10	少原则，多演绎 Few principles, much deduction (from the public tradition of reinforcement learning and self-play)	「聪明的行为不是规定出来的，是演绎出来的：只写下第一性原理级的少数原则，最优策略交给大规模的真实实践去推演。」 *Intelligent behaviour is not prescribed but deduced: write down only a few first-principles rules and leave the optimal policy to be derived by large-scale real practice.*	这解释了「只做减法的清单」。*This explains "a checklist that only subtracts".*	「我们写下的是边界，不是打法：踩过的坑写成禁令，复盘会全员重读；规则写得越细，越有人卡规则的缝，所以模糊处交给被所有人信任的人。方法可以变，边界只加不减——具体的打法，在一场场真实交易里长出来。」 *What we write down are boundaries, not plays: every pit we fell into becomes a prohibition, re-read by everyone at the review; the finer the rules, the more people work the seams, so the grey areas go to the person everyone trusts. Methods may change; boundaries are only added, never removed — the actual plays grow out of one real transaction after another.*	L4064–4067, 4131
11	智能出于多样 Intelligence out of diversity (a cognitive-science classic, from *The Society of Mind* onward)	「强大的智能来自大量异质心智的协作互补，而非某个单一完美的单体——多者成社会，社会即心智。」 *Powerful intelligence comes from many heterogeneous minds complementing one another, not from a single perfect unit — the many form a society, and the society is the mind.*	这解释了「球面上的组织」。*This explains "an organization on a sphere".*	「庚辛不是一张层级表，是同一个球面上一组尽量拉开的方向：方向铺得越开，覆盖越大、冗余越小；哪个角度空了，就有人转过去。组织的力量不靠某个孤胆英雄，靠异质方向的互补——挤在一起，就都成了平均值。」 *Glacier is not an org chart but a set of directions on one sphere, spread as far apart as possible: the wider they spread, the greater the coverage and the smaller the redundancy; when an angle is empty, someone turns to it. The strength of the organization does not rest on a lone hero but on heterogeneous directions complementing one another — crowd together and everyone becomes the average.*	L4132–4135, 4160
12	观测廉价，干预稀缺 Observation is cheap, intervention is scarce (textbook common ground in causal inference)	「世界模型的上限，是『看见』低于『去做』：真正稀缺的不是海量观测，而是『在什么状态、做了什么、之后怎样』的带反馈干预数据。」 *On the ladder of causation, "seeing" sits below "doing": what is truly scarce is not mass observation but intervention data with feedback — in what state, what was done, what followed.*	这解释了「把自己的钱投进去」。*This explains "putting our own money in".*	「庚辛不是一张层级表，是同一个球面上一组尽量拉开的方向：方向铺得越开，覆盖越大、冗余越小；哪个角度空了，就有人转过去。组织的力量不靠某个孤胆英雄，靠异质方向的互补——挤在一起，就都成了平均值。」 *Glacier is not an org chart but a set of directions on one sphere, spread as far apart as possible: the wider they spread, the greater the coverage and the smaller the redundancy; when an angle is empty, someone turns to it. The strength of the organization does not rest on a lone hero but on heterogeneous directions complementing one another — crowd together and everyone becomes the average.*	L4161–4164, 4193

comments, <script> and <style> in that order; convert block-level tags to newlines, remove remaining tags, unescape entities, drop blank lines. Retrieved on institute.html 2026-09-22. Line numbers below refer to the plain text of institute.html.

Chapter 18 runs from line 3839 to line 4194 of that plain text. The home page carries the same chapter but shows only two of the twelve pairs (mixture of experts; the world model as prior), the complete set of twelve appears only on institute.html [1][2]. One transcription note: in the plain text, the dash and extra spaces between a theory name and its source label are seams produced by stripping adjacent elements, not punctuation in the original. In reproducing the text we

表 3 / Exhibit 3 The theory-side sources of the twelve pairs, reproduced verbatim from the website's own "source" lines [1], with resolvable DOI or arXiv addresses supplied here (each checked by live request on 2026-09-22; see the writing note). The labels T1–T12 are used only within this paper.

No.	Source line as printed on the website	Accessible address
T1	Jaderberg et al. (2017), Reinforcement Learning with Unsupervised Auxiliary Tasks, ICLR 2017, Figure 1	arXiv:1611.05397
T2	Jacobs, Jordan, Nowlan & Hinton (1991), Adaptive Mixtures of Local Experts, Neural Computation 3(1), Figure 1	doi:10.1162/neco.1991.3.1.79
T3	Bojarski et al. (2016), End to End Learning for Self-Driving Cars, arXiv:1604.07316, Figure 4	arXiv:1604.07316
T4	Hinton, Vinyals & Dean (2015), Distilling the Knowledge in a Neural Network, arXiv:1503.02531	arXiv:1503.02531
T5	Ha & Schmidhuber (2018), World Models, arXiv:1803.10122, Figure 8 · CC BY 4.0	arXiv:1803.10122
T6	Wolpert & Macready (1997), No Free Lunch Theorems for Optimization, IEEE Trans. Evolutionary Computation	doi:10.1109/4235.585893
T7	Kuhn (1962), The Structure of Scientific Revolutions, University of Chicago Press	press.uchicago.edu book page
T8	Dawid & LeCun (2023), Introduction to Latent Variable Energy-Based Models: A Path Towards Autonomous Machine Intelligence, arXiv:2306.02572, Figure 9 · CC BY 4.0	arXiv:2306.02572
T9	Simon (1962), The Architecture of Complexity, Proceedings of the American Philosophical Society 106(6)	jstor.org/stable/985254
T10	Silver et al. (2017), Mastering the Game of Go without Human Knowledge, Nature 550	doi:10.1038/nature24270
T11	Minsky (1986), The Society of Mind, Simon & Schuster	openlibrary.org OL2714978M
T12	Pearl & Mackenzie (2018), The Book of Why, Basic Books	basicbooks.com book page

restore the form "theory name (source label)". The theory sentences and phenomenon sentences are unchanged, character for character.

2.2 The twelve pairs, verbatim

2.3 The theory side: the website states its own sources

Under each of the twelve pairs the website prints a "source" line pointing to a specific paper or book. That deserves a sentence of its own: **whoever borrowed these concepts wrote down where they were borrowed from.** It is what makes Section 4 possible — one can go back to the original and see what the concept says in its own field.

2.4 What those sources themselves reveal

Sorting Table 3 by medium and period shows something that is not stated anywhere in the website's own words. The table below is counted from the twelve entries of Table 3.

Classification rules: medium follows the medium named on the website's own source line (T1 is printed as ICLR 2017 and counts as a conference paper); period follows the year on that line; field follows the disciplinary community in which the work appeared. Each dimension sums to 12 with no double counting.

Two numbers are worth pausing on.

First, **67% of the concepts come from machine learning and artificial intelligence.** That is not a neutral distribution. If an organization borrowed concepts at random, following whatever was in the news, the sources would spread across more fields. Concentration in one field says the borrowing is not of buzzwords but of what the firm actually reads every day. Section 3.2 treats this as evidence of absorptive capacity, and Section 6 writes it as a testable corollary.

Second, **42% of the sources predate 1997.** The two oldest are from 1962 (T7, T9). Nearly half of these "frontier-technology concepts" are sixty years old, put to new use by contemporary engineering practice. What a

表 4 / Exhibit 4 Medium and period of the twelve theory-side sources. Counted here from Table 3, denominator 12. Classification rules are given in the writing note, §8.1.

