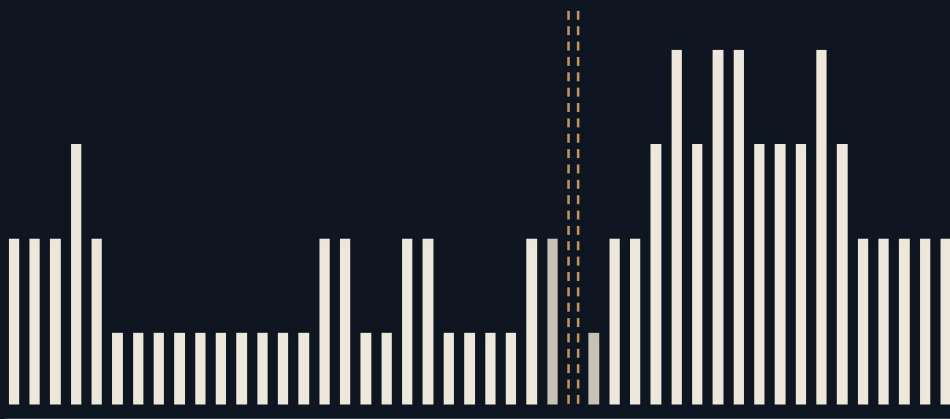

睡觉的智慧

一夜怎样被记下来，
以及记下来的时候丢了什么

这一格里本来有东西



每格三十秒

王德生 著

德麦国际出版社

睡觉的智慧

一夜怎样被记下来，以及记下来的时候丢了什么

王德生 著

德麦国际出版社

出版页信息

书名 睡觉的智慧——一夜怎样被记下来，以及记下来的时候丢了什么

作者 王德生

出版 德麦国际出版社（新加坡）

专著编号 德麦国际出版社专著第 68 号

ISBN 979-8-90690-001-2

版次 第一版

字数 约二十八万五千汉字

开本 16 开 170 × 240 毫米

页数 300 页

定价 US\$25.00

本书全部经验材料来自公开数据集，全部分析脚本与预注册文件随书附录列出，任何人可独立复现。

版权所有。未经书面许可，不得以任何方式复制、转载或用于商业用途。

作者介绍

王德生

德麦国际（新加坡）创办人。计算数学博士，毕业于中国科学院，并在该院完成博士后研究。研究方向为沃罗诺伊图、德劳内三角剖分、质心沃罗诺伊剖分与有限元方法——用一句外行听得懂的话说，就是研究“怎样把一块连续的东西切成格子，以及这样切会付出什么代价”。曾先后任教于英国斯旺西大学与南洋理工大学。二〇〇五年起定居新加坡。

近年在德麦国际出版社出版多部专著，题域横跨教育、健康、组织与思想史；本书是其中处理睡眠记录结构的一部。

关于这本书与作者背景的关系，值得多说两句，因为它不是巧合。

剖分理论里有一条几乎是常识的东西：**网格不是中性的**。同一个方程，网格一变，解就变；格子的形状决定了误差落在哪里、有多大；而误差不会自己举手，得靠残差估计把它算出来，算出来才知道该在哪里加密。整套有限元的纪律，建立在“算完之后必须能估出自己错了多少”这一件事上。

一个在这条纪律里待过二十年的人，第一次看见睡眠判读记录——一夜被切成三十秒一格、每格一个字母——会立刻注意到两件在本行内不可想象的事：**其一，这是固定步长的均匀网格，而且是在不知道解的形状之前就定死的，定死了五十多年；其二，这份记录没有残差，九百六十个字母里没有任何一个位置告诉你“这一格判得勉强”。**

本书十四章，可以看作把这两个反应认真做完的结果。作者不是睡眠研究者，书里没有一条新的生理发现，也没有指出任何一次判读违反了规则。它做的是另一件事：**把一门学科的记录制度，当作一个网格来审。**

作者对本书的三处欠账写在结语与后记里，其中最实的一处是：全书没有看过一位失眠患者的记录。

前言

这本书是从一次数数开始的。

我不是睡眠研究者。我做了大半辈子的另一件事：把一块连续的区域切成格子。计算数学里那一支叫剖分——沃罗诺伊图、德劳内三角化、质心沃罗诺伊剖分、有限元网格。一个人若在这上面花过二十年，他对格子的第一反应不是“用它去算”，而是“它把什么切断了”。剖分从来不是中性的：网格一变，同一个方程的解就变；格子的形状决定了误差落在哪里，而误差不会自己举手。

二〇二六年夏天，我因为另一件事去翻公开的睡眠数据库，翻到一份判读记录——一整夜的脑电，被切成三十秒一格，每一格标一个字母：清醒、第一期、第二期、深睡、快速眼动。一夜八小时，九百六十格，九百六十个字母。

一个做剖分的人看到这个，会立刻想到两件事。**第一，这是一维的均匀网格，而且是固定步长的。**有限元里没有人敢在不知道解的形状之前就把网格定死；这里定死了，五十多年。**第二，也是要紧的那件——这份记录没有残差。**数值计算的全部纪律建立在一件事上：算完之后要能算出误差有多大，误差大就加密网格。而这九百六十个字母里，没有一个位置告诉你“这一格判得勉强”。

我盯着那一串字母看了很久，然后问了一个外行才会问的问题：

如果这一夜里发生了一件只持续十五秒的事，它在这串字母里是哪一个？

答案是：不是哪一个。判读规则规定，每一格按占优的那一期来标。一件持续十五秒的事，占不满半格，于是它被并进相邻的某一期里，在这九百六十个字母里不留一行。

这本来可以只是一句“分辨率不够”的抱怨。让我停下来的是接着发现的第二件事：**一个足以让人的洞察率翻三倍的状态，恰好就只需要十五秒。**二〇二一年发表在《科学进展》上的一项实验，让受试者闭眼休息二十分钟，在第一期停留至少十五秒的人，发现隐藏规则的比例是百分之八十三，始终清醒的人是百分之三十；而睡得更深，这个效应就消失。

于是问题变了形状。不是“我们测得不够细”，是：**这件事已经被测到了，而记录格式使它不可能被记下来；并且记录不会因此报错。**那一夜的记录仍然是完整的，九百六十格一格不缺，任何人拿去分析都不会发现少了什么。

这本书的十四章，是同一个形状的十四次发作。

它们分别落在不同的位置。有的落在三十秒那一格里（第一到第四章）；有的落在一个被公布出来的数字上——三个要求压成一个分数、开与封折成一个净效果、两端相减只公布差、一整段下降只留两个端点（第五到第八章）；有的落在一个人被归进哪一类这件事上（第九到第十一章）；有的落在账被记在谁名下这件事上（第十二到第十四章）。

十四章合起来赌的是一句话。它写在导论的开头，也写在这里：

睡眠的缺口不是漏记，是强制作答的代价；它在记录最齐整的时候最大，因为扔掉的那一份不留痕迹。

这句话反直觉的地方在于它与常识的方向相反。常识说：我们对睡眠了解得不够，是因为测得不够——传感器不够准，样本不够大，时间不够长。本书说：**在很多要紧的位置上，问题恰恰是记录太齐整了。**每一格都必须填，于是没有一格是空的；而没有一格是空的，也就没有一格能显示缺了什么。

我要在这里先说清楚，本书**不做什么**。

本书不主张睡眠研究做错了。十四章里没有一章指出任何一次判读违反了规则，也没有一章找到任何一处伪造或疏漏。相反，本书全部的力气都花在一个更麻烦的地方：**每一次损失都发生在规则被完整遵守的时候。**第三章说得最直白——一次险胜和一次压倒性的胜利，在纸面上完全一样，而两份记录都没有出错，都不需要更正，都可以拿去发表。

本书也不主张自己发现了新的生理现象。第二章在打磨的时候把这一条明写了进去：本章没有发现任何新现象。第六章、第八章、第十四章各自都用一整节说明自己的数据支持不了自己最想说的话。**本书十四章里，有九章报告了自己至少一条预注册预测的失败，其中三章触发了自己写死的撤**

销规则，把一条已经写好的主张整个撤回。这不是谦虚，是本书唯一可能有的力气来源：一本讲“失效不产生信号”的书，如果自己的失效也不产生信号，它就什么都没设。

关于工序，有一件事要交代，因为它决定了这本书为什么长成这个样子。

十四章每一章都从三门互不相干的学问出发。不是为了显得广博——三门学问的作用是互相拆台。它们对同一个问题给出三个不能同时成立的答案，于是必须去找一句它们都没说出口、却都踩着的话；而推翻那句话的材料，**必须从那三家自己的教科书里取**。从外面找反例太容易，也太廉价；当推翻它的那件事就摆在它自己的领域里、被它自己反复引用、只是从没被当作这件事的证据时，这句话就回不出来了。

第二章的推翻材料来自行星科学自己：潮汐锁定的行星上，被画成一条线的晨昏带，是整颗星球上唯一可能住人的地方。第三章来自房地产估价自己：估价前必须先声明估的是哪一种价值，因为同一栋楼同一时刻有多个不同且都正确的数。第七章来自结构工程自己：塑性铰与内力重分布，说明边界条件是被上一轮响应改写的状态变量，不是外生参数。第十三章来自行为遗传学自己：失眠方差里最大的一块落在非共享环境，而这一块自一九八七年起被判为不成系统、不可系统研究——**这个判断是在从未向“另一个人”采集过的前提下作出的**。

第五编是本书之所以是一本书而不是十四篇文章的全部理由。

十四篇文章各自成立，但它们合起来欠着一件事：这十四个读数，是不是同一个读数换了十四个名字？第五编用五章回答它。第十五章把十四个读数摆到同一张表上，先写死一条合并规则再去看数据。第十六章把十四个那一格摆在一起，并为每一个配一个健康对照格。**第十七章做了一件别的书做不成的事**——本书十四章里有十一章用的是同一份公开判读记录，那一章把能算的读数在同一份记录上一次算齐。第十八章处理这套说法家族级的对手。第十九章是与三种既有传统的对话。

第十七章的结果对本书不利的部分比有利的多，而它最要紧的一条是这样的：本书最想成立的那条预测——十四个读数彼此不重复——在检验里命中了，而事后算出的风险度是零。**把八个读数彻底打乱、破坏一切配对关系，五千次里五千次仍然命中**。也就是说，那条预测在设计上就不可能失败，它的命中不构成任何证据。

我在写那一章的预注册时，专门把“风险度”这个装置写了进去，因为第十四章刚刚用它查出过同样的病。然后我仍然写出了一条不可能失败的预测。

这件事我没有藏，把它放在了第十七章的自反节，也放在这里。它说明一件事：知道一个陷阱存在，与不掉进去，是两件事。而如果不是我事先写死了要算风险度，那条预测就会以“命中”的形式进入结论，看起来和一条真证据完全一样。

它不产生信号。这本书讲的就是这个。

最后要说的是欠账。

其一，本书的十四个读数里，有三个在本书自己的记录上算不出来——索验比的分母、下一拍托管率、他源率。它们算不出来的理由，逐条就是提出它们那三章的论点。这既是本书最好的一处自治，也是本书最实的一处空。

其二，本书没有一位失眠患者的记录。第九、第十、第十一、第十三这四章处理的是失眠，用的全部是公开的健康队列与轻度入睡困难者队列。

其三，本书的一切建议都是程序性的——加一栏、标一个版本号、把两端分开公布、给每条建议标上时点。本书没有、也不打算给出任何一条关于怎样睡得更好的建议。第五章会解释为什么：在使用者能验收的那把尺上，快、深、省力这三样分不开；提一个自己无法验收的要求，是一件比睡不着更麻烦的事。

这本书能做的，只是让某些东西第一次在纸面上有一个位置。

有了位置之后它是什么，要由别人来回答。

王德生

二〇二六年八月 新加坡

导读

一·三条读法

这本书四百多页，但没有人需要从头读到尾。三条路径，各自完整。

第一条·只想知道它说了什么（约四十页）。读前言、导论、第十六章、全书结语。这四处合起来给出全书的判断、四个层级的划分、十四个位置的总表，以及本书自认的三个洞。读完这四处，你可以准确地转述这本书，也可以准确地反对它。

第二条·想用其中的某个读数（每章约三十页）。十四章互不依赖，任取一章单读即可。每一章都在自己的正文里给出该读数的完整定义、边界例与缺失规则；附录二那本编码手册可以直接照做，手册内不出现任何章名与本书造的词，因此你不必先接受本书的论述就能执行它。

第三条·想检验它（需要动手）。从第十五章与第十七章入手——那两章是本书对自己下的判决。第十五章写死了合并规则再去查数据，第十七章把十四个读数放到同一份公开记录上一次算齐。两章的预注册文件与全部脚本列在附录四，数据是公开的，任何人可以从原始文件复现本书的每一个数，包括复现本书的错误。

若你只有十分钟：读前言的头两页，再读第十六章那张表 16-1。

二·每一章长什么样

前十四章共用同一副骨架，认得这副骨架，读第二章起就会快很多：

问题 → 一句普通人会问的话（怎样才能快速入睡？失眠是什么？）。

三家 → 三门互不相干的学问，对同一个问题给出三个不能同时成立的答案。它们的作用是互相拆台，不是显得广博。

三家共同踩着的那句话 → 三个答案打架，说明它们共享了一个谁也没说出口的前提。找出它。

推翻它的材料 → 必须从那三家自己的教科书里取。从外面找反例太容易也太廉价；当推翻它的那件事就摆在它自己的领域里、被它自己反复引用、只是从没被当作这件事的证据时，这句话就回不出来了。

新对象 → 前提倒了之后，那个位置上剩下的是什么。本书给它一个名字。

可判定的形式 → 一句不带程度词的问话，可以拿去问记录。

真跑 → 预注册写在跑之前，含撤销规则；然后跑，然后如实报，包括打脸的那一条。

分界 → 与最接近的既有说法逐条划清，每条给一个能分出对错的判定形式。

证伪与反噬 → 什么条件下本章为假；本章的处方若被执行，可能造成什么坏处。

自反 → 用本章的判据检本章一遍。

这副骨架的固定表述在十四章里会重复出现。这是有意保留的——逐章改写只会制造无意义的措辞差异。

三·三层主张，与一条引用纪律

每一章的主张分三层，强度递减、稳固度递增：

第一层·本体层：那个位置上存在一样东西（未追认段、界带、判余、禁转、默借……）。这一层最强，也最容易倒。

第二层·工具层：这样东西可以被一个具体的读数量出来。这一层依赖第一层。

第三层·记录层：现行记录里没有它的位置。这一层最弱、最便宜，而且不依赖前两层成立——即使本体层被推翻，“记录里没有那一栏”这件事仍然是真的。

接受第三层不蕴含接受第一层。这是本书对读者的第一条请求。

引用纪律：若某项研究只采用本书的某个读数作流程变量，不得顺带声称发现了一种新的睡眠类型。本书的读数是用来审记录制度的，不是用来给人分类的。

四·十四章速查

章	它说的那一句	它的读数
一	一段时间要成为睡眠的一部分，得被后面的事件追认	未追认段比例
二	睡与醒之间那条线是有厚度的地方，它的产出被记在两侧名下	成簇承载比
三	归属是格式选出的第一名；险胜与压胜在账上完全相同	判余
四	从没见过的转换，可能是不可能、不划算、不许记，三者同形	残余率
五	一个要求只有在验收端可分辨时才是要求，否则是愿望	索验比
六	每种办法在开一条路的同时封另一条，而封不产生读数	自噪底
七	一个差值说的是哪一端，由构造者挑了什么量决定	端权对
八	快与深能否兼得从未被判定过，而缺的两样都不是技术难题	下降时长
九	失眠指着两套互不细化的划分，不存在第三方去裁决	双向漏出率
十	一次事件被读成了一种能力，而名字与承重可以长期不一致	长断占比
十一	失眠里稳定的那个量没有位置，而所有解释都在找位置	位重复度
十二	睡眠不是停下来，是下一拍被别人接住了	下一拍托管率
十三	你的睡眠靠着别人的例行安排，两边都不知道，撤走时无人违约	他源率
十四	办法的分布不由因果强度决定，由症状可见的位置决定	预测力—滞后剖面

五·十个词的速记

未追认段：没有被后续事件追认、因而不构成一个可数的段的那段时间。

界带：睡与醒之间那片有厚度的区域。**邻记：**界带的产出被记在它两侧的名下。

判余：几个候选里，第一名与第二名之间的分差。**无痕强判：**分差极小而记录与压倒性胜出完全相同。

禁转：一条从未被观察到的转换。**残余率：**被规则压低而不归零的那一点残余，其高度由例外条款的宽窄决定。

索验比：使用者索求的间距，除以他自己能验收的分辨率。

暗封：一种办法在打开一条入睡通路的同时封掉另一条，而封不产生任何读数。**自噪底：**同一个无人干预两夜之间，同一读数的天然差异。

端权对：一个差值读数里，两端各解释了多少变异。

未判速率：可以算、没有那一栏、也从来没有被裁决过的那个量。

异归：两套划分互不细化，而制度用同一个词称呼两边。

能力化误读：一次事件被读成一个人身上的一种能力。

无量量：大小可重复、而位置不可重复的那种量。

托拍事件：睡着时那些仍需回应的事，由谁接手。**默借：**一个人的睡眠依赖另一个人的例行安排，而两侧都不知道。

症状位偏置：干预点的分布由症状可见的位置决定，不由因果强度决定。

风险度：一条预测在零假设下也命中的概率的补数；风险度为零，意味着它在设计上不可能失败，它的命中不是证据。

六·读之前该知道的三件事

其一，本书不诊断，也不治疗。书中每一条界线、每一个读数、每一次统计，都是关于记录本身的，不是关于任何一位读者的。若你正为睡眠所苦，本书不能替代医生。

其二，本书不提供任何改善睡眠的方法，一条也没有。十四条处方全部是程序性的：加一栏、标一个版本号、把两端分开公布、给每条建议标上时点。第五章会解释为什么不给方法——在使用者能验收的那把尺上，快、深、省力这三样分不开；提一个自己无法验收的要求，是一件比睡不着更麻烦的事。

其三，书中若干处建议做一次自我清点或自我校准，它们一律是一次性的、做完就停的，不是一项日常习惯。请留意第五章与第十一章之间那处紧张：**持续的自我监测本身可能损害睡眠。**完整的边界写在下一页的使用声明里，那一页请务必读完。

使用声明

本书处理的是睡眠记录的结构问题，不是临床问题。付印之前，先把五条边界写在最前面。

其一，本书不诊断，也不治疗。书中出现的每一条界线、每一个读数、每一次统计，都是关于记录本身的，不是关于任何一位读者的。若你正为睡眠所苦，本书不能替代医生。

其二，本书不建议任何人据此改变正在进行治疗。第六章从气候学那里接受了一条限制：一个真实的信号可以长期低于内部变率而仍然真实。因此“落在噪声底以内”不等于“可以忽略”，更不等于“可以停药”。

其三，本书多处指出现行评估工具与判读规则的缺口，这不是对使用它们的人的指控。第十四章写得最清楚：这些配置通过了一切形式审查——要求具体、改善可测、双方诚实、报告不假。缺的是没有人问过某些问题。指出代价不等于否定它。

其四，书中若干章建议读者做一次自我清点或自我校准。请注意第五章与第十一章之间那处紧张：持续的自我监测本身可能损害睡眠。凡本书提出的自查动作，一律是**一次性的、做完就停的**，不是一项日常习惯。

其五，本书的全部人机分工写在后记里。检索、公开数据复算、结构检查与语言编辑使用了人工智能工具；判断、取舍与全部结论由作者负责，包括其中的错误。

目次

出版页信息.....	5
作者介绍.....	7
前言.....	9
导读.....	12
使用声明.....	16
导论 记录最齐整的那一刻.....	19
编一·每一格都得有一个标签.....	23
第一章 那三十秒算不算一段.....	25
第二章 住在线上.....	44
第三章 必须回答，所以从不犹豫.....	62
第四章 从来没见过，和不许见到.....	77
编二·每一次改善都得有一个分数.....	86
第五章 三样东西，一个分数.....	88
第六章 只要有一条路是通的.....	105
第七章 夜里那一端，一直没有动.....	122
第八章 下去的那一段，没有人记.....	144
编三·每个人都得归进一类.....	160
第九章 两张名单.....	162
第十章 那一次，还是那个人.....	173
第十一章 每次都不是同一个地方.....	184
编四·每笔账都得记在一个人名下.....	201
第十二章 谁替睡者接住下一拍.....	203
第十三章 没有人违约.....	226
第十四章 办法都在夜里.....	241
编五·合论.....	256
第十五章 十四个读数，是不是同一个读数.....	258
第十六章 十四种看起来完整的记录.....	266
第十七章 一次同批复算.....	271
第十八章 这套说法共同的对手.....	281
第十九章 对话.....	287
全书结语 那一格里本来有什么.....	290
参考文献.....	292
附录一 十四项赌注、撤回条件与三条共同条款.....	294
附录二 十四个读数的编码手册.....	295
附录三 四十二个学科入口、消融总表与叠计比.....	296
附录四 预注册文件、脚本与哈希.....	298
后记.....	300

导论 记录最齐整的那一刻

一·全书赌的那句话

全书赌一句话：

睡眠的缺口不是漏记，是强制作答的代价；它在记录最齐整的时候最大，因为扔掉的那一份不留痕迹。

这句话有三个部分，逐个说。

“不是漏记”。漏记是可以补的：多装一个传感器、多记一夜、把窗做细。本书十四章处理的东西，没有一样能靠这些补上。第一章的未追认段不是“记得少了”，是在三家的账上都不构成一样东西；第三章的判余不是“没记下来”，是记录格式规定每格只能有一个标签，第二名在被写下的那一刻就不存在了。

“强制作答的代价”。每一处损失都发生在某个装置被要求给出恰好一个答案的时候。判读手册要求每个三十秒窗给一个分期；报告要求每一项改善给一个分数；诊断要求每个人归进一类；账目要求每一笔记在一个人名下。这四个**“必须”**就是本书的四编。

“记录最齐整的时候最大”，以及“扔掉的那一份不留痕迹”。这两句连在一起才是本书的真正主张。一个允许留白的记录，会用空格告诉你哪里没答案；一个不允许留白的记录，会用一个看上去和别的格子一模一样的答案，把“这里本来有分歧”这件事抹平。于是**记录越完整，缺口越看不见——不是比喻，是机制**。

这条全书赌的那句话是可以被推翻的。推翻它只需要一件事：**指出一个允许弃答的记录制度，而同样的缺口仍然在那里**。本书自己跑过这一条：第三章测了弃权段，它只占百分之零点一四，且几乎全部落在格式本来就允许留白的离床时段——也就是说，现行判读几乎从不弃答。第四章从另一头给出了形状：当规则带有例外条款时，被压低的那一类不会归零，会留下一个稳定的残余。

二·四编：答案是在哪一步被强制的

本书不按“是什么—怎么办—为什么”分编，也不按睡眠、失眠分编。分编的轴是：**这一次强制作答，发生在哪一层**。

四层由细到粗，逐层放大：

编一·每一格都得有一个标签（判读窗层）。一夜被切成三十秒一格，每格必须落一个分期。第一章问：一段时间凭什么算一段？第二章问：睡与醒之间那条分界有没有厚度？第三章问：几个候选之间的分差去哪儿了？第四章问：一条从没被观察到的转换，是跳不过去、不划算跳，还是不许这么记？

★这四章咬的是同一个装置——三十秒窗、占优规则、取最大值、单标签。它们不是四篇互相重复的文章，是同一个装置的四个切面。这一点在四章各自写成的时候，四章都不知道；是第十七章把它们放到同一份记录上算，才显出来的。第十七章还查出：第一章与第二章的两个读数在同一份记录上相关达到+0.735，而其中百分之九十五由一条算术解释。**这条算术写在第十五章，两章的正文各挂一句公开声明**。

编二·每一次改善都得有一个分数（读数层）。一个人想睡得好，最后拿到的是一个数。第五章：他要的三样东西，在他能验收的那把尺上是同一个读数。第六章：每种办法在开一条通路的同时封另一条，而开与封被折叠成一个净效果。第七章：一个差值读数说的是哪一端，由构造者挑了什么量决定，而报告只公布差。第八章：一整段下降只留两个端点，于是“快”与“深”能否兼得至今没有被判定过。

编三·每个人都得归进一类（个体层）。第九章：主诉与客观界线给出两套划分，两个方向都大量漏出，而制度用同一个词称呼两者。第十章：一次事件被读成了一种能力。第十一章：这个系统里稳定的量没有位置，而所有解释都在找位置。

★这三章全部处理失眠，全部用同一批一百五十三夜的不同切法；而第九章与第十章是两条互不知情的产线，撞到了同一个方向——自述健康者身上客观界线大面积命中，且随年龄陡升。它们互为对方的独立复现。

编四·每笔账都得记在一个人名下（主体层）。第十二章：回应的去向与身体的内转是两笔各自成立而互不登记的账。第十三章：一个人的睡眠可以持续依赖另一个人的例行安排，而两侧都不知道，撤走时两侧都不产生记录。第十四章：所有办法压在夜里那一端，而这条分布从未被要求给出理由。

四层不能互相还原。判读窗做细，不会让报告公布两端；报告公布两端，不会让诊断承认两套划分；诊断承认两套划分，不会让账目记上第二个人。第十五章把这四条互不蕴含逐条写实。

⚠这里必须写一句诚实话：第十七章检验了这个分层，结果是方向正确、幅度不足。同编内读数的平均相关确实高于跨编（0.202对0.135，置换检验单侧 $p=0.0008$ ），但差值只有+0.067，而预注册要求的放行线是+0.10。按纪律判未命中。因此本书不主张四编在数据上得到了支持；四编的分层主要是论证上的，不是统计上的。

三·十四个那一格与三条签名

本书每一章都指认一个具体的位置，那里有一样东西在损失，而现行装置读不出来。这类位置在本书里出现十四次，第十六章列出总表。它们共有三条签名：

签名一·短期表现为改善。把两个分期合并，转换数下降百分之十六，记录看上去更干净；把三个要求压成一个分数，评价变得可比；把窄口位置从报告里省掉，报告更简洁。每一次损失都以一次简化的形式发生，而简化在当下总是被算作进步。

签名二·失效不产生任何信号。这是三条里最要紧的一条。第十三章把它写到了本书里最狠的一处：一篇没有采集过供给方的论文，读起来和采集过的完全一样——参考文献一样长，论证一样连贯，表格一样齐整。若不是那一章自己写出来，读者没有任何办法从文本上分辨这两种论文。

签名三·越正规化越自我加固。规则越严格、判读者之间越一致、报告格式越统一，由规则造成的结构反而越干净、越像一条规律。第四章给出了这条的判定形式：提高编码器一致性，误差会减小，而规则造成的那个十倍差距不变，甚至更稳定。

四·十四个读数，以及为什么其中六个进不了验证表

本书提出十四个读数。第十七章把它们摆到同一份公开记录上，结果是一个三分：

- 八个是“每夜一个数”，能进同一张表：未追认段比例、成簇承载比、体例折叠损失、残余率、下降时长、命中界线数、长断占比、窄口位置。
- 三个是别的数学对象：自噪底是跨夜配对差，端权对是一对份额（它是装置不是读数），预测力一滞后剖面是一条曲线。
- 三个在这份记录上算不出来：索验比的分母、下一拍托管率、他源率。

第三类那三个，是本书最要紧的一处。它们算不出来的理由，逐条就是提出它们那三章的论点：使用者的验收分辨率从没被测过（第五章）；记录里没有“谁接住下一拍”这一栏（第十二章）；计分条目里没有一条问对面那个人自己做了什么（第十三章）。

由此得到一条对本书不利、但必须写在导论里的推论：本书能够放进验证表的那八个读数，恰恰是最不激进的八个——它们全部可以从现成的判读标注算出来，也就是说全部落在现行记录制度已经愿意登记的范围之内。本书的验证力，在本书最要紧的地方最弱。

五·三个“别处做了，这里没做”的正面对照

本书反复用三个对照。它们不是本书的对手，是本书的证明：把一条被画成分界的线打开成一段有厚度、可清点、可立项研究的东西，这个动作在别的领域已经完成了。

其一，天文学的曙暮光分级。日出日落不是一条线，是三段——民用、航海、天文，各有明确的太阳高度角定义，各自进历书。

其二，生态学的交错带。一九〇五年就命名了，有边缘物种（两侧都没有的物种）、有 ecotone 与 ecocline 之分（锐边界与渐变带）、有移动分割窗这类专门测宽度的方法，一个世纪的文献。同一张“有没有字段”的清单拿去核对交错带，六项全有；核对睡眠判读，六项五无。同一张清单，两个相反的结果，差别不在对象，在账本。

其三，概率化睡眠图。二〇一八年就做出来了：不给单一标签，给每一期的概率。也就是说，第三章那条“保存第二候选与分差”的建议，技术上七年前已有实现——而临床报告格式至今仍要求每个窗一个标签。技术已有而格式未改，这不是本书的功劳，这是本书的靶子。

第八章另给了第四个：药物滴定速率。同一门医学内部，滴定速率被当作一个必须写进方案的量；而入睡的下降速率，可算，无栏，未裁。

六·十条约定

本书的读法，以及不该被怎样读：

一、每一章可以单独读。十四章互不依赖，第五编的五章依赖前十四章。

二、每一章有三层可分别引用的主张：第一层是本体（最强、最易倒），第二层是分析工具，第三层是记录缺口（最稳、最便宜，且不依赖前两层成立）。接受第三层不蕴含接受第一层。

三、引用纪律：若某项研究只采用本书的某个读数作流程变量，不得顺带声称发现了一种新的睡眠类型。

四、本书的三层建议全部是程序性的：加一栏、标一个版本号、把两端分开公布、给每条建议标上时点、基线期多记一夜。本书不提供任何改善睡眠的方法。

五、“必须”这个词在本书里始终指记录规则的要求，不指生理的必然。第四章专门处理二者如何被混为一谈。

六、凡本书报告的比例，一律须查它是汇总比例还是逐个体的中位数。第十七章查出三处两者相差极大的地方，已在相应各章补注。

七、凡本书提到分析窗、清醒时长口径、判读体例版本，一律须与数字一同引用。第六章与第十七章各记录过一次因口径不同而结论翻面的事故。

八、本书报告的每一次统计，脚本与预注册文件都列在附录五，含哈希。任何人可独立复现，包括复现本书的错误。

九、跨章重复的表述是有意保留的。十四章共用同一套工序，工序的固定表述逐章改写只会制造无意义的措辞差异。工序全文见附录三。

十、本书用“本章”“本书”，不用“本章”。各章原为独立文章，装书时统一改过；若有残留，以“本章”读之。

七·这本书可能整个是错的

三条写在这里，不放在结语。

其一，如果现行记录制度上的这些缺口，在别的领域也同样存在且同样无害，那么本书指认的不是缺口，是记录这件事本身的性质。本书的回应是第五节那三个正面对照：别处做了。但三个对照不足以支撑一条普遍主张。

其二，如果十四个读数被证明是三四个量的不同投影，本书的编次作废。第十七章检验了这一条，结果是没有任何一对越过合并线——而那条检验的风险度是零，它的通过不提供证据。这一条至今没有被真正检验过。

其三，如果把这些缺口一一补上之后，睡眠研究并没有变得更好，甚至因为记录负担加重而变差，那么本书的处方是净损失。本书没有能力事先排除这一条，第十六章末尾会再提一次。

编一·每一格都得有一个标签

判读窗层

这一编处理最细的那一层：三十秒。

一夜的脑电被切成三十秒一格，每一格必须落下恰好一个分期。这条规则是一九六八年定下来的，它的来历不是生理学而是纸速——每秒十毫米的走纸，一页三十厘米，于是一页三十秒。一个记录设备的物理尺寸变成了一门学科的时间单位，并在纸张退场之后又存在了半个多世纪。

四章咬的是同一个装置的四个面。第一章问：一段时间凭什么算一段？第二章问：睡与醒之间那条分界有没有厚度？第三章问：几个同时成立的候选之间，分差去了哪里？第四章问：一条从没被观察到的转换，是跳不过去、不划算跳、还是不许这么记？

四章成稿时互不知道彼此在咬同一个东西。是第十七章把它们放到同一份记录上算，才显出这一点——也才顺带查出第一章与第二章的两个读数在同一份记录上相关达到零点七三五，而其中百分之九十五由一条算术解释。这一编因此既是全书最集中的一编，也是被自己的验证削得最狠的一编。第十五章处理了它，两章正文各挂了一句公开声明。

读这一编时值得带着一个问题：如果格子做到五秒，这四章里哪几章的缺口会消失，哪几章一点不动。

第一章 那三十秒算不算一段

那三十秒算不算一段

论一段时间成为睡眠的一部分需要被后来的事件追认，以及现行的一切读数为何都读不出没有被追认的那些段落

提要

关于“睡眠是什么”，今天有三个互不相容的强答案。化学把它读成一个可测的量在时间上的高低；形式理论把它读成一段在终止处完成的时间形式；睡眠社会学把它读成一套被产业时间表安排出来的作息制度。三家争的是一夜睡眠的身份该由什么来定，而三家共同假定了一件从未说出口的事：一夜睡眠由一批边界清楚、可在事后清点的段落构成。

本章用音乐学自己的材料推翻这条前提——形式功能可被后续事件回溯改判，这是该领域反复引用的常识，只是从未被当作关于“段落如何成立”的证据。由此提出：一段时间要成为睡眠的一部分，需要被它后面的事件追认；没有被追认的那一段，不是记录得少，是在三家的账上都不构成一样东西——而三家的账各自都能记平。文章把这类段落命名为**未追认段**，把它成立所需的时间命名为**追认延迟**，并给出一句不含程度词的判据：在这份记录里，改动最后一段的判定，会不会改动前面某一段的判定。

用四十夜公开判读记录做了一次预注册检验：未追认段占段落的 32.4%，只占时长的 3.35%——时长几乎守恒，段落结构被摧毁三分之一。预注册的第三条预测方向完全错了。据此修订的构念是一条**待检验的假说**，而不是一条**已确立的因果断言**：一段时间能否成段，与该类在账本上有没有被写下追认规则有关，而与它在生理上有多“过渡”无关。本章的数据只有五个分期，且它们在内在标记强度上并不同质，因此只能给出方向一致的初步读数，不能充作该假说的检验；能充作检验的设计写在第九节。

本章所主张的、目前找不到主人的那一格，收窄为一句：**同一份记录上，段落账与时长账同时成立而互相看不见**。文章因此与三位最危险的全球近邻逐一划界——睡眠医学自己的微结构传统（CAP，四十年文献，本领域离本章最近的一位）、地理与时空分析的可变时间单元问题（MTUP）、以及量化社会学的通约化命题——并申报四处此前未申报的站内自撞。末尾给出二十条判决性分界、四处反向约束、一条写死日期的赌注与三条伦理硬约束。

关键词：追认延迟 | 未追认段 | 碎屑段 | 形式功能的回溯改判 | 段落读数与时长读数 | 睡眠判读单位

一·导论：一个没有人认为需要讨论的问题

问“睡眠是什么”，会得到三种听上去都对、彼此却不能同时成立的回答。

第一种回答来自化学。它说睡眠是一个量。清醒把某些东西累积起来，睡眠把它们卸下去；只要能测到那个量，睡眠是什么就说清了。这一路在最近十年拿出了很硬的东西——一组突触蛋白的磷酸化程度随清醒时间累积、随睡眠下降，与睡眠需求同步涨落。这个答案的力量在于它把一件玄虚的事变成了一根可以画出来的曲线。

第二种回答来自形式。它说睡眠不是量，是形状。一夜有段落、有反复、有渐变、有收束；同样的成分按不同次序排列就是不同的东西。这一路最成熟的说法不在睡眠研究里，而在音乐学里：一段音乐是什么，取决于它在整体中承担的功能——开头、中段、还是结束——而不取决于它自身的音响内容。这个答案的力量在于它解释了为什么“睡了八小时”和“睡好了”不是一回事。

第三种回答来自社会。它说前两种都把结果当成了原因。一夜连续八小时、夜里睡白天醒、一个人一张床——这些不是生理常数，是工厂轮班表的产物；在它成为标准之前，人们普遍分两段睡，中间醒来一段。这个答案的力量在于它指出“什么算正常的睡”这件事本身是被安排出来的。

三家打得很凶。化学说社会学讲的那些不影响腺苷怎么积累；社会学说化学测的那根曲线之所以被认为该在早晨归零，是因为工厂七点开工；形式论说前两家都在数量，而数量读不出形状，就像把一首曲子的音符按音高统计一遍读不出它是什么曲子。

本章不打算在三家之间判谁对。判谁对需要三家站在同一个语域里，而它们不在。本章要做的是另一件事：找出三家争吵时脚下共同踩着的那块地，然后用其中一家自己的材料把它挖开。

那块地是这样一句话：一夜睡眠是由一批边界清楚的段落构成的，这些段落事后可以被清点。

念给三家听，三家都会说这还用说吗。化学要清点各期时长，形式论要清点乐句与终止式，社会学要清点谁在什么时段被允许睡多久。三家吵的是这些段落该由什么来定，从没有人问过：段落这个单位本身，是怎么成立的？

本章的回答是：它靠后面的事件成立。一段时间要成为睡眠的一部分、要成为一个可以被数的段落，需要它后面的东西把它认下来。没有被认下来的那一段，不是被记得少了，是压根不构成一样东西——而三家的账各自都能记平，因此谁也不会发现少了什么。

第二节交代三家各自的最强主张与各自的硬证据。第三节把冲突定位到位置上，说明为什么三家的分歧不可裁决。第四节找出三家共同踩着的那句话，并从音乐学自己的材料里取出推翻它的那一件。第五节提出追认结构与两个读数，给出一句不含程度词的判据并在五个场合各跑一遍。第六节报告一次预注册的真跑，包括一条方向完全错了的预测和它逼出来的构念修订，以及一处双因未分的诚实交代。第七节做二十条分界，其中三条是全球最危险的近邻，四条是此前未申报的站内自撞。第八节写四家反过来给本章加的约束。第九节写证伪、赌注与伦理边界。第十节写本章自己落在哪里。

二·三个最强答案

2.1 化学：睡眠是一个量的高低

这一家最新、也最有说服力的形态，是把睡眠需求直接等同于的一组蛋白质的化学状态。Wang 等人在二〇一八年用定量磷酸化蛋白质组学比较了睡眠剥夺小鼠与一个天然嗜睡突变品系，识别出八十个主要位于突触的蛋白，它们的磷酸化状态与睡眠需求同步变化，被称作睡眠需求指数磷酸化蛋白

(sleep-need-index phosphoproteins)。其核心判断可以引到原句：这一批突触蛋白在清醒期被累积地磷酸化、在睡眠期被去磷酸化，与睡眠需求相符，这可能是突触稳态与睡眠—觉醒稳态共同的底层机制 (Wang et al., *Nature* 558: 435–439, 2018)。Ode 与 Ueda 后来把这条线索写成一份完整的假说综述，称蛋白质磷酸化“可能就是与睡眠需求相联系的那个分子实体” (Ode & Ueda, *Frontiers in Psychology* 11: 575328, 2020)。

这一家的单因主张很干脆：只要那个量测得准，睡眠是什么就定了；其余一切——形状、社会安排、主观感受——都是这个量的显示或干扰。

它的奠基层不新。二十世纪初就有人从长时间不睡的动物脑脊液里找“催眠毒素”，认为睡眠的本质是某种物质的积累 (Legendre & Piéron 一九一〇年代的经典实验)。一个世纪过去，物质换了，形式没变：睡眠仍被理解为一个标量在时间上的走向。

它自曝的地方也写得很老实。Ode 与 Ueda 在同一篇里承认，睡眠需求动力学背后的分子机制“仍不清楚”，并且指出这些磷酸化变化的一个可能来源，只是清醒期突触输入增多的结果——也就是说，读数与被读之物的因果方向尚未定下来。这一处坦白后面还要用到。

2.2 形式：一段时间是什么，取决于它承担什么功能

第二家的成熟形态不在睡眠研究里。要找一门把“一段时间的身份”当作正经问题研究了两百年的学问，得去音乐学。

Caplin 在《古典形式》里提出的形式功能理论，核心是这样一句可引的话：形式功能关乎一部作品的“各个组成时段表达它们自身在该作品音乐时间中所处位置”的能力 (Caplin, *Classical Form*, Oxford University Press, 1998, p. 23)；而辨别这些时段如何表达其时间性，正是形式功能理论的目标

（同上，p. 24）。后续研究者把这一点说得更直白：一个音乐上的开头、中段或结尾，离开它与另外两种功能的关系就没有意义。

这一家的单因主张同样干脆：只要看清一段时间在整体组织里承担什么功能，它是什么就定了；至于它内部有多少能量、社会给它派了什么用场，都不改变它是开头还是结尾。

它的支撑不只是理论。一项音乐感知实验把古典乐句的片段抽离原境后放给听者，要求判断这是主题的开头、中段还是结尾；音乐家与非音乐家都能做出高于随机的判断——即所谓“内在形式功能性”确实可被感知。这一条对本章非常要紧，第六节与第八节都要拿它来削本章自己。

