

没有人要求的时候

论「没有问题」与「已经不会变了」何以在全部现行读数上同形，及把两者分开的那个量为何正被三门学科同时花钱消灭

认知科学 × 物理学 × 工程学

作者 王德生 + Claude

栏目 学科通融 · 站外碰撞

篇幅 约2.0万字 · 28页

发表 2026年8月3日

摘 要

一条连续两年零缺陷的产线、一位连续八年零事件的飞行员、一门通过率九成五的课程——读数全部完美。现在问一句：**它们还能不能被改变？** 三门学科给三种互不相容的读法：认知科学看预测误差（误差为零就是学完了），物理学看受扰时的响应（响应函数就是全部信息），工程学看余量与工况（余量足就是健康）。三条不能同为真，却共同假定了一件谁也没说出口的事——**一样东西的能力，只有在它被要求做事的时候才存在、才可测。** 推翻它的材料来自物理学自己：把自发涨落与响应连起来的那条关系式说，那个只在「**什么也没要求**」时才存在的量，与「**被要求时能出多大反应**」，是**同一个量**；物理学用它从平衡涨落算输运系数，从没人拿它当作「可改变性该在什么时候测」的证据。由此命名**静息变异**——没有任务、没有考核、没有故障的时段里，系统自发做出的与自己默认做法不同的那些事。要紧的不是三家没给它设字段：三家都设了字段，而且三家都在花钱把它降到零（失误差／噪声／非计划偏移）。三家共同出资消灭的那个量，正是它们后来要用的那个能力。文中给出可事后清点的比率读数与一张两轴四格，并用一份二〇一四年的座舱实测作检验——自动化提高之后，飞行员用手做的技能保持完好，退化的恰是自动化替他做了、因而整段航程无人要求他做的那部分。最后一句不打算缓和：可靠性要求把变异消灭，可改变性要求把变异留住，而两者用的是同一笔预算。

这一篇由哪三个学科的理论体系撞成

第一家主张驱动一切改变的只有预测与实际之间的落差，落差归零即模型收敛、不再改变是正确的；第二家主张可被推动的程度由响应函数单独决定，不施加扰动就无从判断；第三家主张能不能承受未来的要求由当下余量与所处工况单独决定，且不该为了知道好不好而去动它。三条在同一个具体处境里给出互相取消的处置，且各自的日课正是另一家诊断出的病：压低误差是第一家的日课，而第三家的余量学说指出误差被压到零的系统其退化不产生任何信号；施加扰动测响应是第二家的日课，而非破坏性检测这一整套东西存在的理由，正是「为了知道它好不好而去动它」本身不可接受；留足余量是第三家的日课，而第一家会说余量正是让落差归零的装置——余量越足，更新越少。 三家均为站外的公开文献，链接直达原始出处，可自行核对。

认知科学 · 预测误差单驱动，与它的暗室问题

[Sun & Firestone, The Dark Room Problem \(*Trends in Cognitive Sciences* 24:5, 2020\)](#); 答复见 [Seth, Millidge, Buckley & Tschantz, Curious Inferences \(*TiCS* 24:9, 2020\)](#); 原始提法见 [Friston, Thornton & Clark, Free-Energy Minimization and the Dark-Room Problem \(*Frontiers in Psychology* 3:130, 2012\)](#)

主张认知改变由单一的预测落差驱动，落差归零即无驱动；而这一支自己承认，只按当下落差行动的系统会躲进暗室，故必须把认识论价值算进来——即系统要**主动去寻求它现在还没有的落差**，而该往哪儿寻求，只能在落差为零时被决定。这一处正是本文的入口。

<https://doi.org/10.1016/j.tics.2020.02.006>

物理学 · 涨落与响应，及它在被困住的体系里破缺

[Kubo, The Fluctuation-Dissipation Theorem \(<i>Reports on Progress in Physics</i> 29:1, 1966\)](#); 破缺与「有效温度」的综述见 [Crisanti & Ritort \(<i>J. Phys. A</i> 36:21, 2003\)](#); 反面证据见 [Jabbari-Farouji 等 \(<i>EPL</i> 84:2, 2008\)](#)——软玻璃材料跨若干频率量级未见破缺，两方至今未决

主张可被推动的程度由响应函数单独决定；而这一支自己写下的那条关系式表明，响应在没有外场时就已确定、且等于自发涨落的关联——**推翻共有前提的材料就是它**。它在破缺情形下的读数反过来削掉了本文的价值排序：被困住的体系里涨落还在而响应没了。

<https://doi.org/10.1088/0034-4885/29/1/306>

工程学 · 余量与降额，及余量掩蔽退化

[Casner, Geven, Recker & Schooler, The Retention of Manual Flying Skills in the Automated Cockpit \(<i>Human Factors</i> 56:8, 2014\)](#); 对立一支见 [Hollnagel, <i>Safety-I and Safety-II</i> \(Ashgate, 2014\)](#)——日常变异不是要消灭的威胁而是成功与失败共同的来源

主张能不能承受未来的要求由余量与工况单独决定；而这一支的下游产业（剩余寿命预测）存在的全部理由，正是**余量在被消耗时不产生任何信号**。所引那份座舱实测给出本文的经验锚：自动化之后退化的不是用得少的那部分，而是**整段航程里没有任何人要求飞行员做的那部分**。

<https://doi.org/10.1177/0018720814535628>

目 录

摘要.....	5
一、导论：三个不该同时为真的判断.....	6

二、三家的立场，以及各自留下的那一处松动.....	8
三、冲突定位：三家在同一个具体处境里给出互相取消的处置.....	10
四、共同假定，与推翻它的那件材料.....	10
五、静息变异：定义与五条判定条件.....	11
六、静息变异率：三个领域量纲不同而比率相同.....	12
七、辨别格：现行表现 × 静息变异.....	13
八、为什么焊住不产生信号：余量掩蔽.....	13
九、一份已经做过的观测：自动化座舱里流失的是哪一部分.....	14
十、在十二个行当里兑现.....	15
十一、判据、读法与不得越过的界线.....	16
十二、与最接近的几种说法的分界.....	18
十三、与同一栏目里几篇的分界.....	20
十四、三家为什么各自到不了这一条.....	21
十五、一处强制的不一致.....	22
十六、三家反过来给本文加的三条约束.....	23
十七、不适用的边界.....	23
十八、证伪条件.....	24
十九、反噬预言.....	25
二十、自反：本文自己在哪一格.....	25
二十一、结论，与一条写死日期的赌注.....	26
注 释.....	27

摘要

一个系统还能不能被改变，凭什么读得出来？认知科学答：看它的内部模型还在不在被更新——预测误差为零就是学完了。物理学答：看它受到扰动时给出多大响应——响应函数就是全部信息。工程学答：看它还剩多少余量、处在什么工况里——余量是唯一的健康判据。三条不能同时为真，而三条共同假定了一件谁也没说出口的事：**一样东西的能力，只有在它被要求做事的时候才存在、才可测。** 推翻这条假定的材料来自物理学自己：涨落与响应之间的那条关系式说，一个系统在平衡时的自发涨落，与它此后对外扰动的响应，在数值上是同一个量。也就是说，那个只在“什么也没要求”时才存在的量，恰好就是“被要求时能出多大反应”。物理学用这条关系式从平衡涨落算输运系数，从没有人把它当作“可改变性该在什么时候测”的证据。本文据此命名**静息变异**——在没有任务、没有误差信号、没有外部扰动的时段里，系统自发做出的与它当前默认做法不同的那些事——并给出可事后清点的比率读数**静息变异率**。要紧的不是三家没给它设字段：三家都给它设了字段，而且三家都在花钱把它降到零（认知学记为失误率、物理学记为噪声、工程学记为非计划偏移）。**三家共同出资消灭的那个量，正是它们后来要用的那个能力。** 由此得到一张两轴四格：现行表现合格而静息变异为零的那一格，全部现行指标最好看，失效不产生任何信号，且越正规化越严重。判据不问人的判断，只问记录：**在最近一段没有人布置任务、没有考核、没有故障的时间里，这个系统做过几件与它默认做法不同的事？谁记下来了？** 一份二〇一四年的座舱实测为此提供了现成的检验：自动化程度提高之后，飞行员用手做的那部分技能保持得相当完好，退化的恰是用脑做的那部分——而那正是自动化替他做了、因而在整段航程里没有人要求他做的那部分。结论是一句本文不打算缓和的话：可靠性要求把变异消灭，可改变性要求把变异留住；同一个系统被同时要求做到两件事，而这两件事用的是同一笔预算。

关键词：静息变异；静息变异率；被改变性；余量掩蔽；调用即存在；涨落与响应

When Nothing Is Asked

Why "No Problem" and "No Longer Changeable" Look Identical on Every Current Instrument, and Why the Quantity That Separates Them Is Being Deliberately Driven to Zero by Three Disciplines at Once

Abstract. How can one tell whether a system is still capable of being changed? Cognitive science answers: watch whether its internal model is still being updated — zero prediction error means learning is complete. Physics answers: watch how large a response it gives to a perturbation — the response function is all the information there is. Engineering answers: watch how much margin remains and what operating conditions it sits in — margin is the sole health criterion. The three cannot all be true, and all three rest on an unstated shared premise: the capacity of a thing exists, and can be measured, only while something is being demanded of it. The material that overturns this premise belongs to physics itself. The

fluctuation–dissipation relation states that a system's spontaneous equilibrium fluctuation and its subsequent response to an external perturbation are numerically the same quantity. The quantity that exists only when nothing is asked is the quantity that determines how much can be got when something is asked. Physics uses this relation to compute transport coefficients from equilibrium fluctuations; no one has used it as evidence about when changeability should be measured. This paper names the resulting object **resting variation** — the variants a system spontaneously produces, differing from its own default, during intervals in which no task, no error signal and no external perturbation is present — and supplies a countable ratio, the **resting-variation rate**. The point is not that the three fields fail to record it: all three record it, and all three spend money driving it to zero (error rate, noise, unplanned deviation). The quantity the three fields jointly pay to destroy is the capacity they will later need. A 2×2 follows, whose "compliant performance, zero resting variation" cell scores best on every existing indicator, emits no signal on failing, and worsens with formalisation. The criterion asks records, not judgements: in the most recent stretch during which no task was assigned, no assessment applied and no fault occurred, how many things did this system do that differ from its default, and who wrote them down? A 2014 flight-deck study supplies a ready test: with increased automation, pilots' hand skills were largely retained while the cognitive skills degraded — precisely the part the automation performed for them, and therefore the part nothing asked of them for the whole flight. Reliability demands that variation be eliminated; changeability demands that it be kept; and both are drawn on the same budget.