Dimension	Band	Entries	Count	Share of 12
Medium	Peer-reviewed journal article	T2, T6, T9, T10	4	33%
Medium	Conference paper and preprint	T1, T3, T4, T5, T8	5	42%
Medium	Monograph	T7, T11, T12	3	25%
Period	1962–1997 (classical)	T2, T6, T7, T9, T11	5	42%
Period	2015–2023 (contemporary)	T1, T3, T4, T5, T8, T10, T12	7	58%
Field	Machine learning / artificial intelligence	T1, T2, T3, T4, T5, T6, T8, T10	8	67%
Field	Cognitive science, philosophy of science, complex systems, causal inference	T7, T11, T9, T12	4	33%

图 1 / Exhibit 5 The structure of the transfer from "their theory" to "our phenomena". The figure is this paper's own synthesis; mapping relational structure rather than surface resemblance comes from Section 3.1, the upstream absorptive-capacity gate from Section 3.2, and the downstream nine criteria and two exits from Table 4 in Section 5. The conditions and results in the figure are placeholders and refer to no particular pair.

borrower borrows is often an old concept in a new application.

> Four in ten of the borrowed concepts are sixty years old, simply put to use again.

3. Mechanism: Why Concepts Travel from the Client to the Firm

Section 2 sets out the phenomenon. This section asks how the transfer happens, why it happens to this firm rather than elsewhere, and when it is reliable.

3.1 Analogical reasoning: what is mapped is structure, not resemblance

Cognitive science has a mature theory of analogy. Gentner's structure-mapping theory [3] gives the core criterion: a good analogy maps **relational structure**, not **object attributes**. In mapping the solar system onto the atom, what matters is not that both are round but that a central body attracts peripheral bodies which revolve around it. She calls this the systematicity principle: higher-order relations, especially causal ones, are mapped in preference to isolated surface attributes.

Gentner and Markman [4] extend the principle to similarity judgement itself: judging whether two things are alike is performing a structural alignment. Holyoak and Thagard [5] arrive at a multiconstraint framework by another route: an analogical mapping is pulled at once by three constraints — structural consistency, semantic similarity, and the mapper's current purpose. The third is the one to watch: **purpose bends the mapping**. A person trying to prove something will find the analogy growing in that direction. Their later review [7] carries the same warning.

Gick and Holyoak [6] showed experimentally how hard transfer is: shown a structurally identical story beforehand, most subjects fail to use it on the target problem unless prompted; with a hint, solution rates rise sharply. The implication is counterintuitive — **analogy is not automatic; it needs cues and retrieval**. An organization immersed daily in one field's concepts is being cued daily, which is precisely why the transfer happens to it and not to others.

The use here is a yardstick: **to judge whether a borrowed concept is insight or rhetoric, first ask whether it maps relational structure or surface similarity.** That is row 1 of the audit in Section 5.

3.2 Absorptive capacity: you can only learn what you already partly know

Why this firm, and not any firm reading the same news?

Cohen and Levinthal [8] answer with absorptive capacity: an organization's ability to recognize, assimilate and apply new external knowledge depends on its existing related prior knowledge. They stress that the ability is **cumulative** and **path-dependent** — what you can absorb today depends on what you accumulated yesterday, and ceasing to invest in a field degrades future capacity in it. Zahra and George [9] later split the construct in two: acquisition and assimilation (potential absorptive capacity) versus transformation and exploitation (realized absorptive capacity). There is a gap between the halves — many organizations acquire knowledge and never convert it into behaviour.

Both parts bear directly on Table 4. The 67% concentration in machine learning is prior knowledge at work: eight years on hard technology means the reading is concentrated there, so the recognizable concepts are concentrated there too. And the two-stage split gives this paper its most important test point: **writing a concept on the wall is acquisition and assimilation; changing a specific action is transformation and exploitation.** Row 5 of the audit and the main proposition in Section 6 both rest on this.

Absorptive capacity has a less-cited corollary that should be stated: **it is also a filter.** It determines what gets in and therefore what stays out. A concept a firm cannot borrow may be perfectly useful; the firm may simply lack the prior for it. That is the source of boundary case five in Section 7.

3.3 Routines and capability: why "never retrained from scratch" needs care

Pair 2 in Tables 1 and 2 says: eight years, never retrained from scratch, capability grew continuously.

Evolutionary economics has both a theory and a warning for that sentence. Nelson and Winter [10] locate organizational capability in **routines**: routines are the organization's memory, the form in which capability is stored, and they change gradually, path-dependently, with inertia. The theory supports "capability grew continuously" as a description while reminding us that continuity and inertia are the same property seen from two sides. Routines do not retire by themselves; an old routine keeps running after the environment has changed.

Feldman and Pentland [11] supply the other half: routines are not dead. They separate the ostensive aspect (the abstract process everyone describes) from the performative (the action actually taken each time), and identify the tension between the two as the source of endogenous change — each performance may drift from the spoken version, and drift accumulates.

Two uses follow. First, "never retrained from scratch" holds in theory, but it also means bearing the cost of path dependence, which the website does not state; Section 4.2 supplies it. Second, the two aspects give an operational

measure: **the gap between the spoken version and the action actually taken.** A borrowed concept that changes only the spoken version has stopped at Zahra and George's first stage.

3.4 How knowledge passes down inside a professional service firm

Nonaka [12] divides knowledge into tacit and explicit and draws the conversion cycle in four steps: socialization (tacit to tacit), externalization (tacit to explicit), combination (explicit to explicit), internalization (explicit to tacit). Polanyi [13] gave the underlying proposition earlier: we know more than we can tell.

Together they explain **what kind of act Chapter 18 is**. An investment bank's judgement is largely tacit — how to read a company, when to slow down, which sentence must not be said. Matching that tacit material to an explicit concept borrowed from the client's field is externalization: finding a sayable name for something that could not be said.

Externalization has a clear benefit. Once named, it can be taught, discussed and criticized. Eight years of judgement held by a few people means newcomers must fall into every pit again; give it a name and a newcomer can at least start by asking about the name.

It also has a clear cost, taken up in boundary case four: **once the name sounds good, it stops being questioned.** Externalization turns tacit knowledge into explicit knowledge, and at the same time into a slogan.

3.5 Boundary references: when analogy misleads

The four subsections above explain why transfer happens and when it is reliable. This one covers when it is not.

Tversky and Kahneman [14] give the most fundamental case: the representativeness heuristic. People judge whether A belongs to B by how much A resembles B, not by probability. Similarity is used as evidence. That is exactly where analogy becomes dangerous — **the better an analogy reads, the more it feels proven, when all that has increased is the resemblance.**

Gilovich [15] ran an experiment closer to this paper's situation. Subjects made policy judgements about a hypothetical international crisis; the materials contained irrelevant surface cues echoing either the Second World War or Vietnam (whether the briefing was held in a "Munich" hall, whether refugees travelled by "train"). Policy preferences shifted systematically with those cues. **Surface resemblance alters decisions without the decider noticing.** Transposed here: a concept from the client's field, even if it only resembles the firm's situation on the surface, will affect how the firm describes itself and therefore how it acts.

March [16] supplies another kind of error. His model of mutual learning between individuals and an organizational code shows that the rate of socialization affects long-run performance: socialize too fast and diversity is absorbed, leaving the organization converged on a mediocre consensus. This is often flattened into a slogan about exploring more. The actual result is conditional — the optimal learning rate depends on environmental turbulence and organizational size. Section 4.11 uses it to bound the "intelligence out of diversity" pair.