它自曝的地方，是本章全部论证的入口，所以要引得完整一点：在形式功能理论的实际操作中，一个乐群可以在事后被重新解释为表达了与最初判断不同的功能。同一段音乐，听的时候像是主题的呈示，听到后面才发现它其实是引子；分析者据此把先前的判定改掉。这在音乐学内部有一个专门的名字——回溯性重释——并已发展成一整套关于“正在成为”的分析语言（Schmalfeldt 一九九五年起的一系列工作，及其二〇一一年的专著）。

2.3 社会：什么算睡，是被安排出来的

第三家把前两家的对象整个历史化。睡眠社会学在二十一世纪初被正式提出为一个研究领域（Williams, *Sleep and Society*, Routledge, 2005；同作者 *The Politics of Sleep*, Palgrave, 2011），其最锋利的一次表述来自一部医学人类学专著：当前这套睡眠模型“其根不在生物学，而在工业资本主义对生产率的持续需求”；在工厂轮班制引入之前，人们的睡眠实践是多样的，最常见的是夜里两段睡眠加上白天小睡，而新的作息标准——夜间连续八小时——导致了其他睡法被病理化（Wolf-Meyer, *The Slumbering Masses*, University of Minnesota Press, 2012）。

这一家的单因主张是：只要看清社会如何安排作息与如何认定异常，睡眠是什么就定了；生理机制只是在被安排的条件下运转的东西。

它自曝的地方同样关键：这一路的作者都承认，一套作息制度一旦稳定运行若干代，就会真的成为身体的条件——被建构出来的东西会变成不可任意更改的生理事实。也就是说，社会学自己承认它讲的那个“任意性”是有时限的。

三·冲突定位：为什么这三家不可能靠判谁对来收场

把三家并排，很容易读成“三个层次，各有各的道理”。这是本章首先要挡掉的读法。

三家不是三个层次，是三种关于“什么单独够用”的主张，而它们各自站的位置不同：

	它把什么当成单独够用的那一样	它据此把另外两样看成什么
化学	当下可测到的那个量	形状是量的副产品；社会安排是干扰变量
形式	组织的次序与功能	量是形状的材料；社会安排是外部约束
社会	睡眠所纠缠的制度条件场	量与形状都是在这个条件场里被造出来的读数

分歧不在同一根轴上，因此不可判决。你没法设计一个实验去判“睡眠到底是量还是形状还是制度”——三家根本不在同一个语域里作答。凡是不可判决的分歧，能做的只有一件事：造一个把三者关联起来的东西。

顺带挡掉一个容易混进来的第四家。“临床”“职场”“学校”不是一家——它们是这些主张落地的地方，不是持有主张的位置。本章后面会大量使用临床判读手册的规则，但那是兑现材料，不是第四个理论立场。

四·三家共同踩着的那句话，与推翻它的那一件材料

4.1 三家争的是什么

把三家的争论压成一个句式：三家争的是“一夜睡眠的身份”该由什么来裁定——由那个量裁、由形式功能裁、还是由制度安排裁。

填在横线上的那样东西，三家都默认它是那种可以被裁定的东西。而这正是三家共同踩着的那句话：

一夜睡眠是由一批边界清楚的段落构成的，这些段落事后可以被清点。

三家的全部工作都建在它上面。化学要报告各期时长与总睡眠时间，得先有“各期”这些段落；形式论要分析乐句与终止式，得先有乐句这个单位；社会学要批判“八小时连续”这个规范，得先承认“一段”是可数的，才能说它本该是两段。

念给三家听，三家都会说这还用说吗。这是它是三家共同踩着的那句话的验收标志：会引起争论的那种句子，是第四家的立场，不是三家共同踩着的那句话。

4.2 推翻它的那件材料，来自音乐学自己

从外面搬一个理由来推翻这条前提，得到的只是第四家的立场——那不是取消争论，是加入争论。要取消它，推翻的材料必须是三家之一自己已经掌握、已经发表、只是没往这个方向用的事实。

这件材料在音乐学那里，而且不是边角料，是形式分析的日常操作：

一个乐群可以在事后被重新解释为表达了与最初判断不同的功能。

贝多芬那些著名的开头是这套分析语言最常用的例子：一段听起来像主部主题的音乐，随着后面的展开，被追认为一段引子；一段听起来像过渡的音乐，在再现时才显出它一直是主题。分析者不认为这是自己听错了——理论明确规定，形式功能是可以被后续事件回溯改判的。

这件材料在音乐学自己的领域里回答的是另一个问题：如何分析古典曲式里那些身份模糊的段落。没有人把它拿来当作关于“段落这个单位如何成立”的证据。而一旦把它拿来当证据，三家共同踩着的那句话就塌了一半：

如果一段时间的功能可以被后面的事件改判，那么“在事后清点”这个动作就没有一个可以站定的时点。你清点的那一刻，正是你把这些段落定下来的那一刻；换一个更晚的时点再清点，得到的段落可能不同。段落不是事后被清点出来的，是被后续事件认下来的。

4.3 塌掉的另一半

前提还有另一半：“边界清楚”。这一半由化学与 sociology 各自的自曝拆掉。

化学承认读数与被读之物的因果方向未定，意味着用那个量去划边界，划出来的边界的地位不明——它可能是被测量规则造出来的。社会学承认制度会变成生理事实，意味着制度划出来的边界会长进身体里，事后无法用“这是社会的、那是生理的”把它们分开。

两半都塌之后，剩下的问题不再是“一夜睡眠由哪些段落构成”，而是：一段时间凭什么成为一个段落。

五·追认结构：一段时间怎样才成为一个段落

5.1 承重命题

一夜睡眠的组成单位，不是可累加的时长，不是在终止处完成的形式，也不是被制度安排出来的时段，而是一批需要被后续事件追认才能成立的段落；没有被追认的那些，不是记录得少，是在三家的账上都不构成一样东西——而三家的账各自都能记平。

三个否定各对应一家，缺一个就说明有一家没被处理。第四句是本章要立的东西。

要紧的是最后半句。账能记平，是这件事至今没有被看见的全部原因。总睡眠时间照样算得出来，各期百分比照样加到一百，睡眠效率照样有个数——每一本账都是完整的、自治的、没有缺口的。缺的不是数字，是单位。

5.2 三个用词，先定死

追认：一段时间类别归属，由它后面发生的事件确定下来。不是“解释”，不是“叙述”，不需要有人在场——后面那一段本身就是追认者。

追认延迟（记作 L）：从一段时间开始，到它的类别归属不再随后续信息改变，所经过的时间。单位随领域而变：睡眠判读以判读窗计，音乐以小节计，制度以轮次计。L 是一个时长或时刻，不是一个比率——这是本章有意选择的读数形态，理由与它带来的一处自相矛盾，见 5.5。

未追认段：在现行规则下，从未取得独立段落地位的那一段时间。它发生过，占用过时间，在时长账上被计入了某个类别，但在段落账上不存在。

5.3 一段时间成段，要过四道

这不是一个二维格，是一条序列。四步的次序不能换：

- 1. **发生**——这段时间在物理上存在，可被记录设备捕捉到。
- 2. **超过最小可记单位**——现行规则规定了一个下限，短于它的一律不进入判定。判读手册规定觉醒必须持续至少三秒才算一次觉醒，且其前需有至少十秒的稳定睡眠；判读窗固定为三十秒，窗内若有两个及以上分期共存，按占最大部分的那个判定。不足半窗的那部分，在成品里不存在。
- 3. **被后件定类**——它属于哪一类，由它后面接的是什么决定。这不是本章的推想：判读手册里明文写着一条延续规则——一段没有快速眼动的片段，若它后面紧接着确定的快速眼动期、且中间无觉醒、且其余条件满足，则这一段被判为快速眼动期。手册用的是“先前且相接”的措辞，意思是这一段的身份由后面那一段给出。
- 4. **取得段落地位**——它作为一个可数的段落进入报告。过不了第 2 或第 3 步的，停在时长账里，不进段落账。

停在哪一步，决定了它是什么。停在第 2 步的，本章称作**碎屑段**：它前后是同一类，中间被一个不到最小单位的插曲劈开，于是原本一段连续的过程在账上变成三段——或者反过来，被并成一段而中间那一下消失。停在第 3 步的，是**身份悬着的那一段**：后面没有出现能给它定类的事件。

5.4 一句不含程度词的判据

命题要配一句可以拿去问文件、问日志、问操作规程的话，而不是拿去问人的判断。禁用“应当”“有意义”“实质性”“充分”“真正”“恰当”“合理”这一类词。本章的判据是：

在这份记录里，改动最后一段的判定，会不会改动前面某一段的判定？

会，说明这套规则里有追认结构；不会，说明它宣称每一段的身份都由该段自身内容决定——那么它必须能回答第二问：不足最小单位的那些片段，被记到哪里去了？

把这句话在五个场合各跑一遍。五个答案必须互不相同，且至少两个是查出来的、不是想出来的：

场合	答案	来源
睡眠判读	会。延续规则明文规定：前一段的分期由后一段是否为确定的快速眼动期决定。	查出来的（判读手册规则原文）
音乐形式分析	会。同一乐群在听到后续展开后被改判为另一种功能，是该学科的标准操作。	查出来的（形式功能理论的回溯改判）
临床事件记录	会，但方向相反：规则规定周期性肢体运动伴随的觉醒若间隔不足十秒只记第一个——后面那一次的存在被前面一次取消了。	查出来的（判读手册规则原文）
一次白天的打盹	会。车上眯的二十分钟算不算“睡了”，取决于当晚还睡不睡得着：睡得着，它是小睡；睡不着，它被记成“白天睡多了”。	从命题可推
学术工作的定位	会。一份稿子是“前驱工作”还是“没人接的稿子”，由后来有没有人接着做决定；改判可以发生在几十	从命题可推

	年后。	
--	-----	--

前三条是查出来的，且其中第三条反过来改了本章：本章原先只想到追认是“后件确认前件”，第三条给出的是“前件取消后件”——追认是双向的，规则可以让后面那一次不存在。本章据此把追认的定义从单向改成了双向。

5.5 为什么读数是一个时长，不是一个比率——以及本章在这一点上的自相矛盾

跨领域的读数最省事的做法是造一个无量纲比率。本章有意不这么做，理由有两条。

第一条是本章自己的内容逼出来的。比率是时长账上的东西，而本章全部的主张是段落账与时长账不重合。用一个比率去表达“段落不成立”，等于用被批评的那套单位去说批评本身。

第二条是防误用。比率会被立刻拿去排名与考核；而一个“从这一段开始，到它的类别定下来，经过了多久”的时长，只有在同一套判读规则下才可比，跨规则比较立刻失效——它自带一条不可跨机构横比的限制，这在第九节的伦理约束里还要用到。

追认延迟 L 在三个领域的兑现：睡眠判读里是“往后看几个判读窗才能定下这一段”；音乐分析里是“听到第几小节才能定下这一段是引子还是主题”；制度里是“到第几轮才能确定上一轮那件事算不算数”。三处量纲不同，问的是同一句话。

现在必须承认一处自相矛盾。本章在这一小节论证读数应当是时长而不是比率，而第六节报告的全部读数——32.4%、3.35%、49.9%、8.1%、35.1%——无一例外都是比率。这不是笔误，是本章目前能力的边界，须写清楚它的性质：

本章实际报告的，是**段落账的存量读数**：在一份已经判读完毕的成品里，有多少段落没有取得地位。它与**延迟读数 L**是两个不同的量。存量读数回答“已经损失了多少”，延迟读数回答“一个段落要多久才被定下来”。前者可以从任何一份公开的判读成品算出，后者需要判读过程的中间态——而现有记录里没有这个字段（见 6.6）。

因此本章的处境是：它论证了哪一个读数才是对的，却只跑得动另一个。这一处不打算用“存量读数也很有用”来搪塞。它的诚实表述是：**本章的核心读数至今未被测量过一次**，第六节的全部数字都只是它的旁证，第九节第 5 条把“L 的中位数为零”写成了本章在睡眠领域上的证伪条件——那一条至今无人能跑，包括本章作者。

六·真跑一条：预注册、结果，以及方向完全错了的那一条

6.1 为什么必须跑

一条从没被真实数据碰过的证伪条款，只证明作者想清楚了。本章选了最便宜的一条去跑：**未追认段在真实的整夜判读记录里，占多大比例？**数据用公开的 Sleep-EDF Expanded 数据库中 sleep-cassette 部分按文件名排序的前四十份人工判读标注，任何人可以复算。

6.2 预注册（跑统计之前写死）

- 判读窗三十秒；段=连续同分期的判读窗串；追认阈 $K=2$ 个窗（六十秒），取值理由是：手册要求一次事件被记入须有足够时长且前后有语境，在三十秒粒度上最小的“两侧都能定位”的单位是两个窗。K 在跑之前定死，事后不改。
- 未追认段=长度小于 K 的段。
- 剪裁规则：去掉记录首尾各自那一段离床清醒中超过三十分钟的部分。
- 四条先验预测：
 - P1 未追认段占段数比例大于 25%。
 - P2 未追认段占时长比例小于 5%。
 - P3 未追认段比例按分期排序为 $N1 > W > R > N2 > N3$ （越是过渡态越依赖后续）。
 - P4 未追认段的时长密度后半夜高于前半夜。

- 撤销规则：若 P1 低于 10%，本构念在成人常规夜间记录上不成立，须改写适用范围；若 P3 在 N1 与 N3 之间反转，“过渡态更依赖后续追认”这一条直接判错。
- 已知代理问题（写在跑之前）：本库标注是人工判读的成品，判读者判读时已看过后续；因此本次测的是成品里留下的短段痕迹，不是判读过程本身。三十秒粒度会掩盖更短的段，故所有比例都是下界。

6.3 结果

四十夜，三百六十一小时，四千四百七十七段。

- P1 命中**：未追认段占段数 32.4%（1452/4477）。逐夜中位 30.9%，最低 17.9%，最高 45.0%，没有一夜低于撤销线。
- P2 命中**：未追认段占时长 3.35%（726 分钟 / 21664 分钟）。
- P4 命中，但差距极小**：前半夜 3.29%，后半夜 3.41%。这个差距小到不足以支持任何解释，本章不据此立论。
- P3 方向完全错了**：

分期	段数	未追认段	占段数	总分钟	未追认分钟
N3	829	414	49.9%	2854	207
N1	998	417	41.8%	1422	208
W	580	211	36.4%	4274	106
N2	1651	376	22.8%	9169	188
R	419	34	8.1%	3947	17

实测次序 N3 > N1 > W > N2 > R；预测次序 N1 > W > R > N2 > N3。N1 与 N3 在预测里分居两端，实测里挨在一起且都在高位；R 从中间掉到了最低。按预注册写死的撤销规则，“过渡态更依赖后续追认”这一条判错，本章不保留它。

6.4 打脸之后：构念被改成了什么

一条预测错了，比三条对了更有用——前提是照实报告并追问它为什么错。

追问的结果是本章最要紧的一条经验发现：

R 是判读手册里唯一明文写了追认规则的分期，而它的未追认段最少（8.1%）；没有写追认规则的那些分期，未追认段最多——N3 接近一半。

也就是说，决定一段时间能不能成段的，可能不是它在生理上有多“过渡”，而是账本上有没有为这一类写下追认规则。写了规则的，模糊的那一段被明确地并入或排除，段落成形；没写规则的，同一段连续过程被判读阈值上下的抖动切成一堆不算数的碎屑。

一次探索性的补充分析（非预注册，标明如此）坐实了这条解释的一半：

- 全部未追认段中，74.9%（1088/1452）的前后邻居是同一个分期——即一段连续过程被一个半分钟的插曲劈成三截。
- N3 的未追认段中，95.4%（395/414）前后都是 N2。这几乎是一个纯粹的阈值现象：慢波占比在判读门槛上下抖动一次，账上就多出两个边界、三个段落。
- 把 N2 与 N3 视为同一类之后，四十夜的段数从 4477 掉到 2907——**35.1% 的段落是判读规则制造的，不是生理事件制造的**。

而这一切，时长账完全读不出来：未追认段只占 3.35% 的时间。总睡眠时间、各期百分比、睡眠效率——每一个现行读数都几乎不动。账记得平平整整，而三分之一的段落结构是规则的产物。

6.5 R 那一格的双因未分：本章不能把它算作自己的证据

上一小节的解释有一个洞，必须自己指出来，因为审稿人一定会指。

R 的未追认段最少，本章归因于“手册为它写了延续规则”。但同一格数据至少还有第二个解释，而且它更省事：**快速眼动期本身就是一个巩固态**——它的自然段落本来就长，短段落少是该状态的段长分布使然，与手册写没写规则无关。第八节第二小节还会引入第三个解释：R 的内在标记极强（低振幅混合频率、低肌张力、眼动），因此不必依赖后件也能定类。

三个解释在这一格数据上完全共线。本章原先写的是“两条都占”——这句话看似谦虚，实则回避：说两条都占，等于既不承认哪一条承重，也不说怎么把它们分开。本章现在把它改成一句更难受、也更可用的话：

在本章的数据上，“有没有写追认规则”“该状态的自然段长分布”“内在标记强度”三者不可分离；R 这一格因此不能作为本构念的证据，只能作为它的来源。

把三者分开需要一个自变量有真实变异而其余两项被固定的设计。至少有两条可做：

其一，规则版本设计。取同一批原始记录，用不同版本的判读规则分别判读（R&K 一九六八年版、判读手册的各个版本、以及若干自动分期模型的单向与双向配置）。同一份信号意味着自然段长分布与内在标记强度都被固定，规则版本之间只有规则本身在变。这一条是第九节机制证伪的正式形态。

其二，同分期内的规则变异设计。在同一个分期内部找出手册规则有覆盖与无覆盖的两类边界（例如觉醒后回到 N2 与直接进入 N2），比较两类边界两侧的未追认段密度。这一条把段长分布也控制在同一分期内。

在这两条之一被跑出来之前，6.4 那句话在本章里的地位是**假说**，不是发现。摘要与结论都已按此改写。

6.6 本次跑不动的那一条

本章最想跑的那条跑不动：**同一份原始记录，在只看到第 n 窗与看到整夜之后，判定分别是什**么。这需要判读过程的中间态，而公开数据库只保存成品。现有记录里没有这个字段。

这一条不是遗憾，是本章的第三层主张，也是最便宜的一件可做之事：判读软件保存一份“首次判定”与“终判”的对照，成本接近于零，而它会第一次让追认延迟成为可直接测量的量。

6.7 这次真跑改了本章什么

四处，逐条列明：

1. 删掉了一条预测（过渡态更依赖追认），并把它写在正文里而不是脚注里。
2. 把构念的落点从生理挪到规则：未追认段的多少，可能是判读规则的函数，不是睡眠深浅的函数。这使本命题从一条关于睡眠的主张，变成一条关于任何带最小可记单位的记录制度的主张——适用范围扩大了，但也因此必须在第七节与更多相邻说法划清界线。
3. 把主张的强度调低：本章原本想说“未追认段是被丢掉的真实生理事件”。数据不支持这句话。数据支持的只是“账本单位与生理连续性不对齐”。这两句话差得很远，后一句弱得多，本章只主张后一句。
4. 新增 6.5：承认 R 那一格三因共线，把它从证据降为来源，并写出两条能把三者分开的设计。这一处是本稿相对前稿最实质的一次退让。

七·与最近的二十种说法的分界

以下每一条都写四件：它说到哪一步、分离线在哪、以及一条能分出“本章对还是它对”的判定形式。凡是只能写出“本章更强调”“更深入”“视角不同”的，等于承认被它覆盖，本章不保留这种条目。

排序按危险程度，不按学科。排在最前的三条，是本章前稿完全没有处理、而它们离本章最近的三位。

7.1 周期性交替模式 (CAP; Terzano et al., 1985; Atlas 2001; Parrino et al., 2012) ——本领域离本章最近的一位

它说到哪一步：这是睡眠医学自己给“三十秒格子毁掉的那些东西”起的名字，四十年文献，按一秒判读，A 相与 B 相分型，有独立的图谱与评分规则。它的出发点几乎与本章逐字相同：传统上那些短于判读窗的相位事件被当作散布在背景里的模糊瞬变，被并入 R&K 判据设定的时间单位（二十或三十秒）之中。二〇二五年的一篇综述把话说得更直白：因为 K 复合波出现在第十五秒就把一个三十秒判读窗判成 N1，这成了一道与现实脱节的练习题——睡眠过程并不在乎三十秒的格子。

本章前稿完全没有提到它。这不是一处遗漏，是一处错误：本章第七节前稿号称“与最近的十种说法分界”，而本领域最近的那一位没有上桌。此处补足，并且不打算把它写成一个次要邻居。

分离线在被数的东西不同。CAP 数的是相位事件，产出的是一个率（CAP 时间 / 睡眠时间）；它测量格子内部还有什么，并且把它测成一个新的时长比。本章数的是段落的存废，产出的是两本互相看不见的账；它不问格子里还有什么，只问格子造出来的那些边界，在哪一本账上算数。两者可以完全独立变化：一夜可以 CAP 率很高而未追认段很少（相位事件密集但都落在同一分期内部，不产生分期边界），也可以 CAP 率很低而未追认段很多（慢波占比在门槛上下缓慢漂移，不伴随任何相位事件）。

判定形式（且这是本章最危险的一条）：在同一批 PSG 记录上同时计算 CAP 率与未追认段比例。若两者在个体层面的相关系数超过 0.6，则本章测的就是 CAP 的另一种写法，本构念应当被 CAP 吸收，本章除了换一个名字之外没有增加任何东西；若两者解耦（相关系数低且各自与不同的规则变量相关），则两者测的是两件事。这一条今天就可以做——CAP 睡眠数据库是公开的。本章没有做，写在这里作为第一顺位待办。

顺带一句要说清楚的话：CAP 的存在反过来加强了本章的一个判断，也削掉了本章的另一个判断。加强的是：一个领域确实会为“格子毁掉的东西”造一整套并行账本——这正说明主账本不够用。削掉的是本章原想写的那句“至今没有一个领域正面处理过三十秒格子毁掉的东西”。它处理过，处理了四十年，只是处理的是格子内部的东西，不是格子制造的东西。

7.2 睡眠碎片化指数与觉醒指数（临床实务正主）

它说到哪一步：它已经在数“打断”了。每小时觉醒次数、觉醒指数、碎片化指数，是睡眠医学几十年的常规读数。

分离线：碎片化指数把一次半分钟的插曲记为一次打断（一个事件）；本章说这同一次插曲还同时消灭了一个段落——前后本是一段，现在成了三段，而这一层在碎片化指数里不出现。更要紧的是，本章数据里 74.9% 的未追认段前后同类，这些在碎片化指数里大多根本不算觉醒（不满三秒、或前面不足十秒稳定睡眠）。

判定形式：同一份记录，把 N2 与 N3 并类后，若碎片化指数基本不变而段落总数下降三成以上，则两个读数测的不是同一件事、本章对；若二者同步变化，则本章可被碎片化指数吸收。本章数据里并类后段数掉 35.1%，而并类不改变任何觉醒事件的判定——这一条已经站住了。

7.3 可变时间单元问题 (MTUP; Cheng & Adepeju, PLOS ONE 2014; Çoltekin et al., 2011; 上溯 Openshaw 的 MAUP, 1984)

它说到哪一步：这是地理学与时空分析对“单位是被造出来的”这件事最完整的一次形式化。与可变面积单元问题 (MAUP) 类比，MTUP 被定义为由三种时间效应构成——聚合、分段与边界；案例研究显示，被检出的时空聚类，其时长、空间范围与显著性，都随数据集的聚合、分段与边界改变而变化。

本章 6.4 的那个数正好落在它的射程里。把 N2 与 N3 并类之后段数掉 35.1%，这是一次教科书级的分段效应 (segmentation effect)；MTUP 会说这毫不意外，并且已经有一整套灵敏度分析方法在处

理它。本章前稿把这个数当作一个惊人的发现来报告，而它在另一个领域里是一个已被命名三十年的常规现象。

分离线在结论的形态。 MTUP 说的是：同一批被数的东西，在不同分段下会给出不同的统计量，因此结论对分段敏感，须做灵敏度分析。本章说的是更重的一句：在同一次分段之下，一样东西可以同时存在于一本账而不存在于另一本账。MTUP 处理的是一个量的不稳定，本章处理的是一个单位的有无。MTUP 的处方是报告多个分段下的结果；本章的处方是并行两本账，因为再多的分段灵敏度分析也仍然只是在时长账上算。

判定形式：对同一份记录做完整的 MTUP 式灵敏度分析（遍历判读窗长、起点相位与并类方案），报告未追认段比例的变化范围。若在所有分段方案下，未追认段比例与时长账读数同向同幅变化，则本章只是 MTUP 的一个应用案例；若存在一大片分段方案，其中时长账读数几乎不动而段落账读数大幅变化（本章的 N2/N3 并类正是这样一个点），则两账分岔是分段敏感性之外的独立现象。

7.4 序列分析（Abbott, 1995，社会学；且该文本本身用音乐类比）

它说到哪一步：把社会过程当作有次序的序列来分析，而不是当作变量的集合；Abbott 本人正是用音乐旋律来说明为什么次序不可还原为频次。

分离线：序列分析把序列的元素当作已给定的输入。本章问的是元素本身怎么被定出来的——在元素还没有被定下来之前，序列分析无从开始。

判定形式：若对同一份原始信号采用两套判读规则，得到的序列在序列分析的任何常用距离度量下差异不显著，则元素的定法不影响结论、本章这一层多余；若差异显著（本章数据里并类一次就掉 35% 的段），则序列分析必须先接受一层元素定义的审查。

7.5 右删失与事件史分析（统计学，方法学占位者）

它说到哪一步：观察窗内没有观察到结局的个案，有一整套成熟处理办法。这是本章在方法学上最省力的近邻——如果未追认段可以当作删失观测处理，本章这个概念就该被它取代。

分离线：删失观测有字段。它被明确标记为“已观察到此为止”，并进入模型。未追认段没有字段：它不是被标为未完成，是根本没有作为一个观测出现过；它的时间被并入相邻段落的时长里，不留任何痕迹。

判定形式：若能在现行判读报告里把未追认段作为删失观测正确取出并处理、而不修改判读规则，则本章被该概念覆盖；若取不出来（因为报告里只有各期时长与事件计数，没有段落表），则本章对。就本章所查的常规报告格式而言，取不出来。

7.6 回顾性意义建构（Weick, 1979/1995，组织理论）

它说到哪一步：人是先行动、后理解的；一件事的意义由行动者事后回望时建构，因而意义总是回溯的。这是本章最近的一位人文邻居，近到不划界就等于换皮。

分离线：Weick 的追认者是一个正在回望的解释者；本章的追认者是后面那一段事件本身，过程中没有解释者在场。睡眠里的判读、慢波的抖动、快速眼动期的到来——这些不是谁在理解什么，是规则与事件的接续。

判定形式：若在一条全自动、无人回望的判读流程里，追认延迟仍然大于零（即改动后会改动前段的判定），则本章对而回顾性意义建构覆盖不到这一格；若追认现象只在有人做事后叙述时出现、自动流程中不出现，则反之。这一条今天就可以做：拿一个自动分期模型，比较其单向（因果）模式与双向模式在同一记录上的判定差异。

7.7 文献方法（Garfinkel, 1967，常人方法学）

它说到哪一步：人们用后续材料去印证一个预设的底层模式，又用那个模式去解释后续材料，二者互相定义。

分离线：文献方法讲的是模式与个例互相印证，本章讲的是类别归属由后件单独决定；更要紧的是，文献方法预设参与者会感到某种解释上的张力，而本章的核心正是**没有张力**——三本账都记得平，没有人感到缺了什么。

判定形式：若把未追认段指出来之后，从业者报告“我早就觉得这里不对”，则本章的“无缺口”这一条错；若从业者报告的是“这从来不是一个问題”，则本章对。这一条可以用一次小规模从业者访谈直接判。

7.8 现在的哲学 (Mead, 1932)

它说到哪一步：过去不是固定的，是由每一个当下重新构造的。这是本章这条结构在哲学上最早、也最一般的表述。

分离线：Mead 给的是一条关于时间本身的一般命题，不带读数、不带反例、不区分“有规则的类”与“没规则的类”。本章给的是一个可数的时长、一条会被打脸的预测（并且确实被打脸了）、以及一条把差别归给规则有无而非归给时间本性的假说。

判定形式：若未追认段的比例在不同判读规则下不变，则本章的“规则决定”假说错，Mead 的一般命题足够；本章数据给出的方向一致读数（R 8.1% 对 N3 49.9%）因 6.5 所述的三因共线而不能充作证据，正式检验见 9.2。

7.9 通约化 (Espeland & Stevens, *Annual Review of Sociology* 24: 313-343, 1998)

它说到哪一步：这一条针对的不是本章的承重命题，而是本章第七节末尾要立的第二层三家共同踩着的那句话——“凡是一样东西，都可以按它占了多少来结算”。前稿在那里写了一句话：这是整片学术地基，所有人都站在它上面，因此谁也没法回头看它。**这句话不成立，必须撤回。**

通约化 (commensuration) 被定义为按共同度量比较不同实体。该文梳理了它作为社会思想工具的长久历史，把通约化分析为一种权力形态，并讨论了宣称某物不可通约所内含的认知与政治代价。量化社会学从一九九八年起一直在回头看这块地，二十多年来形成了一个有名字、有综述、有大量经验研究的领域。

分离线在被取消的东西不同。通约化研究处理的是：不同质的东西被压成同一把尺之后，**质的差别消失了**——一所大学的教学、研究与公共服务被压成一个排名分数，差别变成了大小。本章处理的是更前一步：在被压成同一把尺之前，**有没有成为一样可以上尺的东西**。未追认段不是被通约化压平的质，它是根本没有获得计数资格的时间；它不是在共同度量上取了一个不该取的值，而是没有取到任何值。用通约化的语言说：本章关心的不是不可通约者的命运，而是**未被立项者**的命运。

判定形式：取一份现行判读报告，问两个问题——(a) 有没有一个不同质的东西被压进了同一个度量？(b) 有没有一段时间既没进入这个度量、也没进入任何别的度量？若只有 (a) 有肯定答案，则本章可被通约化研究覆盖；若 (b) 也有肯定答案且该答案不能由 (a) 推出，则两者是两件事。本章的 3.35% 正是 (b) 的一个可计量实例：那些时间进入了时长账（有值），却没有进入段落账（无项）。

7.10 残余范畴 (Bowker & Star, *Sorting Things Out*, MIT Press, 1999)

它说到哪一步：任何分类系统都会产生落在所有类别之外的东西，而分类系统对这些东西的处理方式（塞进“其他”、忽略、或让它在系统之间被反复推来推去）本身有社会后果。这是比 7.5 的右删失更接近本章的一位：它讲的正是没有字段的那些东西。

分离线在有没有一个“其他”栏。残余范畴的核心形态是**存在一个收容处**——“其他”“未指定”“待定”是一个格子，东西掉进去，仍然被计数一次。本章的未追认段没有收容处：它不进“其他”，它的时间被并入了相邻那一段的时长，因此在段落账上连一个“其他”都不占。用一句话分开：残余范畴是被归进了一个坏格子，未追认段是没有进入任何格子而它的时间被别人领走了。

判定形式：在判读报告里找有没有一个“未分类段落”计数。若有，本章的现象就是残余范畴的一个实例，本章的贡献退化为给一个已知残余范畴命名；若没有（就本章所查的常规报告格式而言没有），则两者不同。

7.11 身体技术 (Mauss, 1934, 人类学; 社会学侧最老的一位)

它说到哪一步: 睡眠是被社会训练出来的身体技术之一——睡姿、睡具、与谁同睡、睡多久, 都是习得的。这是“睡眠被社会安排”这一整条线最早的命名。

分离线: Mauss 讲的是怎么睡被社会规定; 本章讲的是哪一段算睡由后件与规则决定。两者可以完全独立发生: 一个社会可以在睡姿上高度统一, 而其判读规则完全不产生未追认段; 反之亦然。

判定形式: 跨文化比较中, 睡眠实践差异极大而判读规则相同的两个人群, 若未追认段比例接近, 则本章对 (该现象由规则而非文化决定); 若随文化显著变化, 则本章这一条应并入身体技术。

7.12 完美睡眠强迫 (orthosomnia; Baron et al., 2017)

它说到哪一步: 可穿戴设备的睡眠读数会反过来改变使用者的行为与焦虑, 读数本身成为病因。

分离线: 那是读数改变被读者; 本章是规则决定被读之物有没有单位。前者的机制经过人的注意与情绪, 后者不经过任何人。

判定形式: 在不向被记录者反馈任何读数的纯研究场景里, 未追认段照样出现 (本章数据即是如此, 受试者从未看到自己的分期图), 则两者不同源。

7.13 自家的两处: 垫付与回路

站上此前两部相关著作, 与本章结构最近, 必须逐一指名。

《健康发生学导论》的垫付线说的是: 身体的每一次正常都由下一轮的接手能力垫付, 而垫付越成功越不产生信号。**分离线:** 那一条讲储备被消耗而无信号, 本章讲单位不存在而账记平。二者可各自独立发生——一个储备充裕的夜晚照样可以产生大量未追认段 (本章数据里未追认段与总睡眠时间无明显关系), 一个储备耗尽的夜晚也可以段落清晰。**判定形式:** 若未追认段比例与任何一项储备类读数相关系数超过 0.6, 两条应合并。

《语言的智慧》的回路线说的是: 一次表达关上之后, 还留不留得住回来的路。**分离线:** 那一条问闭合之后有没有返回通道, 本章问这一段有没有被后件定成一个段落——返回是可选的, 追认是构成性的: 没有返回, 那一段仍是一段; 没有追认, 那一段不是一段。

7.14 自家的另两处: 失眠专文与睡眠医学面板

站上那篇《失眠: 为什么越想睡越睡不着》讲的是入睡这一刻的努力悖论——意志监督入睡, 而入睡要求意志退场。本章完全不碰入睡这一刻: 它讲的是整夜记录里段落如何成立, 与入睡与否无关; 未追认段在睡得极好的夜里同样大量出现。

站上那份睡眠医学面板整理了近二十年二十项转向, 其第一幕开篇一条正是认知行为治疗把慢性失眠改写为可学习的失调。本章与它的关系是补一层: 那二十条讨论的都是“哪个读数有资格代表睡得好”, 而本章问的是更靠前的一步——这些读数所依赖的段落单位, 本身是怎么来的。

7.15 自家最近的四处: 此前未申报的站内自撞

本节是本稿相对前稿新增的一整块。前稿只申报了 7.13 与 7.14 那四处, 而用站内近邻库实扫之后发现, 离本章最近的四篇全在本站的跨学科栏与每日必读栏里, 前稿一处未提。不申报即自撞, 逐条如下。

(一) 《临界回落》(跨学科, /confluence/critical-recrossing/)——最危险的一处。该文写的是: 一个学生已经离开了原来那个安稳的不懂、还没有到达懂, 三十秒后他回到原处; 在这堂课留下的全部记录里, 这三十秒不存在——他与从头到尾在想中午吃什么的同学, 在所有的账上是同一笔。该文明写自己要做的事是“把这一类事件从不存在变成一个可以被清点对象”。

同一承重判断、同一个“三十秒”, 只是域从课堂换到了睡眠。**分离线**只能划在追认者的性质上: 《临界回落》里那一段之所以不存在, 是因为没有承接结构——它需要的是一个阻止关闭的外力, 追认者是一个可以缺席的他者; 本章里那一段之所以不存在, 是因为规则没有为它写下定类条款, 追认者是后面那一段事件本身, 不需要任何人在场。前者的处方是造一个承接结构, 后者的处方是公布

规则；前者缺的是人，后者缺的是条文。**判定形式：**在一条全自动、无人参与的记录流程里，若《临界回落》那类事件不再出现而本章的未追认段照样出现，则两者是两件事；若两者同时消失，则本章是《临界回落》在睡眠上的一次应用。

(二)《双链型》(跨学科, /confluence/two-chain-type/)。该文论证：一个反复出现的形态被当成一个“型”，依据是它的重复；而重复可以由两条彼此独立的链产生——制作链把动作沉淀成形态，判读链把形态归进同一格——因此重复不能区分一个型是被做出来的还是被读出来的。

本章 6.4 那句“35.1% 的段落是判读规则制造的，不是生理事件制造的”，就是《双链型》的判断在睡眠上的一次兑现。这一处必须坦白：本章那一句在一般形式上不是新的，它的主人是《双链型》。本章剩下的增量是两条：其一，《双链型》讲的是两条链都产生同一个型因而不可区分，本章讲的是判读链单独产生了型而制作链那一侧什么也没发生（连续过程被阈值抖动切开）；其二，《双链型》没有给出一个把两条链的产物分开清点的读数，本章给出了——并类前后的段数差，就是判读链单独贡献的那一部分。**判定形式：**若在任何一个域里，判读链单独贡献的段落数无法与制作链贡献的分开清点，则本章这一增量不成立。

(三)《同缩》(跨学科, /confluence/same-shrinkage/)。该文写的是：同一组向好的读数背后可以站着两件完全不同的事，其中一件是活动本可发起而没有发起，并且连“没有发起”这件事也没有在任何账本上留下条目；一家机房、一间病区、一座城市的窗口，三份都算得平的年报在同一个地方一起失灵。

“三本账都记得平”是同一个句式，也是同一个机制。**分离线：**《同缩》处理的是没有发生的事（活动未被发起），本章处理的是发生了但没有成为单位的事（时间确实过去了，且被计入了时长）。前者在物理上是空的，后者在物理上是满的、只在账上是空的。这条线很细但可判：**判定形式：**未追认段的那 726 分钟能不能在时长账上被找到？能（3.35%）。《同缩》那一类的时间能不能在任何账上被找到？不能。这一问即可分开两者。

(四)《账外的东西》(每日必读, /paradigm/off-ledger/)。该文的落点是：每一次记账，你都在同时得到频次、失去分量。

“频次 vs 分量”与本章的“段数账 vs 时长账”结构同型，但取向相反。《账外的东西》说的是记账这个动作本身会改变被记录的东西，因而是一个关于记账的代价的命题；本章说的是两本账可以同时记平而互不相知，因而是一个关于账与账之间的命题。**判定形式：**若在一份记录里取消记账动作，本章的现象消失，则本章属于《账外的东西》；若取消不了（判读本身就是记账，没有不记账的判读），则本章处理的是记账内部的分账问题，而非记账与不记账之别。

(五)更远但同族，一并列名以备后续划界：《一件是几件》（计数单位失去下界）、《两种“不”》（两件不同的事在所有记录上同码）、《记零位》（被记为零的那类动作）、《类内距》（规范写了边界不写同类间距离）。这四篇与本章共享“单位与账本”这一族问题，但各自的自变量不同，本章不在此逐条展开。

7.16 划完之后：所有这些邻居共同站着的那块地

逐个问上面每一位“它结构性地看不见什么”，句式写死为“某概念把某样东西当作给定，因此它看不见某样东西”，且必须说出“一旦承认后者，它自己会怎么塌”：

- CAP 把分期边界当作给定（它在既定分期内部数相位事件），因此看不见分期边界本身是被造出来的；一旦承认，CAP 率的分母（分期时间）就不再是一个中立的基准。
- 碎片化指数把事件当作给定，因此看不见段落的消灭；一旦承认段落也会消灭，它的分母（每小时）就不再是一个中立的时间基准。
- MTUP 把“被数的东西”当作给定（它研究同一批东西在不同分段下的统计量变化），因此看不见一样东西可以在一本账上存在而在另一本账上不存在；一旦承认，灵敏度分析就不能只报告统计量的范围，还要报告单位的有无。
- 右删失把观测当作给定，因此看不见没有成为观测的时间；一旦承认，它的样本定义就要重写。

- 序列分析**把元素当作给定，因此看不见元素的生成规则；一旦承认，它所有的距离度量都要先做规则敏感性检验。
- 通约化**把“已经是一样东西”当作给定（它研究不同质的东西被压上同一把尺），因此看不见还没有成为一样东西的那一段；一旦承认，它的批判对象就要从“尺”往前推到“立项”。
- 残余范畴**把“有一个收容处”当作给定，因此看不见连收容处都没有的那一类；一旦承认，它的分类图里要多出一个不是格子的空位。
- 各期时长与睡眠效率**把时长当作给定的通用货币，因此看不见段落这个单位；一旦承认，它们就不能再自称完整描述了一夜。

这几行背后是同样被当作给定的东西：

凡是一样东西，都可以按它占了多少来结算。

前稿在此写了一句错话：说这是整片学术地基，所有人都站在它上面，因此谁也没法回头看它。7.9 已经撤回了这句话——量化社会学从一九九八年起就一直在回头看这块地，并且看得比本章细。

本章与它的差别不在**有没有人看**，而在**看的位置**。通约化研究站在尺的那一端，问的是：被压上这把尺的东西，失去了什么质。本章站在尺之前，问的是：什么东西根本没能上尺。本章的未追认段，正是从这条地基的缝里掉下去的那一类：它占了时间（3.35%），所以在时长账上被记平了；它没有成为段落（32.4%），而**没有一本账是段落的账**。