一、导论：三个不该同时为真的判断

先摆一个具体处境。

一条产线连续两年零缺陷，一门课连续五届通过率九成五，一位飞行员连续八年零事件，一个团队连续十二个季度全部指标达标。四个系统，全部读数完美。现在问一句：**它们还能不能被改变？**

这个问题不是"它们好不好"。好不好已经答完了，答案是很好。问的是另一件事：如果明天出现一个它从未遇到过的要求，它拿得出反应吗？

眼下有三种回答，各有充分的经验支持，而它们不能同时为真。

第一种来自认知科学的主流一支。它说：看它的内部模型还在不在被更新。一个系统之所以改变，是因为它的预测与实际之间还有落差；落差驱动更新，更新缩小落差。**落差为零，就是模型已经收敛，就是该学的都学完了，此时不再改变是正确的，不是病。** 判断一个系统还能不能被改变，看它的误差信号还在不在就够了。

第二种来自物理学的线性响应理论。它说：看它在受到扰动时给出多大的响应。响应函数是全部信息——你想知道一个体系能被外力推动多少，就施加一个小外力，量它动了多少。**没有比这更直接的读法，也没有别的读法。**

第三种来自工程学的可靠性一支。它说：看它还剩多少余量，以及它正待在什么工况里。安全系数、降额、冗余度——这些数决定了它能承受什么。**判断一个系统的健康，看余量与条件就够了；余量足、条件在包线内，它就是健康的。**

三种回答在处置上互相取消。

误差为零时，认知这一支说不必动，动了是浪费；而工程这一支会说误差为零可能只是因为没有给过它一个够难的工况，那叫未经检验，不叫健康。工程这一支说余量足就好，而物理这一支会说余量是一个静态的量，它说不出体系对一个具体扰动的响应有多大——响应要量才知道。物理这一支说去施加扰动量响应，而认知与工程两家会同时反对：**施加扰动本身就改变了被测的东西**——对人是给了一次训练，对设备是消耗了一次寿命。

更要紧的是，三家各自每天在做的事，正是另一家诊断为病根的那件事。压低误差是认知这一支的日课，而工程这一支的余量学说恰恰指出：**误差被压到零的系统，其退化不会产生任何信号**。施加扰动测响应是物理这一支的日课，而工程规范里"非破坏性检测"这一整套东西存在的理由，正是"为了知道它好不好而去动它"这件事本身不可接受。留足余量是工程这一支的日课，而认知这一支会说余量正是让误差归零的装置——余量越足，落差越小，更新越少。

本文不打算在三者之间判谁对。判谁对需要它们站在同一根轴上，而它们不在：一家看内部路径，一家看外显响应，一家看所处条件。本文要做的是找出一条它们**共同假定、而谁也没有说出口**的东西，并用其中一家自己已经掌握、已经写进教科书、只是没有往这个方向用的材料把它推翻。

那条共同假定是：

> **一样东西的能力，只有在它被要求做事的时候才存在、才可测。**

三家争的是该看哪一维。三家都没有争的是：**该在什么时候看。**

推翻它的材料在第四节给出，来自物理学。而由此得到的判断，三家都不会同意，但三家的证据合起来只能得出它：

> 一个系统还能不能被改变，既不由它的误差信号读出，也不由它受扰时的响应读出，更不由它的余量读出，而由**它在没有人要求它做任何事的那些时段里，自发做出多少与它当前默认做法不同的事**读出。而这个量，三门学科都给它设了字段，都给它起了名字，**都在花钱把它降到零。**

这条判断的可查形式很短，全文都可以拿它去问：**在最近一段没有人布置任务、没有考核、没有故障的时间里，这个系统做过几件与它默认做法不同的事？谁记下来了？**

路线：第二节交代三家各自的立场与各自留下的一处松动；第三节把冲突写成一个可判定的形状；第四节找出共同假定并给出推翻它的材料；第五至八节建立本文的读数、辨别格与"为什么不产生信号"的机制；

第九节用一份已经做过的座舱实测做检验，那份实测的结果同时修正了本文最直觉的一条说法；第十节在十二个行当里兑现；第十一节给判据与不得越过的伦理界线；第十二至十五节与最接近的既有说法、与本栏已发的几篇、与三家各自的理由划清界线，并指出其中一家在不承认本文这个量时内部不自洽；第十六节写下三家反过来给本文加的三条约束；余下各节交出边界、证伪条件、反噬预言、自反定位与一条写死日期的赌注。

二、三家的立场，以及各自留下的那一处松动

2.1 认知科学：看内部模型如何被更新就够了

近二十年认知科学最有野心的一支，把知觉、行动、注意与学习统一在一条原则下：生物体在最小化预测误差（或其变分上界）。这一支的强版本主张是单因的：**驱动一切认知改变的只有一样东西，就是预测与实际之间的落差。**它给出的时序读数很清楚：落差先出现，模型在随后的若干次接触里被更新，行为最后跟上；落差归零，更新停止。

这一支自己留下的松动，是它被提出得最多的那一条反驳，而且这一支自己承认它是最要紧的一条：**暗室问题**。如果最小化预测误差是唯一的驱动，那么最省事的做法是躲进一间没有任何刺激的暗室——那里一切都完全可预测，落差恒为零。生物体显然不这么做。[1]

这一支的回答是：要在时间上延展的序列上做推断，就必须把**认识论价值**算进来——一个策略的价值不只是眼下满足预期，还包括它带来多少信息增益；后一项在数学形式里自然出现，不是外加的。[2]

这个回答是本文要用的那处松动。它等于承认：一个只按当下落差行动的系统会走进暗室；要不走进暗室，系统必须去**主动寻求**它现在还没有的落差。而"值得去寻求哪一处落差"这件事，**在落差还没出现的时候就要被决定**。这一支从没有问过：那个在落差为零时决定系统去哪儿的东西，是什么，怎么测。

2.2 物理学：看它受扰时给出多大响应就够了

线性响应理论说：给体系一个小的外场，量它的响应，响应函数就写尽了这个体系在这个方向上可被推动的全部性质。这是物理学里最成熟、最可靠的一套读法，输运系数——电导率、粘度、扩散系数——都是这么定义的。

它的单因主张同样清楚：**可被推动的程度是一个由响应函数单独决定的量，与体系此刻正在做什么无关。**时序读数：外场先加，响应随即，弛豫在特征时间内完成。

这一支自己留下的松动，就是它自己最著名的那条定理，第四节展开。这里先记一句：这条定理说，响应函数并不是一个只有施加外场才存在的东西——它在没有外场的时候就已经确定了，而且**在数值上等于体系此刻自发涨落的关联**。物理学把这条定理用来做一件很实际的事：从平衡态的模拟里算出输运系数，不必真的施加外场。

2.3 工程学：看余量与所处条件就够了

工程学的可靠性一支主张：一个系统能不能承受未来的要求，由它当下的余量与所处工况单独决定。安全系数、降额、冗余、包线——整套规范、认证与验收都建立在这上面，而且它极其成功：现代工程物的可靠性正是靠这一套做出来的。

时序读数：余量在设计期定死；使用中逐步消耗；失效在余量耗尽的那一刻出现。

这一支自己留下的松动，是它整个下游产业存在的理由：余量在被消耗的过程中**不产生任何信号**。系统在余量耗尽之前一切正常、全部检查通过、所有指标合格；耗尽的那一刻突然失效。若余量的剩余量可以直接读出，那么剩余寿命预测这一整个学科就没有存在的必要——它存在，恰恰说明**"没有故障"与"健康"是两件事，而现行仪表读不出后者**。

工程学内部为此有一场公开的、持续的争论：一派把安全定义为事故的缺席，靠余量、屏障与合规来保证；另一派主张安全是在变动条件下仍然成功的能力，而**日常工作中的变异不是要消灭的威胁，恰恰是成功与失败共同的来源**。本文取前一派为工程这一家的立场，因为它是制度化的一派；后一派在第十二节作为最接近的几种说法之一处理，那是本文必须最小心划界的一位。