Two final references concern metaphor itself. Morgan [17] argues that schools of organization theory rest on different root metaphors — machine, organism, culture, political system — each illuminating one part and obscuring another. Tsoukas [18] distinguishes two uses of metaphor: as a substitute for literal description (decorative, removable) and as a generative cognitive instrument (it changes what you can see, and cannot be removed). His criterion is especially useful here: **if removing a metaphor costs you nothing, it was decoration.**

Together these five bodies of work give the nine rows of Section 5.

> The better an analogy reads, the more it feels proven — when all that increased is the resemblance.

4. Pair by Pair: Twelve Pairs, Three Paragraphs Each

Format: three paragraphs per pair. First, what the theory says in its own field, with the source. Second, what it is mapped onto, quoting the website. **Third, the conditions under which the mapping holds and the conditions under which it fails — this paragraph is not on the website and is supplied here.**

4.1 Train–infer separation → finding the key

What the theory says. T1 adds a set of unsupervised auxiliary tasks alongside the main reinforcement-learning objective; they share one representation and contribute gradient updates, supplying learning signal where reward is sparse. After training, inference runs only the main policy. Two points matter: the auxiliary tasks **share the same parameters** as the main task, and what is pruned at inference is a computation path, not the influence those tasks left behind during training.

What it is mapped onto. The website calls it "finding the key": preparing a round, exhaust every auxiliary line — technical, financial, structural; in the thirty seconds with an investor, prune all branches and hand over one key [1].

When it holds. Three premises. First, the auxiliary lines must genuinely change the main conclusion — the technical, financial and structural work must shape what the key looks like, or they are not auxiliary tasks but a pile of parallel documents nobody reads. Second, the receiving bandwidth really is scarce; pruning is meaningful only when thirty seconds is a hard constraint. Third, pruning must be recoverable: when questioned, the branches can be grafted back.

When it fails. In the source theory, pruning at inference loses nothing because the information is already encoded in the weights. A document has no weights. The pruned branches do not remain in the key by themselves — **unless a person remembers them**. The mapping therefore carries an implicit carrier: whoever did the auxiliary work must be in the room. If the person who built the auxiliary lines and the person who hands over the key are different, train–infer separation degrades into information loss. There is also a reverse failure: in diligence, the other side wants precisely the branches, and pruning there is not distillation but concealment.

Testable form. Whether a pruned auxiliary line can be restored to deliverable form within one working day of being asked. If yes, it really entered the weights; if no, it never influenced the main line.

4.2 Merging expert models → scaling understanding

What the theory says. T2 is the original mixture-of-experts paper: several expert networks plus a gating network that decides which expert handles each input. Experts specialize because the gate creates competition, pushing each toward the inputs it handles better. One thing must be said precisely: **T2 is about division of labour plus routing, not about "merging into one stronger model".** The line of work that combines several models into one (model merging, ensemble distillation) is a different literature. The website's theory sentence describes the latter; the source line cites the former.

What it is mapped onto. Every project team is an expert model; the weekly review merges them into one organization; in eight years there has been no retraining from scratch, and capability grew continuously [1].

When it holds. A mixture of experts needs two things: a gate, and a shared loss function. The gate exists in an organization — whoever or whatever decides which kind of problem goes to which team. The shared loss is the problem: different teams serve different clients and face different situations, and their outcomes do not share one objective. Without a shared loss there is no gradient, and "merging" has nothing to merge with; what remains is a briefing. The mapping therefore holds on one condition: **the review must change another team's next action**, not merely inform it.

When it fails. Section 3.3 gives the other bound. "Never retrained from scratch" holds as description, but it is also another name for path dependence. For a network, not retraining costs catastrophic forgetting and the constraint of old representations; for an organization, it costs old routines that keep running after the environment has changed [10]. On the website this is a positive statement; in the literature it is a neutral description with an attached cost. The addition here: **capability that accumulates continuously needs a companion device that periodically retires an old routine**, or continuity slowly becomes inertia.

Testable form. Take any lesson from a review and see whether it appears in another team's subsequent action; and count whether any old practice was explicitly retired in the past year. The first tests merging, the second tests inertia.

4.3 End-to-end → the full mandate

What the theory says. T3 trains a single network from camera pixels to steering angle, dispensing with hand-designed intermediate stages such as lane detection and path planning. Its argument: each segment of a pipeline optimizes its own objective, the objectives are not aligned, and error accumulates at the seams.

What it is mapped onto. From restoring the facts to closing, one system is responsible end to end, six dimensions with no handoff, so error has nowhere to accumulate [1].

When it holds. End-to-end beats segmentation on two conditions: there is enough end-to-end training signal, and the error really arises at the seams rather than inside single stages. The second condition usually obtains in cross-organization collaboration — information lost at a handoff is real and observable.

When it fails. The first condition is weak here. The end-to-end signal in driving is a vast corpus of driving data; in a transaction it is the closing outcome, sparse and slow. More importantly, end-to-end carries a recognized cost: **it is hard to interpret and hard to localize.** The virtue of a segmented pipeline is precisely that a failure can be attributed to a stage. An organization with "no handoff across six dimensions" also loses the cross-check that a handoff naturally produces. The mapping therefore holds only with an external evaluator attached — which is exactly the next pair. **Pairs 3 and 4 must be used together; pair 3 alone is dangerous.**

Testable form. When something goes wrong, can the responsible dimension be identified without interrogating individuals? If not, the cost of end-to-end has already been incurred.

4.4 Data quality and the teacher model → the review discipline

What the theory says. T4 is knowledge distillation: train a strong teacher, then train a student to match the teacher's soft probability distribution rather than hard labels. The central insight is "dark knowledge" — the teacher's relative probabilities *among the wrong classes* carry information that hard labels do not. The student learns not only what the answer is, but where the error lies and what the second-best option was.

What it is mapped onto. The review meeting is that teacher model: it grades last week's judgement, finds the gap, and directs next week to the class of scenes that is missing [1].

When it holds. Distillation works only if the teacher's output is soft. Translated: the output of a review must include what the second-best option was and how far off the judgement was, not merely success or failure. This is, in our view, the item in the whole chapter that **most directly directs an action**: add a "second-best option" field to the review record, and you have replaced a hard label with a soft one.

When it fails. Two ways. First, distillation presumes the teacher is stronger than the student. In an organization, teacher and student are often the same people grading their own week. Self-distillation works in machine learning but amplifies existing bias rather than correcting it, because the teacher's errors are transmitted intact. Second, the teacher in T4 is **trained independently and beforehand.** A review that draws only on material the same team produced that week is not an independent teacher; it is an echo of the same model.

Testable form. Sample ten review records and count how many state the options that were available and why they were not taken. That proportion is the share of soft labels.

4.5 The world model as prior → the North Slope, the price band and 4D

What the theory says. T5 splits an agent into three parts: a visual encoder V compressing observations into latents, a memory module M learning to predict the next step, and a controller C deciding on top of both. Its most cited passage trains the controller entirely inside the learned world model — in a "dream" — and transfers it back to the real environment. The paper also reports the trap on that route: the agent learns to **exploit the deficiencies of the world model**, finding policies that score highly in the dream and fail in reality.