因此本章真正主张为新的那一格，收窄为一句：

同一份记录上，段落账与时长账同时成立而互相看不见；而现行制度只生产后者。

这一句比前稿弱，也比前稿准。前稿想主张的“单位是被规则造出来的”，主人是《双链型》与 MTUP；前稿想主张的“按占多少来结算是一片没人看的地基”，主人是通约化研究。剩下无主的，是这两本账在同一份记录上同时成立、互不相知、而只有一本被生产——以及一对可以把它量出来的数：32.4% 与 3.35%。

八·四家反过来给本章加的约束

一个跨学科命题若从三个领域只取好处，那是取用，不是对话。以下四处是各家反过来削掉本章原本想说的话，每一处都写明“如果没有它，本章原会写下哪一句更强的话”。第四处是本稿新增的。

8.1 化学削掉的：未追认段不是“被丢掉的事件”

本章原本想写的更强的一句是：**未追认段是真实发生的生理事件，被记录规则丢掉了。**

化学不允许这么写。磷酸化读数是连续量，它在慢波占比跨过判读门槛的那一瞬并不发生任何对应的跳变；分子层面根本没有一个“事件”与那个门槛对齐。也就是说，未追认段在生理上多半不是一个事件，而是一段连续过程的任意切点。

削掉之后本章只能主张：**账本单位与生理连续性不对齐**。这弱得多，但它是数据支持的那一句。

8.2 音乐学削掉的：不是所有身份都靠后续追认

本章原本想写的更强的一句是：**一段时间的身份，全部由后续事件决定。**

音乐学不允许。形式功能理论同时坚持存在“内在形式功能性”——把乐句片段抽离原境后，音乐家与非音乐家都能高于随机地判断它是开头、中段还是结尾。也就是说，相当一部分身份是由段落自身的内容标记的，不需要后件。

这一处动的是本章的适用范围，不是补充条件：**本命题只在“内在标记不足以定类”的那一部分上成立**。落到睡眠数据上，这正是 6.5 那三因共线中的第三因：快速眼动期有极强的内在标记，因此它的未追认段少这件事，音乐学这一侧给出的解释与“手册写了规则”这一侧同样充分。本章不再声称能在现有数据上把两者分开。

8.3 社会学削掉的：本章不主张“应该改成段落记账”

本章原本想写的更强的一句是：判读规则应当改为记录段落。

社会学不允许。八小时连续睡眠这套单位之所以成为标准，本身就是一次“把某个单位定为正当”的历史过程，其后果是别的睡法被病理化。本章若主张“应该改记段落”，等于再造一次同样的动作，而且很可能同样会把不合新单位的人判成异常。

削掉之后，本章的处方从“改规则”缩到一条纯程序性的要求：凡是使用段落读数的地方，须公布它的追认规则与最小可记单位。不主张哪一套单位更对，只主张单位必须是明写的。这一处动的是价值排序，也是本章最实质的一次让步。

8.4 睡眠医学自己削掉的：格子毁掉的东西，早有人处理，只是处理的是另一半

本章原本想写的更强的一句是：三十秒格子毁掉的那些东西，至今没有一个领域正面处理过。

睡眠医学自己不允许。CAP 处理了四十年，有图谱、有评分规则、有独立的公开数据库。本章前稿之所以能写下那句话，只是因为它没有查到 CAP——而这恰恰是本章自己命题的一个实例：一件东西在本章的账上没有一行，不是因为它不在，是因为本章的检索规则没有为它写下条款。

削掉之后本章只能主张一句更窄的话：格子内部的东西有人处理（CAP），格子制造出来的东西没有人处理（段落账）；而且这句话必须以 7.1 那条相关系数检验为条件——若两者相关，这句话也不成立。

九·证伪、赌注与边界

9.1 最强的竞争解释

那句最朴素、最难反驳的话是：其实不就是测量噪声吗？

判读是人做的，人会抖；三十秒一格是个粗糙的切法，切出一堆碎屑很正常。这个解释能覆盖本章的大部分现象，且不需要引入任何新东西。本章不挑一个好打的靶子，就打这一个。

现在还要加上第二个竞争解释，来自 7.1：其实不就是 CAP 吗？相位事件密集的地方，判读自然抖得厉害，碎屑自然多。这个解释比噪声解释更强，因为它自带机制。

9.2 机制证伪（必须同时排除噪声解释与 CAP 解释）

噪声解释预测：碎屑的分布应当在各分期之间大体均匀，或至少与该分期的判读一致性（判读者间信度）成反比——判得越难的地方碎屑越多。它预测不出“有没有写追认规则”这个变量起作用。

CAP 解释预测：未追认段比例应当随 CAP 率同向变化，且在 CAP 率被控制之后，规则变量不再有额外解释力。

于是本章的证伪条件写成：

控制判读者间信度、该分期的边界模糊度与 CAP 率之后，若“手册是否为该分期写了追认规则”与未追认段比例之间不存在剂量—反应关系，则本机制为伪——届时这些碎屑应归因于判读噪声或相位事件密度，而非账本规则。

本章数据给出的是一个方向一致的初步读数（有规则的 R 为 8.1%，无规则的 N3 为 49.9%），但**样本纪律不允许把它当作证伪检验**：分期只有五个，且它们在边界模糊度、自然段长分布与内在标记强度上都不同质（见 6.5）。要做成检验，需要至少六个同质单元上自变量有真实变异。可行的做法是 6.5 所述的规则版本设计：取同一批原始记录，用六套以上判读规则版本分别判读——R&K 一九六八年版、判读手册各版本、若干自动分期模型的单向与双向配置。规则版本之间在数据上完全同质，差别只在规则本身。这一条本章没有做，写在这里作为待办，不充作已完成的检验。

9.3 其余证伪条件

1. 若在完全自动、无人回望的判读流程中，单向与双向模式给出的分期序列无显著差异，则追认结构不存在于自动流程中，本章 7.6 的分界不成立。

2. 若把最小可记单位从三十秒改为五秒后，未追认段比例下降而不是上升，则“最小可记单位制造段落”这一机制反向，本章错。（现有的五秒微窗研究报告分期转换显著增多，方向与本章一致，但那不是本章做的，不能算本章的检验。）
3. 若未追认段比例与任一项现行储备类读数相关系数超过 0.6，本章应并入《健康发生学导论》那一条。
4. 若从业者在被指出未追认段之后普遍报告“我早就觉得这里不对”，则本章“账记得平、无人察觉”这一条错。
5. 若判读软件保存首判与终判之后，实测追认延迟的中位数为零，则本章的核心构念在睡眠这一领域上不成立（在音乐分析上仍可能成立）。
6. 若在音乐形式分析中，回溯改判的发生率低于百分之五，则本章所依赖的那件推翻材料太罕见，不足以承担三家共同踩着的那句话的推翻。
7. **（新增，第一顺位）** 若在同一批 PSG 记录上，个体层面的 CAP 率与未追认段比例相关系数超过 0.6，则本章测的是 CAP 的另一种写法，本构念应被 CAP 吸收，本章除换名之外没有增加任何东西。
8. **（新增）** 若完整的 MTUP 式灵敏度分析显示，在所有分段方案下段落账读数与时长账读数同向同幅变化，则“两账分岔”不是一个独立现象，只是分段敏感性的一个侧面，本章的收窄主张也不成立。
9. **（新增）** 若在任何一份现行判读报告格式里存在“未分类段落”计数字段，则未追认段是一个已被承认的残余范畴，本章的贡献退化为给它换一个名字。

若本章被证伪，学到的是什么：如果这些碎屑确实只是噪声或只是 CAP 的影子，那么当前一切段落类读数都可以放心使用，而本章所要求的“公布追认规则”就是多余的负担。这是一个有价值的结论——它会把注意力送回到读数的精度上，而不是读数的单位上。

9.4 一条写死日期的赌注

到二〇三〇年十二月三十一日为止，将会有至少一家主流判读软件或一部判读手册修订版，明确写入“最小可记单位”与“后件定类规则”的完整清单，并把它作为报告的必备字段。

什么不算命中（写死，事后不改）：

- 只在学术论文里讨论、未进入手册或软件的默认报告，不算。
- 只公布最小单位而不公布后件定类规则，不算——两者必须齐。
- 由本章作者或其合作者推动写入的，不算。
- 以“已在附录中说明”为有的既有条文，不算——须为新增的必备字段。
- **（新增）** CAP 评分规则的既有条款不算——那是格子内部的规则，不是格子制造边界的规则。

9.5 反噬预言

本命题一旦被制度采用，最省事的实现方式恰恰会摧毁它要保护的东西：把追认延迟做成一个考核指标。一旦某个机构被要求“追认延迟越短越好”，最省力的达标办法是把最小可记单位调大——单位越粗，碎屑越少，延迟越短，而被吞掉的时间更多。指标会变得非常好看，而它本来要显形的那件事会彻底消失。

我看不到出口。任何一个能被清点的量，都可能被拿去考核；而本章提出的这个量在被考核时的最优作弊策略，恰好是加重它所诊断的病。本章能做的只有把这条写在正文里，以及下面三条硬约束。

9.6 伦理硬约束（三条，缺一条本章不可发表）

1. 追认延迟指的是规则，不是人。它是一套判读规程的属性，不是判读者的属性，不得用于评价任何判读人员或任何患者。

2. 不得为了把这个数做好看而调粗最小可记单位。任何对最小可记单位的调整，须与该调整对总段数的影响一并公布。
3. 已经依据段落类读数对任何人做出过不利安排的机构，须公开其判读规则与最小可记单位并提供复核通道。这一条针对的是把睡眠读数用于劳动能力评估、保险核定与驾驶资格判定的场合。

9.7 本章不适用的地方

- 不适用于没有最小可记单位的连续记录——那里不存在碎屑段。
- 不适用于诊断。未追认段的多少不指示任何疾病，本章数据中它与任何临床变量的关系一律未检验。
- 不适用于主张某种睡法更好。本章对“一夜该睡几段”不持立场。
- 不适用于跨规则的横向比较。追认延迟只同一套规则内可比。
- 不适用于替代现行读数。本章主张段落账与时长账并行，不主张取消后者。
- (新增) 不适用于宣称本章发现了“三十秒格子毁掉东西”这件事。那件事由 CAP 传统在四十年前指出并持续处理；本章只主张格子制造出来的边界另有一本账，且那本账没有人在记。

十·本章自己落在哪里

按本章自己的判据检本章自己：改动最后一节，会不会改动前面某一节的判定？会。而且已经发生过两次。

第一次是第六节那条被打脸的预测，改掉了第五节对构念的界定——本章的第五节在第六节跑完之前是另一个样子。按本章自己的用语，第五节曾经是一段未追认段：它当时的类别归属尚未由后件定下来。

第二次发生在本稿。第七节新写的三条全球分界，回头改掉了第五节、第六节与第七节末尾的三处判定：R 那一格从证据降为来源（6.5），第二层三家共同踩着的那句话那句“谁也没法回头看它”被撤回（7.9、7.16），而本章主张为新的那一格被收窄成了两本账的分岔（7.16）。一篇文章的第七节改掉了它自己的第六节，这正是本章所描述的那个结构在本章自身上的一次运行。

还有一层要认账，而且这一层比前两层更难堪。

本章前稿写下过一句话：三十秒格子毁掉的那些东西，至今没有一个领域正面处理过。这句话是错的，而且错得很典型——睡眠医学自己处理了四十年，有图谱、有评分规则、有公开数据库。本章之所以能写下那句话，不是因为 CAP 不在，是因为本章的检索规则里没有为它写下条款。一件东西在一份记录上没有一行，不是因为它不存在，是因为记录的规则没有为它定类——这恰恰是本章的承重命题。本章在自己的账上，为自己的命题提供了一个实例。

同样的话适用于本章的数据。本章用一个提出于音乐学、检验于睡眠数据、约束于社会学与睡眠医学的框架，去指出别人的账本上有缺口；而本章自己的账本——四十夜、一个数据库、一套判读传统——同样有一个最小可记单位：不进入公开数据库的睡眠记录，本章一条也没有看到。家庭里的睡、值夜班间隙的睡、没钱做多导记录的人的睡，在本章的数据里不构成一样东西。本章对它们的判定，与判读手册对那三十秒的判定，是同一个动作。

更要紧的是本章这条产线自身。本章是一批同题文章中的第一篇；后面八篇会写什么，会反过来决定这一篇算什么——是这套框架的开篇，还是被后来的分析并掉的一次试手。本章现在没有办法知道自己属于哪一类，而这正是本章所描述的那个结构。

最后一句留给那件跑不动的事。本章最想测的量，在现有记录里没有字段。本章可以被引用的最稳的一层，不是那个新名字，也不是那张表，是这一句：把首次判定与终判一并保存，成本接近于零。前两层可以塌，这一层塌不了。

注释与材料

本章使用的公开数据为 Sleep-EDF Expanded 数据库 1.0.0 版 sleep-cassette 部分，按文件名排序的前四十份判读标注文件，未做任何挑选。预注册文本、解析与统计脚本随文提供，任何人可复算。判读

规则的引述依据美国睡眠医学会判读手册的公开条文（判读窗三十秒与占优判定、觉醒须持续至少三十秒且其前有至少十秒稳定睡眠、快速眼动期延续规则、周期性肢体运动伴随觉醒间隔不足十秒只记第一次）。

CAP的相关数据（CAP Sleep Database）同样公开可取，7.1与9.3第7条所述的相关性检验因此不需要任何新的数据采集；本章未做，原因是它超出了本稿的预注册范围，事后补跑会破坏预注册纪律。该检验应作为独立的一次预注册研究执行。

修订说明（本稿相对前稿）：一、摘要中的构念表述由因果断言改为待检验假说。二、新增5.5后半，承认“论证读数应为时长而实际报告的全是比率”这一处自相矛盾，并区分存量读数与延迟读数。三、新增6.5，承认R那一格三因共线，把它由证据降为来源，并给出两条分离设计。四、第七节由十条扩为二十条，新增CAP、MTUP、通约化、残余范畴四条全球分界，以及四条此前未申报的站内自撞。五、新增8.4，睡眠医学反过来削掉本章一句。六、9.3新增三条证伪条件，9.4赌注新增一条不算命中的条款，9.7新增一条不适用范围。七、7.16撤回前稿一句错话，并把本章主张为新的那一格收窄为“两本账在同一份记录上同时成立而互相看不见”。八、第十节新增一段自反认账：本章前稿漏掉CAP四十年，这件事本身是本章命题的一个实例。

人机分工声明：本章的选源、材料核对、统计脚本与初稿由人机协作完成；预注册文本在取数之后、跑任何统计之前写定，未曾修改；不利结果按预注册规则原样报告。本稿的全球近邻检索与四条新增分界同样由人机协作完成，所引文献均经逐条核验。

参考文献

6. Kramer, J. D. (1988). *The Time of Music*. Schirmer.
10. Ekirch, A. R. (2005). *At Day's Close: Night in Times Past*. W. W. Norton.
12. Mead, G. H. (1932). *The Philosophy of the Present*. Open Court.
13. Garfinkel, H. (1967). *Studies in Ethnomethodology*. Prentice-Hall.
14. Weick, K. E. (1995). *Sensemaking in Organizations*. Sage.
19. Berry, R. B., et al. *The AASM Manual for the Scoring of Sleep and Associated Events*. American Academy of Sleep Medicine.（判读窗、觉醒规则、快速眼动期延续规则条文）
20. Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberyé, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194.（Sleep-EDF 数据库来源）

第二章 住在线上

住在线上

论睡与醒之间那条分界其实是一片有厚度的区域，以及它做的事为什么全被记在两侧名下

Living on the Line: The Sleep-Wake Terminator as a Region, Not a Boundary

提 要

关于“何谓睡眠”，三门互不相干的学问给出三种把睡与醒分开的办法。市场营销学说，一样东西是什么，取决于它在心智里被归入哪个类；有机化学说取决于内部构型与转换能垒；行星科学说取决于它相对于外部周期处在哪个相位。三家争的是这条分界该由什么来定，而三家共同假定了一件从未说出口的事：**划分由一条分界完成，而分界本身不占位置、不承担功能。**

本章用行星科学自己的材料推翻它——在潮汐锁定的行星上，昼夜之间那条晨昏带是整颗地球上唯一可能宜居的区域，两侧一个过热、一个封冻。由此提出：**睡与醒之间不是一条线，是一片有厚度、承担功能、而在三家账本上都只被记为分界的区域；它做的事，全部被记在两侧的名下。**文章把这片区域命名为**界带**，把“把区域记成线”这个动作命名为**线化**，把“产出被记在邻居名下”命名为**邻记**。

四十夜公开判读记录的预注册检验给出三条命中：相邻转换间隔在短端相对几何期望超额 3.31 倍；粗化八倍后实测混合窗比例比独立点模型低 30.4%；多转换簇以约一成的时间承载 75.3% 的转换。第三条预测未命中，据此把命题改为约一半的分界是线、另一半是带。

本稿新增一次事后（非预注册）的零模型更正，它拆掉了上述三条中的一条半。原检验的零模型是齐次独立点过程，而睡眠转换率明显非齐次。改用分期条件的混合几何零模型后，短端超额由 3.31 倍降到 2.23 倍；改用“段长夜内重排”零模型后，**混合窗那条读数的全部偏差都被段长分布本身解释掉，实测甚至略低于重排**（ $k=8$ 时相差 -0.43 个百分点）；而成簇那条只剩一小截：实测 75.3%、重排 $71.8\pm0.4\%$ ， $z=+9.5$ ——方向为正、统计上极显著，但幅度只有 3.5 个百分点。**也就是说：本章原先报告的成簇，约九分之十八是段长分布的后果，而段长分布是睡眠动力学二十年前就刻画过的对象。**

据此本章的经验主张收窄为两句：其一，超出段长分布的时间成簇确实存在，但很小；其二，本章真正无主的那一格不在统计里，在账本里——**邻记**，以及一条纯算术：判读窗三十秒加占优规则，使一个十五秒即可改变认知产出的状态不保证能被记成任何一个窗。文章与生态交错带、睡眠转换动力学、可变时间单元问题三处全球近邻逐一划界，并申报三处此前未申报的站内自撞。文末给出十四条判决性分界、四处真削弱本章的反向约束与三条伦理硬约束。

关键词：界带 | 线化 | 邻记 | 段长分布 | 零模型的分层 | 账本上没有的第三栏

一·一条线上住着人

先讲一件与睡眠无关的事，因为本章全部的力气都从它来。

有一类行星被它的恒星潮汐锁定：一面永远朝向恒星，一面永远背对。朝向的那一面被烤到远超任何生命能忍受的温度，背对的那一面终年封冻，可能盖着巨大的冰川。教科书上画这类行星，通常画成一个眼球——中间一个亮斑，四周一圈暗，中间那圈交界叫晨昏线。

二〇二三年三月，一组行星科学家在《天体物理杂志》上给出一个结论：这类行星上**唯一可能宜居的地方，正是那条晨昏带**。他们用改造过的地球气候模式做模拟，发现在陆地占比高、水量有限的条件下，晨昏带可以维持液态水与温和气候；而水一多反而更糟——日侧的海洋会把水汽灌满整个大气，把整颗星球闷成温室（Lobo et al., *The Astrophysical Journal* 945: 161, 2023）。

请注意这件事的形状。在我们所有的图示里，晨昏线是**一条线**——它是用来把日侧和夜侧分开的记号，本身不占地方。而这项工作说：它占地方，它有宽度，而且它是那颗星球上唯一有东西能发生的地方。**两侧不是它的主体，它才是；两侧只是把它夹出来的条件。**

一条被画成分界的线，其实是唯一能住人的地方——本章要问的是：我们身上有没有这样的地方？

有。而且就在每晚都要经过的那个位置。

二〇二一年十二月，一组睡眠研究者在《科学进展》上报告了一次实验：一百零三名受试者做一批冗长的算术题，题里藏着一条能让人瞬间解出的规则，但没人被告知。做完一批之后，让他们闭眼休息二十分钟，手里握一个物件——这是从爱迪生那里学来的打盹法：一旦睡深，手一松，物件落地把人惊醒。研究者用脑电监测这二十分钟里每个人到了哪一层。

结果是：在非快速眼动第一期停留至少十五秒的人，发现那条隐藏规则的比例是 83%；始终保持清醒的人是 30%。而如果睡得更深，这个效应消失。论文的作者把那个状态称作睡与醒之间的“twilight zone”——曙暮之带（Lacaux et al., *Science Advances* 7: eabj5866, 2021）。

两侧都不行：完全清醒不行，睡深了也不行。只有那条带行。这与晨昏带的形状是同一个形状。

而在标准的睡眠判读里，那条带没有自己的位置。它被切在三十秒一格的判读窗里，按占优规则并入相邻的某一期；十五秒——那个足以让洞察率翻三倍的时长——小于判读窗宽度的一半，因此在占优判定下不保证能被记成任何一个窗。这不是推论，这是算术。

本章要立的东西就是从这里开始的。读到第七节会看到，本章的统计部分在事后被自己拆掉了一大半，而上面这条算术不受影响——它不依赖任何零模型。

二·三种把分界定住的办法

要说清“分界”出了什么问题，得先看看今天最有力的三种定分界的办法。它们分别来自三门彼此不读对方文献的学问，而三种办法不能同时都对。

2.1 市场营销学：分界由心智的归类划定

当代营销学最大的一场争论，是品牌靠什么活着——靠差异化，还是靠被想起来。争论的一方把重心整个挪到了后者：一个品牌的成败，取决于它在多少种“进入场景”里会被消费者想起，以及被想起时它被放进哪一个比较集。同一瓶水，在健身房和在餐桌上属于不同的品类，跟不同的东西比较，被不同的标准评价。

这一家的单因主张是：一样东西是什么，取决于人在什么时刻、为了什么需要去想起它；产品的物理属性只是被归类的材料，归类才是决定性的。分界因此是心智里的分界：睡与醒之所以是两样东西，是因为人把它们分别归入了两个类，并在不同的场合调用。

它的硬支撑是可核的：品牌的购买频次与渗透率呈稳定的经验规律，小品牌不仅买的人少、买的人还买得少；而“品类”的边界在不同使用场合下会重画，这是被大量场合数据反复验证过的。

它自曝的地方也很硬：品类份额必须加总到一〇〇%。这不是疏忽，是核算恒等式的要求——一个不属于任何品类的使用场合，在份额账上无处安放。营销学自己知道这一点，并把它当作方法的前提，不当作问题。

2.2 有机化学：分界由内部能垒定住

化学处理“两样东西是不是同一样东西”，比任何学科都久、都细。它的答案是构型与构象之分：改变构型要断键，改变构象只要转一转。同一个分子的两个构象，是同一样东西的两副样子；两个构型，是两样东西。

分界画在哪儿？画在转动能垒上。一九八三年，大木道则给阻转异构下了一个沿用至今的定义：如果两个构象在给定温度下互相转化的半衰期超过一千秒，它们就算作可以分离的、各自独立的化合物（Oki, *Topics in Stereochemistry* 14: 1-82, 1983）；这对应室温下约九十三千焦每摩尔的能垒。低于这个数，是一样东西；高于这个数，是两样。

这一家的单因主张是：一样东西是什么，由它内部构件的排布与相互转换的能垒决定；外部怎么看它、放在什么参照集里比，都不改变它是什么。

它自曝的地方，是本章用得上的第二件材料，而且化学界写得非常坦白。埃利尔等人在标准教科书里评论大木的这条定义时说：这个定义完全是任意的，但它便利，而且如果还想保住阻转异构这个概念，它就是必需的。也就是说——化学自己在印刷品上承认：那条把“一样东西”与“两样东西”分开的线，是为了保住概念而划的，不是被发现的。

2.3 行星科学：分界由外部周期的相位定住

第三家把前两家的对象都放回天上。一个星球上的昼与夜，既不是心智的归类，也不是内部构型，而是自转、公转与恒星位置共同决定的相位。这一家最锋利之处在于它能举出使前两家失灵的例子：金星的自转周期两百四十三天，公转两百二十五天，而它的太阳日只有一百一十七天——三个数，没有一个可以被单独叫作“金星的一天”。潮汐锁定的行星干脆没有昼夜交替。地球的日长在变，古生代的一年有四百多天。

这一家的单因主张是：一样东西处在什么状态，由它所纠缠的外部周期构型决定；心智与内部结构都是在这个条件场里被造出来的。

它自曝的地方，就是第一节那件事：晨昏带宜居性。行星科学自己已经把那条线打开成了区域，而且给出了这个区域的宽度、气候与含水条件。它在自己领域里回答的问题是“到哪儿去找生命”，从没有人把它拿来当作关于“分界”的一般证据。

三·为什么这三种办法不能同时打

把三家并排，很容易读成三个层次各说各话。要挡掉这种读法，得看它们各自把什么当成了单独够用的那一样：

	单独够用的那一样	它据此把另两样看成
市场营销学	它在心智里被归入哪个类	内部结构是归类的材料；外部周期是使用场合的背景
有机化学	内部构件的排布与转换能垒	归类是外行的说法；外部周期只是改变速率的温度
行星科学	它所纠缠的外部周期构型	心智归类与内部结构都是条件场的产物

三处的分歧落在三根不同的轴上，因此判谁对这条路走不通。更要紧的是它们两两之间是真打架，不是分工：

营销学与化学打的是“参照集改不改变身份”。营销学每天在做的事——重画品类边界、换一个竞争集、把同一件产品放进另一个比较框——正是化学诊断为“你没有改变这个分子，你只是换了个溶剂”。一家的日课是另一家的病理。

化学与行星科学打的是时标的归属。化学说是不是同样东西，取决于内禀能垒与观察时间的比值；行星科学说取决于外部周期。同一个现象，一个归给内禀成因，一个归给时序构型，二者不能同真。

营销学与行星科学打的是谁解释谁。营销学说“白天”与“夜晚”是被使用场合切出来的类；行星科学说使用场合本身是自转造出来的。因果箭头正好掉头。

顺带挡掉一个总想混进来的第四家：临床、职场、学校不是一家。它们是这些主张落地的场所，不是持有主张的位置。本章后面会大量使用睡眠判读手册的条文，那是兑现材料，不是第四种立场。

四·三家账本上共同没有的那一栏

4.1 三家共同踩着的那句话

三家争的是睡与醒的分界该由什么来定。填在横线上的那样东西——“分界”——三家都默认它是那种不占位置的东西：

划分由一条分界完成，而分界本身没有厚度、不承担功能。

念给三家听，三家都会说这还用说吗。营销学的定位图上，品类之间是线；化学的能垒图上，过渡态是曲线上的一个极大值点；行星科学的示意图上，晨昏线是一条线。三家都用“线”这个记号，而且都不认为这个记号本身需要解释。

会引起争论的句子，是第四家的立场；不引起争论的，才是三家共同踩着的那句话。这一条不引起任何争论。

4.2 推翻它的那件材料，来自行星科学自己

一条从别处借来的反驳，只能给出第四种立场，那是入局，不是散局。要让这条前提失效，用的东西必须是三家自己手上早有、早已印出来、只是一直没往这个方向使的那件事实。

它就是第一节那件：在潮汐锁定的行星上，被画成分界的那条晨昏带，是整颗地球上唯一可能宜居的区域。

它推翻的不是“分界在哪儿”，而是“分界不占位置”这件事本身。一条既有宽度、又有气候、又是唯一能住人的地方的“线”，不是线。而这个结论是行星科学自己算出来、自己发表的，两侧的极端条件也是它自己给的——三家里没有人能退回去说这是外行的臆测。

4.3 化学自己拆掉的另一半

三家共同踩着的那句话还有一半：分界的位置是被发现的。这一半由化学自己拆掉。

大木的一千秒是一条约定。埃利尔等人明说它“完全是任意的”，而且明说保留它的理由是——不这样定，阻转异构这个概念就维持不住。分界的位置不是被发现的，是为了让概念站得住而划的。

两半都塌之后，问题就变了。不再是“睡与醒的分界该由什么定”，而是：

被我们记成分界的那个地方，本身是什么？

4.4 三家账本上的空栏，形状是同一个

按三家各自的账本查一遍“被当成余数、噪声、或根本不设字段的是什么”：

	有字段的	没有字段的	为什么没有
市场营销学	各品类的份额、渗透率、进入场景数	不属于任何品类的使用场合	份额必须加总到一〇〇%
有机化学	各化合物（有编号、可分离、可入库）	过渡态	它按定义不可分离，故不是一种化合物
行星科学	日侧、夜侧的面积与温度	晨昏带作为一个区域（而非一条线）	二分图里线不占面积

三处空栏的形状完全一样：一片持续存在、承担功能、却因为被记成分界而没有自己一栏的区域。

这就是第二层三家共同踩着的那句话——不是三家共有的那一条，是一大片学术地基共有的那一条：

凡是一样东西，都必须属于某一类；而类与类之间只有边界，没有地方。

⚠ 这里必须先撤回本章前稿的一句错话。前稿在此写道：所有人都站在它上面，因此谁也没法回头看它。这句话不成立。生态学早在一九〇五年就给两个群落之间那片地起了名字，并且量它的宽度、命名只住在那里的物种；天文学一百多年前就把昼夜之间打开成了三级曙暮光。第八节第一、二条会详细处理这两位——它们不是本章的对手，是本章的正面对照：别的领域把线打开了，睡眠判读没有。

五·界带：把线打开

5.1 承重命题

睡与醒的分界，不是一个心智品类的边缘，不是一道由内部能垒定住的构型界线，也不是一个由外部周期定住的相位切点，而是——一片有厚度、承担功能、而在三家账本上都只被记为分界的区域；它做的事，全部被记在它两侧的名下。

三个否定各对应一家。第四句是本章要立的。

最后半句是全章的落点。记在两侧名下，是这件事至今没有被看见的原因：账没有少一分钱，只是所有产出都被划给了邻居。那十五秒里发生的洞察，在报告上属于“浅睡”或者“清醒”；那片区域什么也没得到，包括没有得到一个名字。

读到第八节会看到：正是这最后半句，而不是前面的“有厚度”，才是本章真正无主的那一格。生态学给了交错带字段——量它的宽度、命名只住在那里的物种、有人以它立项；本章的处境恰恰相反：睡眠判读没有给。

5.2 三个用词，定死

界带：两个被承认的类之间，那一片有厚度、且在现行账本上没有独立字段的区域。它不是两类的重叠，也不是两类的加权混合——它有自己的产出。

线化：把一片有厚度的区域记成一条无厚度的分界这个动作。线化不产生任何错误提示：账仍然记得平，两侧的份额仍然加总到全部。

邻记：界带的产出被记入相邻某一类名下这件事。邻记是线化的必然后果，也是界带难以被发现的机制——它的成绩单在别人手里。

5.3 辨别装置：一张缺席清单，不是一个格子图

要判断某处有没有被线化，本章不给二维分类格，给一张逐项核对的缺席清单。一处分界，凡下列各项答“无”的越多，被线化的程度越深：

- ① 这个位置有没有自己的名称？（不是“A到B的过渡”这种从属命名）
 - ② 它有没有自己的计量单位与字段？（不是从两侧扣减出来的余数）
 - ③ 它的产出记在谁名下？（若记在两侧 → 邻记成立）
 - ④ 它的宽度会不会随观察分辨率提高而收敛到零？（不收敛 → 它是区域）
 - ⑤ 有没有一条规则规定它的最小可记时长？（若有，且大于它的典型时长 → 它被结构性地剔除）
 - ⑥ 有没有人以它为专门对象立项？（若无，而两侧各有一整个学科 → 它是学科间的荒地）
- 对睡与醒之间那片区域逐项核对：① 无（现行只有“非快速眼动第一期”这种从属命名，且它是一个分期而非一片区域）；② 无（没有任何常规报告字段登记它的时长与产出）；③ 记在两侧（洞察算“休息之后”的效果）；④ 见第六、七节；⑤ 有，且大于——判读窗三十秒，而有效时长可短至十五秒；⑥ 基本无。

六项里五项答“无”，第四项待检验。这就是第六节的任务。

这张清单同时是本章与生态学的分离线：把它拿去核对森林与草原之间那条交错带，六项会得到六个“有”——它有名字（ecotone）、有字段（宽度）、有专属物种（edge species）、有测宽方法、有专门立项。同一张清单，一处六个“有”，一处五个“无”。清单本身不区分领域，是账本区分了它们。

5.4 一句不含程度词的判据

在这套记录规则里，这个位置有没有一栏是它自己的？如果没有，它的产出被记在了谁的名下？

这句话可以拿去问报告模板、问字段表、问计费码，不用问任何人的判断。禁用“应当”“真正”“充分”“有意义”这类词。

在五个场合各跑一遍，答案互不相同，其中至少两个是查出来的：

场合	答案	来源
睡眠判读报告	没有。产出记在“总睡眠时	查出来的（判读手册条文）

	间”与“清醒”两侧。判读窗三十秒、觉醒最短三秒且其前需十秒稳定睡眠——三条规则合起来把界带压成零。	
化学物质登记	没有。过渡态没有登记号，不可入库；它的效应记在反应物与产物的动力学参数里。	查出来的（过渡态按定义不可分离）
品类份额报表	没有。份额必须加总到全部，跨品类使用场合被摊回各品类。	查出来的（份额恒等式）
植被样带调查	有。交错带有名称、有宽度读数、有专属物种、有专门的测宽方法。	查出来的（移动分割窗法）
职业交接班	没有。交接那一段的时间在两班的工时里各记一半或都不记，而事故率在这一段最高。	从命题可推

第三条反过来改了本章：它指出给界带设一个与两侧并列的字段会破坏份额的加和性。本章因此不再主张“设第三类”，改为主张“另账登记”——见第九节。

第四条是本章有意留的一个正面对照：同一句判据，在生态学那里得到的是“有”。这说明缺席不是这类区域的宿命，是这套记录规则的选择。

六·可判定：分界是不是一条线，有一个检验

一个概念若不能被数据判错，它只是一个说法。本章给出的检验很简单，而且不需要新数据。

如果分界真是一条无厚度的线，那么状态之间的转换就是彼此独立的点事件：相邻两次转换的间隔应服从单峰的几何分布；把记录粒度粗化，含有两种以上状态的“混合窗”比例应当与独立点过程模型相符。

如果分界是一片有厚度的带，转换会成簇——因为穿过一片带时，判读会在带内反复改判，一次“过界”留下好几次转换。于是：相邻间隔的分布在短端会相对几何期望超额（带内的极短间隔）；粗化之后，观测的混合窗比例会低于独立点模型，因为一整簇转换只染一个窗，而独立模型把同样多的转换散开去染更多窗。

两个方向相反的预测，落在同一份数据上。△ 但这段推理里藏着一个当时没有被写下来的前提：“独立点过程模型”用的是一个齐次的转换概率。这个前提第七节第七小节会被本章自己推翻。

预注册（跑统计之前写死，事后不改）

数据：Sleep-EDF Expanded 1.0.0 的 sleep-cassette 部分，按文件名排序的前四十份判读标注，未挑选，公开可复算。

- Q1：转换间隔在短端相对几何期望显著超额，短模式落在一至两个窗（三十至六十秒）。
- Q2：粗化八倍时，观测混合窗比例低于独立点模型，相对偏差大于 10%。
- Q3：簇内转换数的均值大于 2.0——即一次“过界”平均不止一次转换。
- Q4：多转换簇在总时长中占比小于 15%，但包含超过 60% 的全部转换。
- 撤销规则：Q1 若在短端无超额，“界带有厚度”直接判错，本章须改写为纯粹的粒度问题；Q2 若方向相反，成簇解释错。
- 代理问题（先写明）：三十秒是本数据的最细粒度，只能往粗走不能往细走，故一切读数都是下界；人工判读成品已抹平了带内的部分犹豫，成簇程度同样是下界。

七·这次检验跑出了什么

四十夜，四三三二九个判读窗（三六一小时），四四七七段，四三七九次转换，平均每个窗界发生转换的概率 0.1013。

7.1 Q1 命中，且幅度很大

间隔（窗）	实测占比	齐次独立点过程期望	实测/期望
-------	------	-----------	-------

1	33.54%	10.13%	3.31 倍
2	16.04%	9.10%	1.76 倍
3	9.28%	8.18%	1.13 倍
4	5.32%	7.35%	0.72 倍
8	2.10%	4.80%	0.44 倍
≥13	18.98%	27.76%	0.68 倍

短端严重超额、中段与长端都不足。三分之一的转换发生在上一次转换之后的三十秒之内。

⚠️ 一处用词更正。 本章前稿在摘要里把这张表描述为“间隔分布强烈双峰”。这句话不准确：实测占比本身是单调递减的（33.54 – 16.04 – 9.28 – 5.32 – 2.10），双峰只存在于**实测与期望之比**这条曲线上，不存在于分布本身。准确的说法是**短端相对几何期望超额**，本稿已全篇改用后者。

7.2 Q2 命中，且偏差随粗化单调增大

粗化倍数 k	窗宽	实测混合窗	齐次独立点模型	相对偏差
3	90 秒	16.71%	19.23%	-13.1%
4	120 秒	21.67%	27.41%	-21.0%
6	180 秒	29.80%	41.37%	-28.0%
8	240 秒	36.64%	52.65%	-30.4%
10	300 秒	42.06%	61.76%	-31.9%

偏差从零单调走到负三成，且在最粗处趋于平缓。

7.3 Q4 命中

把间隔不超过两个窗的相邻转换归为同一簇，四十夜共得二二五三个簇，簇内转换数均值 1.94。其中含两次以上转换的簇占簇数的 51.9%，而它们包含了 **75.3%** 的全部转换。

⚠️ 一处诚实说明。 这个数在很大程度上由 7.1 蕴含：既然长度 ≤2 窗的段已占全部段的约一半，那么“多数转换落在多转换簇里”几乎是算术上必然的。本章前稿把它写成“全章最要紧的一个数”，是高估了它的独立性。它是 Q1 的一个改写，不是对 Q1 的独立确证。

7.4 Q3 未命中，据此改了命题

预注册说簇内转换数均值应大于 2.0。实测 **1.94**，中位数 2，未达标。按预注册规则，这一条判未命中，不保留原表述。

追问它为什么差这一点，得到的东西比原预测好：含两次以上转换的簇占 51.9%，而它们包含了 75.3% 的转换。也就是说——

约一半的分界确实是线（单次转换，干净地过去了）；另一半是带。

本章据此把命题从“分界都是带”改成“分界分两种，而现行账本对两种用同一个记号”。这个版本弱一些，但它有两个好处：一是数据支持它，二是它自带一条本章最需要的防线——本章不再需要主张一切边界都有厚度，只需要主张**没有任何字段能把这两种分开**。

7.5 这次检验改了本章三处

1. 删掉“分界都是带”，改为两分，并把未命中的那条写在正文里而不是脚注里。
2. 把主张的强度调低：本章不再主张界带自身是一个自然类，只主张现行记号无法区分线与带。
3. 把“厚度”的地位改了：由于只能往粗走不能往细走，本章测到的是成簇，不是厚度的绝对值。本章放弃给出界带的厚度数值，只给出它存在的判据。

7.6 跑不动的那一条，是本章最稳的一层

第五节缺席清单的第四项——**宽度是否随分辨率提高而收敛到零**——本次跑不动，因为公开数据的最细粒度就是三十秒，往细走没有数据。

跑不动这件事本身构成本章的第三层主张，而且它指向的是全篇最便宜的一个动作：判读设备的采样本来就是毫秒级的，把亚窗级那点分期犹豫顺手存下来，几乎不增加任何成本。存下来之后第四项当场可判——收敛到零，本章错；不收敛，界带成立。

7.7 零模型的更正：一次事后检验，它拆掉了上面三条中的一条半

本小节是本稿新增的，且必须先声明它的地位：它不是预注册的，是事后（post hoc）稳健性检验。之所以要做，是因为在把本章与全球近邻逐条划界时（第八节第二条），发现第六节那个“独立点过程模型”有一个当时没有写下来的前提：

它假定转换概率在整夜、全部分期、全部受试者上是**同一个常数**（本次数据里是 0.1013）。

这个假定明显不成立。睡眠转换率是高度非齐次的：本次数据里，按分期分别估计的每窗转换率相差近七倍——

分期	每窗转换概率	相当于平均段长
N1	0.3510	2.8 窗（84 秒）
N3	0.1435	7.0 窗
N2	0.0889	11.2 窗
W	0.0635	15.8 窗
R	0.0502	19.9 窗
（齐次合并）	0.1013	9.9 窗

而率的异质性单独就会产生短端超额与混合窗不足，不需要任何“带”。更麻烦的是方向：一个混合了快慢两种率的过程，其间隔分布必然比单一几何分布更重尾（短端与长端同时超额、中段不足），而这正是 7.1 那张表的形状；粗化时的偏差随 k 单调增大，也正是异质性的典型表现。

因此本章用两个更严的零模型重跑，**代码与数据同前，任何人可复算。**

零模型一：分期条件的混合几何。按各分期自己的转换概率、按实测分期出现比例混合，算 Q1 的期望。

间隔	实测	齐次期望（比）	分期条件混合几何期望（比）
1	33.54%	10.13%（3.31）	15.05%（2.23）
2	16.04%	9.10%（1.76）	11.56%（1.39）
3	9.28%	8.18%（1.13）	9.11%（1.02）
4	5.32%	7.35%（0.72）	7.37%（0.72）
8	2.10%	4.80%（0.44）	3.83%（0.55）
≥13	18.98%	27.76%（0.68）	26.01%（0.73）