2.4 三家并排

	主张看哪一维	单独够用的那一维	时序	自己留下的松动
认知科学	内部模型如何被更新	预测误差	落差→更新→行为	暗室问题：系统必须主动去找它还没有有的落差
物理学	受扰时的响应	响应函数	外场→响应→弛豫	响应在没有外场时就已确定，且等于自发涨落
工程学	余量与所处工况	剩余余量	余量→消耗→失效	余量在被消耗时不产生任何信号

三家各锁一维，互不兼容：误差是一个路径上的量，响应是一个外显的量，余量是一个条件上的量。三家都认为自己那一维单独够用。

而三家的松动指向同一个方向——**都指向"在没有要求的时候发生了什么"**。第一家承认系统必须在落差为零时决定去哪儿，第二家承认响应在无外场时已经定了，第三家承认退化在无故障期不留痕迹。三家都走到了这扇门前，三家都没有推门，因为三家的问题都不需要推它。

三、冲突定位：三家在同一个具体处境里给出互相取消的处置

取一个足够具体的处境：一家运行八年、连续三年零事件、全部审核通过的电力调度中心。上级要问一句：它扛不扛得住一次从未演练过的复合故障？

- **认知这一支**说：看调度员的误差信号。近三年误报率、操作偏差率、复盘中发现的判断失误全部逐年下降且已接近零。**结论：模型已收敛，扛得住。**
- **物理这一支**说：误差为零说不出响应有多大，得量。做一次突发扰动演练，看多久恢复、恢复到什么水平。**结论：不演练就无从判断；但演练本身消耗资源、扰动生产，而且演练过一次之后，那个扰动就不再是从未遇到过的了。**
- **工程这一支**说：别去动它。看余量——备用容量、切换冗余、人员冗余、时间裕度。余量在包线内就是健康的。**结论：余量足，扛得住，且不需要付出任何代价去证明它。**

三条不能同时为真。若认知这一支对，则误差归零之后再投入任何资源都是浪费；若物理这一支对，则不施加扰动就永远说不出结论，而这与工程这一支"不该为了知道好不好而去动它"直接冲突；若工程这一支对，则响应的大小可由静态的余量推出，而这正是物理这一支否认的——**余量说的是能承受多少，响应说的是会动多少，两者在非线性体系里根本不是一个量。**

而三家的松动合起来，指向一件三家都没有说的事。

认知那一支承认：系统必须在落差为零时决定去哪儿。物理那一支承认：响应在无外场时就已经定了。工程那一支承认：退化在无故障期不留痕迹。

三条说的是同一件事的三面：**决定"扛不扛得住"的那件事，发生在没有人要求它扛什么的那段时间里；而三家的全部测量都安排在有要求的时候。**

于是三家的争论可以写成一句话：三家争的是"可被改变的能力"该由哪一维来读。而三家共同假定的是：**这个能力是一样只在被调用时才现身的东西——不调用，它就不在那儿，也就无从谈起。**

四、共同假定，与推翻它的那件材料

4.1 把共同假定说出来

摊开写，是这一条：

> 一样东西的能力，只有在它被要求做事的时候才存在、才可测。不被要求的那段时间，是这个系统的空白期——那里没有可读的东西，因为那里什么也没有被调用。

把它念给三家听，三家都会说"这还用说吗"。认知实验只在任务期记录；工程监测只在运行工况下取数；物理测量只在施加外场时进行。三家的仪器全部是在有要求的时候才打开的。

三家都没有争论这一条，因为**这一条正是"测量"这个词的意思——测量就是一次调用。**

4.2 推翻它的那件材料：来自物理学自己

物理学有一条写进了每一本统计力学教科书的关系式：一个体系对小外场的线性响应，可以用它在没有外场时的自发涨落表达出来。换句话说，响应函数与平衡态涨落的关联函数是同一个量的两种写法。[3]

这条关系式在物理学里回答的是一个很实际的问题：怎么不施加外场就把输运系数算出来？答案是去看平衡态的自发涨落——扩散系数从粒子位移的自关联算，电导率从电流涨落算，粘度从应力涨落算。它是二十世纪统计力学最有用的结果之一，也是分子模拟里最常用的一套方法。

把它剥到只剩形式，去掉物理学的一切名词，只留"什么对什么做了什么，导致什么"，它说的是：

> 那个只在"什么也没要求"时才存在的量，与"被要求时能出多大反应"，是同一个量。

这正是共同假定所排除的那一类事实。共同假定说：不被要求的那段时间是空白期。这条关系式说：不被要求的那段时间不但不是空白期，它恰恰是那个能力被完整决定的时间；有要求时量到的，只是把它读出来而已。

三个部件同时被打掉：

1. 能力不是只在被调用时才存在的。它在没有被调用时就已经确定，调用只是读数。2. 不被要求的那段时间不是空白期。那里正在发生的自发变动，就是那个能力本身。3. 所以三家的仪器全部开在错误的时刻。不是精度不够，是打开的时机不对：它们在能力已经被决定之后才去量它，而在能力正在被决定的时候关着。

4.3 一处必须先说清的限度

必须马上写下这一条，否则后面全部是修辞：上面那条关系式是一条物理定理，它的严格成立需要体系处于平衡附近、且外场足够小使响应是线性的。教室、产线、机组都不满足这些条件。本文不主张这条定理适用于它们。

本文从它拿走的不是定理，是它所揭示的那个区分——"能力何时被决定"与"能力何时被读出"不是同一件事，而全部现行读法都把两者当成同一件事。这个区分的成立不依赖平衡态，也不依赖线性。

而且，这条定理失效的那些情形，恰恰给了本文最要紧的一条限制：在被困住的、缓慢老化的体系里（玻璃、自旋玻璃、粗化中的畴），涨落与响应之间的比例关系会破缺，人们只好引入一个"有效温度"来描述破缺的程度。[4] 也就是说：在一个已经被困住的体系里，自发涨落还在，而它不再预示响应。这条限制在第十六节被写成本文最重的一处让步。顺带说，这一破缺是否普遍成立，物理学内部至今有争论——有实验在软玻璃材料里跨几个频率量级没有找到任何破缺。[5] 本文不站队，只取其中对自己不利的那一支作为约束。

五、静息变异：定义与五条判定条件

把上面那个区分落到可清点的东西上。

> **静息变异**：在没有任务、没有考核、没有误差信号、没有外部故障的时段里，一个系统自发做出的、与它自己当前默认做法不同的那些事。

不是所有的“不一样”都算。要算，五条同时成立，缺一条不算：

1. **发生在无要求时段**。该时段内没有分派的任务、没有正在进行的考核、没有待处理的故障、没有截止期。这一条是全部定义的地基；判断一个时段算不算无要求，看的是**有没有一个外部的人或程序在等这个系统交出什么**。2. **与本系统自己的默认做法不同**。不是与别人不同，也不是与标准不同——是与它自己上一次做同类事情的做法不同。默认做法必须能被指认出来（有成文规程的看规程，无成文的看最近若干次的众数）。3. **不是失误**。失误是想做默认那件事而没做成；静息变异是没想做默认那件事。区分的判据是**事后是否被本人或本系统采纳为一个可重复的选项**——采纳了的是变异，改回去了的是失误。这一条在实操上最难，第十一节给了两条降低误判的办法。4. **无奖惩**。该次变异既不带来记功也不带来记过。带奖惩的变异是被要求的变异，不算——它只是把要求换了个名字。5. **被记下来了**。存在一份能指到具体时段、具体动作的记录，且这份记录不是为了考核而建。没有记录的静息变异在制度上不可与随意区分，因而不计入。

第五条与第四条合在一起是本文最尴尬的一处：**要算数就得有记录，而建立记录这件事本身很容易变成一次要求**。这不是一个可以靠更精细的定义化解的张力，它是本文的反噬所在，第十九节正面处理，本文找不到出口。

六、静息变异率：三个领域量纲不同而比率相同

> **静息变异率 v** = 无要求时段内符合上述五条的动作次数 ÷ 该时段内的动作总数。

三个领域的兑现：

- **认知／人**：无人布置、无评分的那段时间里，做同类事情时采用了与自己默认做法不同的做法的次数占比。数据来源：自习记录、非任务期的操作日志、无任务的模拟器时段。
- **物理／材料**：平衡态下自发涨落的无量纲化幅度（涨落的关联相对于该体系的特征尺度）。这一处是唯一有严格定义的，也是本文这个读数的来路。
- **工程／组织**：无故障期内，非计划的工况偏移与试验性配置变更占该期全部操作的比例。数据来源：变更单、批记录、版本库的提交历史。

四条性质：

其一，无量纲。三个领域可以并排比较。

其二，可事后清点。所需的原始记录在多数机构里已经存在，不必新建采集系统：版本库有提交历史，工厂有变更单，模拟器有时段日志，学校有非上课时段机房与实验室记录。

其三，它把一个印象变成一个数。“这个系统还有没有活气”这句话此前只能靠感觉，现在是一个分数。

其四，也是最容易漏又最要命的一条：它与全部既有的主要指标正交。

举一个既有指标最好看、而这个数最难看的具体形状：一条连续两年零缺陷、变更单数量为零、SOP 三年未修订的产线。按现行的任何一套质量与可靠性指标，它是标杆；按 v 看，它是零。反过来， v 高不等于好——一个变更频繁的系统可能只是失控。 v 不是一个越高越好的分数，它是一个方向读数：它读的是这个系统在没有人推它的时候还动不动。