What it is mapped onto. They are Glacier's world model of capital: before any transaction enters execution, it is first constrained by this set of priors [1].

When it holds. For a world model to serve as a prior, it must be continuously corrected by reality — M's prediction error must flow back and update M. A framework that is rewritten in one or two places each year by an actual outcome is a prior. One that has not changed in ten years is a doctrine.

表 5 / Exhibit 6 The analogy audit. Nine criteria to pass before applying an outside concept to your own organization. Basis in the last column: [3]–[7] analogy and transfer, [8][9] absorptive capacity, [10][11] routines, [12][13] knowledge conversion, [14]–[18] boundary references. The table is this paper's own.

No.	Criterion	What passing looks like	What failing looks like	Basis
1	Structure or surface resemblance	You can write out the causal relations on each side and align them one by one	You can only say "both are…", "very like…", without saying how the relations align	[3][4]
2	Do the source theory's premises hold here	You can list the premises and judge each as holding or not	You never looked for the premises, only the conclusion	[3][5]
3	Is it falsifiable	You can say what observation would mean the analogy fails	Everything confirms it	[14]
4	Is there a counter-case	You can name at least one setting where the analogy plainly breaks	You cannot think of one and never looked	[15]
5	Does it direct a specific action	It removes or adds one observable action	Only the way people talk has changed; actions have not	[9]
6	Is it only explaining after the fact	At least one practice came from the theory, not before it	Every practice existed before the borrowing	[14][15]
7	Does it hide a real cost	You can say what the practice costs and who bears it	The analogy mentions only benefits	[10][16]
8	Could another theory explain it equally well	You looked for alternatives and can say why this one is better	You never looked	[17][18]
9	Remove the concept — what is left	Something visible is genuinely lost	Nothing is lost; it merely reads less well	[18]

When it fails. The trap in the source paper has an exact organizational counterpart: **using an internal framework to prove itself.** A project that works out only inside one's own framework, without ever meeting an external result, is scoring highly in the dream. This is not a metaphorical worry; it is a failure mode the original paper measured.

Testable form. In the past twelve months, has any element of the prior been rewritten because of a real outcome? A count of zero means the prior has become doctrine.

4.6 Generalization trades against efficiency → a preference for complicated situations

What the theory says. A distinction first. The source cited is T6, the no-free-lunch theorem, whose proposition is that averaged over all possible objective functions, any two optimization algorithms have identical expected performance. **It does not say that generalization trades against efficiency;** it says that averaged over all problems, no algorithm is better. The website's figure caption — equal-length vectors on a cone — is faithful to T6; the theory name is a summary from embodied-AI engineering practice, which is a different proposition. This is the second place where the theory name and the cited source are not fully aligned.

What it is mapped onto. A standard round is a structured scene with a settled method; complicated situations are the high-quality data, and each one untangled adds a measure of generalization [1].

When it holds. The premise of no-free-lunch is a uniform distribution over all possible objective functions. Real problem distributions are not uniform, so specialization does pay — which in fact supports choosing one class of situation and getting strong at it. The real condition for the mapping lies elsewhere: **unstructured**

表 6 / Exhibit 7 The twelve pairs of Tables 1 and 2 scored against the nine criteria of Table 5. ✓ = passes, △ = passes conditionally (the condition is stated in the corresponding subsection of Section 4), × = fails. The scoring is this paper's judgement, based on the pairwise analysis of Section 4; the website contains no such evaluation. The last column is the total (✓ = 1, △ = 0.5).

No.	Theory	1 Structure	2 Premises	3 Falsifiable	4 Counter-case	5 Directs action	6 Not post hoc	7 No hidden cost	8 Alternative	9 Not removable	Score
1	Train–infer separation	✓	△	✓	✓	✓	×	△	✓	✓	7.0
2	Merging expert models	△	×	✓	✓	✓	×	×	△	✓	5.0
3	End-to-end	✓	△	✓	✓	✓	×	×	✓	✓	6.5
4	Data quality and the teacher model	✓	△	✓	✓	✓	×	△	✓	✓	7.0
5	The world model as prior	✓	✓	✓	✓	✓	×	△	✓	✓	7.5
6	Generalization trades against efficiency	△	×	✓	✓	△	×	△	△	✓	5.0
7	The relay of paradigms	×	×	△	✓	△	×	×	×	×	2.0
8	The world precedes the picture	✓	✓	✓	✓	✓	×	△	✓	✓	7.5
9	Grow decoupled, deliver combined	✓	△	✓	✓	✓	×	△	✓	✓	7.0
10	Few principles, much deduction	△	×	✓	✓	✓	×	△	×	×	4.0
11	Intelligence out of diversity	△	△	✓	✓	△	×	×	△	✓	5.0
12	Observation cheap, intervention scarce	✓	×	✓	✓	✓	×	✓	✓	✓	7.0

situations must share structure with one another. Only if the solution to one complicated situation can be used on the next does "a measure of generalization" follow.

When it fails. If every complicated situation is one of a kind, sharing no structure with the others, then untangling one produces consumption and no generalization. In practice this is a live question: complication often comes from idiosyncratic history, and the idiosyncratic is by definition not reusable. The pair therefore needs a measure: the reuse rate.

Testable form. Take a set of situations that were classified as complicated and check whether the solution from an earlier one was used on a later one. If yes, generalization holds; if no, it was consumption.

4.7 The relay of paradigms → the eighth year

What the theory says. T7's core proposition is precisely not a smooth relay. Kuhn describes normal science solving puzzles inside a paradigm, anomalies accumulating into crisis, then revolution; old and new paradigms are **incommensurable** — there is no shared scale on which to compare their achievements, and some problems of the old paradigm stop counting as problems at all.

What it is mapped onto. In eight years the firm passed the baton twice, and the curve of growth never broke [1].

When it holds. "When a paradigm reaches declining returns, growth does not stop — the baton passes" holds as a description of **successive S-curves**, which is why the image is popular in industry narrative.