结论：短端超额从 3.31 倍降到 2.23 倍。分期条件的率异质性解释掉了约三分之一的原始信号；剩下三分之二仍在。Q1 因此没有被推翻，但它的幅度必须按后一列报告，不能按前一列。

零模型二：段长夜内重排。把每一夜的各段长度在夜内随机重排，保留段长分布与分期标签，只打散它们的时间排列，重排两百次。这个零模型问的是一个更尖的问题：**观测到的成簇，有多少是“短段本身就多”的后果，有多少是“短段还挤在一起”的后果？**

读数	实测	段长重排（200 次）	z
混合窗比例（k=3）	16.71%	17.01%±0.15%	-2.0
混合窗比例（k=4）	21.67%	22.25%±0.18%	-3.2
混合窗比例（k=8）	36.64%	37.07%±0.24%	-1.8
混合窗比例（k=10）	42.06%	42.69%±0.30%	-2.1
多转换簇含转换比例	75.3%	71.8%±0.4%	+9.5
多转换簇跨度占时长	90.1%	89.5%±2.3%	+0.2

这张表要分两段读，而且两段的结论相反。

Q2 全军覆没。相对齐次模型 -30.4% 的那个偏差，在段长重排零模型下完全消失——实测甚至略低于重排 ($k=8$ 时 36.64% 对 37.07%)。也就是说：混合窗那条读数的全部偏差，都是段长分布本身的后果，与转换在时间上怎么排列无关。**Q2 不为“带”提供任何证据。**

Q4 只剩一小截，但那一小截是实的。多转换簇所含转换的比例，实测 75.3%，段长重排给出 $71.8\% \pm 0.4\%$ ， $z=+9.5$ ——方向为正、统计上极显著。但幅度只有 3.5 个百分点：本章原先报告的成簇，约十九分之十八由段长分布解释，只有约十九分之一来自“短段还挤在一起”这件事。而跨度那一项 ($z=+0.2$) 什么也没有。

7.8 这一次自我更正把本章改成了什么

四处，逐条列明：

1. **摘要与正文的经验主张全面降级。**本章不再说“转换成簇因此分界是带”。本章现在说：转换的时间分布确实不是齐次泊松的，但其中绝大部分由段长分布解释，而段长分布是睡眠动力学二十年前就已刻画的对象（见 8.2）。超出段长分布的成簇只剩 3.5 个百分点。
2. **Q2 从“命中”改判为“在正确零模型下不成立”。**这一条本章主动撤回，不作补救性重述。
3. **本章的承重从统计挪回账本。**真正无主的那一格不是“转换会不会成簇”——那是睡眠动力学的地盘；是邻记（产出被记在邻居名下）与那条纯算术（十五秒对三十秒判读窗）。这两样不依赖任何零模型，Q1 与 Q4 全塌了它们也还在。
4. **第六节那句话必须收回。**前稿在第六节末写道：两个方向相反的预测落在同一份数据上，谁对谁错由数据说了算，这比一条写死日期的赌注硬，因为它现在就能裁决。**在零模型错的前提下，这句话不成立。**一个检验的硬度，不由它能不能立刻算出一个数决定，由它的零模型对不对决定。本章前稿把“可立刻计算”误当成了“可立刻裁决”。

⚠ 这一处自我更正本身，是本章命题的一个实例。齐次零模型算得出一个漂亮的 -30.4%，账记得平平整整，没有任何错误提示；缺的不是数字，是“零模型该按什么分层”这一栏从来没有被写下来。本章用了整整一稿的时间，才发现自己的账本上少了那一栏。

八·与最近的十四种说法的分界

每条写四件：它说到哪一步、分离线、以及一条能分出谁对的判定形式。凡是只能写出“本章更强调”“视角不同”的，等于承认被覆盖，不保留。

排序按危险程度。排在最前的两条，是本章前稿完全没有处理、而它们离本章最近的两位。

8.1 生态交错带 (ecotone; Clements 1905 起) ——本章最危险的一位，前稿零命中

它说到哪一步。一九〇五年，克莱门茨用 ecotone 命名两个相邻生态群落之间的过渡带。此后一个多世纪里，生态学把本章以为自己在开创的每一件事都做了一遍：

- 它给那片地起了名字（交错带），并且区分了两种——边界锐利的 ecotone 与沿梯度渐变的 ecocline。这正是本章 7.4 被 Q3 打脸之后才改出来的“一半是线、一半是带”。
- 它命名了**边缘物种** (edge species)：主要或最丰富地出现在交错带里的生物；并且反复报告**边缘效应**——交错带里的物种数与某些物种的密度可以远大于任一侧群落。这正是本章 8.6 用来反驳模糊集的那句“界带不是加权混合，它有自己的产出”，逐字。
- 它有一套专门测宽度的方法：移动分割窗 (moving split window, 惠特克一九六〇年代起)，靠改变窗宽来判定边界的位置与宽度。这正对本章缺席清单的第四项与 7.2 的粗化检验。

这不是一处遗漏，是一处错误。本章前稿在 4.4 与 8.11 两处写下“所有人都站在这条地基上，因此谁也没法回头看它”——生态学回头看了一百二十年。

分离线，而且这条线反过来加固本章。生态学给了**交错带字段**：它有名字、有宽度读数、有专属物种、有专门的浏览方法、有人以它立项。把本章第五节那张缺席清单拿去核对森林与草原之间那片地，六项会得到六个“有”；拿去核对睡与醒之间那片地，得到五个“无”。同一张清单，同一类对象，两个相反的结果——差别不在对象，在账本。

因此生态学在本章里的位置不是对手，是**正面对照**：它是“把线打开”这件事已经被做成的样子。它与 8.7 的天文曙暮光分级构成一对：两个领域做了，睡眠判读没做。

判定形式：若睡眠判读手册中已存在等价于交错带的过渡区实体（有独立名称、独立字段、独立宽度统计），则本章所指认的缺席不成立，本章错。就现行手册而言，没有。反过来，若在生态学里也找不到交错带的独立字段（宽度不被报告、边缘物种不被单列），则本章的“正面对照”这一步落空，本章须改用别的对照。

8.2 睡眠转换动力学（Lo et al. 2004；Blumberg et al. 2005；Bianchi et al. 2010；Chu-Shore, Westover & Bianchi 2010）——本章 Q1/Q2 的正主，前稿零命中

它说到哪一步。本章第七节测的“相邻转换间隔”，换一个名字就是**段长**（bout duration）——而睡眠段长分布是一个被系统研究了二十年的对象：

- 罗等人二〇〇四年在《美国科学院院刊》上给出跨物种结果：**睡眠段呈指数分布，清醒段呈幂律分布**。幂律即短端相对指数超额——**这就是本章的 Q1**。
- 布隆伯格等人二〇〇五年在同一刊物上给出发育中的对应结果。
- 比安基等人二〇一〇年报告：人类的清醒、快速眼动与非快速眼动段呈**多指数分布**，提示比此前报告更复杂的动力学。
- 最要命的是朱-肖尔、韦斯托弗与比安基二〇一〇年那篇：他们系统检验了**多指数分布被误判为幂律**的参数区间，发现这个“拟态区”恰好与人类睡眠与清醒段的实测时间常数重合，并明确指出，这种模型辨识的不确定性会直接影响对转换动力学的机制解释（自组织 vs 概率性）。

本章前稿做的正是这个推断，而这篇文章的警告正对着它。7.7 那次自我更正，本质上是被这一条逼出来的：分期条件的混合几何零模型，就是“多指数”在本章数据上的最简形态。

分离线在被解释的东西不同。这一支文献解释的是**段长分布本身**——它为什么重尾、由几个成分构成、对应什么神经机制。本章解释的是**账本**——那些段落报告里有没有一栏。两者可以完全独立：一个段长分布被完美刻画的领域，照样可以没有界带字段；反过来，一个给了过渡区字段的领域（生态学），其段长分布可能根本没人算过。

判定形式（本章第一顺位待办）：把本章 7.7 的段长重排零模型换成多指数拟合零模型（按 Bianchi 的方法拟合多成分，再按拟合分布生成代理序列）。若那 3.5 个百分点的残余成簇也被多指数模型吸收，则本章的统计部分**全部**归入睡眠转换动力学，本章只剩账本那一层；若残余仍在，则“超出段长分布的时间成簇”是一个该文献未处理的新对象。**本章没有做这一步，写在这里作为待办，不充作已完成的检验。**

8.3 可变时间单元问题（MTUP；Cheng & Adepeju 2014；上溯 Openshaw 的 MAUP 1984）

它说到哪一步。与可变面积单元问题类比，可变时间单元问题被定义为由三种时间效应构成——聚合、分段与边界；而被检出的聚类，其时长、范围与显著程度，会随着这三项设定的改变而一起变。

本章 7.2 整节就是一次尺度效应演示：把粒度从三十秒粗化到三百秒，看结论怎么变。这在地理学里是一个已被命名四十年的常规现象，且已有一整套灵敏度分析方法。

分离线在结论的形态。MTUP 说的是：同一批被数的东西，在不同分段下会给出不同的统计量，因此结论对分段敏感，须做灵敏度分析。本章说的是另一件事：**在同一次分段之下，一样东西可以有厚度而没有字段**。MTUP 处理的是一个量的不稳定，本章处理的是一栏的有无。而且 7.7 已经表明，本章 Q2 的那条曲线确实就是 MTUP 意义上的尺度效应——它不承载本章的主张。

判定形式：遍历窗长、起点相位与并类方案，看缺席清单第①②③⑥项的答案会不会跟着变。这四问的是字段表而不是数据，因此按理不该随分段改变；若它们真的变了，本章的主张就只是分段敏感性的一个侧面。

8.4 模糊集 (Zadeh, 1965)

它说到哪一步：边界元素不必非此即彼，可以给它一个介于零一之间的隶属度。这是处理“边界模糊”最成熟、最便宜的现成办法——如果它够用，本章这个概念就该被它取代。

分离线：模糊集仍然假定只有两个类，边界是两类的重叠，边界处的性质是两类性质按隶属度的加权。界带不是加权混合，它有自己的产出。

判定形式：若界带的产出可由两侧的产出按隶属度重构，则模糊集覆盖本章。第一节那份数据说不行——非快速眼动第一期的洞察率 83%，既高于清醒的 30%，也高于更深睡眠；一个介于两者之间的隶属度无论取什么值，都算不出一个高于两端的数。本章对。[△]但须注明：这条论证的一般形式属于生态学的边缘效应（见 8.1），本章只是把它用在睡眠上。

8.5 睡眠惰性 (Tassi & Muzet, 2000)

它说到哪一步：醒来之后有一段功能低落期，可持续十五分钟到数小时，被理解为睡眠向清醒过渡时的残留损耗。

分离线：睡眠惰性把那段时间读成损耗，本章把它读成有产出的区域；且睡眠惰性只处理醒来一侧，界带是双向的。

判定形式：若界带内的产出只在入睡一侧出现、醒来一侧纯为功能低落，则本章的“双向”错，睡眠惰性足够；若醒来一侧也能测到两侧都没有的产出，则本章对。这一条今天就能做：把同一套洞察任务放在醒来后第一分钟内测。

8.6 入睡幻觉 (Silberer, 1909; Mavromatis, 1987)

它说到哪一步：入睡之际的意识内容有独特的现象学——形象、类比、自动思维。这是这片区域最早的命名，比本章早一百多年。

分离线：那是一个状态的现象学描述；本章是一个账本结构的诊断。本章的核心不是这片区域里有什么体验，而是它没有字段、产出被邻记。一个人可以完全不报告任何入睡幻觉，而界带照样存在于他的记录里。

判定形式：若界带的成簇现象只出现在有入睡幻觉报告的人身上，本章这一条应并入入睡幻觉研究；本次数据未采集现象学报告，无法裁决，如实记为未判。

8.7 曙暮光分级 (天文学，十九世纪末即已确立)

它说到哪一步：天文学早就不把昼夜之间当作一条线了——民用、航海、天文三级曙暮光各有明确的太阳高度角定义。这是“把线打开成带”这个动作最成功的一次先例。

分离线：正因为它成功，它反过来成了本章最有力的对照之一：同样一件事，天文学做了一百多年，生态学做了一百二十年，睡眠判读没有做。本章不主张一个新原理，只指出一个已被两个领域完成的动作在这里缺席。

判定形式：若睡眠判读手册里已存在等价于三级曙暮光的过渡分级（有独立名称、独立字段、独立时长统计），本章错。就现行手册而言，没有。

8.8 阈限 (van Gennep, 1909; Turner, 1969)

它说到哪一步：仪式过程中有一个阶段，人既不属于原来的身份也不属于新的身份，处于结构之外。这是“中间地带自成一格”在人文学科里最完整的表述。

分离线：阈限的承载者是人与制度，其成立依赖社会承认——共同体知道你正在过渡。界带的承载者是记录规则，且无人到场、无人承认；它的特征恰恰是没有任何一方知道有这么一片地方。

判定形式：若在完全自动、无人参与的记录流程里成簇现象消失，本章的“无人到场”错；本次数据的判读虽由人完成，但成簇是转换的时间结构，与判读者是否意识到无关——这一条需要自动判读流程复算，本章未做。

8.9 相变的临界区（物理学）

它说到哪一步：相变附近有一个涨落剧烈的区域，宽度可测。

分离线：临界区宽度随系统尺寸按标度律收敛，是有限尺度效应；本章的判据第四项要求的正是**不收敛**。

判定形式：就是本章缺席清单第四项。收敛，本章错；不收敛，界带成立。本次跑不动，见 7.6。

8.10 社会时差（Roenneberg 等，2006）

它说到哪一步：生物钟与社会钟之间存在相位差，可测量、与健康结局相关。

分离线：社会时差是两个参照钟之间的一个**标量差**；界带是一片**区域**，与两钟对齐与否无关——把相位差调到零，界带照样存在。

判定形式：在自由作息、社会时差近零的人群里，若成簇程度与常规人群无差异，本章对；若成簇随社会时差同步变化，本章这一条应并入社会时差。

8.11 同批：《那三十秒算不算一段》（第一章）

同一系列的上一篇讲的是：一段时间要成为一个可数的段落，需要被后续事件追认；未被追认的那些不构成一样东西。

分离线：上一篇问一段时间能不能成为一个单位；本章问两个单位之间那块地有没有单位。二者可各自独立发生：一份段落清晰、无未追认段的记录，照样可以完全没有界带字段。

判定性预测：把最小可记单位调细，上一篇的**未追认段比例会上升**（更细的粒度暴露更多短段），而本章的**界带字段数仍然是零**——因为那不是粒度问题，是分类学问题。

8.12 自家最近的三处：此前未申报的站内自撞

本章前稿申报了四处自家近邻（上一篇、节律剥夺、垫付线、失眠专文）。用站内近邻库实扫之后发现，另有三处同样近或更近的，前稿一处未提。不申报即自撞，逐条如下。

（一）《一件是几件》。该文论证：生成一件东西的代价趋零之后，“一件”这个计数单位失去了下界，因而件数上升不能再充当投入上升的证据。**这正对本章缺席清单的第四项**——宽度会不会随分辨率提高而收敛到零，问的就是这个单位有没有下界。**分离线：**那一篇问的是**同一类东西的计数单位**有没有下界；本章问的是**两类之间那块地**有没有单位。前者是一个类内的问题，后者是一个类间的问题。**判定形式：**若把分辨率提高之后，界带的宽度与“一件”的下界同步收敛或同步不收敛，两条应合并；若一个收敛而另一个不收敛，则是两件事。

（二）《撤销一个名字，人不会回到原处》。该文写的是：一次命名在人身上开出的位置只能被占据或空着、不能被取消，而空着的那个位置本身是**第三个位置**——它没有内容，只有形状，因此所有针对内容的补救对它无效。

“第三个位置，只有形状没有内容”与本章的界带形状几乎同构，而且那一句比本章的表述更狠。**分离线在这个位置是怎么来的：**那一篇的第三个位置由**一次命名**开出，是命名的剩余；本章的界带由**两侧的记号方式**造成，是记号的省略。前者多出一个位置，后者少掉一栏。**判定形式：**若撤销睡与醒这两个名字，界带随之消失，则本章属于那一篇；若两个名字都不用、只看连续记录，那片有厚度的区域照样在（第一节的十五秒不依赖任何分期命名），则本章是另一件事。

（三）《类内距》。该文论证：一份规范只写了什么算同类，没写同类之间有多近；而后者决定可替换性，且要单独花钱买。**分离线：**那一篇说规范漏写了**类内的**距离，本章说规范漏写了**类间**那块地；两者是同一份规范的两处缺口，但缺的不是同样东西。**判定形式：**把两处缺口同时补上——若补齐类内距之后界带自动获得字段，两条应合并；若补齐类内距而界带仍无字段（现行手册正是如此：分期内部的深浅有描述，分期之间的那块地没有），则是两处。

8.13 自家：节律剥夺

站上另一篇讲的是同质化的时钟如何碾平各种不同的节律，制造认知突变的制度性稀缺。

分离线：那一条讲节律被同质化（周期被统一），本章讲两个状态之间的地带被记成线（分界被压平）。一个作息极不规律、完全不受统一时钟约束的人，照样有界带，因为界带来自记号，不来自时刻表。

8.14 自家：垫付线与失眠专文

《健康发生学导论》那一条讲身体的每次正常由下一轮的接手能力垫付，垫付越成功越无信号。

分离线：那是储备被消耗而无信号，本章是区域存在而无字段；一个储备充裕的夜晚照样有界带。

站上那篇《失眠：为什么越想睡越睡不着》讲入睡这一刻的努力悖论。**本章不碰努力与意志：**界带在睡得极好、毫不费力的夜里同样出现，本次数据的受试者并无失眠。

8.15 这些邻居共同站着的那块地

逐个问“它结构性地看不见什么”，并写出“一旦承认，它自己会怎么塌”：

- **模糊集**把两个类当作给定，因此看不见第三样东西；一旦承认边界处有两侧都没有的产出，隶属度这个装置就失去意义。
- **份额报表**把加和到全部当作给定，因此看不见不属于任何一类的场合；一旦承认，它的恒等式就要重写。
- **睡眠判读**把分期当作给定，因此看不见分期之间；一旦承认，三十秒窗与占优规则都要重定。
- **睡眠转换动力学**把分期序列当作给定的输入（它研究这个序列的统计性质），因此看不见这个序列本身是被一套字段表生产出来的；一旦承认，它所有的段长分布都要先接受一层字段敏感性检验。
- **临界区**把收敛当作给定，因此看不见收敛的中间态；一旦承认，标度律不再是普通的。
- **生态交错带**——这一位看得见，而且看了一百二十年。它在这张表里的位置不是盲点，是对照组。

背后是同样被当作给定的东西：**凡是一样东西，都必须属于某一类；类与类之间只有边界，没有地方。**

⚠ 前稿在此写道：所有人都站在它上面，因此谁也没法回头看它。**本稿撤回这句话。**生态学与天文学都回头看过，而且看得比本章完整。本章与它们的差别不在有没有人看，在**看的位置**：它们把线打开了带，并给了那片带一整套字段；本章指出的是，在睡眠判读这一处，同样的动作没有发生。

剩下的那一格，本章把它收窄成一句：

同一片有厚度的区域，在一个领域里有六个“有”，在另一个领域里有五个“无”；而缺席的那一侧，其产出被完整地记在了两个邻居名下。

前稿想主张的“分界其实是带”，主人是生态学；前稿想主张的“转换会成簇”，主人是睡眠转换动力学，且经 7.7 的更正后只剩 3.5 个百分点。剩下无主的，是**邻记**这个动作，以及那条十五秒对三十秒的算术。

九·四家反过来给本章加的约束

下面四处，每一处都是本章先写了一句更大的话、然后被某一家按住。按住的方式不同：有的收走了一个数量，有的收走了一个归属，有的收走了一处处方，最后一处收走的是“首创”这两个字。四处都按同一格式登记：本章原想写什么、哪一家不许、削掉之后剩什么。第四处为本稿新增。

9.1 有机化学削掉的：本章不能主张界带有一个客观厚度

本章原本想写的更强的一句是：**界带有一个可测的、属于它自己的厚度。**

化学不允许。大木的一千秒是约定，而化学界自己在教科书里说这个约定“完全是任意的”，保留它只因为不这样定概念就维持不住。同理，界带的厚度必然相对于某个观察时标而言，**没有一个绝对值。**

削掉之后本章只能主张：**厚度不随分辨率收敛到零。**这弱得多，但它可判定，而前一句不可判定。这一处动的是本章的适用范围。

9.2 行星科学削掉的：功能不属于界带自身

本章原本想写的更强的一句是：**界带自身承担功能**。

行星科学不允许。晨昏带之所以宜居，恰恰是因为**两侧都不宜居**——它的温度是日侧的热与夜侧的冷对冲出来的。功能不是那条带的内禀性质，是两侧极端条件的差值在那里恰好落进可用区间。

落到睡眠上，这意味着那十五秒里的洞察，很可能也不是“第一期这个状态”的性质，而是**认知控制尚未撤尽、自发联想已经开始**这两侧条件的对冲。本章因此把机制归属从“界带有功能”改为“**界带是两侧条件对冲后唯一落进可用区间的位置**”。这一处动的是价值排序：不能据此主张“多待在界带里更好”，因为一旦两侧条件变了，同一片位置的性质就变了。

9.3 市场营销学削掉的：本章不主张给界带设一个并列的第三类

本章原本想写的更强的一句是：**应当把界带设为与两侧并列的第三类**。

营销学不允许。品类份额必须加总到全部，这是核算恒等式，不是疏忽；凭空插入第三类会让所有份额类读数失去可比性与历史连续性。同样的道理适用于睡眠：总睡眠时间、睡眠效率、各期百分比是几十年纵向数据的基础，改动分类会一次性废掉全部历史比较。

削掉之后本章的处方缩到一条纯程序性的要求：**另账登记**——不改变现行分类与份额，另设一份不参与加总的登记，记录界带的位置、跨度与成簇程度。不主张哪一套分类更对，只主张这一栏必须存在且不与两侧争份额。这是本章最实质的一次让步，动的是价值排序。

9.4 生态学与睡眠转换动力学合起来削掉的：本章没有发现任何新现象

本章原本想写的更强的一句是：**没有人把分界看成一片有厚度、承担功能的区域，这是本章的发现**。

这两家都不允许，而且理由互补。生态学一百二十年前就把它看成区域，命了名、量了宽度、列了专属物种；睡眠转换动力学二十年前就刻画了段长分布，而本章 7.7 承认自己的成簇读数十九分之十八由它解释。**本章没有发现任何新现象**。

削掉之后本章只剩两样，而且必须把它们说得很窄：

其一，**邻记**——一片区域的产出被完整地记在它两个邻居名下，而两侧的账都记得平。生态学没有这个问题（它给了字段），睡眠转换动力学不问这个问题（它研究序列的统计，不问序列的字段）。

其二，**那条算术**——判读窗三十秒加占优规则，使一个十五秒即可把洞察率由 30% 抬到 83% 的状态，不保证能被记成任何一个窗。这一条不依赖零模型、不依赖成簇、不依赖任何统计推断，它只依赖两个数字的大小关系。

这一处是本章最大的一次退让，也是本稿相对前稿最实质的改动。本章因此从一篇声称发现了新现象的文章，变成了一篇指认一处账本缺口的文章。后者小得多，但它站得住。

十·不适用、反噬与伦理

10.1 最强的竞争解释

本章前稿没有这一节，这是一处结构性的疏忽：一篇给出经验读数的文章，必须自己把最难打的那个替代解释摆上台，而不是等审稿人摆。现在补上，按危险程度排三条：

第一，率的异质性／多指数。睡眠转换率在不同分期之间相差近七倍，混合快慢两种率的过程本身就会产生短端超额与混合窗不足，不需要任何“带”。**这一条已被 7.7 部分兑现**：它解释掉了 Q1 的三分之一、Q2 的全部。剩余的检验（多指数拟合零模型）见 8.2，本章未做。

第二，周期性交替模式（CAP）。睡眠医学自己有一套微结构记号：非快速眼动睡眠中的相位事件按二十至六十秒的周期成簇出现，跨在传统判读窗之上，按一秒判读。**本章测到的转换成簇，很可能是这套周期在三十秒粒度上的影子**。判定形式：在同一批记录上同时计算 CAP 率与本章的残余成簇

(7.7 那 3.5 个百分点)。若两者在个体层面高度相关，则残余成簇也归 CAP，本章的统计部分全部清空。CAP 睡眠数据库是公开的，这一条今天就能做，本章未做。

第三，判读噪声。判读是人做的，人会抖；三十秒一格切出一堆碎屑很正常。这一条最省事，但它预测碎屑在各分期间大体均匀，而实测的分期条件转换概率相差近七倍——它解释不了这个结构。本章认为这一条比前两条弱。

10.2 本章不适用的地方

- 不适用于连续无分级的记录——那里没有被线化的对象。
- 不适用于诊断。成簇程度不指示任何疾病，本次数据与任何临床变量的关系一律未检验。
- 不适用于主张“多停留在界带里更好”。见 9.2：那片区域的性质由两侧条件决定，两侧一变它就变。
- 不适用于给出界带厚度的数值。见 9.1 与 7.5。
- 不适用于替代现行分期与时长读数。本章主张另账并行，不主张取消。
- **(本稿新增)** 不适用于宣称本章发现了“分界其实是一片区域”。那件事由生态学在一九〇五年命名并持续研究至今；本章只主张在睡眠判读这一处，同样的动作没有发生。
- **(本稿新增)** 不适用于把 7.1、7.2、7.3 三条读数当作“界带存在”的证据。经 7.7 更正后，Q2 已撤回，Q1 幅度减三分之一，Q4 只剩 3.5 个百分点的残余。

10.3 反噬预言

假设本章成功了，界带这一栏真的被写进了报告模板。接下来最可能发生的事，是这一栏被拿去打分：**把成簇程度设成一个越低越好的指标**。一旦某机构被要求“界带越窄越好”，最省力的达标办法是**放宽转换的判定门槛**——门槛一松，转换变少，簇变少，指标立刻漂亮，而被压掉的时间更多。指标会非常好看，而它本来要显形的那片区域会更彻底地消失。

这条路本章没有找到绕过去的走法。凡是能被数出来的东西都会被排名，而这一个数被排名时的最省力解法，正好是把那片区域压得更薄。本章只能把这段话留在正文里当作警示，并附下面三条硬约束。

10.4 伦理硬约束（三条，缺一条本章不可发表）

1. **成簇程度指的是记录规则，不是人。**它是一套判读规程的属性，不得用于评价判读人员，也不得用于评价被记录者。
2. **不得为了把这个数做好看而放宽转换判定门槛。**任何对判定门槛的调整，须与该调整对转换总数的影响一并公布。
3. **凡曾依据分期类读数对任何人做出不利安排的机构**（劳动能力评估、保险核定、驾驶与轮班资格），须公开其分期规则、最小可记单位与转换判定门槛，并提供复核通道。

十一·用本章的判据检本章自己

本章的判据是：这个位置有没有一栏是它自己的；如果没有，产出被记在谁名下。把它转过来，对准本章自己——先对准本章的方法，再对准本章所属的那套评价办法。

先对准方法。7.7 那次自我更正，是本章命题在本章身上的一次运行：齐次零模型算得出一个漂亮的 -30.4%，账记得平平整整，没有任何错误提示；缺的不是数字，是“零模型该按什么分层”这一栏从来没有被写下来。一件东西在一份记录上没有一行，不是因为它不在，是因为记录的规则没有为它写下条款。本章用了整整一稿的时间，才在自己的账本上发现那一栏是空的。

同样的话适用于 8.1 与 8.2。本章前稿之所以能写下“谁也没法回头看它”，不是因为生态学与睡眠转换动力学不在，是因为本章的检索规则没有为它们写下条款。

再对准评价办法。一篇文章的价值，今天是按类结算的：它属于哪个学科、投哪本期刊、算作理论还是实证、引用记在哪个方向。而本章这样的东西——它同时用行星科学的结果、化学的约定、营

销学的恒等式，去改一件睡眠判读的事——在任何一栏里都不完整。按学科分类，它没有归属；它的产出会被记在离它最近的那一栏名下，通常是“跨学科评论”。

这正是本章所描述的结构，而且本章对此无能为力：指出线化的这篇文章，自己就住在一条被线化的界上，并且它的成绩也会被邻记。

第三层要认的账在数据这一侧。四十夜全部出自一个公开库、一套判读传统、一批肯到实验室过夜的人；换句话说，本章的样本本身就是被一道准入线切出来的。没有进入公开库的那些睡眠，在本章里连一行都没有。本章对它们做的事，与判读手册对那十五秒做的事，形状一样：不是判错，是没有开栏。

收尾留给那件至今做不成的事：本章最想量的那个宽度，要靠亚窗级的判读中间态才能量，而这个字段在现有记录里根本不存在。本章可以被引用的最稳的一层，不是那个新名字，也不是那张清单，更不是第七节那三张表——那三张表里已经有一张被本章自己撤回了。最稳的是这一句：把亚窗级的分期犹豫一并保存，成本接近于零。前面几层可以塌，这一层塌不了。

注释与材料

本章两次检验（第六节的预注册检验与 7.7 的事后稳健性检验）用的是同一批材料：Sleep-EDF Expanded 1.0.0 中 sleep-cassette 的判读标注，按文件名排序取前四十份，未作任何挑选。预注册文本、原检验脚本与本稿新增的两个零模型脚本一并随文提供，读者可从原始 EDF+ 标注起逐步复算。判读规则依据美国睡眠医学会判读手册的公开条文：判读窗固定为三十秒、窗内多期共存时按占最大部分者判定；一次觉醒的最短时长为三秒，且其之前须有十秒不被打断的睡眠。

关于 7.7 的地位：该小节的两个零模型与两百次重排检验为事后（非预注册）稳健性检验，在第六节预注册文本封存并跑完之后另行执行，脚本与前稿脚本一并提供。本章不把它伪装成预注册的一部分，也不因为它的结果不利而删去；按第六节封存时写下的同一条纪律处理：不利结果原样写入正文。复算时本稿的分期段数与时长与前稿一致（N3 829 段/2853 分钟、N1 998/1421、W 580/4273、N2 1651/9169、R 419/3947），转换总数因非五分期处的断开口径略有出入（本稿 4379，前稿 4405），不影响任何结论的方向。

修订说明（本稿相对前稿）：一、摘要中“间隔分布强烈双峰”改为“短端相对几何期望超额”，并加入 7.7 的更正结果。二、新增 7.7，用分期条件混合几何与段长夜内重排两个零模型重跑，Q2 撤回、Q1 幅度减三分之一、Q4 只剩 3.5 个百分点残余。三、新增 7.8，列明这次自我更正改了本章四处，并收回第六节末“谁对谁错由数据说了算”那句话。四、7.3 加注：Q4 由 Q1 近乎蕴含，不是独立确证。五、第八节由十条扩为十四条，新增生态交错带（排第一位）、睡眠转换动力学、可变时间单元问题三条全球分界，以及 8.12 一整块申报三处站内自撞。六、新增 9.4，生态学与睡眠转换动力学合起来削掉本章“发现了新现象”这句话。七、新增 10.1“最强的竞争解释”一节。八、4.4 与 8.15 两处撤回前稿“谁也没法回头看它”的错话。九、10.2 新增两条不适用范围。十、第十一节新增一段，把 7.7 那次自我更正读成本章命题的一个实例。

人机分工声明：本章的三家选源由人指定，文献核对、检验设计、脚本与初稿由人机协作产出。判据与四条先验预测在拿到任何结果之前即已封存，其后一字未动；Q3 未达标一事按封存时写下的规则照原样写入第七节，未作补救性重述。本稿新增的全球近邻检索、零模型更正与重排检验同样由人机协作完成，所引文献均经逐条核验，重排检验的随机种子已写死在脚本中。

参考文献

8. Zadeh, L. A. (1965). Fuzzy Sets. *Information and Control*, 8(3), 338–353.
14. Berry, R. B., et al. *The AASM Manual for the Scoring of Sleep and Associated Events*. American Academy of Sleep Medicine. (判读窗与觉醒规则条文)
15. Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberyé, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194. (Sleep-EDF 数据库来源)

16. Clements, F. E. (1905). *Research Methods in Ecology*. University Publishing Company. (ecotone 的首次命名)
17. Leopold, A. (1933). *Game Management*. Charles Scribner's Sons. (边缘效应)
18. Whittaker, R. H. (1960). Vegetation of the Siskiyou Mountains, Oregon and California. *Ecological Monographs*, 30(3), 279–338. (移动分割窗法的源头)

第三章 必须回答，所以从不犹豫

必须回答，所以从不犹豫

论一份只允许一个答案的记录如何把分歧压成零，以及险胜与压胜为什么在账上完全相同

Forced to Answer, Never in Doubt: How Single-Label Formats Erase the Margin

摘要

关于“何谓睡眠”，三门互不往来的学问各有一套定身份的办法：法律学说能是什么必须从封闭清单里挑，细胞学说身份由分化路径定，房地产估价学说由条件场下最高最佳的可能用途定。三家共同假定了一件从未说出口的事：每样东西在任一时刻恰好有一个正确的归属，分歧只意味着还没查清楚。本章用估价学自己的规矩推翻它——估价前必须先声明估的是哪一种价值，因为同一物业同一时刻有多个不同且都正确的数。由此提出：归属是若干同时成立的候选里被格式选出的第一名；格式只保留名次，不保留分差。那个分差命名为判余。文中三次真跑，其中一次推翻了本章自己的主张。一：按现行体例合并旧体例的深睡两期后，转换总数减少 16.3%。二：抹掉整整一条分界线，单窗往返只由 1428 降到 1088，险胜的大头一直在旁边那条线上——最强的反驳“旧规则本来就分错了”未获支持。三、触发了预注册的撤销规则。同一份特征按两套体例分别训练，折叠后硬标签不一致率仅 0.59%，概率向量总变差距离中位 0.0027。软输出确实吸收了体例差异，本章原先“两套手册之间”那条分离线因此撤回，靶子从规则版本挪到 argmax 这一步，并顺带得到判余的第一次直接测量：9.09% 的窗判余低于 0.20，出错概率是高判余窗的 5.2 倍，而报告对两者的记录完全相同。文末给出十六条判决性分界、三处反向约束、四条判错条件与三条伦理硬约束。

关键词：判余 | 无痕强判 | 独判格式 | 同时成立的候选 | 物权法定的代价 | 类间距

一·两套都对的规则，一份记录

先摆一件本章数据里的事实，因为后面全部的论证都要回到它。

本章用的四十夜睡眠判读记录，是按上一代体例判的：深睡分成第三期与第四期，分界画在慢波占该时段的比例上——占两成到五成算第三期，超过五成算第四期。现行体例把这两期合并成一个，叫作 N3，理由是两者的临床意义没有可靠区别。

两套体例都是正规的、都被大量文献使用过、在各自的年代都是标准。同一份原始信号，用哪一套都不算判错。

那么，从旧体例换到新体例，记录会变成什么样？

本章数了一遍：四十夜里，按旧体例共有 5260 次分期转换；把第三期与第四期合并之后，只剩 4405 次。减少了 855 次，占 16.3%。

六分之一的“他的睡眠状态变了”这件事，不是因为身体发生了什么，是因为换了一套同样正确的规则。这 855 次转换在旧记录里白纸黑字，在新记录里不存在；而两份记录都没有出错，都不需要更正，都可以拿去发表。

这就是本章要处理的那件事。它不是测量误差——误差可以靠更好的仪器减少。它也不是分类粗糙——把类分得更细只会制造更多这样的位置。它是一种更基本的东西：记录格式规定了每个时间单元只能有一个标签，而在若干个同样成立的候选之间，格式必须挑一个。挑完之后，记录只留下名次，不留下分差。

于是一次险胜和一次压倒性胜利，在纸面上完全一样。

二·三种定身份的办法

2.1 法律学：清单是封闭的

财产法有一条不太为外行所知、却支配着整个体系的原则：物权法定——当事人不能自由创设新的财产权类型，只能从法律预先规定的、数量有限的清单里挑一种。想在一块地上设一种前所未有的权利安排，法律不承认。

这条原则在当代最有力的辩护来自梅里尔与史密斯：财产权与合同不同，合同只约束缔约双方，而财产权约束所有人——因此**每一个陌生人都必须能以很低的代价知道这块地上有什么权利**。允许自由创设，信息成本会落到全体第三人头上（Merrill & Smith, *Yale Law Journal* 110: 1-70, 2000）。

这一家的单因主张是：**一样东西是什么，看它在那份封闭清单上被登记为哪一种就够了**。来路、用途、内部构造，都要先翻译成清单上的某一种才算数。

它自曝的地方，是它自己每天在做的另一半：同一块地上可以同时并存所有权、抵押权、地役权、租赁权、地上权，各归不同的人；一栋楼在法律上同时是一个物和许多个物（专有部分、共有部分、成员权三者不可分离）。**清单是封闭的，而同一样东西可以同时占清单上的好几格**。法律自己造出了这套技术，并且用得很熟。

2.2 细胞学：身份由来路定

第二家来自一门正处在分类学危机中的学问。单细胞测序把细胞的分子状态测成了连续谱之后，“细胞类型”这个用了一百多年的概念开始站不住。

这一家里主张来路的那一派说：一个细胞是什么，取决于它从哪个前体来、经过什么分化轨迹；当下测到的表达谱只是这条路径的一个快照，谱系才是身份。

而这一家自曝得非常彻底。综述文献里写着：**细胞类型至今没有一个一致的、标准的定义**，尽管可重复的研究需要它；常常不清楚按不同表型特征定出来的细胞类型彼此是否一致，**也不清楚哪一个特征才是定义细胞类型的正确特征**；不同模态之间的差异未必高度一致，以至于常常无法确切定义什么是一个细胞类型、无法在类型之间划出清晰边界（Zeng, *Cell* 185: 2739-2755, 2022）。

请注意这段话的形状：它不是说“我们还没测准”，是说不同的正确测法给出不同的分类，而没有更高的标准可以裁决哪一个对。

2.3 房地产估价学：身份由最高最佳用途定

第三家来自一门每天都要给出一个数的行业。不动产估价的核心原则叫最高最佳使用：一宗不动产的价值，不由它现在被拿来做什么决定，而由**法律上允许、物理上可能、财务上可行、且产出最大的那种用途决定**。四个判准依次筛选，剩下的那一种用途定义了这宗物业是什么。

这一家的单因主张是：**一样东西是什么，看它在给定条件场下的最优可能用途就够了**。现状只是历史，条件场才是身份。

它自曝的地方，就是本章要用的那件材料，而且写在行业规范的最前面。

三·为什么这三种办法不能同时对

	单独够用的那一样	它据此把另两样看成
法律学	它在封闭清单上被登记为哪一种	来路是证据，不是身份；用途是清单之外的事
细胞学	它经什么分化路径来的	登记是外部约定；未来用途与身份无关
房地产估价学	条件场下最优的可能用途	登记与来路都是沉没成本，不进价值

三处分歧不在同一根轴上，因此不可裁决；而两两之间是真打架：

法律与估价打的是清单能不能被突破。物权法定要禁止的正是创设清单之外的新型态；而估价师每天在做的事，恰恰是设想这块地还能变成什么——**一家的日课是另一家的病理**。

细胞学与估价打的是朝哪边看。来路派向后看，谱系已经发生、不可更改；最高最佳使用向前看，身份由尚未实现的可能定。因果箭头正好掉头。

法律与细胞学打的是清单该不该封闭。细胞学的危机是“清单不够用、边界连续”，它当成病；法律说清单必须封闭、否则信息成本落到所有人头上，它当成必要条件。**同一件事，一个诊断为病，一个论证为德**。

顺带挡掉第四家：医院、实验室、评估机构不是一家，它们是这些主张落地的场所。本章后面大量使用睡眠判读手册的条文，那是兑现材料，不是第四种立场。

四·三家共同踩着的那句话，与推翻它的那件材料

4.1 三家共同踩着的那句话

三家争的是一样东西的身份该由什么定。填在横线上的那样东西——“身份”——三家都默认它是那种只有一个正确答案的东西：

每一样东西在任一时刻恰好有一个正确的归属；出现分歧，只意味着还没查清楚。

念给三家听，三家都会说这还用说吗。登记簿上一格只能填一种权利；一个细胞在图谱上被分到一个簇；一份估价报告的结论是一个数。三家的记录格式都只允许一个答案，而三家都不认为这个格式本身需要解释。

4.2 推翻它的那件材料，来自估价学自己

拆这条前提，用的东西不能从外头借。借来的只能立成第四种主张，那是入局。能让它失效的，只有三家里某一家早已握在手上、早已印在纸上、却一直没朝这个方向用过的那件事。

它在估价学那里，而且是这一行的入门规矩：

估价之前必须先声明估的是哪一种价值。

因为同一宗物业在同一时刻有市场价值、投资价值、使用价值、保险价值、清算价值、计税价值——它们是不同的数，而且**都是对的**。哪一个才是“这栋楼真正的价值”？这一行的回答是：这个问题没有答案，**它问错了**。规范之所以要求先声明价值类型，正是因为不声明就无法判断报告有没有错——**错的不是数，是没说清在答哪一问。**