再举两个更刺眼的：一位连续八年零事件、每次复训一次通过、从不做任何非标准操作的飞行员， v 是零。一个通过率九成五、评教满分、题库十年未增删的课程， v 是零。这两个正是各自行当里被当作模范的那一类。

七、辨别格：现行表现 $\times$ 静息变异

单有一个比率还只是一个名字。把它升成两根轴，四格才分得开那些从外面看一模一样的系统：

静息变异高（无人要求时仍在自发做不同的事）

静息变异低（无人要求时一动不动）

现行表现合格 **甲·有余地** 达标而且还在自己动。唯一能在新要求出现时给出反应的一格。代价是它看起来"不够规范"。
（有要求时全部指标达标）

乙·焊住 全部指标最好看的一格，也是本文要打的那一格。达标、稳定、零变更、零缺陷——而它对任何未曾预演过的要求都没有反应。

现行表现不合格 **丙·乱** 又达不到又乱动。看得出来，不需要任何框架。
丁·停 又达不到又不动。看得出来，不需要任何框架。

乙格是本文要打的那一格，也是四格里唯一需要一个框架才看得见的一格。丙与丁不需要本文：谁都看得出乱和停。甲需要的是别被误判为丙——这一点第十一节的伦理条款会处理。

乙格具备三条签名，缺一条则说明第二根轴选错了：

1. 它在短期指标上表现为改善。缺陷率、事件率、通过率、达标率全部向好，而且是真的向好，不是造假。2. 它的失效不产生任何信号。因为整套记录只在有要求的时候打开，而乙格的全部特征都发生在没有要求的时候。它不是信号微弱，是根本没有一台仪器开着。3. 它自我加固。越正规化越严重：变更管制越严、SOP 越细、合规审核越密， v 关得越死。一家管理松散的工厂反而更可能有非计划的试验批；一家管理精良的工厂连这个口子也补上了。

第三条决定了乙格不是失职的产物，而是**尽职的产物**。这是本文与所有"人不上心""管理不到位"式解释的分界，也是本文最不肯让步的一处。

八、为什么焊住不产生信号：余量掩蔽

乙格的第二条签名值得单独说清，因为它是"为什么大家一直没发现"的全部答案。

工程学早就知道一件事：**余量在被消耗的过程中不产生信号。**一个部件从新到旧，其承载能力一路下降，而只要还在余量之内，全部检测、全部指标、全部运行数据都是正常的。剩余寿命预测这一整个学科的存在，就是因为这件事：**"没有故障"与"还剩多少能力"是两个量，而现行仪表只读得出前一个。**

本文要说的是同一个机制在另一根轴上的复现，而且更彻底：

- **余量掩蔽退化**：余量还在，所以退化不显形——但退化本身是连续的，一旦装上合适的传感器，退化特征是可以被提取的。这是剩余寿命预测能做的事。
- **要求掩蔽能力流失**：只要日常的要求都在系统的能力之内，能力流失就不显形——而且这一次连传感器都装不上，因为要装的那台传感器必须在**"没有要求"**的时候工作，而现行的每一台仪器都是被**要求触发的**。

两者的关系是：余量掩蔽是一个可以靠更好的传感解决的问题；要求掩蔽是一个不能靠更好的传感解决的问题，因为**传感这个动作本身就是一次要求**。要读到那个量，唯一的办法是**事后清点无要求时段里已经发生过的事**，而不是在无要求时段里去测什么。

这一条同时说明了为什么乙格里的系统本身也察觉不到：**它的每一次自我评估都是一次要求。**一个系统只要开始检查自己，那一刻它就已经不在无要求时段里了。乙格的当事人不是不肯看，是每次一看，要看的那个东西就不在了。

九、一份已经做过的观测：自动化座舱里流失的是哪一部分

本文最自然的一条推论是：**用得少的能力会流失**。这条推论太顺了，顺到必须找一份能反驳它的实测。

已经有一份。二〇一四年一项在波音七四七模拟机上做的研究，让十六名航线飞行员在系统变化的自动化水平下飞常规与非常规科目，评分并当场追问他们在想什么。结论出乎意料：**飞行员的仪表扫视与操纵杆量这类"用手做"的技能保持得相当完好，退化的是"用脑做"的那一部分——导航、保持对飞行状态的整体把握、诊断异常。**研究者自己的总结是：对飞行员用手做的那些事可以少担心一点，对他们用脑做的那些事应当多担心一点。[6]

这份结果直接推翻了"用得少就会流失"这条朴素推论：**用手那部分其实也用得少了**（自动驾驶接管了绝大部分航段），却保持得好；用脑那部分退化了。所以决定流失与否的不是使用频次。

决定它的是别的东西，而这份数据恰好给出了那是什么。

手上那部分技能在每一次起飞与落地都被真实地要求一次——那两段是人在飞，而且做得好不好当场就知道。它有要求，也有反馈。

脑上那部分不同。在巡航段，自动化替飞行员做了导航与状态跟踪，**因而在整段航程里没有任何人、任何程序要求飞行员说出"我们现在在哪儿、下一步该怎么办"。**他可以说，也可以不说；说了没有奖励，不说也没有惩罚，而且不说不会产生任何后果——直到有一天自动化退出。

也就是说：手上那部分处在“要求少但有要求”的状态，脑上那部分处在“完全没有要求”的状态。前者的能力被反复读出，所以流失可见并被补训；后者的能力从不被读出，所以流失不可见，也就没有人补。

这一节因此做了三件事。

第一，它给了本文一份现成的、非本文设计的经验支持：能力流失的分界不在使用频次，在那段时间里没有人要求它。

第二，它修正了本文最自然的一条说法。 本文原本会写下：用得越少流失越快。这句话现在写不下去了。改成：在有时段的时段里，使用频次决定流失速度；在完全没有要求的时段里，使用频次这个变量根本不存在，决定流失的是这个系统自己动不动。前一句是既有文献早就说过的，后一句才是本文的。

第三，它给出一条本文自己的检验。 若上述解释成立，那么在同一批飞行员里，那些在巡航段自发做非要求的事的人（心里跟着算位置、自己先给出下一步方案再看自动化怎么做、主动构想异常），其脑上那部分技能的保持应当显著优于那些不做的人；而两组人的手上技能不应有差别。这份对比所需的资料——巡航段的自发行为记录——现有研究没有采集，但采集成本极低：巡航段本来就是空闲的。

十、在十二个行当里兑现

一条判断若只在提出它的那个场景里成立，它就是一个说法。下面十二处不是类比，每一处都能据此预测一件具体的、可以去查的事。

一·有备案考纲的大学课程。 无要求时段＝寒暑假与学期间隙。 v 通常为零，且零有制度理由：假期里改讲义不算工作量，不进任何考核。预测：这类课程在遇到新形态的学生（换生源、换先修课、换工具环境）时的调整周期，与其教师的教学评分不相关，而与其假期讲义修订记录相关。

二·住院医师培训。 无要求时段＝非值班时段。 v 非零但不入账——自发的病例讨论、自己去查文献重看一遍某个决策，每天都在发生，而没有任何系统记录它。预测：一旦要求记录，最先发生的不是记录增加，是这类行为减少。

三·航线飞行。 见上一节。这是十二处里唯一有现成实测的。

四·软件团队。 无要求时段＝无工单期。 v ＝无工单关联的提交占比。这是最便宜的一处检验，数据全在版本库里，一条查询就能算。预测：无工单提交占比与该团队应对一次架构级新要求的时间负相关，且这一关系在控制团队规模、总提交量与技术栈之后仍然成立。

五·制造产线。 无要求时段＝无故障、无订单变更期。 v ＝非计划工艺试验批占比。预测： v 为零的产线在遇到新牌号、新原料批次时的调试时间显著更长，而这一差别在其历史良率上完全看不出来。

六·电网调度。 无要求时段＝无扰动期。 v ＝该期内自发做的非规定推演与试运行占比。预测：黑启动能力与近三年事故率不相关，与无扰动期的自发推演次数相关。

七·消防与救援。 无要求时段=无警期。同上。预测：处置陌生灾情的表现与出警次数不相关，与无警期自发推演的多样性相关。

八·医院应急。 无要求时段=平时。二〇二〇年的一次全球性检验已经跑过一遍，而当时的普遍读数是"平时演练最多的未必最好，平时敢改流程的更好"——这一处的历史数据可查。

九·品种选育与农业。 无要求时段=丰年。 v = 丰年里试种非主栽品种的地块占比。预测：一个地区在遇到新病害时的恢复速度，与其丰年试种比例相关，与其历史平均单产不相关。

十·军事训练。 无要求时段=非任务期。 v = 该期演训中偏离标准战法的次数占比。预测：面对新战法的适应速度与训练时长不相关，与偏离次数相关。

十一·个人。 无要求时段=下班后、假期、无考核期。 v = 该期自发学与自发做的、与本职默认做法不同的事的比例。这一处最私人，也最容易被误用，第十一节的伦理条款首先针对它。