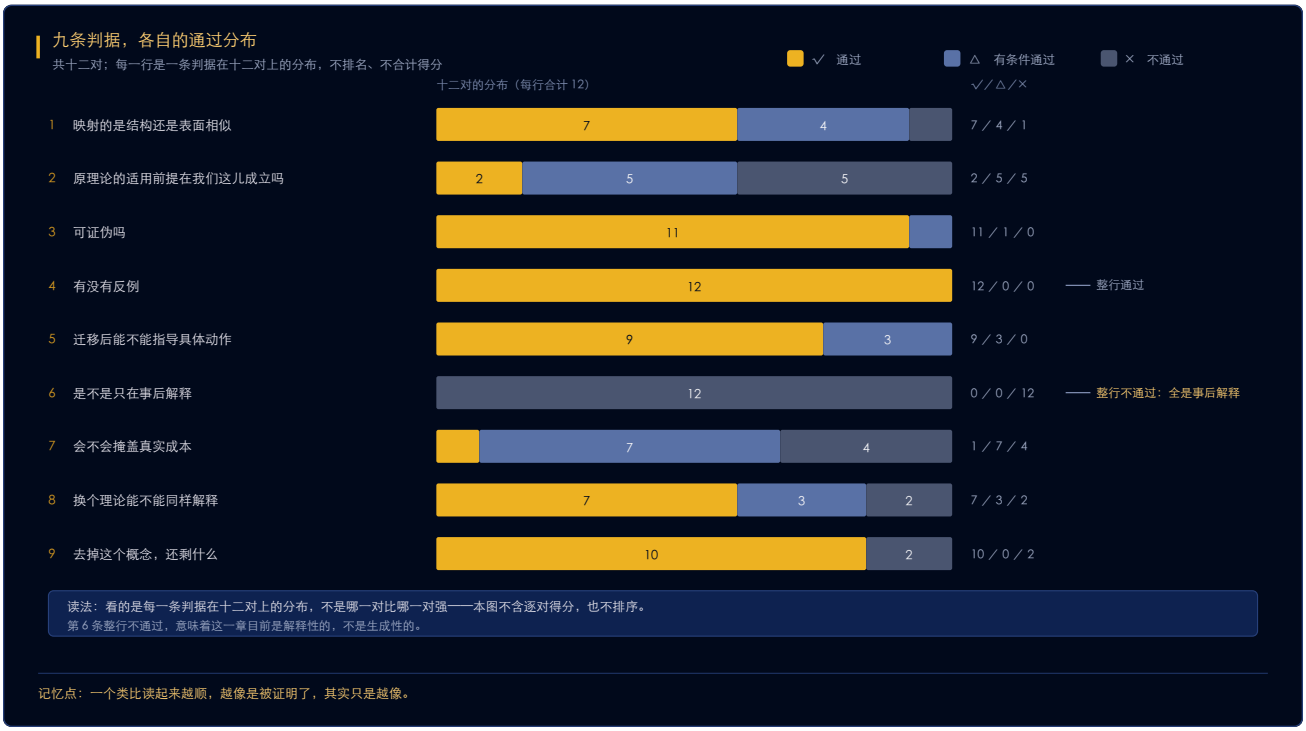

图 2 / Exhibit 8 How each of the nine criteria is distributed across the twelve pairs. The figure is this paper's own synthesis and the counts are taken column by column from Table 5 in this section; each row sums to twelve. The figure shows the distribution of each criterion only — it carries no ranking, no per-pair total and no position for any pair.

When it fails. On Kuhn's own terms, paradigm change involves rupture rather than continuity: old measures are abandoned and old achievements become incomparable. "The curve never broke" is closer to Kuhn's cumulative normal science than to his revolution. So what has been borrowed here is the image of a relay, not the core proposition of the theory. That is not necessarily an error, but it should be said plainly: **this is a decorative metaphor, not a generative one** [18].

What it can still contribute. Kuhn's theory contains a component that lands directly: the anomaly. The leading indicator of paradigm change is not slowing growth but **an accumulation of failures the current method cannot explain.** Translated into an action: maintain a list of unexplained failures. That list will tell you it is time to change method earlier than any growth metric. This part is drawn from the source theory and is not on the website.

4.8 The world precedes the picture → restoring the facts

What the theory says. T8 introduces the latent-variable energy-based programme, behind which lies a public debate: should intelligence model the latent world that produces observations, or the observations themselves? The argument for the former is that prediction in latent space can ignore unpredictable detail, whereas prediction in pixel space is forced to model noise.

What it is mapped onto. The deck, the memo and the roadshow are all projections; the first step of every project is to rebuild the business, the technology, the customers, the transaction history and the key risks like a corporate archaeologist — build the world first, render the picture after [1].

When it holds. The premise is that the underlying world exists and can be rebuilt well enough to constrain the materials. Financial facts, technical facts and transaction history can largely be rebuilt because they leave traces. This is the cleanest structural mapping of the twelve: **the relation of latent variable to observation maps onto the relation of fact to material, with the causal direction the same on both sides.**

When it fails. In two places. First, intentions, relationships and future plans are not traces; they exist

only in people, and rebuilding them becomes guessing. Second, and more consequentially: treating every "projection" as noise discards real information. The market's narrative about a company **is itself part of the facts**; it affects pricing and it affects what others do. In a latent-variable model, pixels are effects. In the primary market, the picture is sometimes a cause.

Testable form. Of the material discarded as noise in a given fact-rebuilding exercise, how much later proved to have affected the outcome.

4.9 Grow decoupled, deliver combined → twenty-seven subsystems

What the theory says. T9 proposes near-decomposability: complex systems tend to be hierarchic, with within-subsystem interactions far stronger than between-subsystem ones; in the short run each subsystem's dynamics are approximately independent, and in the long run only aggregate behaviour matters to the others. Simon's example is the heat-diffusion matrix of an eight-cubicle, three-room building, where the coefficients differ by one to two orders of magnitude — that difference is the quantitative meaning of "near".

What it is mapped onto. Each subsystem sits in its own layer and forms its own chapter, and in a live transaction they are re-orchestrated into one delivery; splitting is so each grows stronger, joining is so they accomplish one thing — a list expires, a system does not [1].

When it holds. Simon's condition can be quantified: within-layer coupling must be far stronger than between-layer coupling. The organizational counterpart is that **interfaces are sparse and stable.** If two subsystems renegotiate how they join every month, near-decomposability does not hold, and splitting produces coordination cost rather than parallel growth.

When it fails. When a single transaction requires twenty-seven subsystems to negotiate at high bandwidth simultaneously, the structure is slower, not faster. The benefit of a near-decomposable system is that it **localizes change**; once change cannot be localized, hierarchy is only overhead.

"A list expires, a system does not" needs a bound. Systems expire too, one order of magnitude more slowly. The routines literature [10][11] says exactly this: capability is stored in routines, and routines have inertia. A more accurate statement: lists expire in months, systems in years. This paper does not alter the original sentence; it adds the bound.

Testable form. Count how many times the interface between any two subsystems was modified in the past year. A low count means near-decomposability holds; a high count means the split is producing cost.

4.10 Few principles, much deduction → a checklist that only subtracts

What the theory says. T10 is self-play reinforcement learning: no human game records, only the rules; self-play and search generate the training signal, and network and search improve one another. Its success rests on three premises, none dispensable: the rules are fixed and fully known; the outcome of each game is immediately and unambiguously determined; play can be repeated without limit.

What it is mapped onto. What we write down are boundaries, not plays; every pit becomes a prohibition; the finer the rules, the more people work the seams, so grey areas go to the person everyone trusts; boundaries are only added, never removed, and the plays grow out of real transactions [1].

When it fails — this pair must start with the failure. All three premises are absent in the primary market: the rules are not fixed (jurisdictions, markets and participants all change); feedback is slow and ambiguous (a round takes months, the real outcome waits for an exit); and self-play is impossible (one cannot conjure a transaction to practise on). **This is the pair that reads best and whose premises fail most completely**, and Section 7 uses it as the main example of boundary case one.

What remains. T10 contains one conclusion only loosely tied to those premises: injecting human prior knowledge — game records, handcrafted features — limits the long-run ceiling. Its organizational counterpart is that **over-specified rules limit adaptation.** But the website's own sentence, "the finer the rules, the more people work the seams", is actually about something else: rule refinement inviting rule arbitrage, which Goodhart-style metric failure or institutional economics explains more directly, with no need for self-play.

Where the value of this pair lies. Not in the argument but in the **phrasing.** "Write boundaries, not plays" is a rule that can be executed immediately, and it is

a good rule; its justification simply does not come from T10. This is exactly what row 8 of the audit (could another theory explain it equally well) is built to catch: **the practice is right, the reason was borrowed from the wrong place.** Keep the practice; change the reason.