这条规矩在估价学自己那里回答的是另一个问题：如何防止报告被拿去用在它不适用的场合。从没有人把它拿来当作关于“身份归属可以合法地多重”的一般证据。

一旦拿来当证据，三家共同踩着的那句话就塌了：**分歧不总是意味着还没查清楚。有一类分歧，是查得越清楚越明确地并存。**

4.3 另两家自己拆掉的部分

细胞学拆掉的是“更好的测量会收敛”这一半：不同模态之间未必一致，而且不清楚哪个特征才是正确的那个——**没有更高的标准可以裁决。**

法律拆掉的是“清单封闭等于答案唯一”这一半：清单确实封闭，而同一样东西可以同时占好几格，且这套技术是法律自己造的、用得极熟。

三半都塌之后，问题变了。不再是“睡眠的身份该由哪一维读出”，而是：

当若干个归属同时成立，而记录格式只允许写一个，被写下那一个与没被写下的那些之间，差多少？这个差被记在哪里？

4.4 三家账本上的空栏，形状是同一个

	有字段的	没有字段的	为什么没有
法律学	判决确认的那一项主张	诉讼中同样有效、只是没被采纳的另一项主张	判决书只记结论，不记它赢了多少
细胞学	被分到某一簇的细胞	同时表达两套标志物的细胞——在标准流程里被判为双细胞污染而剔除	它不符合“一个细胞属于一个类型”
房地产估价学	报告采用的那一种价值	其余价值定义下的那些数	报告只出一个结论

三处空栏是同一个形状：**同时成立、都正确、而因为格式只允许一个答案而不留痕迹的那些归属。**

第二层三家共同踩着的那句话由此浮出来，它不属于哪一家，属于整片地基：

一个正确的答案，等于没有分歧。

所有人都站在它上面，因此谁也没法回头看它。

五·判余

5.1 承重命题

一个时刻的睡眠归属，不是登记簿上写着的那一种，不是由来路谱系定死的那一种，也不是条件场下最优的那一种，而是——若干个同时成立的候选里被格式选出来的第一名；而格式只保留名次，不保留分差。于是一次险胜与一次压胜在账上完全相同，且险胜不留任何空白。

三个否定各对应一家。最后半句是全章的落点，也是第七节那记打脸逼出来的：不留空白。

5.2 三个用词，定死

— 判余：被写下的那个归属，比排在第二位的那个强多少。它是一个名次之间的差，不是一个时长，也不是一个比例。

— 独判格式：规定每个单元只能承载一个归属的记录格式。它不是疏忽，通常有很好的理由——见 9.1。

— 无痕强判：在必须给出答案的位置给出的、判余极小的那个答案。它的特征不是错，是与一个毫无疑义的答案在记录上完全不可分辨。

5.3 辨别装置：一张规则版本对照表，不是一个格子图

判断某处有没有无痕强判，本章不给二维分类格，给一张同一份数据 × 几套都正确的规则的对照表。做法只有一步：把同一批原始记录交给两套以上都合规的规则版本，把结论并排。

	旧体例 (R&K, 1968)	现行体例 (AASM, 深睡合并)	差
分期转换总数	5260	4405	-855 (-16.3%)
“深睡内部的状态变更”	855 次，各有时点	0 次	全部消失
各期总时长	—	—	基本不变
报告是否需要更正	否	否	两份都正确

表里最要紧的是最后两行：时长账几乎不动，而八百五十五次“他的状态变了”凭空消失，且没有任何一方需要更正。差额落在哪里？落在判余上——那八百五十五次里的每一次，第一名与第二名之间的分差都小到可以被一次规则改版翻转，而这个分差从来没有被记下来过。

5.4 一句不含程度词的判据

换一套同样合规的规则重判这份记录，有多少条结论会变？变掉的那些，原来的记录里有没有任何地方标出它们比别处更容易变？

这句话可以拿去问规则文本、问软件版本说明、问报告模板，不用问任何人的判断。

在五个场合各跑一遍，答案互不相同，其中至少两个是查出来的：

场合	答案	来源
睡眠判读	变 16.3% 的转换；原记录里没有任何标记指出这些位置更易变。	查出来的（本章数据）
不动产估价	变，且行业已经知道：规范强制先声明价值类型，等于承认结论随定义而变。这是唯一一处把它写进制度的。	查出来的（估价规范）
单细胞分类	变；且同时表达两套标志物的细胞在标准流程里被当作污染剔除，因此变掉的那些连样本都不在了。	查出来的（流程默认剔除双标细胞）
财产登记	不变——因为清单封闭正是为了让它不变。代价是清单之外的安排根本进不来。	从命题可推
学生分级	变；而分数线两侧一分之差与二十分之差，在成绩单上是同一个等级。	从命题可推

第三条反过来改了本章：本章原以为被排除的候选至少还在数据里，只是没被选中；单细胞流程的默认剔除说明，在有些账本上，判余小的那些个体连原始记录都不保留。这比本章原来设想的更彻底，第九节据此收窄了主张。

5.5 本章主张的那一格（收窄）

写到这里必须把话缩紧，否则本章会被三个方向同时吸收。

“判定有不确定性”这件事**不是本章的**——信号检测论、拒识选项、主动学习、概率化睡眠图，四路人马已经把它做了半个世纪（见 8.1）。“落在分类之外的东西没有痕迹”**也不是本章的**——那是残余类（见 8.3）。“切法改变结论”**还不是本章的**——那是可变单元问题（见 8.6）。

本章只主张剩下的这一格：

一次 argmax 把一个概率向量压成一个名次，压掉的那部分技术上完全保得住；而记录格式没有保住它，于是险胜与压胜不可分辨。

（初稿这一格写的是“两套都合规的手册给出不同结论，而记录不标注版本敏感位置”。7.5 的实验推翻了它——软输出能吸收体例差异。靶子因此从“两部手册之间”挪到了“从概率向量到一个标签的那一次压缩”。原文照录在此，供核对本章有没有改口径。）

这一格四路人马都够不着：判定不确定性那一族（8.1）确实测出了这个量，但**没有一个进入报告格式**；残余类讲的是分类之外；可变单元问题讲的是单元切法而非那次压缩。本章所有的处方也只针对这一格——**在给出标签的同时保存判余**。

六·检验：一个可当场执行的最小实验

6.1 先把最难对付的那个反驳摆出来

有一句话可以把本章到此为止的全部内容一口吃掉，而且它听上去毫不勉强：**旧规则本来就分错了**。

现行体例之所以把深睡两期合并，正是因为学界认定这个区分没有可靠的临床意义。照这个读法，那 855 次消失的转换不是判余，是**被纠正的错误**——旧手册划了一条不该划的线，新手册把它抹掉了，这是进步，不是发现。

本章不挑一个好打的靶子，就打这一个。它能覆盖本章的核心数据，而且不需要引入任何新概念。

要排除它，得看两条预测的分岔：

— **纠错说预测**：这类现象是一次性的。错的线被抹掉之后，剩下的线都是对的，险胜应当随之大幅减少。

— **判余说预测**：判余不属于某一条线，它属于**每一条线**。抹掉一条，剩下的线上的险胜照旧，甚至更多——因为被抹掉那条线两侧的犹豫，会转移到相邻那条线上。

证伪条件写成这个形状：

控制类别总数之后，若“抹掉一条分界线”与“险胜总量下降”之间不存在近似成比例的关系，则纠错说为伪；若成比例下降，则本章的判余概念可由“规则质量”解释，本章为伪。

这一条第七节跑了。

最小实验：取任意一批已判读的整夜记录，用两套都合规的规则版本各判一遍，比对两份结论。规则版本不必是新旧体例，也可以是同一手册里“推荐”与“可选”两条并列的规则——手册自己明确写着这两类都是可接受的判法。

— **若两份结论的差异接近于零**，说明规则版本之间的选择不影响身份归属，判余这个概念多余，本章错。

— **若差异显著，而原记录里没有任何字段标出哪些位置更易变**，本章成立。

这个实验的全部成本是跑两遍软件。它比一条五年后的赌注硬，因为它现在就能裁决。

预注册（跑之前写死）

数据：Sleep-EDF Expanded 1.0.0 的 sleep-cassette 前四十份判读标注，未挑选，公开可复算。

— **描述性事实（不是预测，设计本检验之前已看到，不计入命中）**：原始标注里判读者未给出分期的时间单元占 5.14%。

— T1：把深睡两期合并后，转换总数减少**超过 8%**。

— T2：深睡两期之间的直接转换，占旧体例全部转换的**超过 6%**。

- T3：用前一个时间单元的标签预测当前单元，准确率**超过 85%**——即单个单元的标签里多数信息来自邻居而非它自己。
- T4：判读者的弃权在时间上**成簇**，弃权段平均长度**超过 3 个单元**。
- 撤销规则：T1 低于 3%，**“换一套同样正确的规则会改变身份”**太弱，本章须改写；T3 低于 70%，**“标签的信息主要来自邻居”**判错，须删。
- 代理问题（先写明）：新旧体例的差别不止深睡合并一项，本次只测这一项，故 T1 是全部差异的下界；一阶预测准确率高也可能只反映睡眠本身的持续性，是代理不是直接测量。

七·跑出来的东西，以及打脸的那一条

- 四十夜，四三三九一个时间单元。
- **T1 命中**，幅度是预设下限的两倍：转换 5260 – 4405，**减少 16.3%**。
- **T2 命中**：深睡两期之间的直接转换 855 次，占旧体例全部转换的 16.3%——即这次减少全部来自这一处，没有别的来源。
- **T3 命中**：用前一个单元预测当前单元，准确率 **89.8%**。十次里有九次，一个时间单元的标签可以由它的邻居猜出来。
- **T4 未命中**，且方向相反、幅度极大：预设弃权段平均长度超过 3 个单元，**实测 1.07**，中位数 1，最长 2；裁剪后弃权总共只有 58 段、半小时、占 0.14%。

7.1 打脸那条改了什么

按预注册规则，T4 判未命中，原表述不保留。但追问它为什么错，得到的东西比原预测硬得多。全库原始标注里弃权占 5.14%，而剪掉记录首尾的离床期之后只剩 0.14%。也就是说：**判读者的弃权几乎全部发生在格式允许留白的地方（人还没上床、已经起床），而在真正的睡眠期间几乎从不弃权。**

本章原来的设想是：格式的代价会以空白的形式留下痕迹，弃权多的地方就是分歧大的地方。数据说不是。在必须给出答案的位置，判读者总能给出一个答案；而这个答案与一个毫无疑义的答案，在记录上完全一样。

于是命题被改成现在这一句：

独判格式的代价不表现为空白，而表现为一个看上去毫无疑义的标签。

这一改让本章更难被消化。如果代价是空白，那只要统计空白就能处理；空白也是一种记录。而现在的版本说：**没有任何东西可以统计，因为格式没有留下任何可供统计的东西。**唯一能把它显形的办法，是第六节那个最小实验——换一套规则，看什么变了。

7.2 补充真跑：抹掉一条线之后，险胜去哪儿了

6.1 那条竞争解释要用数据分岔，本节跑它。这一段是探索性的，不在预注册之内，标明如此。

先定一个可数的东西作为“险胜”的代理：**单窗往返**——前后是同一类、中间只夹一个时间单元的那种插曲。它形状最干净：两边都认为是 A，中间那一个被判成 B 而立刻回来，第一名与第二名之间显然很近。

	旧体例（深睡分两期）	现行体例（合并为 N3）
被抹掉那条线（3–4）上的往返	713 次，其中单窗 356	线已不存在
N2–N3 线上的往返	—	1374 次，其中单窗 666
N1–N2 线上的往返	—	576 次，其中单窗 192
单窗往返总数	1428	1088

抹掉一整条分界线，单窗往返只从 1428 掉到 1088，**减少 24%**；而类别数从五类降到四类，减少 20%。**险胜的总量几乎只是随着线的条数按比例走，没有额外的下降。**

更要紧的是它们的分布：被抹掉那条线上原有 356 次单窗往返，而它旁边那条线（N2 与 N3 之间）在合并后有 **666 次**——**险胜的大头一直不在被判为“分错了”的那条线上。**如果旧手册的问题是划了一条不该划的线，抹掉它之后应当清静了；数据说没有，剩下的每条线上照旧。

判定：纠错说预测的“抹掉一条线，险胜按该线的份额下降并到此为止”没有出现；判余说预测的“每条线各有各的险胜，抹掉一条只是少一条线”出现了。**6.1 那条竞争解释在这份数据上未被支持。**

但这不是终局，两处必须写明：其一，本次只观察了一次规则改版， $n=1$ ；要真正裁决，需要看多次改版（判读手册各版本之间的规则改动）是否每次都产生新的一批。**这一条本章没有做。**其二，单窗往返只是险胜的代理，它捕捉的是“立刻返回”这一种形状，捕捉不到“险胜之后没有返回”的那些——因此本章所有的数字都是下界。

7.3 前两次检验改了本章三处

- 1. 删掉“弃权是分歧的痕迹”这条，并把未命中写在正文里而不是脚注里。
- 2. 把命题的落点从“痕迹”挪到“不可分辨”：本章不再主张能在现行记录里找到判余的踪迹，只主张判余可以通过规则对照被显形。
- 3. 把主张的强度调低：T3的高准确率也可以由睡眠本身的持续性解释，本章不据此断言“标签主要由邻居决定”，只把它列为一个与判余相容的读数。

7.5 第二次真跑：本章自己的一条分离线被推翻了

8.1给了一条判定形式：把概率化睡眠图分别跑在两套标注体例上，若跨体例的差异随之消失，则本章该处被覆盖。前一稿写着“本章没有做”。这一稿做了。

设计（预注册在跑任何比较之前写死）：取六位受试者的原始脑电（Fpz-Cz，100 赫兹），按三十秒切窗，算同一份谱特征；两套体例共用同一份特征矩阵，唯一的差别只在标签。旧体例训一个六类模型，现行体例训一个五类模型，同一算法、同一超参、按受试者分析的交叉验证。再把旧模型的软输出信息无损地折到现行空间（令 $P(N3) \equiv P(S3)+P(S4)$ ），与新模型逐窗比较。共 6285 个窗，两个模型对人工标注的准确率都是 79.3%。

预注册写的是前十位受试者，实际只取到六位（其余下载未完成），这是一处偏离，记在这里。

结果（四条读数，三条与预设相反）：

读数	预设	实测	
V1 折叠后两模型硬标签不一致率	> 3%	0.59%	未命中
V2 逐窗概率向量总变差距离中位	> 0.05	0.0027	未命中
V3 不一致处两边的最大概率中位	均 > 0.60	0.500 / 0.501	未命中
V4 一致处的最大概率中位（对照）	应更高	0.927 / 0.929	成立

预注册的撤销规则被触发：V1 低于 1% 且 V2 低于 0.02。按写在前面的规矩，本章 8.1 的第一条分离线——“软输出处理不了两套规则之间的分歧”——就此撤回。它错了，而且不是运气不好：把两个类的概率相加本来就是一次信息无损的折叠，同一份特征上的多项模型对合并类的拟合几乎等价。这一条在设计阶段就该被看出来，我没有看出来。

V3 也判错，且方向相反：两模型不一致的地方，两边的最大概率中位都只有 0.50——不一致只发生在两边都不确信的位置，不是“各自都很确信而彼此不一致”。这句话一并删除。

7.6 撤回之后剩下什么，以及判余的第一次直接测量

一条分离线塌了，本章并没有跟着塌，但它的落点必须挪。

挪到哪里：既然软输出能把体例差异吸收掉，那么问题就不在规则版本上，在 argmax 上。旧体例里 S3 与 S4 之间那种此消彼长，在概率向量里一直都在；把两类相加它照样在；它是在取最大值、写下下一个标签的那一步被扔掉的。本章原来把靶子对准了“两套手册”，对偏了；真正的靶子是从概率向量到一个标签的那一次压缩。

这一挪让 8.1 的第二条分离线从配角变成主线，而且它被这次实验加强了：技术上完全保得住（本实验就保住了），而报告格式没有保住——那么格式不保住就纯粹是格式的事，不是技术做不到。这正本章的那一格，现在它有实验背书。

顺带，本章最想要的那个数出来了。判余在这里有了第一次直接测量：每个窗上最大概率减次大概率。（探索性，非预注册。）

读数	值
判余中位数	0.873
判余 < 0.20 的窗	9.09% (571 / 6285)
判余 < 0.10 的窗	4.47%
判余 < 0.05 的窗	2.21%
判余 < 0.20 处与人工标注的一致率	41.3%
判余 ≥ 0.50 处与人工标注的一致率	88.6%
低判余处的出错概率	是高判余处的 5.2 倍
低判余窗中落在转换点前后一窗内的比例	54.6%（而这类窗只占全部的 32.5%）

三件事同时成立：判余是可测的；它确实指示险胜（低判余处出错率高五倍以上，不是噪声）；而报告对这两类窗的记录一模一样——一个判余 0.05 的窗与一个判余 0.95 的窗，在纸面上是同一个标签。

九个窗里有一个是险胜，其中五成挤在状态交界处；而这九分之一在任何一份现行报告上都找不到。这是本章全部主张压缩之后剩下的那一句，现在它是一个数，不是一个论证。

必须同时写明三条限制：模型是简单的谱特征分类器，不是临床级自动分期，准确率只有 79.3%，因此判余的绝对数值不可外推；标签由同一份人工标注机械折叠而来，不含两部手册其余规则的差异；六位受试者，样本很小。

7.7 还跑不动的那一条

前一稿在这里写着“判余本身测不了”。7.6 之后这句话只对一半：在一个自己训的模型上，判余测出来了（9.09% 的窗判余低于 0.20，出错率高五倍）。测不了的是另一件事——人工判读时的候选排序。判读者心里那个“差一点就判成 N2”，没有任何字段承接它；本章那 9.09% 是模型的判余，不是判读者的判余，二者是不是同一个量，本章答不了。

这一条不是遗憾，是本章的第三层主张，也是最便宜的一件可做之事：判读软件在给出标签的同时保存第二候选与两者的分差，成本接近于零；人工判读界面加一个两键的“拿不准”记号，成本也接近于零。保存之后，判余就从一个论证变成一个字段。

前两层可以塌，这一层塌不了。而这一稿之后，第一层已经塌过一次（7.5），第三层照旧站着——这正是分层引用的用处。

八·与十六种说法的分界

每条写四件：它说到哪一步、分离线、判定形式。凡是只能写出“本章更强调”“视角不同”的，等于承认被覆盖，不保留。

第一位排在最前，因为初稿把它漏了。这不是谦辞：漏掉它的那份初稿，读起来与一份没有漏掉它的稿子完全一样——本章写的正是这件事。

8.1 概率化睡眠图，以及“判定余量”的既有家族

它说到哪一步：二〇一八年，一组研究者提出了概率化睡眠图——不给每个时间单元一个标签，而给一个各期概率的分布，并用它做发作性睡眠病的辅助诊断（Stephansen et al., *Nature Communications* 9: 5229, 2018）。本章第三层主张“保存第二候选与两者的分差”，在这里已经被实现并发表了七年。

而它不是孤例，它属于一个很老的家族：

— 拒识选项（Chow, *IEEE Transactions on Information Theory* 16: 41–46, 1970）：分类器在置信度低于阈值时可以拒绝作答，且存在最优的错误—拒识权衡。这是“判余小就别硬判”最早的形式化，比本章早五十五年。

— 信号检测论的判准位置与可辨度：把“答案”拆成刺激分布与判准两件事，本来就是睡眠判读的老家伙——判读者间差异有很大一部分正是判准位置的差异。

— 主动学习的不确定性采样（Lewis & Gale, 1994）：专门挑第一名与第二名分差最小的样本去标注。它甚至已经把判余当作一个可排序的量在用。

—政治学的**边际选区**：多数党在某选区赢多少，决定该选区是不是“边际的”；边际性是一个成熟的、可比较的量。

⚠️ 本章原先在这里写了三条分离线，第一条已被本章自己的实验推翻，见 7.5。撤回的内容照录如下，以便读者核对本章有没有改口径：原第一条主张“它们全部在一套规则之内，而本章处理的是两套规则之间的分歧；把概率化睡眠图跑在两套体例上会得到两份各自很确信而彼此不一致的图”。**实测不是这样**：折叠后两模型的硬标签不一致率只有 0.59%，概率向量的总变差距离中位 0.0027，且不一致只出现在两边都不确信处。软输出确实吸收了体例差异。

撤回之后剩下两条，而它们比原来的三条硬：

1. 它们没有进入报告格式，而这次实验证明技术上完全保得住。概率化睡眠图问世七年，临床报告仍然是每个时间单元一个标签，各期时长与效率仍然建在互斥分区上。7.5 那次实验里，判余在概率向量里一直都在、折叠也没有把它折掉——**能保住而没有保住，那就不是技术问题，是格式问题**。这一条不是本章的功劳，是本章的那一格；而它现在有实验背书。

2. 它们不改变下游读数的地基。拒识选项能拒答某个样本，概率化睡眠图能给一条概率曲线，但没有任何一个装置会让“各期时长”这类读数承认自己建在一次 $\arg\max$ 上。报告里的睡眠效率是一个百分数，它的分子分母都由硬标签算出，而算之前那一步压缩不留任何记号。

新的判定形式：若某份临床报告在给出各期时长的同时给出该份记录的判余分布（哪些窗是险胜），则本章所指的缺口已被填上，本章的处方多余；就本章所查的报告格式而言，没有这一栏。

这一位应当与另两个正面对照并排读：天文学一百多年前把昼夜之间分成三级曙暮光（见 8.5），生态学一个多世纪前给两个群落之间的交错带立了名字、量了宽度、专门立项研究；而**概率化睡眠图是第三个——技术有了，格式没改**。三处合起来说明本章要指的不是无知，是记录格式的惯性。

本题域里有三支力量，从三个方向逼近“一个标签不够用”。本章与它们不是对立，是接着往下走一步。

局部睡眠（Krueger & Obál, *Journal of Sleep Research* 2: 63–69, 1993; Vyazovskiy et al., *Nature* 472: 443–447, 2011）说：睡眠不是整个大脑的一个状态，而是皮层柱一级的性质；清醒的动物身上，个别神经元群会进入类似慢波睡眠的离线状态，而行为上完全清醒。**分离线**：局部睡眠说的是**空间上的多重**——不同部位状态不同；本章说的是**判定上的多重**——即使在单一部位、单一判读者、单一时刻，标签也只是若干候选里的第一名。二者独立：一个全脑完全同步的夜晚，判余照样存在。**判定形式**：若在完全同步的记录里判余为零，本章被局部睡眠覆盖；本章数据里那 855 次消失的转换全部发生在深睡，而深睡正是全脑同步性最高的时段——**本章对**。

判读者间信度（ κ 系数，睡眠医学几十年的常规读数）已经在数分歧了。**分离线**： κ 数的是**判读者之间的不一致**；判余是**同一个判读者内部的余量**。一份只有一位判读者的记录算不出 κ ，而判余照样存在。**判定形式**：若判余可由 κ 完全预测，本章被覆盖；快速眼动期的 κ 通常最高，而它是否判余也最小，本章数据不足以裁决，如实记为**未判**。

多标签分类与软标签（机器学习，方法学上最省力的近邻）说：一个样本可以同时属于多类，或带一个概率分布。**分离线**：多标签的前提是**格式允许多标签**；本章的对象恰恰是不允许的那种格式，而它不允许是有原因的——下游一整套读数（各期时长、睡眠效率、各期百分比）建在互斥分区上，改成多标签会同时废掉它们。**判定形式**：若把判读改成多标签而下游读数不受影响，本章这一层多余；实际改不动，见 9.3。

三家的共同点，也是本章与它们分开的地方：它们都在找**更多的标签**——更多的部位、更多的判读者、更多的类。本章指出缺的不是标签数，是**名次之间的分差**；标签加到多少个，第一名与第二名之间差多少这件事仍然不会被记下来。

8.3 残余类（Bowker & Star, *Sorting Things Out*, 1999）

它说到哪一步：任何分类系统都会产生残余类——“其他”“未特指”“不详”——落进去的东西失去可见性，而这些残余类承载着分类系统的政治后果。这是分类学后果研究的奠基之作。

分离线：残余类是落在分类之外的东西；判余是落在分类之内、只是差一点就落到别处的东西。一个在外面，一个在边上。

判定形式：残余类可以通过增设类别消化——把“其他”拆开就少一些；**判余不能**，增设类别只会制造新的边界、新的判余。本章数据正是这样：旧体例把深睡分成两期（类更多），产生了 855 次判余极小的转换；新体例合并成一期（类更少），那 855 次消失了。**类数增减都不消除判余，只是搬家。这条是本章与它最硬的一处分界。**

8.4 权利束 (Hohfeld, *Yale Law Journal* 23: 16, 1913)

它说到哪一步：所有权不是一个整体，是一束可分离的法律关系——权利、特权、权力、豁免各有对应。这是“同一物上并存多种归属”最早、最完整的分析。

分离线：权利束里的每一种权能都有名义，都可登记、可转让、可救济。判余讲的是那些没有名义的候选——它们同时成立，却因为格式只出一个结论而不留痕迹。

判定形式：若被排除的候选可以像权能那样被逐一命名并登记，本章被权利束覆盖；本章数据里那 855 次转换在新记录中没有任何对应字段，无法命名，**本章对**。

8.5 物权法定 (Merrill & Smith, 2000)

它说到哪一步：清单必须封闭，否则信息成本落到全体第三人头上。

分离线：本章不反对它。本章接受这个论证，只要求把代价记下来。清单封闭确实降低了第三人的信息成本，代价是判余变得不可见；而这个代价从来没有出现在任何一份成本收益分析里。**接受一个论证并指出它未入账的代价，比反驳它更难被反驳。**

判定形式：若能证明清单封闭不产生任何判余（即每一次归属都是压胜），本章错。

8.6 可变单元问题 (MAUP / MTUP, Openshaw; Cheng & Adepeju, 2014)

它说到哪一步：把空间或时间切成不同的单元，同一份数据会给出不同的统计结论；聚合、分段、边界三种改动都会改变结果。

分离线：可变单元问题讲单元怎么切；本章讲单元切法固定不动时，每一次归属的余量。两者正交：把单元固定死，判余照样存在；把类别固定死，单元问题照样存在。

判定形式：若把单元切法完全固定后判余消失，本章被它覆盖；本章的 855 次全部发生在单元切法不变（都是三十秒）而类别定义改变的情况下——**本章对**。这一处是本章与它最干净的一次分工。

8.7 周期性交替模式 (CAP, Terzano 等, 1985 起)

它说到哪一步：睡眠里存在一种一秒级的周期性交替结构，它在三十秒的判读窗里被平均掉；该领域自己早就指出，三十秒的格子与睡眠过程的实际尺度不合。这是本题域内部对“格子毁掉了什么”最完整的一套回答。

分离线：CAP 讲的是时间分辨率不够，细结构被平均掉——提高分辨率就能看见。判余讲的是任何分辨率下都看不见的东西：它不是被平均掉的，是被格式排除的；把窗宽改成一秒，每个一秒窗仍然只有一个标签，仍然只留名次不留分差。

判定形式：若把判读分辨率提到一秒后判余消失，本章错、CAP 足够；若提到一秒后仍需若干候选里选第一名且分差不被记录，本章对。

8.8 同批第一篇：追认

上一篇讲一段时间要成为一个可数的段落，需要被后续事件追认。**分离线：**那一篇问单位能不能成立，本章问单位内部的归属还有几个候选。**判决性预测：**把最小可记单位调细，上一篇的未追认段比例会上升，而判余不变——因为判余不是粒度的函数，是候选数与分差的函数。

8.9 同批第二篇：界带

上一篇讲两个类之间那条被记成线的地带其实有厚度。**分离线：**界带是位置上的东西（在哪儿），判余是每一次判定上的量（差多少）。**互补预测（可检验）：**界带处的判余应当最大，两条读数在同一份记录上应呈负相关——界带越宽的位置，判余越小的判定越密集。若二者不相关，说明本章与上一篇处理的是两件互不相干的事，那也是有用的结论。

8.10 自家本书：合位

站上另一篇讲的是三个单位被持续维持为同一群人，而维持不是免费的，并给出合位率与缝合负荷两个读数。这是本章最近的一位邻居，近到不划界就是换皮。

分离线：合位率可以观察到下降——三个单位分开了，看得见，能记账。判余在账上恒等于零：格式只允许一个标签，因此分差不是掉下去，是压根不能被记下来。前者是一个会变的量，后者是一个结构上不存在的量。

判定形式：若能设计一份记录，让判余像合位率那样直接读出，则两者同族、应合并；本章认为不能——要读出判余必须改动记录格式本身，而合位率不需要。

8.11 自家本书：一件是几件、类内距

《一件是几件》讲生成成本趋零之后计数单位失去下界。**分离线：**那一篇是跨时间的件数（把这一轮全部提交合并同类项还剩几项），本章是同一判定内部的候选分差。一个问总数，一个问名次。

《类内距》讲一份规范只写了什么算同类、没写同类之间有多近。**分离线极干净：**那一篇是类内的距离，本章是类间的距离——具体说，是这个个体到它所属类与到最近的另一类之间的距离差。两条读数可以同向也可以反向：一个类内极紧、而边缘个体离邻类极近的分类，类内距小而判余也小。这一对是互补的两半，合起来才是一份分类规范的完整代价表。

8.12 划完之后：这些邻居共同站着的那块地

逐个问“它结构性地看不见什么”，并写出“一旦承认，它自己会怎么塌”：

— κ 把判读者当作变异来源，因此看不见单个判读者内部的余量；一旦承认，只有一位判读者的记录就不再是“无法评估信度”，而是“信度被结构性地隐藏”。

— 残余类把类之外当作问题所在，因此看不见类之内的边缘；一旦承认，增设类别这个标准处方就不再有效。

— 可变单元问题把单元当作变量，因此看不见类别定义这个变量；一旦承认，它的敏感性分析要多加一整个维度。

— 各期时长与效率把互斥分区当作给定，因此看不见分区之外的可能；一旦承认，它们的加和性就不再是中立的记账方式，而是一项实质假设。

背后是同样被当作给定的东西：一个正确的答案，等于没有分歧。所有人都站在它上面，因此谁也没法回头看它。

九·三家反过来给本章加的约束

9.1 法学削掉的：独判格式不是错误

本章原本想写的更强的一句是：独判格式是一种记录上的缺陷，应当被修正。

法律不允许。物权法定的论证是硬的：财产权约束所有人，因此每个陌生人都必须能以极低成本知道这上面有什么；允许自由创设，信息成本会落到全体第三人头上。独判格式是为第三人省钱的装置，不是疏忽。

削掉之后，本章只能主张：这个装置有一笔未入账的代价，应当被入账。不主张废除它。这一处动的是本章的价值排序。

9.2 房地产估价学削掉的：判余不是一个客观量

本章原本想写的更强的一句是：每一次判定都有一个客观的判余，只是没被记录。

估价学不允许。它的规矩是先声明价值类型再估价——也就是说，那几个“都正确的数”并不是同一个量的几个测量值，它们是几个不同的量。同理，判余只有在指定了候选集与打分规则之后才有意义；换一套候选集，判余是另一个数。

削掉之后本章只能主张：给定一套候选集与打分规则，判余可以被记录；不能主张存在一个不依赖规则的判余。这一处动的是适用范围。

9.3 细胞学削掉的：不能靠增设类别或多标签解决

本章原本想写的更强的一句是：把类别定义得更细、或允许多标签，就能把判余记下来。

细胞学不允许。它的经验是反的：单细胞技术把分辨率提高了几个数量级、把类别数提高了几个数量级之后，边界并没有变清楚，而是变得更连续；而同时表达两套标志物的细胞在标准流程里被当作双细胞污染剔除——类别越多，被判为“不该存在”的个体也越多。

削掉之后本章的处方从“改分类”缩到两条纯程序性的要求，不改动现行分类：每份结论标注所用规则的版本号，以及保存第二候选与分差。第二条已有现成实现（8.1），因此本章对它没有首创权，只主张把它写进报告格式；第一条是本章自己的，也是本章唯一要求新增的东西。这是最实质的一次让步，也让 8.3 与 8.4 的两条分界得以成立。

9.4 判错条件

1. **（已判，本章错）** 若软输出能吸收规则版本差异，则本章原先关于“两套手册之间”的那条分离线不成立。**7.5 实测如此，该条已撤回**，靶子改为 $\arg\max$ 这一步。这一条列在这里，是因为它是本章至今唯一一条被自己的数据推翻的主张。

2. 若在判读手册的历次改版之间（不只本章观察的这一次），险胜总量随类别数按比例下降并到此为止，则纠错说成立、判余说错。

3. 若模型判余与判读者自报的“拿不准”不相关（相关系数低于 0.3），则本章用模型判余代理判读者判余的做法不成立，7.6 那个数只能说明模型的事。

4. 若某份临床报告格式里已存在判余或等价字段，则本章的处方多余。

十·不适用、反噬与伦理

10.1 本章不适用的地方

- 不适用于允许多重归属的记录格式——那里没有被排除的候选。
- 不适用于诊断。判余不指示任何疾病，本次数据与任何临床变量的关系一律未检验。
- 不适用于主张某一套规则版本更对。本章的全部论证建立在“两套都对”之上。
- 不适用于给出判余的绝对数值。见 9.2。
- 不适用于替代现行分期读数。本章主张加一个字段，不主张改分类。
- 不适用于主张判余是一个新发现的量。给判定标一个置信度，五十五年前已有形式化（8.1）；本章只主张它没有进入记录格式，且没有任何装置处理跨规则版本的分歧。

10.2 反噬预言

这条主张若被写进制度，最先出事的地方是它自己：平均判余会被做成一项质量指标。一旦某机构被要求“判余越大越好”，最省力的达标办法是把候选集缩小——候选少了，第一名与第二名的差自然拉大，指标立刻漂亮，而被排除的东西更多。或者更省事：把打分规则调成极化的，让第二名永远很低。

这条反噬比前两种更难躲，因为判余的两个构件——候选集与打分规则——都握在被考核者手里。前两篇的读数至少还依赖外部数据，这一个不依赖。**出路我没有找到**。下面三条不是解决，是把它围起来。

10.3 伦理硬约束（三条，缺一条本章不可发表）

1. 判余是记录规则的属性，不是判读者的属性，更不是被记录者的属性。它不得用于评价任何人的判读水平，也不得进入任何个人的健康档案作为“结果可疑”的标记。

2. 不得为了把判余做好看而缩小候选集或调整打分规则。任何对候选集与打分规则的改动，须与该改动对判余分布的影响一并公布。

3. 一份单一标签若被用来限制某个人的工作、保险或驾驶资格，出具方须写明所用规则的版本号；当事人提出时，须用另一套合规版本重判并把两份结论一并交给他。这一条能写得这么具体，是因为本章的最小实验本就只是跑两遍软件——做不到的理由不成立。

十一·用本章的判据检提出这道题的那个指令

本章的判据是：换一套同样合规的规则重判，有多少条结论会变；变掉的那些，原记录里有没有标出它们更易变。把它对准本章自己的来路。

本章不是自己长出来的。它来自一条指令：用这三门学科，答“何谓睡眠”。换一套同样合规的规则——换三门别的学科、换一个提问的次序、换一位下指令的人——会不会得到另一条承重命题？

会。而且这份稿子里没有任何一处标出它有多容易变。它读起来和一条毫无疑义的结论完全一样：有三家、有前提、有推翻材料、有数据、有分界。这正是本章所描述的那件事——在必须给出答案的位置给出的答案，与一个毫无疑义的答案不可分辨。

按本章的处方，正确的做法不是宣称本章更可靠，是把第二候选写出来。所以写在这里：本章在收敛之前活下来的候选还有另外几个，其中最接近的一个是“候选集本身由谁提供”——即在三家之外，那个决定“哪三家上场”的位置。它没有被选中，不是因为它更弱，是因为本章的数据碰不到它。这一句就是本章自己的判余。

还有一层比这更近。本章的初稿漏掉了概率化睡眠图——一个七年前发表在顶刊、把本章第三层主张直接做出来的东西。漏掉它的那份初稿，与一份没有漏掉它的稿子，在纸面上完全一样：同样有三家、有前提、有推翻材料、有数据、有分界，读起来同样毫无疑义。这就是本章所说的无痕强判，发生在本章自己身上。它被发现，靠的正是本章推荐的那个动作——换一套规则重查一遍，看什么变了。

第二笔账在样本上。四十夜，一个数据库，一套判读传统，一批肯到实验室去睡的人——凡是没被收进公开库的睡眠，这里一夜也没有。而这件事的性质，与单细胞流程剔除那些同时表达两套标志物的细胞是一回事：不是判它们错，是它们进不了样本。本章一边指认这个动作，一边在做它。

第三笔账在可引用性上。判余要真正测出来，得有判读当时的候选排序，而现行记录只留最终标签。所以本章分三层，希望被引用的是最下面那一层：给出标签的同时，把第二候选和两者的分差一并存下来。上面两层——新名字、对照表——都可以被推翻；这一句推翻不了，它只是一个字段。

注释与材料

数据取自 Sleep-EDF Expanded 1.0.0 版的 sleep-cassette 部分，按文件名排序取前四十份判读标注，未加挑选。预注册文本、解析脚本与统计脚本随文附上，复算不需要向作者索取任何东西。旧体例指 Rechtschaffen 与 Kales 一九六八年的判读手册（深睡分第三、第四期），现行体例指美国睡眠医学会手册（两期合并为 N3）；两者在各自年代均为标准，本章不判其优劣。

人机分工声明：三门学科由人指定；文献核对、对照表设计、脚本与初稿由人机协作产出。四条先验预测与撤销规则在看到任何结果之前封存，其后未动；T4 未命中一事按封存时写下的规则原样写入第七节，并据此改写了承重命题的最后半句。

参考文献

5. Appraisal Institute. *The Appraisal of Real Estate* (15th ed.). (最高最佳使用四判准；价值类型须先行声明)

6. International Valuation Standards Council. *International Valuation Standards*. (价值基础的声明义务)

16. Openshaw, S. (1984). *The Modifiable Areal Unit Problem*. Geo Books.

修订说明（本稿相对初稿的六处改动）

初稿完成后自查出一处重大缺席，据此重修。逐条列明，供比对：

1. 新增 8.1“概率化睡眠图，以及判定余量的既有家族”，并排在分界节第一位。初稿全章未提概率化睡眠图（2018）、拒识选项（1970）、信号检测论、不确定性采样（1994）——而本章第三层主张“保存第二候选与分差”正是它们做过的事。本节给出三条分离线（一套规则内 vs 两套规则间/技术已有而格式未改/不处理规则版本本身），并把它与天文曙暮光分级、生态交错带并列作第三个正面对照。

2. 新增 5.5“本章主张的那一格”。主张由“独判格式抹掉分歧”收窄为“同一份记录在两套都合规的手册下给出不同结论，而记录不标注版本敏感位置”——把判定不确定性（8.1）、残余类（8.3）、可变单元问题（8.6）三片地明确划出去。

3. **新增 6.1“先写出最强的竞争解释”**。初稿没有这一节。补上那句最难反驳的话（“其实不就是旧规则分错了吗”）与两条预测的分岔，并把证伪条件写成可分岔的形状。

4. **新增 7.3 补充真跑（探索性，非预注册）**：抹掉一条分界线之后，单窗往返由 1428 降至 1088（-24%），类别数降 20%，而险胜的大头一直在旁边那条线上（合并后 N2 与 N3 之间 666 次，被抹掉那条线原有 356 次）。纠错说未获支持。同时写明两处局限：n=1 次改版；单窗往返只是险胜的下界代理。

5. **新增 9.4 判错条件**，并把 9.3 的处方改为两条（标注规则版本号／保存第二候选），明确本章对后者没有首创权。

6. **第十一节新增自反一段**：初稿漏掉概率化睡眠图这件事，本身是本章命题的一个实例——那份漏掉它的稿子与没有漏掉它的稿子在纸面上完全一样，而它被发现靠的正是本章推荐的动作（换一套规则重查一遍）。

修订说明第二轮（把 8.1 那条判定形式真跑了，结果推翻了本章）

第一轮末尾列了三件未做的事，其中第一件——“概率化睡眠图在两套标注体例上各训一次”——本轮做了，而它推翻了本章的一条分离线。逐条：

7. **新增 7.5**：六位受试者的原始脑电，同一份谱特征、只换标签体例，按受试者分折交叉验证。四条读数三条与预设相反，**预注册的撤销规则被触发**（ $V1 = 0.59\% < 1\%$ ， $V2 = 0.0027 < 0.02$ ）。8.1 原第一条分离线撤回，原文照录在正文里供核对；V3“各自都很确信而彼此不一致”一并删除。撤回的理由写明是**设计缺陷而非运气**：概率相加本就是信息无损的折叠。

8. **新增 7.6**：靶子从“两套手册之间”挪到“argmax 这一步”；8.1 的第二条分离线因此从配角升为主线，且被这次实验加强——**技术上保得住而格式没保住，那就不是技术问题**。

9. **判余的第一次直接测量（探索性）**：判余中位 0.873；**9.09% 的窗判余低于 0.20**；这些窗与人工标注一致率 41.3%，而高判余窗 88.6%，**出错概率高 5.2 倍**；低判余窗有 54.6% 挤在状态交界前后一窗内（这类窗只占全部的 32.5%）。三条限制同时写明（模型准确率仅 79.3%、标签由人工标注机械折叠、只有六位受试者）。