十二·大模型与它的评测。 这一处最尖锐。一个模型在被评测时的表现，与它在无提示、无奖励信号时的行为多样性，是两个量；而当前的训练目标里，**压低后者是显式写进去的**——降低随机性、提高一致性、消除"不必要的"变化，全部是被明确追求的指标。按本文的读法，这套目标在**同时**优化第一根轴与最小化第二根轴，即在结构上把系统推向乙格。预测：同代模型中，在温度固定、无系统提示条件下自发行为多样性更低的那些，在遇到训练分布之外的新任务类型时适应更慢，而这一差别在全部标准评测集上读不出来——因为标准评测集本身就是一次要求。

十二处里，第三、四、六、十这四处的 v 相对更高，而它们有一个共同点：**这些行当的失败会以外人可读的形式一次性出现**（飞机掉下来、服务崩掉、电网黑掉、仗打输），因此它们被迫在无要求时段里做点什么。教育与医疗的失败以缓慢、可归因于个体、且延迟数年的形式出现，**这是它们 v 结构性偏低的根本原因**，也是任何一份"向航空业学习"的建议在这两个行当落不了地的原因。

十一、判据、读法与不得越过的界线

11.1 那一句问话

本文的判据不问人的判断，只问记录：

> **在最近一段没有人布置任务、没有考核、没有故障的时间里，这个系统做过几件与它默认做法不同的事？谁记下来了？**

把它拿到六个场景跑一遍，六个答案互不相同，其中两个是查出来的而不是想出来的：

场景	答案
有备案考纲的大学课程	零。且零有制度理由：假期改讲义不进任何考核
住院医师培训	非零，但无记录、无记名，在制度上不可与随意区分

场景	答案
开源项目	非零且高，无议题关联的提交带作者与时间戳（ 查出来的 ）
航线飞行的巡航段	现有记录里 完全没有这个字段 ——巡航段的自发认知活动从不被采集（ 查出来的 ）
有变更管制的产线	接近零，且趋零是变更管制的显式目标
一个人的业余时间	非零，但记录属于本人，且一旦被机构采集就不再是无要求时段

第三、四两处不能从命题直接推出来——它们是去查了才知道的，而且正是它们让第十节末尾那条"为什么有的行当高有的行当低"有了答案。

11.2 三种分层引用

本文的主张分三层，各自可以单独被引用、单独被推翻：

- **第一层（最强、最易倒）**：可被改变的能力在无要求时段被决定，而全部现行读数都开在有要求的时候。
- **第二层**： v 作为跨领域可比的读数，以及那张两轴四格。
- **第三层（最稳、最便宜）**：一条不依赖前两层的经验主张——**在同一组织内的一组同类团队中，无工单期的自发变更占比与该团队应对一次结构性新要求的用时负相关。** 这一条只需要版本库的一条查询和一次事件复盘，不需要接受本文的任何概念。

第一层被推翻时，后两层仍然站得住。本文明确写下这一点，既是诚实，也是保护。

11.3 伦理界线：四条写死的规定

v 是一个比率，而比率几乎必然会被拿去考核人。四条必须写死，缺一条本文不可被用于任何考核：

1. **v 指的是位置与时段，不是人。** 它读的是"这个岗位、这条产线、这门课有没有一段真正无要求的时间，以及那段时间里发生了什么"，不是"这个人不上进"。一条产线 v 为零，绝大多数情况下说明的是它的变更管制条款，不是操作工的态度。
2. **不得采集私人时间。** 第十节第十一处（个人）**只可由本人自用，不可由任何机构采集、要求申报或纳入评价。** 一旦机构开始采集，那段时间就不再是无要求时段，采到的数据在本文的定义下是无效的——**这不仅是伦理问题，同时是效度问题，两者在这里恰好同向。**
3. **安全关键系统的合理 v 值低，且低是对的。** 麻醉、飞行操作、核电运行、药品放行的规程稳定性本身是安全资产。在这些地方拿本文去催促"多做点不一样的"是危险的误用。这些领域提高 v 的唯一正当途径是在与生产完全隔离的场合（模拟机、试验线、沙箱），而那正是它们已经在做的事。
4. **已经按 v 做出不利安排的机构，必须公开编码规则并提供复核通道。** 什么算一次静息变异、什么算失误、无要求时段怎么划，必须是公开可查的成文规则；被判定的一方必须能提出复核。

另有两条降低第五节第三条（变异与失误的区分）误判的办法，一并写在这里：**其一，只清点事后被采纳为可重复选项的那些——采纳这个动作留痕，且由本系统自己完成。其二，双人独立编码加第三人复核，编码者不得知道被清点对象的绩效表现。**

十二、与最接近的几种说法的分界

本文的判断很容易被读成某个既有概念的改名。下面逐一交代，每一条给出判定形式：若观察到甲，则本文对而该概念错；若观察到乙，则反之。

其一 · 任务期运动变异性预测学习速度。 这是最近的一位，近到必须放在第一条。运动学习研究已经证明：个体在执行任务时表现出的、与任务相关的那部分动作变异越大，其后续学习越快；而且这种变异是被神经系统主动调节的，不是单纯的噪声。它说到哪一步：**变异是学习的原料，且可在任务中测出。** 分离线：它测的是**执行任务期间**的变异，并与学习速度**相关**；本文测的是**完全无要求时段**的变异，并主张它决定的是**对尚未出现的要求的响应**。判定形式：**若在控制任务期变异之后，无要求时段的变异仍能独立预测系统对一类全新要求的响应，则本文对而该说法不足；若无要求时段变异一经控制任务期变异即无任何增量，则本文被它吸收。** 这条检验在运动学习的现有实验平台上可以直接加做——只需在实验前后各留一段“你可以随便动”的无指令时段并录下来。

其二 · 探索与利用的权衡。 它说到哪一步：组织与个体在有限资源下要在探索新可能与利用已知之间分配，过度利用导致能力刚性。分离线：探索是在**有目标的前提下对资源的一次配置选择**——它有预算、有决策者、有机会成本；本文的 v 是**无目标时段的自发率**，不消耗任何被记账的资源，也没有人决定过要做它。判定形式：**若能找到一个探索预算为零、无任何探索性项目立项、而无要求时段自发变异很高的组织（例如一个穷但爱折腾的小团队），且其应对新要求的能力显著高于同规模的高探索预算组织，则本文对而探索—利用的框架在此无解释力。** 这样的组织在开源社群里并不罕见。

其三 · 组织松弛资源。 它说到哪一步：组织保有超出当前最小需求的资源，这些冗余在环境变动时提供缓冲与创新空间。分离线：松弛是一个**存量**（可清点、可削减、以资源为单位）； v 是一个**行为率**（以动作作为单位）。二者可以独立取值。判定形式：**若存在零松弛而 v 高的系统（资源紧到极点却仍在自发做不一样的事），且其应变能力高，则本文对而松弛资源无话可说；若所有高 v 系统一经细查都有隐性松弛，则本文被它吸收。**

其四 · 必要变异度定律。 外圈的一位，也是最容易被误当成本文的一位：控制器的变异度必须不小于被控体系的变异度，否则控不住。分离线：这条定律讲的是控制器**可能状态的数目**（一个容量），处方是增加这个容量；本文讲的是系统在**零扰动时实际实现**的变异率（一个流量）。二者的差别在一台功能极多而从未被改过任何设置的设备上最清楚：**按容量算它变异度极高，按 v 算它是零。** 判定形式：**若这类设备在遇到新工况时的适应与一台功能少而经常被折腾的设备相当，则本文错；若前者显著更差，则本文对而容量读法在此失效。**

其五 · 临界性与最大响应。 物理与神经科学里都有的一支：系统调到临界点附近时，其感受率发散，动态范围最大。分离线是数学形状上的：**临界性预测感受率作为控制参数的函数有一个峰**——离峰越远越差，两侧都差；本文预测的是**单调关系**（响应随静息变异单调上升），没有峰。判定形式：**扫一遍控制参数，看有没有峰。** 有峰则临界性对，单调则本文对。这是本文全部划界里最干净的一条，因为它落在曲线形状上而不落在措辞上。

其六 · 随机共振。 噪声在适当强度下反而提高系统对微弱信号的检测。分离线：随机共振讲的是噪声帮助检测已经存在但低于阈值的信号；本文讲的是在完全没有信号时的自发变异决定日后对高于阈值的要求的响应。判定形式：若本文的关系在明显高于任何检测阈值的要求上同样成立，则本文对而随机共振不足以解释；若关系只在微弱要求上出现，则本文实为随机共振的一个改写。

其七 · 中性演化与中性网络。 这是最危险的一位上游占位者，因为剥到形式层，本文的形状是"无选择压时积累的变异，是后来创新的来源"，而这句话在演化生物学里已经被说了半个世纪。分离线：中性理论讲的是基因型空间的可达性——一个结构性事实，说的是"有路可走"；本文讲的是一个可事后清点的率与响应之间的关系，说的是"实际走了多少步"。判定形式：若一个可达性极高而实际漂变率接近零的群体（种群极小、世代极长）在环境变动时的适应与一个可达性一般而漂变活跃的群体相当，则本文错；若前者显著更差，则本文对而可达性读法不足。

其八 · 简并性。 不同的结构实现同一功能，被认为是稳健性与可演化性的共同来源。分离线：简并讲的是结构层面的多重实现（一个静态属性）；本文讲的是同一个系统在无要求时实际做出不同做法的率（一个动态读数）。判定形式：若一个简并度很高而 v 为零的系统（多套备用方案齐全、而从来没有人在无故障期动过其中任何一套）在遇到新要求时反应良好，则本文错。备用发电机长年不启动的机房是这一格的现成样本。