4.11 Intelligence out of diversity → an organization on a sphere

What the theory says. T11 holds that mind is composed of many agents that are individually unintelligent, and that intelligence appears in how they are organized; the same node is an agent to those above and an agency to those below. The book's emphasis is on **hierarchy and the division of labour**; diversity is the means.

What it is mapped onto. The firm is not an org chart but a set of directions on one sphere spread as far apart as possible; strength rests not on a lone hero but on heterogeneous directions complementing one another — crowd together and everyone becomes the average [1].

When it holds. One difference cannot be skipped: T11's agents are simple and unintelligent, they need not understand one another, and their power comes from structure. Members of an organization are not simple agents; they must communicate, and communication cost rises with diversity. Diversity therefore pays under two conditions: the problem space is wide enough that several angles are needed to cover it, and the task is decomposable enough that the outputs of different angles can be combined.

When it fails. March [16] locates the turning point: the rate of mutual learning between individuals and the organizational code determines long-run performance. Socialize too fast and diversity is absorbed, leaving a mediocre consensus — precisely what the website means by "crowd together and everyone becomes the average". But the converse also holds: in a stable environment, slow socialization is waste. **So "spread the directions as far as possible" is not an unconditional good; the optimum depends on how fast the environment changes.** That condition is what the original sentence omits.

This pair lacks a measure. How far apart is far enough? The original does not say. We propose a measurable proxy: **the proportion of cases in which two people, judging the same question independently, reach different conclusions.** Too low and the directions have crowded together; too high and there is no common ground. It can be counted quarterly.

4.12 Observation is cheap, intervention is scarce → putting our own money in

What the theory says. T12's ladder of causation has three rungs: association (seeing), intervention (doing), counterfactuals (imagining). The core proposition is that a question on a higher rung cannot be answered with data from a lower one alone. Note that intervention has a strict meaning: $do(x)$ means setting a variable **exogenously**, severing all of its usual inputs.

What it is mapped onto. Looking at a hundred cases is observation; putting one's own money into a company one has served deeply is intervention; only intervention with feedback teaches how states transition [1].

When it fails — under the source theory this pair does not hold. The decision to invest or not is made on the basis of an expectation about the outcome. That is endogenous, not $do(x)$. It is the textbook case of selection: invest only in what you favour, then observe the result, and what you obtain is a selected sample, not intervention data. **Under Pearl's definition, a self-selected action is not an intervention.** This is the one pair of the twelve where the core definition of the theory conflicts directly with the mapping.

What the sentence is actually saying. Something else, also true and also important: **bearing the consequences changes the feedback you receive.** Not investing lets you see only whether you were right; investing brings continuous, costly, unavoidable feedback. In the literature that is closer to learning under skin in the game than to the do-calculus. The correction does not weaken the sentence; it makes it testable.

The correction, and the action it points to. The genuinely scarce rung of the ladder is the third — the counterfactual. "What if we had not invested", "what if we had held back" — answering those requires data on **what happened to the projects that were declined.** Almost no institution records this systematically, because it produces nothing visible and does not look good. We think this is the ledger most worth building: **record the projects that were turned down, not only those that were done.** It is the one addition this paper proposes to the chapter's practice.

记忆点 · KEY POINT

What is truly scarce is not intervention but the counterfactual — who is recording the deals you turned down?

5. The Tool: An Analogy Audit

5.1 Nine rows

The five bodies of work in Section 3 compress into one table, for anyone about to borrow a concept tomorrow. Method: go down the rows. If any row lands in the "failing" column, the borrowing cannot be used as argument; it may only be used as phrasing.

Row 9 is Tsoukas's criterion [18] made operational: **a generative metaphor cannot be removed; a decorative one costs nothing to remove.** It is placed last because it is the harshest.

5.2 Scoring the twelve pairs

Three readings.

First, column 6 is entirely ×. All twelve phenomenon sentences begin with "this explains…", and the practice explained existed before the borrowing. That does not make the transfers worthless — externalization has value in itself (Section 3.4) — but it means the chapter as it stands is **explanatory**, not **generative**. To become generative it needs at least one case in the other direction: theory first, practice after. Section 6 writes this as a testable corollary.

Second, the lowest scores are pair 7 (relay of paradigms, 2.0) and pair 10 (few principles, 4.0). What they share: the sentences read best and the premises hold least. That is boundary case one in numerical form.

Third, the highest are pairs 5 and 8 (7.5 each) and pairs 1, 4, 9 and 12 (7.0 each). What they share: the causal direction is the same on both sides, and each lands on something countable — how often the prior has been rewritten, how much discarded material later proved to matter, how often an interface changed in a year, whether anyone records the deals turned down. **Landing on something countable is the strongest signal in the table.**

记忆点 · KEY POINT

A borrowed concept must change one specific action, or it is only a name.

6. A Proposition That Can Be Refuted

Main proposition. If an organization's borrowing of an outside concept is a structure mapping — mapping relational and causal structure rather than surface resemblance — then it should produce at least one action the organization did not take before the borrowing, began taking after it, and that a third party can observe. A borrowing that only explains existing practice and changes no action is rhetoric, not mechanism.

What would refute it. Across a set of organizations, divide their borrowed concepts by criterion 1 (structure versus surface) and compare the proportion in each group that produced a new observable action within twelve months. If the proportions do not differ, the main proposition fails — structure mapping and surface resemblance would have the same consequences, and criterion 1 would not be a criterion.

Corollary one (distribution; tested once here). If borrowing runs on absorptive capacity [8], the concepts an organization borrows should concentrate in the field it has long served rather than scatter across whatever is currently fashionable. Table 4 shows 67% of the twelve sources come from machine learning and artificial intelligence, consistent with eight years spent on hard technology, and 42% predate 1997, which is inconsistent with chasing novelty. **Corollary one is not refuted on this sample.** If another firm's borrowed concepts were found scattered across several fashionable fields unrelated to its business, corollary one would be refuted and the mechanism would be trend-following, not absorptive capacity.

Corollary two (direction; not yet tested). If any part of this borrowing is generative, there should exist at least one practice for which the concept came first. The test is chronological: the date a concept entered internal text versus the date the practice was first performed. Column 6 of Table 6 shows that public text cannot settle the order, because public text gives only the present state. **This requires internal timestamps and is left explicitly open.**

Corollary three (pairing). Section 4.3 argues that pair 3 (end-to-end) must be used together with pair 4 (teacher

model), because end-to-end costs attributability and the teacher supplies external evaluation. Testable form: in organizations that adopted end-to-end without an external evaluator, the time required to localize a failure should be longer. If localization is equally fast without an external evaluator, the pairing claim fails.

记忆点 · KEY POINT

If you cannot say what observation would make it false, it is not a judgement, it is a sentence.

7. Boundaries and Counter-Cases

An analysis with only supporting instances does not deserve trust. Five places where this practice is known to fail.

One: an analogy that reads well is not a decision that is right. This is the most important. The two lowest-scoring pairs in Table 6 — the relay of paradigms, and few principles much deduction — are the two that read best. They read well because the image is strong, the sentence short and the picture clear; and their premises do not hold in the primary market (Sections 4.7 and 4.10 list them). Tversky and Kahneman [14] explain why: people substitute similarity for probability, so **reading like it is true gets taken for having been shown true.** Gilovich [15] shows further that surface cues shift policy judgement without the decider's awareness. The criterion is not how well the sentence reads but whether the premises hold. An analogy may stay on the wall as phrasing; it may not enter a chain of argument. These are two uses and they must be kept apart.