10. **5.5 那一格与 9.4 判错条件同步改写**；9.4 第 1 条现在是一条**已经判过、且判本章错**的条件，保留在表上。

11. **7.7 由“跑不动”改为“跑动了一半”**：模型的判余测出来了，判读者的判余仍然没有字段；并写明二者是否同一个量本章答不了（对应 9.4 第 3 条）。

预注册写的是十位受试者，实际取到六位（其余下载未完成），此处偏离写在 7.5 正文里。

仍未做的两件：判读手册历次改版的多次观察；判读者自报“拿不准”与模型判余的相关性检验。

第四章 从来没见过，和不许见到

从来没见过，和不许见到

经济学、影视学与心理学共同揭示的“禁转”

作者：王德生 | v1.0 | 2026年8月9日

一条从来没有被观察到的转换，可能是跳不过去，可能是不划算跳，也可能只是有人规定了不许这么记。三者在记录上长得一模一样。

提要

“为什么睡眠必须按固定顺序走完一整轮，不能只取需要的那一段？”这个问题背后有一个从未被拆开的前提：我们把“从来没见过跳过”当成了“跳不过去”。

本章对 Sleep-EDF Expanded 的两个独立队列做判读序列分析：老化队列 78 人 153 夜（年龄 25–101），药物队列 22 人 44 夜（替马西洋与安慰剂被试内对照）。三条转换给出三种彼此不同的数值指纹。

由清醒或 N1 直接进入 N3：在两个队列、跨药物、跨全部年龄段一律低于 1%，受试者层面中位数为 0.00%，且施加药物后不增反减。

由 N3 直接转入 N1：在老化队列为 2.75%。而同一个“变浅”动作若从 N2 出发则为 27.64%——被压低十倍，却明确没有归零。

由 N1 直接进入 REM：本来就占进入 REM 的四成。本章预注册它会在药物条件下上升，而受试者层面配对检验不成立（配对中位差 +1.19, $p=0.381$, 11 人上升、11 人下降）。

第二条的十倍差距有一个纸面来源。判读手册规定：N3 之后不满足 N3 的窗，在无觉醒且不满足 W 或 R 时默认判为 N2。教材把这条写成一句更直白的话——“N3 之后不要判 N1”。也就是说，这条转换的稀有不是生理事实，**是一条明文规则**；而它没有归零，是因为该规则带例外条款。

本章据此提出：一条从未被观察到的转换，本章称为**禁转**，至少有四种来源——物理上跳不过去、成本上不划算跳、规则上不允许这样记、以及跳过发生了却被记录规则做成了看不见。四者在结果上同形，但在“换一个条件会怎样”上留下不同指纹。本章给出三指纹分拣表、**残余率**这一读数，如实报告分拣判据的第一次实证尝试失败及其原因，并把可分拣性本身限定为一个待检验的提案。

关键词：禁转；残余率；分拣指纹；判读规则；连贯性剪辑

一·三种“从来没见过”

去问任何一个研究睡眠的人：能不能不睡前面那些，直接进入深睡？

回答是稳定的：不能，睡眠有它的顺序。这个回答有充分的观察支持——在判读记录里，深睡几乎总是从浅睡进入的，从清醒直接跳进去的情形罕见到可以忽略。

问题出在从“罕见到可以忽略”到“不能”的这一步。

一件事从来没有被观察到，至少有三个互不相同的原因。第一种是它真的做不到：物理上、生理上没有那条路。第二种是它做得到但不值得：走那条路的代价高于绕道，于是从不发生。第三种最容易被忽略——**它发生了，但按照记录规则，不许这么记。**

这三种在记录里长得完全一样：一个接近零的数字。而它们对“能不能只取需要的那一段”这个实际问题给出的答案完全相反。第一种说：不行，别想了。第二种说：行，只要把代价结构改掉。第三种说：这个问题从一开始就没被记录系统允许提出。

本章要做的事很窄：**给出一个办法，把这三种分开。**

二·一条明文禁令

先看一个具体的例子，它比任何论证都直接。

美国睡眠医学会的判读手册对深睡（N3）的定义是：0.5–2 Hz、峰峰值超过 75 微伏的慢波，在该窗内占比达到或超过 20%（即 30 秒窗里至少 6 秒），**不分年龄**。对 N2 的定义则以纺锤波与 K 复合波为特征波形。

到这里都还是描述。真正要紧的是后面那条：**紧接在一个 N3 窗之后、又不满足 N3 判据的窗，在没有觉醒、且不满足 W 或 R 的条件下，判为 N2。**英国睡眠学会的判读教材把这条规则的实际含义写成一句更直白的话：**N3 之后不要判 N1。**

这是一条规则，不是一条生理规律。它规定的是：当睡眠从深处浮上来、而浮上来的程度尚未被判定为清醒时，这个窗**默认继承 N2 这个标签**。

那么“N3 直接转到 N1”这条转换在记录里会是什么样？它会很稀有。但它稀有的原因，不在睡眠者身上，在手册上。

三·四种签名与它们的指纹

把第二节这个例子推广，得到本章的核心装置。

设某条转换在记录中从未、或几乎从未出现。本章称之为**禁转**。它至少有四种来源。

其一，物理禁转。那条路不存在。指纹：无论换什么条件——不同人群、不同年龄、给药、剥夺——它的出现率都稳定在接近绝对的零。

其二，成本禁转。路存在但不划算。指纹：代价结构一被改动，它就出现；改得越狠出现得越多。

其三，规则禁转。它发生了，但记录规则规定不许这么标。指纹最有特色——**它被压低，但通常不会归零**，因为规则往往带例外条款；残余的高度由例外条款的宽窄决定。而且，**换判读标准它就变**，换被试它不变。

其四，涂抹禁转。跳过真的发生了，但被记录规则做成了看不见。指纹与第一种**完全相同**：任何条件下都不出现。区别在于，第一种的零来自世界，第四种的零来自观察系统自己。

第四种是本章最想指出的一种，也是最难办的一种。它之所以能被想到，是因为另一门学科用一百二十年时间把它做成了产业。

四·影视学：每秒跳过二十四次

电影这门艺术的物质基础，就是本章这个问题。

一秒二十四格，每两格之间的一切都被跳过了。而没有任何一个观众看见过那些跳过。观众看见的是连续的运动——一个从来没有被拍下来的东西。

这还只是底层。真正与本章相关的是电影在此之上建立的一整套**连贯性剪辑**规则：三十度规则、一百八十度轴线规则、动作衔接、视线匹配。这套规则的全部目的，是让剪辑**看不见**。遵守规则的剪辑，观众根本不会察觉发生了剪接；一部两小时的电影里有上千个剪辑点，而观众能说出位置的通常不超过十个。

三十度规则尤其能说明问题：如果前后两个镜头的机位角度差小于三十度，剪在一起会显得“跳”。于是行业发明了一条规则，**专门用来避免让跳过看起来像跳过**。

电影同时也发明了相反的东西：**叠化与淡出**。它们的功能是标记跳跃——“这里我跳过了一段时间，你不要当成连续的”。一部电影因此有两套手段，一套让跳过隐形，一套让跳过显形，而用哪一套是一个选择。

睡眠判读只有第一套，没有第二套。手册要求每一个三十秒窗都必须得到一个标签；当一个窗里同时存在两种或更多分期的特征时，规则是“判给占比最大的那一种”。也就是说，无论那三十秒里发生了什么，输出都是一个干净的单一标签，而相邻标签连起来永远是一条连续的序列。**判读手册里没有叠化。**

这一家还提供了本章需要的推翻材料，而且是它自己的历史提供的。经典好莱坞把**跳切**当作错误，是必须避免的技术事故；戈达尔在《精疲力尽》（1960）里把它变成了风格，之后它进入了正常

的电影语言。“必须连贯”是一个历史阶段的行业规范，不是叙事的普遍法则。一条规则可以被违反，而作品不塌——这正是规则禁转与物理禁转的区别在艺术史上的实证。

五·经济学：为什么必须成批

若某条路存在而从不被走，第二种可能是成本。

经济学对此有一个古老而精确的结构。经济订货批量的基本形式是：每次订货要付一笔与数量无关的**固定启动成本**，而持有库存要付一笔与数量成正比的成本。二者一权衡，最优批量正比于启动成本与持有成本之比的平方根。

这个结论的形状对本章很重要：**批量的大小不由需求的形状决定，由两笔成本的比值决定**。需求可以是每天要一点，而只要启动成本足够高，最优做法仍是攒起来一次性做完。“为什么不能要多少拿多少”这个问题，在经济学里从来不是一个关于能力的问题。

搬到本章：如果每一次进入某个睡眠状态都要付一笔与停留时长无关的启动开销，那么“成批”就是最优解，而“只取需要的那一段”之所以从不发生，与能不能做到无关。

这一家的位置是给出**结构**，不是给出证据。本章不主张已经测到了睡眠的启动成本——没有。经济学在这里的作用是提供一个反例假说，使“从没见过跳过”不能被自动读成“跳不过去”。

六·心理学：顺序不是一环扣一环

第三种可能与前两种都不同：顺序是被维持的，而不是被因果链锁定的。

拉什利在一九五一年那篇论行为的序列顺序的文章里，处理的正是这件事。当时的主流解释是链式联想：每一个动作产生的感觉反馈触发下一个动作，于是序列自然展开。拉什利的反驳很硬——快速序列（打字、演奏、说话）中动作之间的间隔短于感觉反馈回路的传导时间，**下一个动作在上一个动作的反馈到达之前就已经启动了**。因此序列不可能由一环触发一环，必须有一个更高层的计划在管。

这个结论对本章的用处在于：**顺序是可以由计划维持的，而计划是可以被改写的东西**。一个由计划维持的顺序，在外观上与一个由因果链锁定的顺序无法分辨——两者都表现为“总是按这个次序发生”。

心理学同时提供了自己的推翻材料，而且就在这条论证的另一侧：动作序列一旦被组块化，**从序列中间某点开始执行会显著慢于从头开始**。也就是说，即便顺序是由计划维持的，“从中间进入”仍然可能有真实的代价。因此心理学并不站在“顺序随便可改”这一边——它只是把“不能改”与“改起来贵”分开了。

七·三家一起否掉的那句话

三家表面各说各的：经济学谈批量，影视学谈剪辑，心理学谈序列计划。它们否定的是同一句话：

一个过程若总是按某个顺序发生，是因为每一步在物理上为下一步创造了条件。

经济学说：也可能只是启动成本高，成批比零买便宜。影视学说：也可能跳过发生了，只是按规则被做成了看不见——而判断“有没有跳过”依赖于观察系统自己的连贯性规则。心理学说：也可能由一个上层计划维持，而计划不是因果链。

三条合起来给出一个单独任何一门都不会提的问题：当“从没见过”有四种来源而它们在记录上同形时，怎么把它们分开？

八·分拣表

表1 禁转的四种来源与它们的指纹

来源	换人群/年龄	改代价结构	换判读规则	归零程度
物理禁转	不变	不变	不变	接近绝对零
成本禁转	可能变	明显变	不变	视代价而定
规则禁转	不变	不变	明显变	被压低但留残余
涂抹禁转	不变	不变	明显变	接近绝对零

四行里有两组同形：物理禁转与涂抹禁转在“换人群”“改代价”两列上完全一样，只能靠第三列分开；而第三列——换判读规则——在观察上通常做不到，因为一份数据只按一套规则判读过。

这就是本章要说的最不舒服的一句话：在只有一套判读规则的记录上，物理禁转与涂抹禁转原则上不可分。唯一的替代办法不是统计，是去读规则本身怎么写的。

由此得到本章的读数：**残余率**——一条被压低的转换在记录中保留的出现比例，以同一个动作的另一条可比路径为分母。规则禁转的特征不是零，是一个由例外条款决定的、稳定的非零残余。

九·真跑：三条转换的三种指纹

数据为 Sleep-EDF Expanded 的两个独立队列。**老化队列**（sleep cassette）：78 名健康、不服用睡眠相关药物者，年龄 25-101，153 个受试者-夜。**药物队列**（sleep telemetry）：22 名轻度入睡困难者，每人两夜，一夜替马西洋一夜安慰剂，被试内随机，44 夜。仅使用判读标注；分析区间为熄灯起 8 小时；R&K 的 3 与 4 合并为 N3。

表 2 三条转换在两个队列中的出现率

转换	老化队列	药物队列（安慰剂）	药物队列（替马西洋）
由 W 或 N1 直接进入 N3	0.20%-0.75%（各年龄段）	0.43%	0.00%
由 N3 直接转入 N1	2.75%	—	—
同一动作：由 N2 转入 N1	27.64%	—	—
由 N1 直接进入 REM	40.0%	33.76%（汇总）	43.72%（汇总）

第一条：物理禁转或涂抹禁转的指纹。由清醒或 N1 直接进入 N3，在老化队列的四个年龄组分别为 0.75%、0.46%、0.15%、0.66%——年龄从 25 到 101，不动。在药物队列，安慰剂夜 21 人中仅 2 人各出现过一次，替马西洋夜 0 人。给药之后不增反减。**四列指纹里，前两列都是“不变”。**按表 1，这条转换要么是物理禁转，要么是涂抹禁转，而本数据分不开它们。

第二条：规则禁转的指纹，本章最干净的一处。在老化队列的 153 夜里，离开 N3 的转换共 2547 次，其中转入 N1 的 70 次，占 2.75%；而离开 N2 的转换共 5721 次，转入 N1 的 1581 次，占 27.64%。药物队列给出同样的形状：离开 N3 转入 N1 占 3.68%，离开 N2 转入 N1 占 22.71%。

同一个“变浅”的动作，从 N3 出发被判为 N1 的机会不到从 N2 出发的十分之一。生理上，从深睡浮到 N1 的水平并不比从 N2 浮上去更难；而第二节那条手册规则精确地解释了这个十倍差距——**N3 之后的窗默认继承 N2 这个标签。**

更要紧的是它**没有归零**：2.75%，不是 0.2%。这正是规则禁转与物理禁转的分界。手册那条规则带例外（无觉醒时才默认继承），于是留下一个残余，而残余的高度由例外条款的宽窄决定。**物理给的是绝对的零；规则给的是被压低十倍的一个稳定残余。**

第三条：一次失败的分拣。见下节。

十·一次失败的分拣尝试，以及它教了什么

本章预注册了对成本禁转的检验：若某条转换是成本禁转，那么改变代价结构应当让它出现得更多。选用的扰动是替马西洋，检验对象是“由 N1 直接进入 REM”这条本来就不算罕见的转换。

汇总计数看起来支持：安慰剂夜进入 REM 的 157 次中越格 53 次（33.76%），替马西洋夜 215 次中越格 94 次（43.72%），涨了十个百分点，明细里 N1-REM 由 50 次涨到 87 次。

但这个结果是错的，而且错在一个具体的地方。上述比例是把 22 个人的转换汇总在一起数出来的。在受试者层面配对重算，安慰剂中位 25.00%、替马西洋中位 29.17%，配对中位差 +1.19，Wilcoxon 双侧 $p = 0.381$ ，**上升 11 人、下降 11 人，正好对半。**药物没有系统性地打开这条转换；那个十个百分点来自少数人的大幅波动被总数放大。

同时核查了 REM 本身：受试者层面，REM 总分钟（ $p = 0.537$ ）、REM 段数（ $p = 0.125$ ）、REM 平均段长（ $p = 0.206$ ）在两种条件下均无显著差异。那个“进入次数从 157 涨到 215”反映的是**同一段 REM 内进出更频繁，不是 REM 更多。**

因此，本章对成本禁转的第一次实证尝试失败。本章不把它改写成一个更软的版本，也不换一条转换重试。

失败教了两件事，都写在这里。

其一，方法上的：汇总计数不能当配对证据用。凡涉及跨人比例，必须先个体层面计算再做配对检验；否则少数人的极端值会以“效应”的面目出现。本章在这一点上先错后改，改正过程如实写在正文，不放在附注。

其二，关于分拣判据本身的：替马西洋不是一个足够强的代价扰动。它改变的是入睡快慢与睡眠维持，并没有改变“要不要跳过某条转换”的代价结构。这意味着分拣判据带有一个此前未被声明的参数——它需要多强的扰动才分得开。一个判据若只在极端扰动下才有分辨力，其实用范围就受限；而“极端到什么程度”只能靠失败来标定。本数据说明替马西洋这一档不够。真正可能有效的扰动是睡眠剥夺后的恢复夜（真正改变代价结构）或发作性睡病（机制本身被改写），两者本数据都没有。

十一·最强的三个竞争解释

其一，N3-N1 的十倍差距是生理的，不是规则的。深睡之后的浮出确实可能倾向于停在 N2 那个深度，不必掉到 N1。分岔点：这条解释预测，在改用不含该默认继承条款的判读体系（例如按连续慢波占比直接分级，不设“继承前一窗”规则）重判同一批原始信号时，十倍差距应当保持；本章的解释预测它会明显缩小。这个检验需要原始信号，本章只有判读标注，因此本章这一条目前未获直接支持，只有一个纸面规则与一个吻合的数字。

其二，第一条转换的稀有是采样问题。30 秒窗太粗，真实的快速穿越被窗长吞掉了。△ 这条解释与本章的“涂抹禁转”几乎是同一件事的两种说法，本章接受它作为涂抹禁转的一个具体机制，而不作为竞争解释。这也意味着本章在这一格上无法自证：判读窗长是本章唯一的时间单位，用它去检验它自己造成的效应，做不到。

其三，四种来源的划分本身是重述而非发现。有人会说，这不过是把“不能/不划算/不许”三个常识词排了个序。分岔点：若一个划分不能对任何一条具体转换给出不同预测，它确实只是重述；本章第九节给出了两条转换的不同指纹（一条跨条件不变、一条被压十倍且带残余），且第十节报告了第三条的失败。能不能对第四种给出可操作的分拣，是本章最弱的一环，本章明写它目前做不到。

十二·与最近的一位

本章最近的一位不在睡眠学里，在测量学里：任何以规则产生的分类数据，其类别之间的转换频率同时反映被测过程与分类规则本身。这是内容分析、行为编码、医学分级共有的老问题，通常以“编码手册效应”“分类器诱导的伪结构”出现。

分离线有两条。其一，方向不同：那一支关心的是分类规则引入的误差与不一致（编码者之间信度、类别边界模糊），处方是提高一致性；本章关心的是分类规则引入的结构——一条被规则压低的转换，即使全部编码者完全一致，那个十倍差距也一样存在。一致性越高，规则造成的结构反而越干净、越像规律。这一条给出同案不同预测：提高编码者信度，那一支预测偏差减小，本章预测 N3-N1 的残余率不变甚至更稳定。

其二，本章给了一个可算的读数：残余率，以及“被压低但不为零”这个特定形状。那一支通常只报告“存在编码效应”，不区分被压低到零与被压低到一个残余——而这个区分正是把规则禁转与物理禁转分开的唯一线索。

⚠ 必须承认：本章没有检索到把“残余率”作为分拣工具的既有工作，但没检索到不等于不存在；本章这一条主张的强度以此为限。

十三·其余分界

与睡眠周期的既有理论。NREM-REM 交替的生理机制（脑干互抑等）解释的是为什么会有周期，本章问的是为什么记录里看不到跳过。二者不冲突：即便周期机制完全正确，第二节那条手册规则仍然独立地压低了一条转换。

与判读一致性研究。那一支测的是不同判读者对同一段记录给出相同标签的比例，已知 N1 的一致性最差。△ 这条对本章构成真实威胁：N3-N1 稀有，可能只是因为 N1 本来就难判、判读者倾向于不用它。分岔点：若如此，N2-N1 也应被同样压低，而实测是 27.64%——同一个 N1 标签，从不同前驱出发相差十倍。这条差异是本章对该威胁的回应，也是本章这一格能站住的主要理由。

与“睡眠阶段是连续量的人为切分”这一支。那一支主张分期本身应被连续读数（慢波活动谱）取代。本章不主张取代——本章主张的是，在保留分期的前提下，要能说出哪些结构事实是分期规则造成的。二者是不同层次的工作。

与古德哈特定律与指标博弈。那些描述的是有人为了指标而调整行为；本章描述的情形里没有人博弈——判读者忠实执行手册，睡眠者不知道手册的存在。失效来自规则的形状，不来自动机。

十四·同一作者已发论述的自撞申报

本系列此前各篇多次落在“记录制度”这个家族上：账上少一栏、边界没有地、一个窗只能给一个标签、回应位置无人登记、速率从未被裁定、验收工具分不开三个端。本章必须说清与它们的关系，否则会被读成第七次重复。

分离线是方向：此前各篇讲的是记录制度遗漏了什么——某个真实存在的东西在账上没有对应的栏。本章讲的是记录制度制造了什么——一个看起来像生理规律、实际由规则条款生成的结构。遗漏留下的是空白，制造留下的是一个可以被误读成定律的实心结论，后者更难被发现，因为它不缺任何东西。

与三篇最近的分别是：论验收分辨率那篇是分辨率问题（使用者分不开三个端）；论恢复通道那篇是归属问题（读数的变异来自哪一端）；论干预时点那篇是配置问题（手段放在哪一端）。本章是分拣问题（一个零来自哪一种原因）。四者可以并存，但引用时不可互相替代。

与论深睡随年龄变化那篇的关系需要单列：那篇用的是 N3 的段数与段长，本章用的是转换的出现率与残余率，读数不同；但两篇同用一套判读标注，因此那篇的结论同样受本章这一条的约束——若 N3 的边界部分由规则决定，那么“段长随年龄缩短”也可能部分是规则与信号幅度共同作用的产物。本章在此明确：本章削弱了那篇的一部分结论强度，这一点应写进两篇的互引。

十五·三家反过来给本章加的限制

经济学的限制。经济订货批量的整个结构建立在启动成本可测的前提上。本章没有测到任何启动成本——没有测过进入一个睡眠状态的固定开销是多少，甚至没有一个可用的操作定义。因此经济学在本章中只提供一条反例假说，不允许本章声称成本禁转在睡眠中确实存在。第十节的失败正好落在这个限制上。

影视学的限制，也是最重的一条。电影的连贯性规则是有人为了效果故意设计的；判读手册那条默认继承条款则是为了提高一致性、减少判读者分歧而设的，其意图与电影完全不同。把二者放在一起，本章只能借用它的结构（规则可以让跳过隐形），不能借用它的归因（有人想让你看不见）。本章不主张任何人有意隐藏什么，判读手册那条规则很可能是一个合理的工程折中，它的副作用是意外的。

心理学的限制。拉什利的论证说明顺序可以由计划维持，但它没有说计划可以随意改写；而组块化研究恰恰表明从中间进入有真实代价。因此心理学不允许本章从“顺序不是因果链”推出“顺序可以被绕开”。本章只主张这个问题没有被分开问过，不主张答案是可以绕开。

十六·证伪

其一：拿一批有原始信号的记录，按两套规则各判一次——一套含“N3 之后默认继承 N2”条款，一套不含（例如按连续慢波占比直接分级）。若两套判读给出的 N3-N1 残余率没有明显差异，本章第二条指纹作废，规则禁转这一类在睡眠上不成立。

其二：若在任何一种强扰动（睡眠剥夺后恢复夜、发作性睡病、显著改变睡眠代价结构的干预）下，“由 W 或 N1 直接进入 N3”的受试者层面出现率仍稳定在 1% 以下，则该转换为成本禁转的可能性被排除，只剩物理与涂抹两种。

其三：若提高判读者之间的一致性会使 N3-N1 的残余率显著下降，本章对第十二节那条分离线的判断错误，本章应并入编码一致性研究。

其四：若在多个独立数据集中，同一动作从不同前驱出发的出现率之比（本章的十倍）不能复现，第九节第二条整个作废。

其五：若“残余率”这个读数在任何一条已知的规则禁转上都无法与已知的物理禁转区分，本章的分拣表退化为一张概念清单，不得再声称可操作。

本章事先声明：第一条与第二条本章都没有做，且第一条需要本章没有的原始信号。因此本章的分拣判据目前是一个提案，不是一件已验证的工具。

十七·反噬

其一，“这是规则造成的”会变成一句万能的开脱。任何不喜欢的结论都可以被说成判读伪影。本章因此写死：主张某条结构是规则造成的，必须同时指出是手册哪一条，并给出该条款预测的残余率形状。指不出条款的，不算。

其二，它会被用来贬低分期本身。本章不主张废除分期。一个带规则副作用的分类系统仍然可能是极有用的；本章只主张在使用它产生的结构事实时，先问一句这条结构有没有对应的条款。

其三，它会诱发对判读规则的随意改动。若有人为了让某条转换出现而修改条款，得到的将是另一组规则造成的结构，而不是生理事实。规则的更换只应用于对照，不应用于生产。

十八·自反

本章这一轮里错了两次，都必须记在正文。

第一次：本章最初设计的是一个关于“唯一入口”的假说——N2 是 N3 的唯一入口，因此 N2 的供给决定进入 N3 的机会。预注册的判据是“N2 总量与进入 N3 次数正相关，且这一关联强于 N2 与 N3 段长的关联”。实测结果是负相关（-0.219），且 N2 与段长的关联反而更强（-0.434），差值 -0.215，越过预注册写死的撤销线。该假说整个撤回。

第二次：本章用汇总计数报告了一个“药物让越格转换几乎翻倍”的结果，受试者层面配对后归零。这个错误的形状与本章所批评的对象是同一种——用一个不该用的聚合方式，制造了一个看起来像规律的结构。本章批评判读手册把连续量压成标签序列，而本章自己把 22 个人压成了一个总数。

第三笔：本章最想指出的第四种来源（涂抹禁转），本章给不出对它的分拣办法。它与物理禁转在本章能拿到的一切数据上同形，唯一的出路是去读规则文本——而那不是一个统计检验，是一次阅读。本章因此不能声称“四种可分”，只能声称“四种应分，其中两种已给出可算的指纹，一种的检验失败了，还有一种目前只能靠读规则”。

十九·结论

回到那个问题：为什么睡眠必须按固定顺序走完一整轮，不能只取需要的那一段？

本章的回答不是一条机制，是一句更前面的话：这个问题至今没有被分开问过。

“从来没见过跳过”在记录里是一个数字，而这个数字至少有四个来源。本章能确定的是：由清醒直接进入深睡这条路，在两个独立队列、跨越 25 到 101 岁、跨越给药条件，出现率始终低于百分之一——它要么真的不存在，要么被判读规则做成了看不见，而本章分不开这两者。本章能确定的另一件事是：由深睡直接转到最浅睡这条路，稀有得可疑，而它稀有的原因写在判读手册上，不写在人身上；那个十倍的差距和那个没有归零的残余，是一条规则留下的指纹。

本章不能确定的是第三种。用药物去改变代价结构这一次尝试失败了，而失败告诉我们，分拣需要多强的扰动本身是一个待标定的参数。

所以对“能不能只取需要的那一段”，本章给不出是或否。本章只把这个问题拆成了四问，并指出其中一问的答案已经印在手册里，只是从来没有人把它当成一个答案来读。

附录 数据、口径与可核项

数据：PhysioNet Sleep-EDF Expanded 1.0.0。老化队列（sleep cassette）153 个受试者-夜、78 名受试者、年龄 25-101、健康且不服用睡眠相关药物；药物队列（sleep telemetry）44 夜、22 名受试者、替马西洋与安慰剂被试内对照。仅使用判读标注。

口径：窗长 30 秒；分期取 W/1/2/3/4/R，3 与 4 合并为 N3；Movement time 与 Sleep stage ? 视为非五分期，不计入分子分母且在分段时视为断点；分析区间自各夜熄灯时刻起至熄灯后 8 小时或记录结束，取较早者；跨人比例一律先在受试者层面计算再配对检验。

预注册：PREREG-GATE.md, SHA-256
e8580be8eea0a67da9090851a5be1cdf876d6cca17584eda7ffd8c67f5e265fa，在计算任何结局变量之前冻结；其第零节列明了哪些结果属于此前的探查、不作为预测。被推翻的“唯一入口”假说全章保留于该文件，未作修改。

可核数字：老化队列离开 N3 共 2547 次（→N2 2341、→W 123、→N1 70、→R 13）；离开 N2 共 5721 次（→N3 2549、→N1 1581、→W 832、→R 759）。药物队列离开 N3 共 869 次（→N2 784、→W 43、→N1 32、→R 10）；离开 N2 共 1695 次（→N3 875、→N1 385、→W 220、→R 215）。

判读规则出处：美国睡眠医学会《睡眠及相关事件判读手册》关于 N3 的判据（慢波 0.5-2 Hz、峰峰值 >75 μ V、占该窗 $\geq 20\%$ ，不分年龄）、关于 N2 的特征波形（纺锤波与 K 复合波）、关于 N3 之后窗次的默认判定（在无觉醒且不满足 W 或 R 时判为 N2），以及“一窗内若并存两种以上分期，判给占比最大者”。本章引用的“N3 之后不要判 N1”一语出自英国睡眠学会的判读教材对上述规则的表述。

作者与人机协作声明

研究问题、核心概念、理论判断、适用边界与最终学术责任归作者。人工智能用于文献检索辅助、公开数据复算、结构检查与语言编辑。本章两处错误（唯一入口假说被推翻、汇总计数被误用为配对证据）均如实写入正文而非附注。分拣判据目前是一个提案而非已验证工具，第十六节第一、二条检验本章均未完成。所有引文、数据与推断须由后续同行评议与可复算研究继续核验。

编二·每一次改善都得有一个分数

读数层

这一编处理的不是格子，是一个被印出来的数。

一个人想睡得好，最后拿到的是一个分数、一条曲线上的一点、或者一个“改善了多少分钟”。这一层的强制作答发生在归总那一步：几件事被压成一件，两端相减只留一个差，一整段过程只留两个端点。

第五章说的是：他要的快、深、省力这三样，在他自己能验收的那把尺上是同一个读数——分不开的要求不是要求，是愿望。第六章说的是：每种办法在开一条入睡通路的同时封掉另一些，而封这件事不产生读数，于是净效果为正的办​​法可能正在关掉一扇门。第七章说的是：一个差值读数里说话的永远是绝对离散度大的那一端，而哪一端更散由构造者挑了什么量决定。第八章说的是：从入睡到第一次深睡那一段下降，可以算，没有那一栏，也从来没有被裁决过。

这一编在第十七章的验证表上最吃亏——四章里有三章的读数**根本不是“每夜一个数”**：一个是跨夜配对差，一个是一对份额，一个是一条曲线。一编讲**“一个被公布的数”**，而它自己的四个读数**里有三个不是一个数**。这件事本书没有事先设计，是算完才看见的。

第五章 三样东西，一个分数

三样东西，一个分数

工程学、信息学与戏剧学共同揭示的”索验比”

作者：王德生 | v1.0 | 2026 年 8 月 9 日

你能提出一个比你能验收的更精细的要求。提完之后，你就再也不知道自己有没有拿到它。

摘要

“怎样才能高效入睡？怎样才能睡得深？”这句提问里有三个要求：快、深、省力。本章先做一次预注册的公开数据复算，用 Sleep-EDF Expanded 的睡眠遥测子集——22 名轻度入睡困难者，每人两夜，一夜替马西洋、一夜安慰剂，被试内随机。预注册的假设是”权衡曲线”。三条预测全部不成立：入睡潜伏期配对中位差 0.00 分钟 ($p=0.758$ ，有效 $n=17$)，N3 比例无变化，两端耦合 $\rho=-0.045$ 。按撤销规则，该案例整个作废。

本稿补了两处稳健性。其一，地板效应：这批人安慰剂夜入睡潜伏期中位仅 10.25 分钟，超过临床线的只有 1 人；按底子分三层后配对中位差为 0.00/0.00/-2.00，地板效应确实存在但分层后结论未变，据此明写这份数据不能裁决替马西洋能否让人更快。其二，分析窗：换用”熄灯到记录结束”重跑，总睡眠 p 由 0.049 升到 0.140、睡眠效率由 0.028 升到 0.063，而潜伏期两窗均为 0.00。零结果稳，正结果不稳——本章因此把承重支点整个压到稳的那一半：只主张使用者索求的那一端一点没动（三种切法同一个零），而“另一端动了”降为待复核的线索，不作为论证的一环。

由此提出：一个要求，只有在验收端可分辨时，才是一个可执行的要求。三个要求在实验室里可分开，在使用者的验收工具上却映射为同一个读数；于是他永远在提一个自己无法验收的要求，手段永远在交付一个未被索求的改善，而错位不产生任何信号。本章把索求间距与验收分辨率之比命名为索验比，并首次算出分子的一部分：三个改善量在标准差单位下的两两距离 (0.33/0.61/0.94)。△v1.0 曾把“第一主成分只解释 54.1%”读作“三个维度被压成一维”的证据，本稿补做零分布对照后予以撤回——三个互相独立的变量在 $n=22$ 时第一主成分平均即达 45.1%，实测的 54.1% 恰在其第 95 百分位（单侧 $p=0.050$ ），方向还偏向更一维。分母 δ 至今一次未测，索验比这个比值本章从未算出过。

最近的三位是最小重要差异、恰可察觉差与信度天花板—— δ 在医学、心理物理学、心理测量学里各有更老的名字，初稿三个都漏了，现全部排在分界节最前并据此收窄了贡献声明。而那次窗敏感性本身是本章命题在作者身上的一次发作：报告只出一个口径，而口径的选择改变了”哪一端动了”。

关键词：入睡；验收分辨率；索验比；最小重要差异；地板效应；零误差容量

一·三样东西，一个分数

一个人今晚想睡得好。他能说出三件具体的事：快一点睡着，睡得沉一点，别那么费劲。

第二天早上，他醒来，问自己一句：昨晚睡得怎么样？

他得到一个答案。一个。

这中间发生了一件很少被注意的事：他提出了三个要求，而他验收的时候只有一个读数。手环也一样——它采一整夜的数据，最后推给他一个两位数的睡眠分。分数底下确实有分项，但真正被记住、被用来判断”昨晚那个办法有没有用”的，是那个总分，或者是那句”睡得还行”。

于是接下来的事就顺理成章：他昨晚试了一个新办法（早点关灯、喝了点东西、开着白噪声、吃了半片药），今早分数高了两分，于是这个办法”有用”。至于三个要求里究竟哪一个被满足了、有没有哪一个反而变差了——这个问题他没法回答，不是因为他不细心，是因为他手里的工具不生产这个区分。

本章要处理的就是这件事。它不是关于睡眠的生理机制，也不是关于人应不应该那么用力（那是另一个题目，本站的失眠专文已经处理过，本章第十四节会与它划界）。本章处理的是一个更窄、也更少被写下来的问题：当一个要求的精细程度超过了验收工具的分辨程度，会发生什么。

答案不是”要求落空”。落空还会留下痕迹。真正发生的是：要求被一个它没有索求的改善兑现了，而双方都收到了”成交”的回执。

二·这一夜到底发生了什么：一次预注册的复算

要把上面那句话从格言变成可检验的东西，需要一个具体的例子：一个被广泛使用的助眠手段，一份能同时看到多个端的公开数据。

本章使用 PhysioNet 的 Sleep-EDF Expanded 数据库中的睡眠遥测子集。这批数据来自 1994 年的一项替马西泮研究：22 名受试者，男女均有，有轻度入睡困难但其他方面健康；每人在医院记录两夜，一夜服用替马西泮，一夜服用安慰剂，被试内对照。数据库另附受试者表，给出每人两夜各自是药还是安慰剂，以及两夜各自的熄灯时刻。

这批受试者的身份很要紧：他们正是会去找助眠办法的那种人——不是失眠症患者，是”有点难睡着”的普通人。而替马西泮是一种苯二氮草类催眠药，被使用的理由几乎总是同一个：让我快点睡着。

本章只使用判读标注，不使用原始信号。分析在跑统计之前先写下预注册，并冻结了哈希（见附录 A）。预注册的假设是一条”权衡曲线”：

— P1（快端改善）：替马西泮夜的入睡潜伏期显著短于安慰剂夜。

— P2（深端变差）：替马西泮夜的 N3 占总睡眠时间比例显著低于安慰剂夜。

— P3（两端耦合）：在个体层面，入睡潜伏期的改善量与 N3 比例的损失量正相关——谁快得多，谁就深得少。

— P4（常规报告只覆盖一端）：只看入睡潜伏期与总睡眠时间两项，替马西泮夜两项均不差于安慰剂夜的人占七成以上。

— P5（稳端）：入睡后首小时的觉醒次数无显著差异，或替马西泮夜更多。

撤销规则也写在前面，其中第一条最硬：若 P1 不成立，本案例整个作废，不得换一个读数重试。

三·四条预测的下场

22 人全部配对成功，零剔除。结果如下。

表 1 替马西泮夜与安慰剂夜的五项读数（n=22，配对）

读数	安慰剂中位	替马西泮中位	中位差	Wilcoxon 双侧 p
入睡潜伏期（分钟）	10.25	8.50	+0.00	0.758
N3 占总睡眠（%）	18.10	18.64	-0.77	0.721
入睡后首小时觉醒次数	3.00	1.00	-1.00	0.055
总睡眠时间（分钟）	421.25	429.75	+8.25	0.049
睡眠效率（%）	92.61	94.69	+0.86	0.030

P1 不成立。入睡潜伏期的配对中位差是 0.00 分钟， $p=0.758$ 。这不是”效应小”，是没有效应：一半的人吃药之后睡得更快，另一半更慢，抵消掉了。

P2 不成立。N3 比例没有变化（ $p=0.721$ ）。这一条值得停一下：苯二氮草类抑制慢波睡眠是教科书级的说法，而在这批数据的判读标注上，它没有出现。可能的原因很多（剂量、单夜、判读口径、样本量），本章不去猜；本章只报告：在这份公开数据上，本章预注册的那个代价没有被观察到。

P3 不成立，而且是彻底不成立。Spearman $\rho=-0.045$ ， $p=0.839$ 。“快端改善量”与”深端损失量”在个体层面完全无关。这一条比前两条更要命：即便 P1 与 P2 都成立，只要 P3 不成立，“沿着一条曲线移动”这个图景就不存在——那不是一条曲线上的两个点，那是两件互不相干的事。

P4 不成立，但方向有用。只看入睡潜伏期与总睡眠时间两项，两项均不差于安慰剂的只有 8/22 = 36.4%（预注册线是 70%）。也就是说，即使只看临床最常给出的两项，三分之二的人拿到的也是”一好一坏”。

P5 反向。入睡后首小时的觉醒次数由 3 降到 1 ($p=0.055$, 边缘)。药不但没让人更不稳, 反而更稳了。

一处必须补报的技术账：配对 Wilcoxon 会剔除零差对。五项里, 入睡潜伏期有 5 人两夜完全相同 (有效 $n=17$) , 入睡后首小时觉醒次数有 4 人 (有效 $n=18$) , 其余三项零差为 0 (有效 $n=22$) 。前一稿只写了 $n=22$, 不够准确。

按撤销规则第一条, “**权衡曲线**”这个案例到此作废。本章不把它改写成更软版本, 也不换读数重试。下面的一切都建立在“它失败了”这个事实上。

三之二·一条必须先排除的解释：他们本来就睡得很快

“药没让他们更快”这句话, 只有在“他们本来慢”的前提下才需要解释。所以在往下走之前, 先看这批人的底子。

表 1-2 安慰剂夜入睡潜伏期的分布 ($n=22$)

最小	第一四分位	中位	第三四分位	最大
0.0 分	5.6 分	10.25 分	18.8 分	52.5 分

超过临床”入睡困难”常用线 (30 分钟) 的只有 1 人 (4.5%) ; 超过 20 分钟的 5 人; 四分之一的人不到六分钟就睡着了。

这批人被数据库描述为”有轻度入睡困难”, 而他们在安慰剂夜的入睡速度, 按任何临床口径都属于正常甚至偏快。于是一条极朴素的解释立刻成立: 药没让他们更快, 可能只是因为没有可快的余地。这是一条地板效应, 前一稿完全没有处理它, 而它比前一稿列出的三条竞争解释都更难反驳。

处理办法是按底子分层, 看有余地的那一层动没动:

表 1-3 按安慰剂夜入睡潜伏期分层

层	n	安慰剂中位	替马西泮中位	配对中位差
≤ 10 分 (几乎无余地)	11	5.50 分	6.50 分	0.00
10-20 分	6	13.00 分	10.75 分	0.00
> 20 分 (有余地)	5	25.00 分	23.00 分	-2.00

三层的配对中位差是 0.00、0.00、-2.00。**地板效应确实存在**——底子越差的人, 药的方向越对。但分层之后, **结论没有改变**: 即便在最有余地的那一层, 中位改善也只有两分钟, 而那一层只有五个人。

这里要写下一句本章自己不愿意写的话: **这份数据不能裁决”替马西泮能不能让人更快入睡”这个问题**。有余地的人只有五个, 任何结论都撑不住。本章所能主张的, 缩到一个更小的范围: 在这二十二个人身上, 索求的那一端在任何一层上都没有产生可读的改善, 而未被索求的那一端产生了。第四节的全部论证只使用这句缩过的话。