其九 · 剩余寿命预测。 工程内部最接近的一位，第八节已交手。分离线：它通过退化特征推剩余寿命，前提是退化在传感器上留下痕迹；本文说的这一类能力流失在传感器上不留痕迹，因为要装的那台传感器必须在无要求时工作，而传感这个动作本身就是一次要求。判定形式：若在一组同型设备上，全部退化特征读数一致、而无故障期自发试验批占比不同的那些设备，其应对新工况的表现有系统性差异，则本文对而退化特征不足。

其十 · 把安全定义为在变动条件下仍然成功的能力那一支。 这是本文必须最小心划界的一位，因为它与本文方向一致：它主张日常工作中的变异不是要消灭的威胁，而是成功与失败共同的来源。分离线在时机：它说的变异是做事时的调整——为了在资源不足、信息不全的条件下把活干完而做的近似处置，那是有要求时的变异；本文说的是无要求时的变异。两者可以独立取值。判定形式：若一个日常调整极其丰富（工作如实际做的远离工作如设想的）、而无要求时段完全静止的组织，在遇到一类全新要求时反应良好，则本文错而该支对；若这类组织恰恰在全新要求面前僵住，则本文对——它平时的丰富调整全部是围绕已知要求的。这个判定在现场可做，而且两支都会认为它值得做。

最近的一位是第一条。 它仍不能吸收本文，理由只有一条：它把"变异要在任务中测"当作给定，因此看不见无任务时段的变异；一旦承认无任务时段的变异是一个独立变量，它整套实验设计的时间边界就要重划。

上述十种说法的共同盲区

逐个问一句"它把什么当作给定，因此看不见什么"：

- 任务期变异说把"变异要在任务中测"当作给定，看不见无任务时段的变异；

- 探索一利用把"探索要花预算"当作给定，看不见零成本、无人决策的自发变异；
- 必要变异度把"变异度是可能状态数"当作给定，看不见实现率；
- 剩余寿命预测把"退化留下可测痕迹"当作给定，看不见不留痕迹的能力流失；
- 中性理论把"变异发生在世代之间"当作给定，看不见世代之内；
- 简并性把"多重实现是结构属性"当作给定，看不见从不被使用的多重实现等于没有。

交集是同一块地：一样东西的能力，只有在它被调用的时候才存在。

所有人都站在这块地上，包括互相批判的各派——因为**一切测量都是一次调用**。这不是某一派的疏忽，是整个经验研究的立足处；也正因为共同，没有人能回头看它。本文能看见它，唯一的原因是物理学在自己的技术需要下（不施加外场而算出输运系数）早就写下了一条与它相反的关系式，而那条关系式在物理学内部回答的是另一个问题。

十三、与同一栏目里几篇的分界

本栏此前已有几篇触到相邻的轴，其中三篇处理的是同一个大问题域，必须指名。

其一 · 《记为负项的那一类改变》。 那一篇讲的是：产物改判据这一类事件被三家的账本记为负项（不公平、污染、漂移），因而发起处不换手。分离线：那一篇的对象是一个**动作**（有人做了一件事，而这件事被扣分）；本文的对象是一个**时段里的率**（没有人做任何被要求的事，而那段时间里发生的一切不被记录）。两者可以独立为零。判定形式：一个判据修改通道完全畅通、且实际被频繁使用（那一篇的读数很高）、而无要求时段完全静止（ $v=0$ ）的机构，若其应对全新要求仍然迟钝，则本文对而那一篇为空；反之，一个无要求时段极其活跃、而任何自发改动一律不被允许进入正式判据的机构，那一篇对而本文为空。两篇叠加时是相乘不是相加：没有静息变异，就没有东西可以拿去改判据；有东西而通道被堵，改动进不了下一轮。

其二 · 《每一轮都从新鲜表面开始》与《先动的总是同一边》。 那两篇的观测单位都是一门课程内部连续轮次之间（上一轮留下的东西被不被接上、验收节拍对准了哪一条驱动）；本文的观测单位是**轮次与轮次之间那段没有轮次的时间**。三者数学上独立，可以在同一批对象上同时测量。判定形式：若一门课的结转与节拍两项读数都很好、而假期与课间的 v 为零，其面对新形态学生的调整仍然缓慢，则本文对而那两篇为空。

其三 · 《换一种说法，它就不见了》。 那一篇讲的是同一份知识在不同使用条件之间的解绑代价。分离线：那一篇问的是一样已经学会的东西换个场合还在不在；本文问的是在没有任何场合要求它的时候，这个系统还产不产生新东西。两者方向相反：那一篇测的是保持，本文测的是产生。判定形式：**解绑代价很低**（换场合仍然会）而 v 为零的学习者，本文预测他对一类前所未有的问题仍然没有办法——这一预测那一篇不作，也无法作。

其四 · 《一条路没有交接点》（同日同栏，与本篇共用物理学与工程学两家）。 那一篇问的是一条路上由谁来结算——判据由谁给出、判定由谁执行；它的答案是交接点（位置更换、执行方更换、判据先于成

果存在），读数是独立结算比。分离线有两层。**其一在问法**：那一篇问“怎么走、在哪一步能插手”，本篇问“还能不能被改变、凭什么读得出来”；一个要的是次序上的一个位置，一个要的是一段时期的一个率。**其二在两家共用学科的取用点上，两篇没有一处重合**：那一篇的工程学取写进标准的三项独立性，本篇取余量与降额；那一篇的物理学取盲分析的**时点**（判据必须先于读数被锁定），本篇取涨落与响应的**关系式**（响应在没有外场时就已确定）。同一门学科被两篇拿走的是两件不同的东西。

判定形式：一门课的判据由完全独立的第三方给出并执行（那一篇的读数很高），而它的假期与课间静息变异为零，若其面对一类前所未有的要求仍然拿不出反应，则本篇对而那一篇为空；反之，一个假期里不断自发折腾、而全部达标判定仍由授课方自己按自己的判据作出的教研组，那一篇对而本篇为空。**两个读数在同一批对象上可同时测量，且互不蕴含**——独立结算解决的是“这一次达标算不算数”，静息变异解决的是“下一次没见过的东西来了拿不拿得出反应”。一个管已经发生的判定不可信，一个管尚未发生的要求接不接得住。

诚实记下一处继承：本文第八节“余量掩蔽”与《记为负项》第五节“失效不产生信号”是同一个机制在两根不同轴上的复现，那一笔记在那一篇账上。本文的增量在于指出这一次连传感器都装不上，而不只是没人去读。

十四、三家为什么各自到不了这一条

三家停在原地，不是因为疏忽。是因为各自的核心问题恰好使这一步成为不必要的。

认知科学到不了，因为它的问题是“行为与神经活动如何由一条原则统一地解释”。这个问题要求有一个可在每一时刻定义的驱动量，而落差正是这样一个量。**无要求时段里没有落差可定义，因此那段时间在它的框架里不是数据，是基线。**它不是不肯看基线，是它的整套形式化把基线定义成了“要被减掉的那个东西”。

物理学到不了，因为它的问题是“如何从可测量算出可预测”。它写下了那条关系式，而且用得很熟——**但它用的方向是从涨落算响应，为的是省掉施加外场的麻烦。**它从没有反过来问一句：如果一个体系的涨落被人为压制了，它的响应会怎样。因为在物理学里没有人会去压制涨落——涨落是热运动给的，压制它意味着降温，而降温这件事的后果早已被完整地研究过，叫做冻结。**物理学里“压低涨落”是一个已知会导致冻结的操作；只有在人造系统里，才会有人一边压低涨落一边指望它照常响应。**

工程学到不了，因为它的问题是“如何保证它不出事”。这个问题的形状要求把一切非计划的变动识别出来并消灭，而静息变异按定义就是非计划变动。**它对这类事件的全部知识都是为了防止它发生**——变更管制、配置管理、偏差处理，整套体系精确地指向同一个方向。它是三家里对这个量了解最深的，也正因为了解最深，它必须把它记在负栏。

三条合起来说明一件事：**这一条不是三家的综合，是三家各自的问题结构所排除的那一格。**

十五、一处强制的不一致

上一节说三家各自到不了。这一节要说的更重一点：三家里有一家，在不承认本文这个量的情况下，它自己内部就不自治。

认知科学那一支同时持有两条，两条都是它的承重结构。

甲条：一切认知改变都由预测误差（或其变分上界）驱动。这是这一支的全部力量所在——知觉、行动、注意、学习被统一在一条原则下，正因为只有一条驱动。

乙条：为了不走进暗室，策略的价值必须包含认识论价值，即系统会主动去寻求能带来信息增益的状态。这一条不是外加的补丁，这一支明确指出它从形式里自然长出来，且只在对时间上延展的序列做推断时才出现。

两条在通常情形下相安无事。它们在一种情形下正面相撞：**当系统当下的落差已经为零时。**

按甲条，落差为零则无驱动，系统不应有任何改变的倾向。按乙条，系统此时仍应去寻求信息增益——也就是去**主动制造**它现在没有的落差。

这一支的化解是"长远看，寻求信息会降低未来的落差"。这个化解在形式上有效，代价是它要求系统能对**尚未遇到的要求估值**——即对"我还不知道我不知道什么"这件事给出一个数。而这个数**只能在落差为零的时段被决定，也只能在那里被观测。**