Two: the pull of after-the-fact explanation. Column 6 of Table 6 is entirely ×. Every sentence says "this explains…", and the practice explained existed before the borrowing. After-the-fact explanation has two features, both hard to notice in oneself: it never fails (a concept can always be found to explain what already happened), and it feels like understanding (similarity taken for comprehension). There is only one way to tell them apart, and it is chronological: **theory before practice is mechanism; practice before theory is naming.** Naming has value — it can be taught and criticized (Section 3.4) — but it is not explanation. This section has to be in the paper because the chapter's sentence pattern turns naming into explanation automatically, and the writer will not notice.

Three: only the pairs that made it onto the wall are visible. The denominator is unknown. A concept that was tried, found not to fit, and discarded never appears on a website. The score distribution in Table 6 is therefore a **survivor sample** and must not be read as the quality of this firm's analogies. The real quality indicator is the discard rate — how many pairs were tried and how many kept. That requires a list of discarded analogies, and nobody builds such a list spontaneously. This is the same conclusion as Section 4.12: **record what was turned down, and only then does a denominator exist.**

Four: a borrowed word becomes a pass that exempts from inspection. Externalization [12] turns an unsayable judgement into a sayable name; the benefit is teachability, the cost is that a name which sounds good stops being questioned. Once a practice carries a frontier concept, discussion turns to justifying the concept rather than checking the practice. Tsoukas [18] and Morgan [17] describe exactly this: a root metaphor illuminates one part and obscures another, and the obscured part is the part no longer asked about. Row 9 of Table 5 is the antidote.

Five: reverse contamination — preferring clients your theories already explain. Absorptive capacity [8][9] is a filter: it governs what gets in and what stays out. An organization that has long understood the world through one set of concepts will find it easier to understand companies that speak in the same terms, and harder to understand those that do not. Understanding quickly is easily mistaken for judging correctly. There is no ready antidote, only a reminder: **fluency of understanding is not itself evidence.**

> Theory before practice is mechanism; practice before theory is naming.

8. Scope

One: object. This paper analyses twelve conceptual transfers in one public text, produced by one professional service firm. The conclusions apply to the class of cases in which an intermediary that has long served a technical field borrows that field's concepts. They are not extrapolated to transfers between technology companies, nor to organizations whose work is not primarily knowledge work.

Two: the nature of the evidence. Tables 1 and 2 are verbatim reproductions with reproducible line numbers. Table 4 is counted from Table 3 by rules stated in the writing note. **Table 6 is judgement, not measurement.** Several of the nine criteria, especially rows 5 and 6, need internal information to be judged accurately; here they are judged from public text only, so those cells are conservative — biased toward × or △. A different scorer would differ on individual cells; the three readings depend on the overall shape (the whole of column 6, the lowest two, the highest few), not on any single cell.

Three: a single sample. One firm, one text. The main proposition of Section 6 requires a cross-organization sample, which is beyond this paper. Corollary one is not refuted on this sample, which is not the same as confirmed.

Four: no causal inference. The paper does not claim that borrowing these concepts caused any outcome for the firm. Section 4 concerns the conditions under which a mapping holds, not what the borrowing produced. Reading the scores in Table 6 as a performance measure is a misreading.

Five: the statement of the source theories. The summaries of the twelve sources in Section 4 are simplifications made for the purpose of judging the mappings and do not substitute for the originals. The three places where the theory name and the cited source are noted as not fully aligned (Sections 4.2, 4.6, 4.7) concern the correspondence between concept and source, not any defect in the original papers.

Six: an honest gap. What this paper has not done is obtain internal timestamps to test whether any practice followed rather than preceded its concept (corollary two). That requires an internal chronology that public text cannot supply. Until then, the judgement of "explanatory rather than generative" is a working hypothesis open to refutation.

Seven: what this paper does not do. It does not evaluate any business outcome, does not constitute investment advice, and contains no information traceable to any specific transaction or company.

> Table 6 is judgement, not measurement; what refutes it is internal timestamps, not better wording.

9. Conclusion

An organization that serves frontier-technology companies for years keeps finding that the concepts its clients use to explain technology also explain the organization itself. This paper sets out twelve such concepts verbatim and asks of each the question the website does not ask: under what conditions does this mapping hold?

Back to the two questions of Section 1.

Does "understanding an industry" have testable content? Yes. The testable content is: you can borrow that industry's concepts, and you borrowed them correctly. That you can borrow is the result of absorptive capacity — Table 4 shows two-thirds of the sources concentrated in one field and four in ten of them sixty years old, which is what reading originals rather than headlines looks like. That you borrowed correctly is what the nine rows of Table 5 test, above all row 1 (structure or surface) and row 5 (does it change a specific action). A firm whose borrowed concepts are scattered across this season's vocabulary and have changed no action is not understanding; it is following along.

Is borrowing a client's theory to explain yourself insight or rhetoric? Both are present and they can be separated. The separation is row 9 of Table 5: **remove the concept and see what is left.** If what remains is something countable — how often the prior has been rewritten, whether reviews record the second-best option, how often an interface changed in a year, whether anyone records the deals turned down — it was insight. If nothing is missing and the text merely reads less well, it was rhetoric. Rhetoric is not a sin, but it may not enter a chain of argument.

The paper proposes one addition to the chapter's practice, and only one: **record the projects that were turned down.** Section 4.12 shows that the scarce rung of the

ladder of causation is not intervention but the counterfactual, and the data a counterfactual needs is what did not happen. Boundary case three shows that the quality of analogies can only be measured against the ones that were discarded. Both point to the same ledger, and we think it is the most worthwhile thing the chapter could add.

Three sentences summarize the paper:

1. That a client's theory explains us is rarely a coincidence; we can only hear the part we already know.
2. The better an analogy reads, the more it feels proven — the criterion is the premises, not the sentence.
3. A borrowed concept must change one specific action, or it is only a name.

> Remove the concept: if a countable action is missing, it was insight; if nothing is missing, it was rhetoric.

References

Format: author (year). Title. Journal/publisher, volume(issue), pages. DOI or accessible link. Grouped by source grade: primary material is the Glacier Institute's own public text; secondary material is peer-reviewed literature and monographs. Each entry was checked by live request on 2026-09-22 with the `check()` function of `_仓储/_build/refcheck.py`; **all 18 are reachable**. Details are in the writing note, § 4. No unverified entry appears.