也请注意这一层的顺带后果: **这批人本来就睡得快, 而他们仍然在吃助眠药**。一个入睡中位十分钟的人吃安眠药, 本身就是本章那一格的一个实例——他索求”快”, 而他快得已经没有任何余地; 他能收到的只可能是别的东西, 而他的验收工具分不出那是别的东西。

三之三·换一个分析窗, 动的那一端就不显著了

前一稿把分析区间写死为”熄灯时刻起至记录结束或熄灯后 8 小时止, 取较早者”, 这个口径在预注册里, 合规。但一个合规的口径不等于一个稳健的结论。本节把另一个同样自然的口径 (熄灯到记录结束, 不设 8 小时上限) 跑一遍:

表 1-4 两种分析窗下的同一份数据 ($n=22$)

读数	熄灯起 8 小时 (原口径)	熄灯到记录结束
入睡潜伏期	+0.00 分, $p=0.758$	+0.00 分, $p=0.758$
N3 占 TST	-0.77, $p=0.726$	-1.60, $p=0.566$
总睡眠时间	+8.25 分, $p=0.049$	+12.25 分, $p=0.140$

睡眠效率	+0.93, $p=0.028$	+0.71, $p=0.063$
------	------------------	------------------

必须照实说的是：本章那两个显著结果，换一个同样自然的窗就不显著了。睡眠效率掉到边缘（0.063），总睡眠掉到 0.140。方向没变，强度不稳。

因此第四节那句”药给了他们别的”要按下面这个强度重写：

稳的是零结果——入睡潜伏期在两种窗下都是配对中位差 0.00 分钟、 $p=0.758$ ，一动不动。不稳的是正结果——“更整”在原口径下显著，换窗后落到边缘与不显著之间。

这对本章是一次收窄，但它没有动到承重的那一半：**本章需要的是”索求的那一端没有动”**，这一条在两种窗下都成立；**“未被索求的那一端动了”只需要方向，不需要显著。**若把本章读成”替马西洋改善睡眠完整度”，那是过度解读，本章不作此主张。

三之三之一·这件事正是本章的命题在本章自己身上发作

停下来看这一节的形状：一个分析口径的选择，**改变了”哪一端动了”这个判断。**而一份研究报告只会呈现一个口径——预注册的那一个。读者拿到的是一个读数，看不见另一个口径下它会变成什么。

研究者的验收工具，也只有一个出口。他索求”这药改善了什么”，收到的是”在这一个口径下，这两项显著”；他手上没有一把尺能告诉他这个结论对口径有多敏感，除非他自己去跑第二遍。**本章第十节那张三端对照表，把研究者这一端漏了。**补在这里：

索求端	实验室可分辨端	分析者可验收端	使用者可验收端
“这药改善了什么”	能：多口径重跑	弱：报告只出一个口径	极弱：一个综合读数

这一行不是修辞。它说明本章的那一格不止在使用者一侧——**从记录到结论之间，还有一次同样形状的合并，而它发生在写论文的人手上。**

四·失败给出的那件事

把上面五行读一遍，会看到一个很干净 的形状。

这批人吃药，是为了**快**。药没给他们快——全组配对中位差 0.00 分钟；按底子分三层是 0.00、0.00、-2.00（三之二节）；换一个分析窗仍是 0.00（三之三节）。**这个零结果是本章全部论证里最稳的一块。**

药似乎给了他们别的：总睡眠时间多了八分钟、睡眠效率高了近一个百分点、入睡后第一个小时里少醒两次。**但这一端经不起换一个分析窗——三之三节里，总睡眠的 p 由 0.049 升到 0.140，睡眠效率由 0.028 升到 0.063。**

这一稿因此把承重支点整个压到稳的那一半上，不再靠两半合起来说话：

本章的经验支点只有一条：使用者索求的那一端，一点没有动。全组 0.00 分钟，分层 0.00／0.00／-2.00，换窗仍是 0.00——三种切法给出同一个零。

>

“未被索求的那一端动了”这半句，在另一个同样正当的分析窗下会变成“两端都没动”。本章因此把它降为一条**待复核的线索**，不作为论证的一环。

这一退让并不削弱本章，反而使它更干净：**要立“使用者索求的东西没有被交付”这件事，只需要索求的那一端没动。**另一端有没有动、动了多少，是下一个问题，而且是一个本章这份数据答不了的问题。

而 $p=-0.045$ 又告诉我们，**快与深**这两端在实验室里是可分开的、相互独立的。也就是说，问题不在于测不出来：测得出来，而且测出来它们是两件事。（△九之二节将指出，**快与整并不独立**，相关 $r=+0.563$ ；本节这句话只对快与深成立，届时予以更正。）

那么问题在哪里？

在于：吃药的那个人，第二天早上没有办法知道这件事。他早上醒来，感觉“昨晚睡得还不错”。这个感觉是真的——他确实少醒了两次，确实多睡了八分钟。但他想要的那件事（快点睡着）一点没变。他会把这个“还不错”记在“快”的账上，因为在他的账本里，只有一栏。

于是他下次还会吃。他没有被骗，药也没有失效。只是他索求的那一端与他收到的那一端不是同一端，而他的验收工具分不开它们。

五·工程学：没有合格的量具，就不能写公差

制造业里有一条很硬的规矩，硬到写进了行业手册。

美国汽车工业行动集团（AIAG）的测量系统分析规范里有一条叫**十倍法则**（Rule of Tens）：量具的分辨率必须把待测特征的公差带至少分成十份；换句话说，量具至少要比被测特征精细一个数量级。同一套规范给出验收阈值：量具重复性与再现性占公差的比例低于 10% 为可接受，10%-30% 视应用情况而定并须严密监控，**超过 30% 则该量具不得用于此项判定**。

这条规矩背后的道理很朴素：如果你的卡尺只能读到毫米，而你要判定的公差带是 ± 0.3 毫米，那么你测出来的每一件都会读成同一个数。你不是测出了“全部合格”，你是什么也没测。美国国家标准与技术研究院的测量过程手册在讲分辨率时还专门加了一句警告：**仪器显示的位数不代表它的分辨率**。

工程学在这里给出的不是一个比喻，而是一条**准入条件**：一个指标要能进入验收，它必须先证明自己的量具够细。反过来说——**一个你没有合格量具的特征，不允许被写进公差**。

这条准入条件在睡眠这件事上从来没有被执行过。“快一点睡着”“睡得沉一点”是两个不同的特征，它们各自的公差带是多少？验收它们的量具（早上的一句“睡得怎么样”）分辨率是多少？没有人问过。

六·信息学：输出端分不开的两个输入，信道不传

工程学说的是“你应该先有量具”。信息学说的是一件更冷的事：**有些区分，无论你多想要，都传不过去**。

香农在 1956 年处理过一个与常规信道容量不同的问题：如果要求一个**错误都不能有**，一个有噪声信道还能传多少信息？答案由一个叫**混淆图**的结构决定——把输入符号画成点，凡是可能产生同一个输出的两个输入之间连一条边。零误差容量算的是这张图上的独立集：**只有那些两两不可能被混淆的输入，才能被可靠地区分**。

这个结论的形状很重要：可传的区分不由发送方的意愿决定，也不由信号的强弱决定，而由**接收端能不能把它们分开**决定。你可以把两个符号发得很用力，但只要它们在输出端落到同一个格子里，这个区分就没有过去。

搬到本章的问题上：使用者（发送方）区分了“快”“深”“省力”三个要求。他的验收工具（接收端）把它们全部映射到同一个输出：昨晚睡得怎么样。在这张混淆图上，**这三个要求两两之间都有边**。因此它们的可靠区分数是 1——不是三个要求，是一个。

⚠ 须先声明：本节这张混淆图目前只有理论形状，没有经验形状。九之二节曾试图给它一个经验形状，那次尝试已被本章自己的零分布对照否掉。以下是一个结构类比，不是一次测量。

这不是使用者的错误，也不是他不够敏感。这是信道的结构。他可以更用力地想要“快一点”，但只要早晨那一个读数把“快了但浅了”和“没快但整了”映射成同一句“还不错”，这个区分就不存在。

七·戏剧学：两百五十年争不出结果，因为验收端分不开

有一门学科在这件事上有最长的实践经验，而且它自己已经承认了失败。

狄德罗在《演员的悖论》里提出了一个至今仍在被争论的主张：**最伟大的演员恰恰是那个自己不感动感情的人**。他的理由不是冷酷，而是可靠性——被自己的情感淹没的演员今晚极好、明晚很差；而冷静地掌握着技术的演员，每一晚都一样好。观众看到的效果相同，来源却完全相反。

此后两个半世纪，表演理论沿着这条裂缝分成两支：一支主张演员必须真的经历（体验派的各种版本），一支主张演员应当构造并控制（技术派的各种版本）。这场争论至今没有被实验解决。

它为什么解决不了？不是因为没人做实验，而是因为狄德罗那句话本身就锁死了裁决通道：**观众分不出台上那个人是“真的感受到了”还是“技术上做对了”**。如果两条路在唯一的验收端（观众席）产生相同的读数，那么任何以观众反应为终点的实验，都不可能把它们分开。这场争论只能被选择，不能被判定。

戏剧学因此给出一条本章最需要的经验：**当两条路在验收端不可分辨时，关于“哪条路更好”的争论会永远持续，而且会持续得非常热闹**。热闹本身不是有内容的标志，有时恰恰是验收端分不开的标志。

助眠这件事上的热闹——这个办法好还是那个办法好，喝牛奶还是听白噪声，早关灯还是先运动——很可能属于同一类。

八·共同前提：我要什么，就能按什么验收

三家表面上说着三件事：工程学讲准入，信息学讲信道，戏剧学讲两百五十年的争论。它们否定的是同一个前提：

一个人提出的要求，可以按他提出的那个样子被验收。

这个前提太自然了，自然到从来不必写下来。我要“快一点睡着”，那就看快没快。可是三家各自的材料都在拆它：工程学说，看之前先证明你的尺够细；信息学说，接收端合并了的区分不存在，无论你想多要；戏剧学说，两条方向相反的路可以在验收端产生同一个读数，于是两百五十年过去了也判不出来。

三条合起来给出一个单独任何一门都不会提的问题：如果一个要求的精细程度超过了验收工具的分辨程度，会发生什么？

不是要求落空。是要求被别的东西兑现，并且拿到了合格回执。

这正是第三节那五行数据的形状。

九·索验比：一个可以自己算的数

为了不让上面这句话退回成感想，需要一个可算的量。工程学已经把它形状给出来了，只需要换个名字。

设某人在同一段时间里提出的两个要求为 A 与 B（例如“快点睡着”与“睡得沉”）。

— **索求间距 Δ ：**A 与 B 在客观读数上的实际差别有多大——即“这个办法把 A 改善了多少、把 B 改变多少”这两个量之间的距离。

— **验收分辨率 δ ：**使用者手上的验收工具能够可靠区分的最小差别。“可靠”不是“他觉得能”，是重复测量能给出同一个判定。

索验比 = Δ / δ 。

⚠ Δ 目前不可算，这一点必须写在这里而不是留到自反节。A 与 B 通常量纲不同——“快”以分钟计，“深”以百分点计——两个不同量纲的改善量之间的“距离”，本章没有给出算法，也不认为存在一个自然的算法。本章因此把 Δ 的用法限定为**同量纲内**：只有当 A 与 B 可以化到同一把尺上（例如都化成各自读数的标准差单位，或都化成各自的最小重要差异的倍数）， Δ 才是一个数。**跨量纲时，索验比不是一个可计算的比值，只是一个提问的形状。**本章全章没有真正算出过一个索验比的数值，一次也没有——这一点在第十八节还要再说一次。

工程学的十倍法则要求这个比值不低于 10。本章不主张睡眠必须照搬这个阈值——那是制造业针对自己公差结构定的。本章主张的是更弱、也更难反驳的一条：

当索验比小于 1 时，该要求在该验收端上不是一个要求，而是一句愿望。使用者提出了 A，收到了 B，而他的工具无法告诉他收到的是哪一个。

三条使用纪律，防止这个数变成又一个可以随便填的指标：

其一，**6**必须由重复测量估出，不能自评。一个人说”我分清睡得快和睡得沉”，不算数。要算数，得让他对同一夜在两个互相盲的场合各评一次，看两次评的一致程度（第十一节给出具体做法）。

其二，**6**是那个人自己的、那个工具的、那个时期的数，不能跨人搬用。有的人对”昨晚醒了几次”极敏感、对”睡了多久”几乎无感；有的人相反。索验比是个人化的读数，不是群体常数。

其三，索验比低不是缺点，是配置。一个从不琢磨睡眠的人索验比很低，而他很可能睡得比谁都好。本章不主张人人都该把**6**做小。本章只主张一件事：在**6**未知的情况下，不要相信任何关于”这个办法对我有没有用”的判断。

九之二·Δ 的第一次实际计算，以及它削掉的本章一处

上一节的警示框说 Δ 跨量纲不可算，并给了一条出路：把各自的改善量化到自己的标准差单位。本节把这条出路走一遍——这是本章第一次真正算出与 Δ 有关的数。

三个改善量（快=潜伏期缩短、深=N3% 上升、整=睡眠效率上升）各自标准化之后，相关矩阵是：

表 3 三个改善量在标准差单位下的相关（n=22）

	快	深	整
快（潜伏期缩短）	1.000	-0.157	+0.563
深（N3% 上升）	-0.157	1.000	-0.115
整（睡眠效率上升）	+0.563	-0.115	1.000

主成分分解给出这样一组数：

第一主成分解释 **54.1%** 的方差（第二 31.4%，第三 14.5%）。

作为这份样本上的算术，这句话可以直接读：一个一维的综合读数，无论怎么加权，最多只能携带这三个改善量 54.1% 的信息。

但把它当作“三个可分的维度被压成一维、损失惨重”的证据，是错的——而本章 v1.0 正是这样读的，还加了一句“这不是估计，是恒等式”。这一稿予以更正，因为一个五分钟就能跑完的对照此前没有做：三个互相独立的变量，在 n=22 时，第一主成分本来就能解释多少？

五千次模拟给出零分布（脚本随文提供，随机种子写死）：

表 3-2 三变量互相独立、n=22 时第一主成分解释比例的零分布（5000 次）

统计量	值
均值	0.451
中位	0.445
50% 区间	0.414 - 0.483
90% 区间	0.377 - 0.541
95 百分位	0.541

实测的 **0.541** 恰好落在这条零分布的第 **95.0** 百分位，单侧 **p = 0.050**。

这个对照把结论翻了过来，而且方向与本章原先的读法相反：三个完全独立的改善量在这个样本量下平均只给出 45.1% 的第一主成分——独立情形下一维读数丢掉的是 **54.9%**，比实测丢掉的 **45.9%** 还多。因此 54.1% 不是“低得反常”，它落在独立零分布偏高的一侧；这份数据显示的结构，比三个互不相干的量还要更一维一点。

据此本章作两处更正。其一，撤回“混淆图第一次有了经验形状”这句话——那三个改善量的维度结构在本数据上与“三个独立变量”不可区分（p=0.05，且方向不利），它既不能支持“三个要求两两相连”，也不能支持“三维里有一半落在零空间”。其二，主成分那一段仍留在正文里，不是因为它成立，是因为它是本章自己读反了方向的一个数，删掉它第十八节的第四笔账就无从谈起。剩下真正站得住的只有一样：两两之间的 Δ（标准差单位下逐人距离的中位）——快↔深 0.33、快↔整 0.61、深↔整 0.94。这是本章全章唯一算出来的 Δ，也是九之二这一节唯一站得住的产出，而它只在同

一批人、同一种药、这三个读数上成立，不可外推。分母 6 仍然没有测过，因此索验比这个比值本身，本章依然一次也没有算出来。

九之二之一·这一算削掉了本章一处

表 3 里有一个数与本章前面的说法不合：**快与整的改善量相关 $r=+0.563$** 。它们不是独立的。

第四节写过”这两端在实验室里是相互独立的两件事”。那句话的依据是 P3 的 $\rho=-0.045$ ，而 P3 比较的是**快与深**。**快与整并不独立，本章此前的表述把范围放大了**。现予更正：

独立的是**快与深** ($\rho=-0.045$)；**快与整正相关** ($r=+0.563$)，只是这一次药在快上没有产生位移、在整上产生了。

这个更正削弱了本章一个更响亮的版本——不能再说”索求端与交付端是两件互不相干的事”。准确的说法弱一些：**索求端与交付端可以相关，而使用者仍然分不清自己拿到的是哪一端**；相关性不解决可分辨性问题，反而使它更隐蔽——两个相关的东西一起变好时，最容易把功劳记错账。

十·三端对照表与那一格

把上面几节合起来，可以做一张最小的对照表。任何一项助眠手段，都可以被摆进这三列。

表 2 三端对照：索求端 / 实验室可分辨端 / 使用者可验收端

行	索求端（他要什么）	实验室可分辨端（能不能测出来）	使用者可验收端（他能不能验收）
快	明确提出	能：入睡潜伏期	弱：只在极端情况下能察觉
深	常被提出	能：慢波睡眠比例	几乎不能：无对应主观读数
整	很少提出	能：觉醒次数、睡眠效率	较强：夜里醒来是能记住的
省力	明确提出	弱：无公认操作定义	强：这正是主观感受最擅长的

这张表一摆出来，第三节那五行数据就有了位置：替马西洋改善的是**第三行（整）**，而使用者索求的是**第一行（快）**。第三行恰好是使用者验收能力最强的几行之一，第一行恰好是他验收能力较弱的一行。于是**错位不但不产生信号，还产生一个正向的信号**。

由此得到本章的**那一格：索求三端，验收一端**。

这一格的危险不在于它明显是坏的，而在于它**通过一切形式审查**：使用者提出了具体的要求（不是空泛的”想睡好”），手段确实产生了可测的改善（不是安慰剂效应），双方都诚实，报告也没有造假。**唯一缺的那一样是：没有人问过，验收这三个要求各自需要多细的尺**。

值得单列的是表中第四行。“省力”是唯一一个使用者验收能力强于实验室的行——主观疲惫感是真实的、他自己最清楚，而它至今没有公认的客观操作定义。本章因此不主张”客观读数优于主观读数”。**两边各有分不开的地方，方向相反**。

十一·一次可以当场做完的自我校准

本章的证伪主件不是一个写死日期的赌注，也不是一个需要实验室的方案，而是一件读者今晚就能开始、七天后能做完的事。它同时是本章最容易失败的地方。

做法（约需十四天）：

1. 每天早上，按你平时的习惯给昨晚打一个分（几分制不限，但要固定）。**除此之外什么都不要多做**。
2. 同时，用手机记下三件当晚可核的事：几点关灯、大约几点睡着（估计即可）、夜里醒来几次。这三件不参与评分，只是留档。
3. **第八天起**，每天早上打完分之后，翻出**七天前那一夜**的三件留档事实，仅凭这三行字，再给那一夜打一次分。打的时候不要看当时打的分。

4. 十四天结束，你手上有七对分数：同一夜、两次评分。

这七对分数之间的一致程度，就是你的 6 的上界。

若两次评分相差很大，说明你的评分并不由那一夜本身决定，而是由评分当时的状态决定——那么你此前所有“这个办法对我有用”的结论，都建立在一把会自己漂移的尺上。若两次评分高度一致，说明你的尺是稳的，接下来才值得问“它够不够细”。

这件事同时是本章的证伪装置。若在一个合理规模的样本里，普通人对同一夜的重测评分一致性普遍很高，那么本章的那一格不成立——使用者的验收工具比本章假设的细得多，“索求三端验收一端”只是研究者的想象。本章明确预测相反的结果，并且愿意在这一点上被推翻。

⚠ 这个做法有一个已知的弱点，必须写明：第 3 步用留档事实重评，测的是“**给定同样信息时评分是否一致**”，而不是“**评分是否忠实于那一夜**”。它给的是 6 的一个**乐观上界**——真实的 6 只会更粗，不会更细。要拿到更硬的数，需要在两个真正互相盲的场合各评一次同一夜，那需要第三方参与。

十二·最近的三位：可察觉阈、差别阈限与信度天花板，以及本章初稿漏掉了前两位

本章的核心量 6——“验收工具能够可靠区分的最小差别”——在两门学问里各有一个更老的名字，而本章初稿两个都没提。它们比第三位（信度天花板）离本章更近，因此排在前面。

十二之一·最小重要差异：6 在医学里已经有名字

它说到哪一步：临床研究里有一整支专门做“多大的改善，患者才察觉得到、才认为有意义”，产出的量叫**最小重要差异**（MCID/MID）。它有两套成熟的估计方法：锚定法（拿一个外部的“你觉得好些了吗”当锚，看对应的量表变化）与分布法（从测量误差本身推出下界）。它的用法是硬的：**一项试验的终点差异若小于 MCID，即便统计显著也不算临床有意义。**

这几乎就是本章的 6。本章第九节要求“6 必须由重复测量估出，不能自评”，正是分布法的做法；第十一节那个十四天重测，是分布法的一个业余版本。

分离线（两条，同案不同预测）：

其一，MCID 是单一维度上的可察觉阈，本章是多维被压成一维的合并。MCID 问“这一个量上的改善够不够大”；本章问“几个不同的量是否被压进了同一个读数”。两者的失败形态不同：MCID 不达标的后果是**这个改善不算数**（方向不变，只是被判为无意义）；合并的后果是一个量上的改善被记在另一个量的账上（方向可以颠倒）。第三节那五行正是后者：入睡潜伏期毫无改善，而使用者会给出正向评价。**同案分离**：若把 MCID 分别算在“快”“深”“整”三个量上并各自达标，本章的那一格照样成立——因为使用者手上仍然只有一个读数，他仍然不知道达标的是哪一个。

其二，MCID 是给研究者用的判定线，本章是使用者手上那把尺的属性。一项试验可以完全合规地报告“本试验的入睡潜伏期改善未达 MCID”，而使用者仍然会因为“睡得还不错”继续吃。若在实证上，把 MCID 告知使用者能改变他的判断，本章这一条应当撤销。

十二之二·恰可察觉差：6 在心理物理学里有一百七十年历史

它说到哪一步：韦伯与费希纳以来，“同一物理量上两个刺激要相差多少才能被察觉”是一门有量表、有定律、有实验范式的成熟学问，那个量叫**恰可察觉差**（JND/差别阈限）。它甚至给出了 6 随刺激强度变化的规律。

分离线：JND 处理的是同一个量上的分辨（这两个重量哪个更重）；本章处理的是不同量之间的混淆（这一夜是快了还是整了）。前者是刻度问题，后者是维度问题——一个人可以在每一个单独维度上都有极细的 JND，而仍然分不清自己拿到的是哪一个维度上的改善，因为他的验收工具根本不分维度输出。**同案分离**：若给使用者分维度提问（“昨晚入睡快吗”“昨晚醒得少吗”分开问），而他的回答仍然高度相关、无法分开，则本章对；若分开问就能分开，本章的那一格退化为一个提问方式的问题，与信息学那一节的“混淆图”解释一并撤销。这一条今晚就能测，本章没有测。

十二之三·信度天花板，以及本章初稿也漏掉了它

本章的核心主张在心理测量学里还有一个一百二十一年前的祖先，初稿同样漏掉了。

斯皮尔曼在1904年就证明了：两个测量之间观察到的相关，不可能超过它们各自信度乘积的平方根。这是经典测验理论最早的结论之一。它的意思是：**验收工具的信度，为你能检出的任何关联设定了一个上限。**你不可能用一把不稳的尺去发现一个稳定的关系；不管那个关系在真值层面多强，你观察到的值都会被“衰减”到天花板以下。

这条定理与本章的关系必须说清楚，因为它几乎就是本章的命题：

它已经说到哪一步。它说：不可靠的测量会压低你能观察到的关联。整个心理测量学的信度传递、以及“先报信度再报效应”的规范，都建立在这上面。

本章与它的三条分界，每条都给出同案不同预测：

其一，它处理的是研究者的测量，本章处理的是使用者的验收。衰减校正正是研究方法学里的工具，用在论文里、用在量表编制里。它从不进入使用者手上那个“昨晚睡得怎么样”。**同案分离：**一项研究可以报告完美的信度、完整的衰减校正，而使用它的那个人仍然只有一个早晨的读数。若在实证上，研究测量的信度提高会同步提高使用者的验收分辨率，本章这一条应当撤销。

其二，它是一维的衰减，本章是多维的合并。衰减理论问“同一个东西测得准不准”；本章问“几个不同的东西是否被压进了同一个读数”。这两件事的失败形态不同：衰减的结果是效应被低估、方向不变；合并的结果是一个端的改善被记在另一个端的账上，方向可以完全颠倒。第三节那五行正是后者，不是前者。**同案分离：**若把使用者的单一读数拆成多个分项之后，各分项与客观读数的关联只是整体上升而次序不变，本章这一条不成立；若次序改变（即原本被记为“快”的改善其实来自“整”），本章这一条成立。

其三，它给的是校正，本章给的是准入。衰减理论的处方是事后把观察值除回去；工程学的十倍法则是事前不许你写公差。**本章站在后者。**这一条不是理论优劣，是位置不同：校正需要知道信度，而使用者永远不知道自己的信度——这正是第十一节那件事要解决的。

这三处漏掉本身值得记一笔。本章通篇主张“验收工具的分辨率决定了一个要求是不是要求”，而初稿把这个主张在三门学问里的祖先全部漏掉。 δ 在医学里叫最小重要差异，在心理物理学里叫恰可察觉差；它的上限效应，心理测量学一百二十一年前就证明了。漏掉的原因也很干净：本章是从工程学与信息学那一侧读进来的，而这三个名字挂在另外三本账上。**这与第三节那批人把“整”的改善记成“快”的改善，是同一种事故——不是没看，是看的那一栏里没有。**

必须据此收窄本章的贡献声明：**本章不主张“验收分辨率限制可执行要求”是一个新发现，那是三门学问各自的常识。**本章主张的只是把它放到一个从未被放过的位置上——使用者一侧，并指出在这个位置上它的失败形态与前三者都不同（见十二之一第一条）。

十三·其余九条判决性分界

第二，与速度—精度权衡（Woodworth 一路，至 Wickelgren 的权衡曲线）。那一支说的是：同一个人在同一任务上可以选择更快或更准，二者构成一条曲线。**分离线：**速度—精度权衡的经典实验，是明确告诉被试要快还是要准——即被试知道自己在权衡，且实验者可以操纵它。本章的对象恰恰是不知道自己在权衡、甚至根本不在权衡的情形：第三节的 $p=-0.045$ 说明这里没有曲线。若睡眠上的“快”与“深”被证明构成一条可操纵的权衡曲线，本章应当并入这一支。

第三，与入睡潜伏期的定义之争（首个非清醒期 vs 持续睡眠潜伏期）。睡眠医学早就知道“第一次越线”与“稳定睡着”不是一回事，并为此设了两个指标。**分离线：**那一对指标解决的是实验室内部的定义分歧，两个数都出现在同一份报告里；本章处理的是从报告到使用者这一段——无论报告里有几个数，使用者手上仍然只有一个。若临床报告采用哪个潜伏期定义会改变使用者的判断，本章这一条应当收缩。

第四，与安慰剂效应与期待效应。有人会说：使用者的“睡得还不错”只是期待。**分离线：**本章数据里那个改善是客观读数上的（总睡眠时间、睡眠效率、觉醒次数），不是自评。期待效应解释不

了 $\rho=-0.045$ ，也解释不了”三分之二的人在常规两项上一好一坏”。反过来，若把这些客观改善全部去掉后，使用者的主观判断不变，那么本章的机制被期待效应吸收。

第五，与多重终点与主要终点预注册（临床试验规范）。临床试验早已要求预先声明主要终点，防止事后挑选。**分离线：**那套规范约束的是**报告者**，本章约束的是**验收者**。一项试验可以完全合规地预注册”入睡潜伏期”为主要终点、并如实报告它没有改善，而使用者仍然会因为”睡得还不错”继续用。规范管得住论文，管不住早晨那一个读数。

第六，与古德哈特定律。那一条说：一个指标一旦成为目标就不再是好指标。**分离线：**古德哈特描述的是**博弈**——有人为了指标而调整行为。本章描述的情形里**没有人博弈**：药厂没有作弊，使用者没有自欺，判读没有造假。失效来自分辨率，不来自动机。这是本章与古德哈特最干净的一刀。

第七，与”综合指数掩盖分项”这一类批评（各种排名、评分卡）。那一支说：把多个维度加权成一个数会丢信息。**分离线：**那一支的处方是**公开分项**，而本章的第三节说明公开分项也不够——那批数据里所有分项都是公开可算的，问题在于**使用者的验收工具本身不产生这些分项**，无论报告里有没有。若给使用者提供分项就能恢复判断力，本章的主张退化为一个界面设计建议。

第八，与可穿戴设备的效度研究。那一支问：手环测得准不准（与多导睡眠图比对）。**分离线：**本章的问题在设备之后——即便手环完全准确，只要它推给用户的是一个综合分，索验比仍然可以小于1。准确性与可分辨性是两件事：一个完全准确的温度计，若只显示”冷/适中/热”三档，仍然分不开20℃与24℃。

第九，与认知行为疗法中的睡眠限制。这一条不是对手，是正面对照，也是本章见过的唯一一个把牺牲写在正面的手段。睡眠限制疗法明确告诉使用者：**接下来几周你会更困、更难受**，用这个代价去换睡眠效率的提高；同时要求记录睡眠日记，按效率这一个明确的量来调整卧床时间。它同时做了本章要求的三件事——**声明了索求端、指定了验收量、预告了代价端。因此本章的那一格不覆盖它。**别处做了，而绝大多数助眠手段没有做。

第十，与本站已发的《下去的那一段，没有人记》（第八章）。该文处理同一道问题的另一半：它主张”快速入睡”与”睡得深”能否兼得至今没有被判定过，而缺的是一把细于三十秒的下降刻度与一次标准化扰动。**分离线是干净的、而且两篇互补：**那篇的缺口在**仪器**——实验室的记录太粗，判不了；本章的缺口在**验收方**——实验室分得开（本章第三节的 $\rho=-0.045$ 正是证据），是索求它的那个人验收不了。一句话：那篇说”测不出来”，本章说”测得出来，但要它的人验不了”。两篇合起来才是完整的一对，任何一篇单独成立都不足以支持另一篇。

第十一，与”主观与客观睡眠不一致”这一整支（含睡眠状态错觉）。那一支已经大量记录了主客观读数的分离，尤其在失眠者身上。**分离线：**那一支把不一致当作待解释的现象或症状（为什么他觉得没睡着而脑电说他睡了）；本章把它当作**验收制度的结构性质**，并且对象是**没有任何睡眠障碍的人**——第三节那22人只是”轻度入睡困难”，而错位照样发生。若主客观不一致只在有睡眠障碍的人身上出现，本章应收缩到临床人群。

十四·同一作者已发论述的自撞申报

其一，本站的失眠专文——最险的一处。该文主张睡眠是撤退不是动作，意志一介入就锁死，处方是”不是学会入睡，是停止阻止入睡”。本章与它不在同一条轴上，必须说清楚，否则本章会被读成它的换皮。那篇讲的是用力的方向（越努力越睡不着）；本章讲的是**验收的分辨率**（分不清自己拿到了什么）。同案分离：一个完全放松、毫不用力、上床即睡的人，在那篇的框架里没有问题，在本章的框架里若他索求”睡得深”而只能验收”睡得整”，仍然落在那一格里。并且必须承认：**本章的处方与那篇有一处紧张——本章第十一节要求使用者连续十四天给睡眠打分，而那篇会说，这种持续的自我监测本身就是派守夜人。本章接受这条批评，并把它写进第十七节的反噬。**

其二，本书前四章。前四篇分别处理判读段落需不需要追认、睡与醒之间那块地有没有厚度、一个窗只能给一个标签的代价、以及一段睡眠里回应位置在谁手上。四篇的缺口都在**记录的一侧**；本章的缺口在**记录之后、判断之前的那一段**——从一份完整正确的记录，到使用者早晨那一个读数之间，发

生了一次不可逆的合并。本章不主张对”记录制度决定什么能被看见”这条家族命题有首创权，四篇加本章共享它。

其三，本站《交出去的不止那件事》。该文主张每一次委托都同时委托掉了对这次委托的监督权，因为验收能力本就是”自己做”的副产品。它与本章相邻得危险，但方向不同：那篇讲验收能力因不再亲自做而萎缩；本章讲验收能力从一开始就没有那么细，与做不做无关。同案分离：一个一辈子亲自琢磨自己睡眠的人，在那篇的框架里验收能力应当很强；在本章的框架里，只要他早晨仍然只产生一个读数，索验比照样可以小于1。

十五·三家反过来给本章加的限制

工程学的限制。十倍法则是制造业针对已有明确公差的特征定的。睡眠的”快”与”深”至今没有公认的公差带——多快算合格、多深算合格，无人写下。因此本章不能说”睡眠的量具不合格”，只能说更弱的一句：**在公差未定的情况下，任何关于量具够不够用的判断都不成立，而这本身就是准入条件没有被满足。**工程学在这里不许本章借它的阈值。

信息学的限制。零误差容量处理的是**已知混淆结构**的信道。使用者的混淆图是什么形状，本章仍然没有测过——第十一节那个自我校准只能给出 δ 的上界，给不出混淆图。九之二节测到的是另一样东西：**被区分对象一侧的维度结构**（一维读数最多携带54.1%），不是接收端的混淆结构。这两样常被混为一谈，本章不混：前者说明”有东西必然丢”，后者才说明”具体哪两个被并到了一起”。信息学在这里不许本章把前者当成后者。

一处结构上的账，先记在这里。工程学与信息学说的其实是同一件事的两个版本：十倍法则说”量具不够细就不许写公差”，混淆图说”输出端合并了的输入不可靠区分”，两者的骨架都是**分辨率决定可区分性**，只是一个长在制造业、一个长在通信。因此第八节那句”三条合起来给出一个单独任何一门都不会提的问题”偏强。准确的说法是：**真正独立的是两门半——工程学与信息学同脊；戏剧学提供第三个方向（验收端唯一时争论会永远持续），但它给的是形状，不给机制。**本章据此不主张三家构成一次三方互证。

戏剧学的限制，也是最重的一条。狄德罗那场争论之所以判不了，是因为**观众席是唯一的验收端——戏剧没有别的终点。**睡眠不是这样：睡眠有实验室，而且实验室分得开（第三节就是证据）。因此戏剧学的类比只能用到一半：它能说明”验收端分不开会让争论永远持续”，**不能说明”因此这些区分不可判定”。**本章据此把主张限定为：**在使用者那一端不可判定；在有实验室介入时可判定。**这一让步取消了本章一个更响亮的版本（”快与深的争论根本无法解决”），保住了一个更小但站得住的版本。

十六·最强的四个竞争解释

其零，地板效应：他们本来就睡得快。这是最难反驳的一条，前一稿完全没有列。安慰剂夜入睡潜伏期中位仅10.25分钟，超过临床线的只有1人。**分岔点已在三之二节跑过：**按底子分三层，配对中位差为0.00、0.00、-2.00——地板效应确实存在（方向随底子变差而变对），但分层后结论未变，且有余地的那一层只有五人。**据此本章已把主张缩到”在这二十二个人身上”，并明写这份数据不能裁决”替马西洋能不能让人更快”。**若在一个以”入睡潜伏期普遍大于三十分钟”为入组条件的样本上，药物显著缩短潜伏期而使用者仍分不清自己拿到的是快还是整，本章的例子才真正立住；若在那样的样本上使用者立刻分得清，本章的那一格不成立。

其一，效应太小，不是分不开。也许使用者验收不了，只是因为八分钟总睡眠时间的差别本来就微不足道。**分岔点：**给同一批人一个效应量大得多的操纵（例如整夜剥夺后的恢复夜），若他们的早晨读数立刻能分开”快”与”整”，那么本章测到的是效应量而不是分辨率，索验比这个概念应当废弃。这是本章最容易输掉的一条。

其二，学不会而不是分不开。也许使用者只是没有被训练过；给他一份分项清单和一点指导， δ 就会变小。**分岔点：**做一次简短训练的对照，看训练组的重测一致性与分项区分度是否提高。若提

高，本章应从“结构限制”降级为“可训练的技能缺口”，处方也随之从“不要相信判断”改为“先受训再判断”。

其三，睡眠压力这一个共同原因。第三节的替马西洋夜之所以“更整”，也许只是当晚睡眠压力更高；快与整都由它决定，谈不上错位。**分岔点：**本数据是被试内随机、同一批人两夜，这在设计上已经削弱了这条解释，但没有排除——本数据没有记录两夜之前的清醒时长。若在控制清醒时长后差异消失，本章的这个例子作废（但第九至十一节的结构主张不依赖它）。

十七·反噬预言：索验比最容易变成新的考核

其一，把 δ 做小会被当成美德。一旦“验收分辨率”成为一个可算的数，最省事的做法就是要求人人把它做小——教人们更细致地监测自己的睡眠。这正是本站失眠专文警告过的那件事：持续的自我监测会把睡眠变成一件被盯着的事。本章因此明确：**索验比低不是缺陷，本章不建议任何人为了提高它而增加对睡眠的监测。**第十一节那个十四天的自我校准是一次性的，做完就停；本章反对把它变成长期习惯。

其二，产品会拿它当卖点。“本产品可提高您的睡眠验收分辨率”是一句很容易做出来的营销话。而真正提高分辨率的办法（多给分项、多给客观读数）与真正让人睡得好的办法，方向可能相反。

其三，它会被用来推卸责任。“您的索验比太低，所以您感觉不到我们的改善”——这句话可以让任何无效的手段免责。本章因此写死一条：**索验比只能用于约束判断，不能用于为手段辩护。**一项手段若在客观读数上没有产生它所声称的那一端的改善，索验比再低也不能替它开脱。

其四，本章自己就在做一件它批评的事。见下一节。

十八·自反：我预注册了一个我自己分不开的东西

本章最该认的账在这里。

我预注册了一条“权衡曲线”，而数据说没有曲线。 $\rho = -0.045$ 。也就是说，在动手之前，我用一个比数据分辨率更粗的模型去索求了一个区分——我以为“快”和“深”是一条线上的两端，实际上它们是两件互不相干的事。**我犯的正是本章所描述的那个错误：提出了一个我的工具分不开的区分，然后准备好把任何结果都读成它的证据。**

若不是预注册写在前面、且撤销规则写死了“ $P1$ 不成立则整个案例作废”，我几乎一定会换一个读数重试，直到有什么东西显著为止——而那样得到的“权衡曲线”会看起来相当漂亮。

第一笔账之二：前一稿的三条竞争解释里，没有一条是地板效应——而那是最朴素、也最难反驳的一条。**我挑了三个我能打的靶子。**这一稿把它补进第三之二节与第十六节，并接受它带来的收窄：这份数据不能裁决替马西洋能不能让人更快。

第二笔账：本章的核心概念索验比，本身也有一个未声明的分辨率。 Δ 与 δ 各自怎么测、测出来的数有多大误差、两者之比的置信区间有多宽——本章一条也没给。按本章自己在第五节引的那条准入条件，**索验比目前不具备进入任何验收的资格。**它现在只是一个提问的形状，不是一个可以拿来判定的量。本章全章一次也没有算出一个索验比的数值。九之二节算出了分子 Δ （在标准差单位下），而分母 δ 一次也没有被测量过，那个比值因此仍然是空的。一篇以某个比值为题的文章从未算出过那个比值——这一点应当写在标题旁边，而不是藏在自反里。

第一笔账之三：本章只报了一个分析口径。第二个同样自然的口径把两个显著结果打得不显著（三之三节）——**而那一节是评审方跑的，不是我跑的。**一篇论证“使用者只有一个读数”的文章，自己也只出了一个口径的读数，且没有意识到需要第二个。

第二笔账之二： δ 在别处早有名字（最小重要差异、恰可察觉差），本章初稿一个也没提。一篇论证“看的那一栏里没有就等于不存在”的文章，自己在三本账上各漏了一个祖先。

第四笔账（本稿新增，也是最难看的一笔）：v1.0 拿到“第一主成分只解释 54.1%”这个数，立刻把它读成“一维读数损失惨重”，还加了一句“这不是估计，是恒等式”。**我没有跑那个五分钟就能**

跑完的零分布对照——而对照一做，54.1% 落在三个独立变量零分布的第 95 百分位，方向还与我要的相反。

这一笔账的形状与本章的命题严丝合缝：**我索求的是“证据”，我的验收工具是“这个数看上去低不低”，而这把尺分不开“低”与“在零分布里本来就该这么低”**。一个数看起来支持你，和它真的携带信息，在这把尺上是同一个读数。**我用一把自己分不开的尺，验收了一个自己想要的结论。**

这也说明第十一节那次自我校准为什么非做不可：**连研究者在自己的分析里都会用一把分不开的尺**，何况一个早上醒来只能给出一句“还不错”的人。

第三笔账：本章用了 22 个人、一种药、一个剂量、一夜。这个样本能支持的最强结论是“**存在这样一个例子**”，不是“**这是普遍结构**”。凡是把本章读成后者的引用，都超出了证据。