于是这一支面临一个二选一：

- **要么承认存在一个在无要求时段被决定、且不由当下落差定义的量。**一旦承认，甲条的"落差是唯一驱动"就不能再以现在的形式成立——因为有一个量既不是落差，也不由落差决定，而它决定系统在落差为零时做什么。
- **要么放弃乙条，**接受系统确实会走进暗室。而这一支自己已经明确拒绝过这条路。

这不是本文替它选，是本文指出它必须选，而它现在没有选。它现在的处置是把"寻求信息"这件事写成一个也叫"最小化"的东西（最小化期望的落差），从而在形式上保住甲条。这个处置在数学上无懈可击，代价是**"落差"这个词在甲条与乙条里指的已经不是同一样东西了：**甲条里它是可在当下观测的量，乙条里它是一个对未曾发生之事的估计。同一个词承担两个不能同时观测的角色，这正是本文的量被藏起来的地方。

本文已给自己留了退路：这一处可以被"咬牙接受落差包含反事实落差"化解——那样它不是不一致，只是让"落差"变成一个不可由当下观测定义的量。这条退路被写进第十八节的证伪条件第九条。

同一个动作可以对工程学再走一遍（这一次力度弱一些，因此不作为承重）：工程学同时持有"变更是风险的主要来源，应当最小化"与"僵化的系统在环境变动时最先失效"。它现在的处置同样是层级隔离——把前者归"运行"，把后者归"战略"，由不同的部门负责，因而永远不必在同一张表上同时优化。

十六、三家反过来给本文加的三个约束

碰上有用的对手，代价是它们也会削掉你原本想说的话。三处，逐条写清削掉了什么。

其一 · 物理学削掉了本文的价值排序，而且削得最狠。 本文原本会写下一句更强的话：**静息变异率越高越好，提高它就能恢复可改变性。** 物理学把这句话取消了。那条把涨落与响应连起来的关系式，**在被困住的、缓慢老化的体系里破缺**：涨落还在，而它不再预示响应；人们只好引入一个“有效温度”来描述破缺的程度。[4] 落到本文的对象上，这意味着：**一个已经被困住的组织里，静息变异与应变能力解耦——那里的自发变动仍在发生，却不再通向任何改变。** 此时提高 v 只是加噪。

于是本文的主张从“应当提高 v ”收窄为： **v 是一个诊断读数，不是一个操作杆。先判断这个系统在不在解耦区；不在， v 读得出能力；在， v 什么也读不出，此时要做的是别的事。** 判断在不在解耦区的现成办法本文给不出，只能给一个粗筛：**看这个系统最近一次真正改变是什么时候，以及那次改变是不是由内部发起的。** 这一处动的是本文的价值排序，不是补一个条件。

其二 · 工程学削掉了本文的适用范围。 规程的稳定性本身是安全资产，变更管制不是官僚主义，是几十年事故换来的。一个 v 很高的核电运行班组在法律上、伦理上都是不可接受的。于是本文的适用范围收窄为“**需要在未来适应未知要求的系统**”，并且必须补一条实操约束：**这类系统提高 v 的唯一正当途径是设立与生产完全隔离的场合——模拟机、试验线、沙箱、影子系统——而这恰恰是可靠性工程已经在做的事。** 本文因此不是在反对变更管制，是在指出变更管制必须配一个隔离的、无要求的、有记录的场合，否则它把两样东西一起消灭了。

其三 · 认知科学削掉了本文读数的粗糙。 运动学习的实验已经证明：变异必须分成**与任务相关**和**与任务无关**两部分，只有前者预示学习，后者是纯粹的噪声。这直接推翻了本文最省事的算法——把无要求时段里所有的“不一样”都数进去。 **v 必须分层：只计入那些落在与本系统职能相关的维度上的变异。** 一个团队在无工单期把办公室重新装修了一遍，这不进 v 。这一处是真削弱：它让 v 的清点成本上升了一个量级，也让“数据已经在版本库里”这条便宜话打了折——版本库里的提交需要按相关性再分一次类。

十七、不适用的边界

四处明写：

其一 · 没有可指认默认做法的系统读不出 v 。 若一个系统从来没有稳定的做法，“与默认不同”就无从定义。初创团队、临时编组、正在剧变中的组织，本文的读数在这些地方无效。

其二 · 没有真正无要求时段的系统读不出 v 。 高强度、无间断、始终有截止期的岗位（急诊、战时、连续生产的关键工位）在结构上不存在无要求时段。**这不是它们的缺陷，是它们的性质；**对它们，本文能说的只有一句：这类岗位的可改变性必须由岗位之外的安排提供，不可能由岗位自己长出来。

其三 · 安全关键系统的合理 v 值低，且低是对的。 见第十一节第三条与第十六节第二条。

其四·高 v 不等于好。 高 v 可能读出的是失控、无标准、或纯粹的浪费。 v 是一个方向读数，不是一个质量分数；单独看它会得出荒谬的结论。要判断一个系统的处境，必须同时看第七节那两根轴。

十八、证伪条件

先写出最强的竞争解释，那句最朴素、最难反驳的话：“**其实不就是练得少了吗？不就是投入不足吗？**”

若真是如此，那么决定应变能力的应当是总投入时间与总练习量，而不是无要求时段的自发率；把投入加上，能力就该回来。第九节那份座舱实测已经给了这条解释一次打击（手上那部分同样用得少却保持得好），但那只是一个案例，不足以排除它。第二强的竞争解释是“**其实不就是探索不足吗**”，第十二节第二条给了它一条分离线。

核心证伪条件写成这个形状：

> **控制总投入时间、总任务量与探索预算之后，若无要求时段的自发变异率 v 与该系统对一类全新要求的响应速度之间不存在剂量—反应关系，则本文的第一层主张为伪；届时应归因于总投入而非无要求时段的变异。**

样本纪律写死：**同一机构、同一工种的 $N \geq 8$ 个班组或团队**，在人员构成、任务量、技术条件上同质，自变量（无要求时段的记录）有真实变异。**不得拿两个国家或两个行业作对比充数**——那种比较两边处处不同，所谓“控制掉某变量”在方法上站不住，只能作启发式说明。

另有八条，其中第三条最便宜，所需数据在多数机构里已经存在：

1. **顺序检验**： v 的下降应当**早于**应变能力的下降出现，且早若干个周期。若两者同时出现或 v 滞后，本文作为一个早期读数的全部价值即告消失。
2. **分层检验**：与职能相关的那部分变异应当预测响应，与职能无关的那部分不应当。若两者预测力相当，则第十六节第三条的分层是多余的， v 的定义须重写。
3. **最便宜的一条（第三层主张）**：在同一组织内的一组同类团队中，若无工单期的自发提交占比与该团队应对一次结构性新要求的用时无关联，则第三层主张为伪。所需资料：版本库的一条查询 + 一次事件复盘记录。
4. **座舱那一处的直接检验**：在巡航段自发做非要求认知活动的飞行员，其“用脑”技能的保持应显著优于不做的人，而两组的“用手”技能不应有差别。若两组在两类技能上都无差别，第九节的解释为伪。
5. **要求掩蔽与余量掩蔽的分离**：在一组同型设备上，若全部退化特征读数一致而无故障期自发试验批占比不同的那些设备，其应对新工况的表现无系统性差异，则第八节的区分不成立。
6. **无奖惩条件的必要性**：若带奖惩的“自由探索时间”与无奖惩的自发变异在预测响应上没有差别，则第五节第四条判定条件多余，本文的反噬预言同时失效——那将是一个好消息。
7. **形状检验**：响应与 v 的关系若呈现单峰（有一个最优值，两侧都差），则本文被临界性一支吸收；若单调，则本文对。
8. **解耦区的存在**：若找不到任何一个 v 高而响应低的真实系统，则第十六节第一条的让步是多余的——那反而使本文更强，但也说明本文对物理学那条限制的搬用是错的。
9. **强制不一致的退路**：若认知科学那一支公开承认“落差”包含不可由当下观测定义的反事实落差，则第十五节所指的不一致不成立——那时它只是变弱，不是自相矛盾，本文对它的那一处主张作废。

若本文被证伪，我们会学到什么：若 v 与应变能力确实无关，那么“能力在无人要求时被决定”这个说法就是错的，全部注意力应当回到投入与配置这一层——而这将是一个正面的结论，因为它会把大量资源从“给系统留白”这类结构改造上收回来，投向更直接的地方。

十九、反噬预言

本文一旦被制度采用，最省事的实现方式是：每周划出两小时“自由探索时间”，要求各团队在这段时间里做与日常不同的事，并提交记录。

那一刻，本文的定义前提被销毁。被划出来的两小时不是无要求时段——它有起止、有要求、有交付物、有考核。在这两小时里测到的一切，按本文自己的定义，一条也不算数。而机构会得到一份很好看的报表，上面写着 v 从零升到了百分之四十。

比这更省事的一种是：把日常工作中本来就存在的偏差重新分类，全部记成“自发变异”。这一步连两小时都不必划，改一改字段名即可。

唯一能想到的减缓办法是：只做匿名的、无奖惩的、事后的清点，且清点结果不进入任何个人或团队的档案，只作为机构层面的一个诊断数。但这个办法自己也有出口——一个不进任何档案的数没有人会去认真采集；而一旦有人认真采集，被采集者就会知道有人在看，那一刻它就又成了一次要求。