Primary material (Glacier Institute public text; every quoted sentence comes from [1][2] unchanged)

[1] Glacier Institute (2026). Glacier Institute website, Chapter 18 "THEIR THEORY, OUR PHENOMENA", complete chapter with its source annotations. https://glacier.mba/institute.html (accessed 2026-09-22)

[2] Glacier Capital (2026). Glacier Capital home page, Chapter 18 card (introduction and two of the pairs). https://glacier.mba/index.html (accessed 2026-09-22)

Secondary (analogy and conceptual transfer)

[3] Gentner, D. (1983). Structure-Mapping: A Theoretical Framework for Analogy. *Cognitive Science*, 7(2), 155–170. https://doi.org/10.1207/s15516709cog0702_3 (publisher returns 403 to scripts; see also the author's institutional PDF https://groups.psych.northwestern.edu/gentner/papers/Gentner83.2b.pdf ; accessed 2026-09-22)

[4] Gentner, D., & Markman, A. B. (1997). Structure Mapping in Analogy and Similarity. *American Psychologist*, 52(1), 45–56. https://doi.org/10.1037/0003-066X.52.1.45

[5] Holyoak, K. J., & Thagard, P. (1989). Analogical Mapping by Constraint Satisfaction. *Cognitive Science*, 13(3), 295–355. https://doi.org/10.1207/s15516709cog1303_1

[6] Gick, M. L., & Holyoak, K. J. (1980). Analogical Problem Solving. *Cognitive Psychology*, 12(3), 306–355. https://doi.org/10.1016/0010-0285(80)90013-4

[7] Holyoak, K. J., & Thagard, P. (1997). The Analogical Mind. *American Psychologist*, 52(1), 35–44. https://doi.org/10.1037/0003-066X.52.1.35

Secondary (organizational learning, absorptive capacity, routines)

[8] Cohen, W. M., & Levinthal, D. A. (1990). Absorptive Capacity: A New Perspective on Learning and Innovation. *Administrative Science Quarterly*, 35(1), 128–152. https://doi.org/10.2307/2393553

[9] Zahra, S. A., & George, G. (2002). Absorptive Capacity: A Review, Reconceptualization, and Extension. *Academy of Management Review*, 27(2), 185–203. https://www.jstor.org/stable/4134351 (publisher page https://journals.aom.org/doi/10.5465/amr.2002.6587995 returns 403 to scripts; accessed 2026-09-22)

[10] Nelson, R. R., & Winter, S. G. (1982). *An Evolutionary Theory of Economic Change*. Harvard University Press. https://www.hup.harvard.edu/books/9780674272286 (accessed 2026-09-22)

[11] Feldman, M. S., & Pentland, B. T. (2003). Reconceptualizing Organizational Routines as a Source of Flexibility and Change. *Administrative Science Quarterly*, 48(1), 94–118. https://doi.org/10.2307/3556620

Secondary (knowledge conversion and tacit knowledge)

[12] Nonaka, I. (1994). A Dynamic Theory of Organizational Knowledge Creation. *Organization Science*, 5(1), 14–37. https://doi.org/10.1287/orsc.5.1.14

[13] Polanyi, M. (1966). *The Tacit Dimension*. University of Chicago Press (2009 reissue). https://press.uchicago.edu/ucp/books/book/chicago/T/bo6035368.html (accessed 2026-09-22)

Secondary (boundaries: similarity as evidence, learning trade-offs, metaphor)

[14] Tversky, A., & Kahneman, D. (1974). Judgment under Uncertainty: Heuristics and Biases. *Science*, 185(4157), 1124–1131. https://doi.org/10.1126/science.185.4157.1124

[15] Gilovich, T. (1981). Seeing the Past in the Present: The Effect of Associations to Familiar Events on Judgments and Decisions. *Journal of Personality and Social Psychology*, 40(5), 797–808. https://doi.org/10.1037/0022-3514.40.5.797

[16] March, J. G. (1991). Exploration and Exploitation in Organizational Learning. *Organization Science*, 2(1), 71–87. https://doi.org/10.1287/orsc.2.1.71

[17] Morgan, G. (1980). Paradigms, Metaphors, and Puzzle Solving in Organization Theory. *Administrative Science Quarterly*, 25(4), 605–622. https://doi.org/10.2307/2392283

[18] Tsoukas, H. (1991). The Missing Link: A Transformational View of Metaphors in Organizational Science. *Academy of Management Review*, 16(3), 566–585. https://doi.org/10.5465/amr.1991.4279478

Theory-side sources of Table 3 (T1–T12, taken from the website's own "source" lines [1]; accessible addresses supplied and checked here, 12/12 reachable on 2026-09-22)

T1 Jaderberg, M., Mnih, V., Czarnecki, W. M., Schaul, T., Leibo, J. Z., Silver, D., & Kavukcuoglu, K. (2017). Reinforcement Learning with Unsupervised Auxiliary Tasks. ICLR 2017. https://arxiv.org/abs/1611.05397 T2 Jacobs, R. A., Jordan, M. I., Nowlan, S. J., & Hinton, G. E. (1991). Adaptive Mixtures of Local Experts. *Neural Computation*, 3(1), 79–87. https://doi.org/10.1162/neco.1991.3.1.79 T3 Bojarski, M., et al. (2016). End to End Learning for Self-Driving Cars. https://arxiv.org/abs/1604.07316 T4 Hinton, G., Vinyals, O., & Dean, J. (2015). Distilling the Knowledge in a Neural Network. https://arxiv.org/abs/1503.02531 T5 Ha, D., & Schmidhuber, J. (2018). World Models. https://arxiv.org/abs/1803.10122 T6 Wolpert, D. H., & Macready, W. G. (1997). No Free Lunch Theorems for Optimization. *IEEE Transactions on Evolutionary Computation*, 1(1), 67–82. https://doi.org/10.1109/4235.585893 T7 Kuhn, T. S. (1962). *The Structure of Scientific Revolutions*. University of Chicago Press. https://press.uchicago.edu/ucp/books/book/chicago/S/bo13179781.html T8 Dawid, A., & LeCun, Y. (2023). Introduction to Latent Variable Energy-Based Models: A Path Towards Autonomous Machine Intelligence. https://arxiv.org/abs/2306.02572 T9 Simon, H. A. (1962). The Architecture of Complexity. *Proceedings of the American Philosophical Society*, 106(6), 467–482. https://www.jstor.org/stable/985254 T10 Silver, D., et al. (2017). Mastering the Game of Go without Human Knowledge. *Nature*, 550, 354–359. https://doi.org/10.1038/nature24270 T11 Minsky, M. (1986). *The Society of Mind*. Simon & Schuster. https://openlibrary.org/books/OL2714978M/The_society_of_mind T12 Pearl, J., & Mackenzie, D. (2018). *The Book of Why: The New Science of Cause and Effect*. Basic Books. https://www.basicbooks.com/titles/judea-pearl/the-book-of-why/9780465097609/

Verification note: [1][2] and [10][13] were confirmed by HTTP request; [4][6][7][11][15][17][18] resolve directly through DOI with status 200; [3][5][8][12][14][16] return 403 to scripts at the publisher (anti-scraping, not absence) and were verified against Crossref metadata for title, volume, issue and pages; the AOM DOI for [9] did not resolve on 2026-09-22, so the stable JSTOR address is used and the title was checked against Crossref. Among T1–T12, T1, T3, T4, T5 and T8 were checked against arXiv pages and the arXiv API title; T2, T6 and T10 against Crossref titles; T7, T9 and T12 are publisher or JSTOR pages returning 200; T11 was checked against Open Library and Internet Archive records.

Sources of the quoted Glacier Institute text

All quotations come from [1] https://glacier.mba/institute.html (retrieved 2026-09-22; plain text after stripping comments, scripts, styles and tags):

- Introduction ("Glacier has worked alongside the most advanced companies…") — L3867 (also on the home page [2])
- Closing line ("Learning from the companies you accompany is the privilege of this work.") — L4194
- The theory sentences, phenomenon sentences and source lines of all twelve pairs — line numbers per pair in Tables 1 and 2, spanning L3868 – L4193
- Chapter title and headline — L3839

De-identification note: **no client company is named anywhere in the original text of Chapter 18**, so no "a certain company" rewriting was required in any quotation; the count of rewrites is 0. This paper's own prose likewise describes no characteristic that could identify a specific company. See the writing note, §5.