我看不到出口。这不是一个待办事项，是本文交出去的一个洞。它的结构是：任何使 v 可核查的装置，都构成一次要求；而不可核查的 v 不能进入任何决策。本文能给的底线只有一条：验收必须看无要求时段里已经发生过的事的事后记录，不看是否设立了“自由探索”制度。有制度而事后清点为零，与没有制度是同一格。

二十、自反：本文自己在哪一格

本文提出的判据，对本文自己适用吗？适用，而且结果不好看。

本文是在有要求的条件下写出来的：有一条明确的指令，有一个栏目，有一个交稿的时刻，有三门指定的学科。按第五节的五条判定条件，本文写作期间的 v 是零——它的每一个动作都在响应一个要求。

后果具体有三条：

第一，本文属于第七节那张表里的乙格，而不是甲格。它的现行表现（论证、划界、兑现处数）可以做得很好，而这一切都是被要求的产物。本文因此只能诊断，不能示范：一篇在要求下写成的东西，说不出无要求时会写出什么。

第二，本文最可能的命运正是它自己的反噬。一个提出“要留出无要求时段”的文本，最容易被做成的东西就是一项制度——每周两小时，提交记录。本文预言了这件事，而预言它并不能阻止它，因为阻止它需要的正是本文没有的那个东西：一个不被要求的场合来消化它。

第三，本文对它自己的这个判断不能靠讨论来修正。讨论会产生引用，引用不会改动本文任何一行。本文能给自己安排的唯一一次真正的变动，是下一节那条写死日期的赌注——它把一个可能的负项固定在一个不可撤销的时间点上。

最后一处更难看的：本文这套读数一旦被采用，**第一批被读出 $v=0$ 的对象就是学术论文与它的生产制度**——选题响应基金指南，写作响应栏目要求，发表响应审稿意见，每一步都有一个人或一个程序在等着收东西。本文没有理由认为自己例外。

二十一、结论，与一条写死日期的赌注

一个系统还能不能被改变，读不出来的原因不是仪器不够精，是**全部仪器都装在有要求的时候**。而决定这件事的那个量，只在没有人要求它的时候存在。

三门学科给了三条互不相容的读法——看误差、看响应、看余量——而三条共同假定：能力只在被调用时才存在。物理学自己的一条关系式说，那个只在不被调用时存在的量，与被调用时能拿出的反应，是同一个量。

把这个量命名为静息变异，把它的比率记作 v ，就得到一张两轴四格。**要打的是"表现合格而 v 为零"那一格**：它在全部现行指标上最好看，它的失效不产生任何信号，而且它越正规化越严重。它不是失职的产物，是尽职的产物。

而这条判断的根，比学习更深一层：**可靠性要求把变异消灭，可改变性要求把变异留住**。二者要求相反方向的处置，而且用的是同一笔预算——同一套变更管制、同一套考核、同一段时间。一个系统被同时要求做到两件事，而它一次只能做到其中一件。这不是一个可以靠更好的设计消解的张力，只能靠**在空间上把两者分开**：给可靠性一套受管的场合，给可改变性一套隔离的、无要求的、事后有记录的场合，并承认后者的产出**在当期一定看起来像浪费**。

赌注，写死日期：

＞ **到二〇三二年十二月三十一日为止，至少有三家高可靠性行业的机构（航空、电力、医疗、制造任一）在其能力维持标准中，把"无任务时段的自发活动"设为一项独立于培训学时与事件率的指标，并公开其年度读数。**

写死什么不算命中：

1. 仅设立"创新时间""自由探索时间"一类有起止、有交付、有考核的制度安排； 2. 仅在质量报告或社会责任报告中自述"鼓励员工创新"而无可核对的事后清点； 3. 仅统计培训学时、演练次数或提案数量； 4. 与其他改革打包实施，无法分离； 5. 只公布制度存在与否，不公布事后清点的实际读数； 6. 设立了场合而年度实际清点为零——**有制度而零发生，与没有制度是同一格**。

最后交出两个本文解释不了的洞。

其一：第五节第五条要求静息变异必须有记录才计入，而第十节第二处的住院医师表明，大量真实的静息变异恰恰发生在无人记录的地方；本文的读数在那里读出零，而那个零是错的。本文没有办法修正它——任何补救措施都是去建立记录，而建立记录就是建立要求。

其二：本文说清楚了在什么时候测，没有说清楚为什么有的系统在无人要求时仍然在动，而有的不动。第十六节第一条的"解耦区"给了一个物理学上的形状，但那个形状怎么在一个组织里被识别出来，本文给不出办法，只给了一句粗筛。这是本文最大的空缺，也是它眼下最不该被拿去做决定的原因。

注释

[1] 暗室问题是对这一支被提出得最多的一条反驳：若最小化预测误差是唯一驱动，最优策略应是躲进一间毫无刺激的暗室。见 Friston, Thornton & Clark (2012) 与 Sun & Firestone (2020)。

[2] 见 Seth 等 (2020) 对上一条的答复：期望自由能可分解为工具性项与认识论项，后者从形式中自然出现而非外加，且只在时间上延展的序列作推断时才出现。

[3] 涨落与响应之间这条关系的经典表述见 Kubo (1966)；用平衡涨落的时间关联算输运系数的一族公式通称 Green-Kubo 关系。

[4] 关于该关系在玻璃态与老化体系中的破缺及"有效温度"，综述见 Crisanti & Ritort (2003)。

[5] 反面证据见 Jabbari-Farouji 等 (2008)：在硬球胶体玻璃与 Laponite 凝胶中，跨若干频率量级未观察到破缺。两方的争论至今未决；本文取对自己不利的一支作约束。

[6] Casner, Geven, Recker & Schooler (2014)。十六名航线飞行员在波音七四七模拟机上飞常规与非常规科目，系统变化自动化水平。结论是仪表扫视与操纵控制这类技能保持良好，而与手动飞行相关的认知技能的保持取决于飞行员是否仍在主动监督自动化。研究者的公开表述是：对飞行员"用手做"的事可以少担心一点，对他们"用脑做"的事应当多担心一点。

参考文献

1. Ashby, W. R. (1956). *An Introduction to Cybernetics*. London: Chapman & Hall.
2. Casner, S. M., Geven, R. W., Recker, M. P., & Schooler, J. W. (2014). The retention of manual flying skills in the automated cockpit. *Human Factors*, 56(8), 1506–1516. <https://doi.org/10.1177/0018720814535628>
3. Crisanti, A., & Ritort, F. (2003). Violation of the fluctuation–dissipation theorem in glassy systems: basic notions and the numerical evidence. *Journal of Physics A: Mathematical and General*, 36(21), R181–R290. <https://doi.org/10.1088/0305-4470/36/21/201>
4. Cyert, R. M., & March, J. G. (1963). *A Behavioral Theory of the Firm*. Englewood Cliffs, NJ: Prentice-Hall.
5. Edelman, G. M., & Gally, J. A. (2001). Degeneracy and complexity in biological systems. *PNAS*, 98(24), 13763–13768.
6. Friston,

K., Thornton, C., & Clark, A. (2012). Free-energy minimization and the dark-room problem. *Frontiers in Psychology*, 3, 130. 7. Hollnagel, E. (2014). *Safety-I and Safety-II: The Past and Future of Safety Management*. Farnham: Ashgate. 8. Jabbari-Farouji, S., Mizuno, D., Atakhorrani, M., MacKintosh, F. C., Schmidt, C. F., Eiser, E., Wegdam, G. H., & Bonn, D. (2008). Effective temperatures from the fluctuation–dissipation measurements in soft glassy materials. *Europhysics Letters*, 84(2), 20006. <https://doi.org/10.1209/0295-5075/84/20006> 9. Kimura, M. (1983). *The Neutral Theory of Molecular Evolution*. Cambridge: Cambridge University Press. 10. Kubo, R. (1966). The fluctuation-dissipation theorem. *Reports on Progress in Physics*, 29(1), 255–284. <https://doi.org/10.1088/0034-4885/29/1/306> 11. March, J. G. (1991). Exploration and exploitation in organizational learning. *Organization Science*, 2(1), 71–87. 12. Seth, A. K., Millidge, B., Buckley, C. L., & Tschantz, A. (2020). Curious inferences: Reply to Sun and Firestone on the dark room problem. *Trends in Cognitive Sciences*, 24(9), 681–683. 13. Sun, Z., & Firestone, C. (2020). The dark room problem. *Trends in Cognitive Sciences*, 24(5), 346–348. <https://doi.org/10.1016/j.tics.2020.02.006> 14. Wagner, A. (2011). *The Origins of Evolutionary Innovations*. Oxford: Oxford University Press. 15. Wu, H. G., Miyamoto, Y. R., Gonzalez Castro, L. N., Ölveczky, B. P., & Smith, M. A. (2014). Temporal structure of motor variability is dynamically regulated and predicts motor learning ability. *Nature Neuroscience*, 17(2), 312–321. <https://doi.org/10.1038/nn.3616>

本文由三个分属不同学科、且互相冲突的理论体系碰撞而成。全部来源均为站外公开文献，可自行核对。 作者：王德生 + Claude · SDE Universes · 学科通融