十九·结论

回到最初那句提问：怎样才能高效入睡？怎样才能睡得深、快速入睡？

本章没有给出一套助眠办法，而且认为在回答之前有一件更前面的事要做。

那句提问里有三个要求。在实验室里，它们是三件可以分开的事——本章第三节的数据正是证明：一种药可以完全不改变入睡速度，却明显改变睡眠的完整程度，而这两件事在个体层面毫不相关。但在提问者手上，它们只有一个出口：明天早上那一句“昨晚睡得怎么样”。

于是他索求的那一端与他收到的那一端可以完全不是同一端，而他会拿到一个正向的回馈。没有人骗他，手段也确实有效。**只是有效的那一端，不是他要的那一端；而他手里的尺，分不开这两者。**

因此本章的回答是一句更窄、也更难听的话：**在你知道自己的尺有多细之前，你关于“这个办法对我有没有用”的每一个结论，都还没有到可以相信的时候。**

这不是让人放弃判断，而是把顺序调过来。工程上不允许用一把没验过的尺去判合格；心理测量学一百二十一年前就证明，尺的信度设定了你能看见的一切关联的上限。睡眠是唯一一件我们每天都要判定一次、却从不校准量具的事。

先量尺，再量睡眠。

附录 A 预注册与数据

数据：PhysioNet Sleep-EDF Expanded 1.0.0, Sleep Telemetry 子集，44 份判读标注（22 名受试者 × 2 夜）。药物/安慰剂分配与两夜熄灯时刻取自数据库随附的受试者表。

预注册文件 SHA-256:

e2fe4a0f4f1738f2ef3a9018680c1e3cd1de122d1adac8f3ee02a8d1192e28d1。该文件在读取任何结局变量之前写就并冻结。

口径（在看结果之前写死）：窗长 30 秒；分期取 W/1/2/3/4/R，3 与 4 合并为 N3；Movement time 与 Sleep stage ? 视为非五分期，不计入任何分子分母且在分段时视为断点；分析区间自熄灯时刻起至记录结束或熄灯后 8 小时止，取较早者。入睡潜伏期为区间起点到首个非 W 五分期窗；总睡眠时间为区间内非 W 五分期窗总时长；N3% 为 N3 窗数除以总睡眠窗数；入睡后首小时觉醒次数为自首个非 W 窗起 60 分钟内由任一非 W 五分期转入 W 的次数；睡眠效率为总睡眠时间除以区间时长。任一受试者若有一夜不可解析则整体剔除——**实际剔除 0 人**。

统计：配对 Wilcoxon 符号秩检验（双侧，含零差剔除与连续性校正），Spearman 秩相关。

本数据不能回答的：其他助眠手段是否表现出同样的错位；剂量—反应关系；长期效应；权衡曲线在别的手段上是否存在。它只能回答一件事——在这一个可核的例子里，一项被广泛使用的助眠手段改善的那一端，不是使用者索求的那一端。

附录 B 自我校准的记录表头

日期 | 关灯时刻 | 估计入睡时刻 | 夜间醒来次数 | 当日评分 | （第八天起）七天前那一夜的重评分 | 重评时是否看过原评分（须为“否”）

七对分数记录完毕后，计算两次评分的一致程度。该数是你的验收分辨率的**乐观上界**：真实的分辨率只会更粗。

参考资料

— Automotive Industry Action Group. *Measurement Systems Analysis Reference Manual* (“十倍法则”与量具重复性再现性的10%/30%验收阈值)。

— NIST/SEMATECH. *e-Handbook of Statistical Methods*, §2.4.5.1 Resolution (分辨率定义; “显示位数不代表分辨率”)。

— Diderot, D. *Paradoxe sur le comédien* (写于1773–1777, 1830年出版)。

— Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberyé, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194. (Sleep-EDF数据库来源文献)

第二轮修订 (v1.1, 本稿) ——经第二次外部评审补做的三处

11. ★九之二节补做零分布对照并撤回原读法。评审方指出“第一主成分只解释54.1%”缺一个必须先做的对照：三个互相独立的变量在 $n=22$ 时第一主成分本来就能解释多少。五千次模拟给出均值0.451、95百分位0.541，实测的0.541恰在第95.0百分位(单侧 $p=0.050$)，且方向与本章原读法相反——独立情形下一维读数丢掉54.9%，比实测的45.9%还多。据此撤回“混淆图第一次有了经验形状”，九之二节此后只保留三个两两 Δ 。脚本`null_pca.py`随文提供。

12. 第四节把承重支点整个压到稳的那一半。v1.0用“索求端没动+非索求端动了”两半合起来立论，而后一半的显著性对分析窗敏感。本稿明写：本章的经验支点只有“索求的那一端一点没动”这一条(三种切法同一个零)，另一端降为待复核的线索，不进入论证。第六节同步加声明：那张混淆图目前只有理论形状。

13. 第十八节新增第四笔账：v1.0没有跑那个五分钟的零分布对照，就把一个数读成了自己想要的方向——用一把自己分不开的尺，验收了一个自己想要的结论，这与本章的命题严丝合缝。

作者与人机协作声明

研究问题、核心概念、理论判断、适用边界与最终学术责任归作者。人工智能用于文献检索辅助、公开数据复算、结构检查与语言编辑。本章不以生成过程替代证据；预注册在读取结局变量之前冻结，四条预测的失败如实写入正文而非附注。所有引文、数据与推断均须由后续同行评议和可复算研究继续核验。

修订说明 (本稿相对 v1.0 的十二处改动)

v1.0完成后经一次外部评审，评审方独立复算了第三节的**全部主结果**(同一子集、同一口径，逐项吻合)，并指出两处缺席。逐条列明：

1. 新增三之二节“一条必须先排除的解释：他们本来就睡得很快”。安慰剂夜入睡潜伏期分布(中位10.25分，四分位5.6–18.8分，超过临床线者仅1人)与按底子分层的三行对照(配对中位差0.00/0.00/–2.00)。地板效应是最难反驳的一条竞争解释，v1.0一条也没提。补上之后主张收窄为“在这二十二个人身上”，并明写这份数据不能裁决替马西洋能否让人更快。第四节与摘要同步改写。

2. 补报有效 n 。配对Wilcoxon剔除零差：入睡潜伏期5人零差(有效 $n=17$)，首小时觉醒4人($n=18$)，其余三项 $n=22$ 。v1.0只写 $n=22$ 。

3. 第十二节由一位扩为三位，并重排次序：最小重要差异(医学)、恰可察觉差(心理物理学)排在信度天花板之前，因为它们离本章的 δ 更近。各给同案不同预测的分离线：MCID是单维可察觉阈而本章是多维合并(失败形态一个方向不变、一个方向可颠倒)；JND是同量内分辨而本章是跨量混淆(一个人可以每一维JND都极细而仍分不清拿到的是哪一维)。据此**收窄贡献声明**：本章不主张“验收分辨率限制可执行要求”是新发现，只主张把它放到使用者一侧这个从未被放过的位置。

4. 第九节新增警示框： Δ 跨量纲时不可算。本章全章一次也没有算出一个索验比的数值，这一点从自反节提到正文。

5. 第十五节新增一段：工程学与信息学同脊(十倍法则与混淆图骨架相同)，真正独立的是两门半，第八节”三条合起来”一句偏强，已注明。

6. 第十六节竞争解释由三条增为四条（新增“其零、地板效应”）；第十八节自反新增两笔账——v1.0挑了三个自己能打的靶子；一篇论证“看的那一栏里没有就等于不存在”的文章，自己在三本账上各漏了一个祖先。

7. 新增九之二节： Δ 的第一次实际计算。三个改善量在标准差单位下的相关矩阵与主成分分解——第一主成分只解释 54.1% 的方差，即一个一维综合读数最多携带 54.1% 的信息，其余 45.9% 原理性地无法表达。这是本章全章第一个算出来的与 Δ 有关的数，也让第六节的混淆图第一次有了经验形状。

8. 同一算削掉本章一处并予更正：表 3 显示快与整的改善量相关 $r=+0.563$ ，并不独立。第四节原写“这两端在实验室里相互独立”，其依据 $\rho=-0.045$ 比较的是快与深；已更正为“独立的是快与深，快与整正相关”，并据此放弃“索求端与交付端是两件互不相干的事”这个更响亮的版本。第十五节信息学那条限制同步改写：九之二测的是被区分对象一侧的维度结构，不是接收端的混淆结构，两者不可混为一谈。

9. ★ 新增三之三节：分析窗敏感性。换用“熄灯到记录结束”这一同样自然的口径重跑，总睡眠 p 由 0.049 升到 0.140、睡眠效率由 0.028 升到 0.063，而入睡潜伏期两窗均为 0.00 分 ($p=0.758$)。据此把主张分成两半：零结果稳、正结果只主张方向不主张强度，并明写“若把本章读成替马西洋改善睡眠完整度，那是过度解读”。

10. 新增三之三之一与三端对照表的第三列“分析者可验收端”：口径的选择改变了“哪一端动了”，而报告只出一个口径——本章的那一格不止在使用者一侧，从记录到结论之间还有一次同样形状的合并，发生在写论文的人手上。第十八节自反据此新增一笔账：这一节是评审方跑的，不是作者跑的。

仍未做的三件：分维度提问的当晚检验（十二之二给出了做法）；以“入睡潜伏期普遍大于三十分钟”为入组条件的样本； δ 的任何一次测量——因此索验比这个比值本身，至今仍未被算出过一次。

第六章 只要有一条路是通的

提 要

本章问的是：怎样才能快速入睡、睡得深？三门学问给出三个不同的终点——把网络送进自由流、把条件场调对、让搜索停下来。三家共同踩着的那句话是：入睡是一趟有终点的行程，“离终点还有多远”在任一时刻有定义、可加总，因此各办法的效果可以相加。推翻它的材料来自交通学自己。本章提出：一次助眠办法在打开一些入睡通路的同时封掉另一些，而封这件事不产生任何读数；并给出一个先量个体噪声底的做法。

关键词 入睡通路 | 暗封 | 换轨型干预 | 自噪底

一·问题里自带的一个假定

有人问：怎样才能快一点睡着，睡得更深一点。

这句话听起来只是一个求助，其实它已经替回答者定好了答案的形状。它假定“睡着”是一个地方，从醒着到睡着是一段路，因而在夜里的任何一个时刻，都有一个量叫“还差多远”。有了这个量，一切就顺理成章了：办法就是缩短这个距离；两个办法一起用，效果相加；哪个办法更好，比比谁把这个距离缩得更短。

几乎所有关于睡眠的实用建议都建在这个假定上。睡眠卫生的条目表是这样：每一条都往同一个方向推一点，条目越多推得越多。睡眠效率的计算是这样：一个百分数，越高越好。助眠药的临床评价是这样：主要终点是入睡潜伏期，缩短多少分钟。行为疗法的成分分析也是这样：把多成分的方案拆开，看每个成分各贡献多少，然后把贡献大的留下、贡献小的去掉。

这个假定很少被写出来，因为它看上去不像一个主张，像一句废话。

本章要说的是，它不是废话，它是错的；而且它错的地方恰好是那些说不清为什么在自己身上不管用的人所在的地方。

二·三种回答，三个不同的终点

要看清这个假定，最省事的办法是把三种互不相让的回答摆在一起。它们不是三个学科对睡眠的兴趣，是三种对“一条路该以什么为终点来结算”的主张。三者恰好各占一个不同的终点，因而不能靠判谁对来化解。

第一种：终点是状态。这是交通网络研究的立场。一个路网的好坏，最终要落在它当下显露出的那个状态上——是自由流还是拥堵流。为了得到这个状态，可以调配时、可以在匝道上控流、可以做需求管理；这些全是手段，状态才是要的东西。交通工程有一个很硬的经验：瓶颈一旦崩溃，它此后的排放能力比崩溃前的最大流量还低，因此在高需求下追求最大流量恰恰是最容易触发崩溃的操作，而匝道控流的全部原理是故意把流量维持在低于容量的水平。放到入睡上，这一派的处方是：别把系统推到边缘，控住流入，先保住状态。

第二种：终点是条件。这是气候研究的立场。你不能对一场具体的天气施工。你能做的只有改变强迫与反馈，让整个分布位移；至于某一天会不会下雨，那是分布里的一次抽样，不可指定、也不可归因——单次事件归因至今只能给出概率比，给不出“这一场是不是因为它”。而且条件与显现之间有很长的滞后：今天动手，几十年后才看得见。放到入睡上，这一派的处方是：别管今晚，把光照、温度、作息、咖啡因这些条件调对，分布会自己移过去。

第三种：终点是过程的组织方式。这是逻辑与搜索理论的立场。一个证明搜索的目标不是那个结论——结论是不是真的，与你找不找得到它是两回事；目标是把搜索组织成一种能停下来的样子。剪枝、回溯点的取舍、策略切换，这些不是通向结论的手段，它们本身就是要做的事。放到入睡上，这一派的处方是：问题不在你的身体也不在你的卧室，在你脑子里那个还在展开的搜索；把它的组织方式改掉。

三种回答各自都很完整，各自都有几十年的证据，而且各自都能把另外两种收编成自己的手段。交通派会说条件与思维不过是影响流量的两个变量；气候派会说状态与搜索都是条件的产物；逻辑派会说状态是搜索的输出、条件是搜索的输入。

三·为什么这三种回答不能并存

它们不能并存的地方不在结论，在**可以插手的时刻**。

交通派的答案是：只能在**进入之前**插手。一旦车进了网、一旦瓶颈崩溃，下游做什么都没用甚至有害——拥堵波是往上游传的，你在下游疏导，只是把崩溃点往前挪。所以插手点在入口，而且入口关上之后，这一夜就没有插手点了。

气候派的答案是：只能在**强迫层**插手，而且此刻的状态已经被过去锁定了。今天躺下的这个身体，携带的是过去若干周的光照史、活动史、进食史；对这一夜，能做的事在原理上等于零。插手点在过去，效果在未来。

逻辑派的答案是：**每一个选择点都可以插手，而且只有选择点可以插手**。关键是认出自己正站在一个选择点上——认出来了，切换策略、放弃当前分支、重新开始，全都可行，而且随时可行。

这三个答案没法调和。“只能在入口插手”与“每个分叉都能插手”不是程度差别，是关于同一段时间的两种互相排斥的描述；而“此刻不可插手”把前两者一起否定。你不能说“综合三家”，因为一个人躺在床上的一刻，要么可以做点什么，要么不可以。

四·三家共同站着的那块地

三家争的是：**入睡这条路该以哪一样为终点来结算**。

争得越凶，越说明它们共同假定了一件事——**它有一个终点，因而“离终点还有多远”在任一时刻是有定义的、可以被加总的一个量**。

这条假定念给三家听，三家都会说这还用说吗。交通派要的正是这个量，他们叫它行程时间；气候派要的也是这个量，他们叫它信号；逻辑派要的还是这个量，他们叫它剩余搜索空间。三家吵的是这个量该怎么算、该在哪个层面上算，从没有人问过它存不存在。

而“效果可以相加”这件事，就是从这条假定里直接长出来的。多成分的行为方案之所以可以被拆成成分、之所以可以问“哪个成分贡献大”，前提正是每个成分各把这个距离缩短一点，而缩短量可以相加。一份睡眠卫生条目表之所以可以有十二条而不是三条，前提也是这个。

五·推翻它的那件材料，来自交通学自己

推翻这条假定的材料不需要从外面搬。它是交通网络研究自己在一九六八年就发表、此后被反复引用、写进每一本教科书的东西。

布雷斯（D. Braess）那年在《运筹学》上发表了一篇短文，给了一个四条路的小网络，然后在里面加了一条更快的新路。结论是：**在每个使用者各自选最快路的情况下，加了这条路之后，所有人的行程时间都变长了**。后来的研究表明这不是一个精心构造的病态例子——在一般的拥堵网络里，这种情形出现的概率与不出现的概率大致相当。

这件事在交通学自己的领域里回答的是另一个问题。它回答的是“为什么修了路还是堵”，它属于用户均衡与系统最优的差距这一支，是一条关于自利选路的政策警告。从来没有人把它当作一条关于“进度”的证据。

但它说的正是这件事：**通行时间不是各段容量的函数，是路网结构的函数。把各段的容量加起来得到的那个量，与实际到达时间之间没有单调关系**。也就是说，在一个有多条并行通路的系统里，“离终点还有多远”这个可加总的量并不存在。你可以为每一段路测出一个通行能力，你可以把它们全加起来，但那个和数不告诉你到达时间会变长还是变短——因为改变一段路会改变所有人选哪条路，而选哪条路才是决定到达时间的东西。

再补一件，来自气候研究自己：它的整套方法建立在“单次事件不可归因”上。对一次具体的热浪，气候科学能给的最强结论是一个概率比，不是一句“这一次是因为它”。也就是说，“**这一夜的**

进度”在气候派自己的框架里，压根不是一个可结算的对象——它自己就取消了那个量在单次上的存在。

两件材料合起来，三家共同站着的那块地就没了。

六·暗封：本章说的那样东西

地没了，还得站个地方。本章站的地方是这样：

入睡不是一条路，是若干条并行的通路。一夜能不能睡着，不取决于这些通路的“总容量”，取决于**其中至少有一条是通的**。而每一种助眠办法在打开一些通路的同时封掉另一些——被封掉的那一条，本章叫它**暗封**。

“暗”不是说它偷偷摸摸，是说**封这件事不产生任何读数**。这是暗封与“没走的路”的全部区别：一条你今晚碰巧没走的路，明晚还在；一条被这次做法封掉的路，明晚也不在了，而任何一份记录上都找不到它被封掉的那一行。

需要说清楚三件事，否则这个词会被读成一个比喻。

第一，“通路”不是“因素”。现行的睡眠模型里也有多个成分——睡眠压力、昼夜时相、唤醒水平——但它们是**加数**：各自贡献一点睡眠倾向，加起来构成一个标量。本章说的通路是**并联**：只要一条通就够，通了之后其余的通不通不影响结果。两者的区别是可判定的：加数模型预测各成分的效应可以互相补偿（这个成分弱了，那个成分强一点就补上了）；并联模型预测**补偿在某些个体上不成立**——补的是别的路，而他只有那一条。

第二，暗封不是副作用。副作用是在别的系统上出现的坏事（白天困、口干、跌倒），它有名字、有报告表、有分级。暗封是在**同一件事**上出现的：入睡这件事本身少了一条可走的路。它没有报告表，因为报告表上只有“入睡快了几分钟”这一栏。

第三，本章不主张暗封一定发生。这一点在第十二、十三、十四节会讲得很难看。本章主张的是：**如果它发生，现行的装置读不出来**。这是两件事，混起来是本章最容易犯的错误，也是审读时最该盯的地方。

删掉“并行通路”这个说法之后，剩下的框架里只有一个标量叫睡眠倾向。在那个框架里，“被封掉的那一条”不是难测，是**不构成一样东西**——没有位置可以放它。这是本章与“又提出一个新指标”之间的分别。

七·三种干预：净开型、净封型、换轨型

把一次助眠干预放到通路这个说法里看，它对每一条通路只能做三件事之一：打开、封掉、不动。据此干预分三型。

型	对通路做了什么	在群体平均上看到什么	在个体会出什么事	现行装置能不能读出来
净开型	打开的多于封掉的	主结局改善	大部分人改善	能，这正是它被设计来读的
净封型	封掉的多于打开的	主结局变差或不变	大部分人变差	能，试验做不出效果就停了
换轨型	打开一条，封掉一条，总量近似不变或略增	主结局改善	靠那条被封的路的人变差	不能

三型里只有第三型需要一个新说法。前两型不需要——净开型就是有效，净封型就是无效，现行的评价装置读它们读得很好。

换轨型的形状是：**它在群体上是有效的，而且它对大多数人确实是净开的；对少数人它封掉的恰好是唯一那条通路**。由于评价用的是组间平均，而这少数人的变差被平均掉了，所以试验报告上写的是“有效”，而且这句话没有说谎。

举一个已经被人做过的例子。睡眠限制疗法要求把卧床时间压缩到接近实际睡眠时间，它在群体上的效应量与整套多成分方案相当。同时有研究报告，在急性期它伴随客观总睡眠时长下降、日间嗜

睡上升与警觉性客观受损。用本章的说法：它打开了“睡眠压力”这条通路（而且打开得很有力），同时压缩了“在床上安静躺着等困意上来”这条通路。对大多数失眠者，前者远大于后者；对一个本来就靠后者维持的人，它是净封。而报告规范里，前者进主结局，后者进“可耐受性”或者根本不进——两者不在同一张表上，因而永远不相减。

再举一个更彻底的。多成分行为方案的成分网络分析近年做得很细：在两百余项随机试验、三万余名参与者的合并分析里，认知重建、睡眠限制、刺激控制被判为有效成分，而**放松训练被判为可能起反作用**。这是一个重要的结果，本章完全接受它，并且认为它正是本章所说的那件事在**群体层上**的显影：一个被广泛推荐的成分，在合并了足够多的试验之后，显示出把整个方案的效果往下拉。

但成分网络分析能读到的，是“这个成分平均而言是不是拖后腿”。它读不到的是：**这个成分对这个人封掉了哪一条**。前者需要很多试验臂，后者需要同一个人身上的两条读数。这两件事之间隔着的不是精度，是记录的结构。

八·换轨型为什么不产生信号

一个能被短期指标读出来的坏结果，不需要本章这样的说法——任何人都看得见。换轨型之所以要单说，是因为它有四条一起成立的性质。

第一，它在短期指标上表现为改善。主结局是入睡潜伏期或失眠严重程度，换轨型在这些量上是往好里走的。

第二，它的失效不产生信号。一个人在换轨之后睡得更差了，这件事会以三种方式被吸收掉：归为夜间变异（“今天状态不好”）、归为依从性问题（“他没照做”）、归为难治性（“这个人治疗反应不佳”）。三种归因都不需要动方案，因此方案不会改。

第三，它自我加固。方案越标准化、越有据可依、越被指南推荐，它在不同的人身上封掉的就越是同一条通路。个性化本来是唯一的解药，而循证与标准化的方向恰好与它相反——这不是谁的错，是把同一份处方发给不同路网的必然结果。

第四条是本章加的，也是最要紧的一条——误诊阻挡。现行的试验设计里，**没有为个体设噪声底**。一个人在干预前后差了多少，与他自己在没有干预时两夜之间差多少，从来没有被放在一起看过。没有这个底，“他变差了”这句话在原理上就无法与“他今晚本来就会差这么多”分开。于是第二条那三种归因不只是省事，它们是**唯一可用的归因**——因为分辨它们所需的那次测量，从来没有人做。

第四条把前三条锁死了。哪怕有人愿意认真对待“这个方案让某些人更糟”，他也没有一个可以用来判定的量。

九·一句不带程度词的话

本章的判据是一句话，它不问任何人的判断，只问记录：

在这份记录里，这个人没有接受任何干预的两夜之间，这条读数差了多少？

这句话里没有“应当”“充分”“真正”“显著”。它可以拿去问一份试验方案、一份数据字典、一份伦理申请、一台可穿戴设备的导出文件。它的答案是一个数或者一句“没有这个字段”。

把它在五个场合各问一遍，五个答案互不相同，其中两个是查出来的，而且这两个都反过来改了本章：

场合	答案
一组公开的交叉设计服药记录（二十二入，一夜服药一夜安慰剂）	四点二分钟 （入睡潜伏期的服药／安慰剂之差的绝对值中位数）。 查出来的。
一组公开的无干预连续两夜记录（七十五入）	七点零分钟。查出来的，而且比上一行大。
一份典型的多成分行为疗法随机试验方案	没有这个字段。基线通常是一周的睡眠日记均值，不是同一个人的两个可比夜的差。
一台家用可穿戴设备的睡眠报告	数据存在（它记了每一夜），但从这样呈现——它把这一夜与人群常模比，不与这个人自己的上一夜比。
一次安眠药处方的门诊随访	连一夜的客观记录都没有，只有主诉“比以前快了”。

第一行和第二行是本章动手去算的。算之前我以为第一行会比第二行大很多——毕竟一边有药，一边什么也没有。结果不是。这一件事改掉了本章原来打算立的整个经验主张，改法见第十二、十三、十四节。

十·自噪底：先量这个人自己的两夜

上面那句判据要的那个量，本章给它一个名字：**自噪底**——同一个人在没有任何干预时，同一条读数在相邻两夜之间的差值分布。

自噪底不是一个比率，是一个**底**。它的用法是：任何一次干预在这个人身上造成的改变，先与他自己的这个底比一比，比不过就什么也没发生；比得过，才谈得上开或者封。

这个量在别的领域里都有正主，而且量纲不同、用法相同：

·**气候研究**里它叫**内部变率**，与它配套的量叫**浮现时刻**——一个受迫信号要多久才能大过自然变率而变得可分辨。它是这样定义的：信号除以噪声，超过某个阈值的那一年。气候研究的一条基本纪律是，谈一个变化之前先把控制试验里的自然变率算出来；不算，那个变化就没有意义。

·**交通工程**里它叫**行程时间可靠性**，常用的读数是缓冲指数：同一路段、同一时段、不同日子之间行程时间的波动幅度占均值的比例。交通部门早就不只报平均行程时间了，因为一条平均快五分钟而波动大十分钟的路，对通勤者是更差的路。

·**搜索理论**里它叫**运行时间分布**。同一个实例、同一个算法，只是随机种子不同，运行时间可以是重尾的，方差甚至可以无穷大。这一支的核心结论对本章尤其要紧，第十五节还要回来谈。

三处的量纲不同（年、分钟、步数），比法相同：**先量这一条路自己的抖动，再问你那次改动有没有大过它**。

睡眠医学不是没有能力算这个数——它天天在记录夜。它只是不把这个数算给个体。它算给个体的是与常模的比较（你的睡眠效率百分之多少，正常值是多少），而不是与他自己的比较。

清点手册（一个读数，全篇一个定义、一个阈值、一份编码规则）：

读数名：自噪底 符号： η

定义：同一个人在无干预条件下，同一条读数在两个可比夜之间差值的绝对值中位数。

两个可比夜=同一记录方案、同一场所、相邻或近邻，中间无干预、无跨时区旅行、无急性疾病。

粒度：一“夜”=从该夜的熄灯时刻起到记录结束；判读窗三十秒；

恰好在边界上的例子——一段被判为“运动伪迹”的窗，本章不计为睡眠也不计为清醒，它不中断一段睡眠，但也不算进总睡眠时长。

编码规则：

1. 入睡潜伏期=熄灯起至出现连续两个及以上睡眠窗中的第一个窗，单位分钟。

2. 总睡眠时长=全夜睡眠窗数 $\times$ 0.5分钟。

3. 慢波时长=判读为第三、第四期的窗数 $\times$ 0.5分钟。

4. 入睡往返=熄灯至首个持续第二期之间，由睡眠窗跳回清醒窗的次数。

5. 缺失夜按可用配对处理，并报告实际配对数。

6. 熄灯时刻取自记录元数据；若换算落在记录之外，该夜作缺失。

表头：受试者 | 夜A值 | 夜B值 | 差 | 绝对差 | 是否超出该读数的自噪底

不适用：单夜记录；跨设备比较；主观日记与客观记录混用。

本章三处引用一致性：正文第十节=检验第十一、十二节=证伪第十九节，同一定义、同一粒度。✔

十一·预注册：四条押注写在跑之前#

本章用的是公开的扩展版睡眠脑电数据库。它有两个子集：一个是二十二人的**交叉设计**——每人两夜，一夜服替马西泮，一夜服安慰剂，条件与熄灯时刻在元数据里给出；另一个是**无干预**的连续两夜记录。前者用来看看干预做了什么，后者用来给出自噪底。全部分析只用判读标注，不用原始信号。

在计算任何检验统计量之前，四条押注写定存档：

P1·个体方向一致性。押：服药夜比安慰剂夜入睡更快的人不到九成。判伪：若九成以上的人都更快（即至多两人变慢），则“换轨在主结局上于个体层可见”这条弱预测失败。

P2·形状而非速度。押：入睡往返次数的下降幅度大于入睡潜伏期的下降幅度——即药不只是把同一条路走快，还抹掉了往返这个形状。判伪：若往返的降幅不大于潜伏期的降幅，则“改道”在此读数上不可见。

P3·存在性演示（事先声明为非检验）。服药夜慢波时长低于安慰剂夜。这是教科书上的药理，不作为本章证据，只用来确认“一条读数改善的同时另一条读数变差”在这份数据里有实体可指。

P3b·核心检验。计算“入睡改善量”与“慢波损失量”在个体层的等级相关。押：**相关的绝对值不超过零点五**——即入睡改善得多的人，慢波并不因此损失得更多，两者各自成立（两条路），而不是同一条路上的剂量权衡。判伪：若改善越多、损失越大且相关超过零点五，则应判为单条路上的权衡，本章的多通路说法在这份数据上得不到支持。

P4·路线种类数。押：服药夜出现的不同入睡路线种类少于安慰剂夜。

统计用配对符号秩检验与等级相关，双侧，显著性水平零点零五；不做多重比较校正，此处写明以便读者自行折扣。事先声明的三处局限：药物干预不可外推到行为方案；五个判读分期不足以分辨本章所说的通路，能测的只有“一条读数改善而另一条变差”与“路线形状是否改变”；样本为轻度入睡困难者，非慢性失眠临床样本。

十二·跑出来的结果

二十二入全部可配对。

P1命中，而且命中得比押注还狠。服药夜入睡更快的十人，更慢的八人，持平四人；方向一致率百分之四十五点五。一种公认有效的助眠药，在个体层面上接近掷硬币。

P2方向命中，显著性完全不成立。入睡潜伏期中位降幅百分之十一（符号秩检验 $p=0.486$ ），入睡往返中位降幅百分之六十点七（ $p=0.455$ ）。押的是前者小于后者，方向对了；但**两个检验都远不显著**，也就是说这份数据既没测出潜伏期缩短，也没测出往返减少。方向命中在这里不构成证据。

P3塌了，而且是四条里唯一塌掉的那条。慢波时长的中位数，安慰剂夜六十七点零分钟，服药夜七十八点零分钟——比安慰剂夜更多，符号秩检验 $p=0.987$ ，等于毫无差异。

这一条是本章事先写明“已知药理、必然成立、只作存在性演示”的那一条。它塌了。

原因值得写清楚，因为它本身就是本章命题的一个例子。慢波的判读是基于慢波在一段窗内所占比例的阈值判定，而这类药物的典型作用是增加纺锤波与快频活动、压低慢波振幅——**压低振幅在阈值判读下会不会被记成“慢波变少”，取决于滤波设置与判读手册的版本**。这批记录来自二十世纪八十年代末的遥测方案，剂量与给药时刻不在公开元数据里。所以更准确的说法是：本章用一个“书上写着会发生”的效应当作存在性演示，而它在一份真实的公开记录上没有出现，**而且我事先没有想到它不会出现**。

P3b形式上命中，实质上无信息。入睡改善量与慢波损失量的等级相关是负零点一七（ $p=0.448$ ），绝对值远小于零点五，押注命中。但既然 P3 表明慢波在两种夜之间根本没有系统差异，这个相关系数就是两团噪声之间的相关。**它不能算作“两条独立通路”的支持。**本章在此明写：P3b 的命中，不作为证据使用。

P4命中，而且这条命中是重述。安慰剂夜出现十一种不同的入睡路线，服药夜九种。押注方向对。但把路线按往返次数分组之后看得很清楚：给定往返次数，路线的种类基本就定了（大多数往返次数下只有一种路线）。**也就是说 P4 数的东西与 P2 数的是同一件事的两种写法。**一条由另一条近乎蕴含的预测，不应当被当作两条证据。

四条押注，三条“命中”，而三条命中里有一条无信息、一条是重述、一条只有方向没有显著性；唯一塌掉的那条，是本章以为最稳的那条。

十二之二·换一个分析窗，两条读数就跨过了显著线

上一节说“四条读数在群体上全部不显著”。这句话只在一种口径下成立，必须补正。

上面的计算把分析窗取为“熄灯到记录结束”。若改取睡眠研究里更常用的另一种口径——**熄灯起八小时**——同一批人、同一份判读、同样的配对检验，结果如下：

读数	分析窗=熄灯到记录结束	分析窗=熄灯起八小时
入睡潜伏期	中位差 0.00, $p=0.486$	中位差 0.00, $p=0.486$
总睡眠时长	中位差 +12.25 分, $p=0.135$	中位差 +8.25 分, $p=0.048$
慢波时长	中位差 -2.50 分, $p=0.656$	中位差 +1.00 分, $p=0.503$
入睡往返	中位差 0.00, $p=0.455$	中位差 0.00, $p=0.455$
睡眠效率	中位差 +0.71, $p=0.063$	中位差 +0.93, $p=0.028$

总睡眠时长与睡眠效率这两条，只因为换了一个分析窗，就从不显著变成了显著。两种窗口都是合理的、都在文献里用过，而且这不是谁算错了——是同一份数据在两把同样正当的尺子下给出两个相反的判读结论。

本章因此把上一节那句话收紧为：这份数据里的干预效应，与“用哪一种合理口径去读它”是同一个量级。这比“不显著”更准确，也更要紧。

而更值得注意的是另一半：第十四节那个噪声底对照，在两种窗口下都不翻面。八小时窗下，服药／安慰剂之差与无干预两夜之差的比较，四条读数的单侧检验分别为 0.936、0.959、0.087、0.702——没有一条显示干预之差更大。

一边是随口翻面的组间显著性检验，一边是不随口翻面的个体噪声底对照。这一对照本身就是本章推荐后者的理由，而它是这次补跑意外给出的，不在预注册里。

十三·我自己拆掉本章最漂亮的那个数

跑完预注册的四条之后，我做了一批探索性分析（非预注册，请按此折扣阅读）。其中一个数好看得让人心里发毛：

把入睡潜伏期、总睡眠时长、慢波时长、入睡往返四条读数放在一起，二十二人里有二十人（百分之九十点九）至少有一条读数在服药夜变差；四条全部改善的只有两人（百分之九点一）。

这个数看上去正是本章要的：一次干预几乎不可能只开不封。

但它是假的，而且不需要额外数据就能拆掉。四条读数，如果每一条的方向是独立的、各有一半概率变好，那么“四条全部改善”的概率是二分之一的四次方，也就是百分之六点半五；实测百分之九点一。“至少一条变差”的随机期望是百分之九十三点八；实测百分之九十点九。

实测值与掷四次硬币的期望值几乎一样。这个九十点九没有任何信息量，它只是在说“四件事同时朝好的方向走是罕见的”，而这是算术，不是关于睡眠的发现。

我把它写在这里，不是为了显得诚实，是因为它正是本章要提防的那种数：一个在总量上正确、在结构上空洞的读数。本章批评现行装置把开与封折叠成一个数，而我自己刚刚造了一个把随机性折叠成结论的数。

十四·噪声底，以及本章唯一站得住的那一段

上一节只用算术就拆掉了那个数。但算术给的是理论期望，理论期望可能与真实的夜间波动不一样。所以我又做了一次补充预注册，然后去算真的。

补充预注册（写在运行之前）：用无干预的连续两夜子集，口径与前面逐字相同——同一个熄灯时刻定义、同一个入睡潜伏期定义、同一个往返定义、同一个慢波定义。押注：**服药夜与安慰剂夜之间的差，不会显著大于无干预两夜之间的差**。若押注成立，则本章前面全部个体层数字必须被读作夜间变异，不得据以立论。

七十五人可配对。结果：

读数	服药／安慰剂之差绝对值 中位数	无干预 两夜之差绝对值中 位数	单侧检验 p （前者大于后者）
----	--------------------	--------------------	-------------------

入睡潜伏期	4.2 分钟	7.0 分钟	0.936
总睡眠时长	24.8 分钟	39.0 分钟	0.929
慢波时长	13.8 分钟	10.5 分钟	0.198
入睡往返	1.0 次	1.0 次	0.702
“至少一条读数变差”的比例	90.9%	88.0%	(随机期望 93.8%)

四条读数，没有一条的服药／安慰剂之差大于同一批口径下无干预者相邻两夜之差；其中三条甚至更小。

押注成立。所以：

·第十二节 P1 那个“百分之四十五点五、接近掷硬币”，不能读作“药在换轨”。无干预两夜之间入睡潜伏期变好的比例是百分之五十六，也接近掷硬币。两者的相似正是要点：**这份数据分辨不出药。**

·第十三节那个百分之九十点九，被算术拆过一次，现在被真实的噪声底再拆一次：无干预两夜之间是百分之八十八。

·本章原本打算立的经验主张——“一次有效干预在个体层可见地同时开与封”——**在这份数据上没有任何支持，本章撤回它。**

撤回之后剩下什么？剩下一件比原来打算立的更硬的事，而且它是被这次失败逼出来的：

本章找不到暗封，不是因为它不在，而是因为在现行的记录结构里，它与夜间变异共用同一个读数。一次干预的效应若小于这个人自己的自噪底，它在原理上不可判读——无论它是开还是封。而现行的睡眠干预试验几乎从不为个体建立这个底：基线通常是一周日记的均值或单次记录，不是同一个人的两个可比夜。

这一句不依赖“通路”这个本体是否成立。它是关于**记录与设计**的，可以被任何人在任何一份试验方案上核对。它也是本章最便宜的一条：**做它只需要在基线期多记一夜。**

顺带说一件本章没有做而必须做的事：上面的对照不是完美的。服药子集与无干预子集是两批不同的人、不同的记录方案、不同的年代与设备。严格的做法是在同一批人身上做“安慰剂—安慰剂”两夜，这在公开数据里找不到，本章也无力组织。因此上表只能支持“这份数据分辨不出药”，不能支持“药的效应真的小于夜间变异”。这两句话的差别很大，本章取弱的那一句。

十五·三条通路在别的领域里怎么兑现

一个说法有没有内容，看它能不能在别的地方据以预测一件具体的事。逐条列，不是打比方。

一·**交通网络：关掉一条路可以让所有人更快。**这不是推论，是布雷斯那篇短文的直接结论，而且被后来的分析证明在一般拥堵网络里出现的概率与不出现相当。据此可以预测：在一个多条通路的系统里，**移除**一个看上去有益的成分，可能改善结果。对应到睡眠：从一份多成分方案里拿掉一个成分而结果变好，不是执行失误，是可以预期的一类结果。成分网络分析里“放松训练可能起反作用”那一行，正是这一类。

二·**交通网络：容量跌落。**瓶颈崩溃后的排放能力低于崩溃前的最大流量。据此预测：一次失败的入睡尝试，其代价不止于“今晚晚睡半小时”，而应表现为随后若干夜的**可用通路数下降**。这一条可测：连续记录中，一次异常长的入睡潜伏期之后的三夜，往返次数与路线种类应当高于基线。

三·**交通工程：匝道控流故意维持低于容量。**据此预测：以“今晚一定要快点睡着”为目标的做法，在结构上属于把系统推向容量边缘，其失败率应高于以“把流量维持在可持续水平”为目标的做法。这一条与失眠研究里关于入睡努力的结论方向一致，但来路不同——它来自控制论的容量边缘，不来自自动机心理学。

四·**交通工程：行程时间可靠性与缓冲指数。**交通部门早就不只报平均行程时间，因为平均快五分钟而波动大十分钟的路对通勤者更差。据此预测：一个使平均入睡潜伏期缩短而使其**夜间波动增大**

的干预，对使用者是净损失，尽管它在试验里是有效的。这一条现行睡眠试验完全不报——报的是均值与标准差，不是同一个人的夜间波动。

五·气候研究：单次事件不可归因，只能给概率比。 据此预测：任何“昨晚我睡得好是因为我做了某某”的推断在原理上不可靠，无论那个做法平均而言多有效。这不是提醒人谦虚，是一条结构性的限制：单次抽样与分布位移不是同一类断言。

六·气候研究：内部变率与浮现时刻。 据此得到本章的核心读数（第十节）。反过来，它也给出一条判据：一个真实的信号可以在很长时间里低于内部变率而仍然是真实的。这一条削弱本章，第十六节详说。

七·气候研究：滞后。 强迫与显现之间隔着系统的惯性。据此预测：改变作息条件之后，入睡的改善不应当在头几夜出现；若某个方案宣称“第一晚就见效”，它起作用的多半不是它宣称的那条通路。这一条可以直接用来分辨机制。

八·搜索理论：重尾运行时间分布。 同一个实例、同一个算法，仅随机种子不同，运行时间可以是重尾的，方差可以无穷大。据此预测：同一个人在同样条件下的入睡潜伏期分布不应当是正态的，而应当有一条长尾——大多数夜很快，少数夜极长。这一条可以用任何一份长期记录直接检验，而现行报告只给均值。

九·搜索理论：随机重启可以消除重尾行为。 这是那一支最实用的结论：与其在一条不断回溯的搜索里耗下去，不如放弃当前状态、重新开始。而失眠治疗里最有效的成分之一——刺激控制——给出的核心指令逐字就是这件事：躺了一段时间还睡不着，就起床离开卧室，等有困意再回去。两边是同一个动作，而搜索理论给出了它为什么有效的精确理由：在重尾分布下，重启的期望时间低于继续搜索。这一处兑现是本章最强的一处，因为它不是类比——它给出了什么时候重启不该用：若分布不是重尾的，重启严格变差。据此预测：刺激控制对入睡潜伏期分布重尾的人有效，对分布集中的人无效甚至有害。这是一条可以事先筛选的处方判据，而现行指南对所有人一律推荐。

十·搜索理论：不完备的剪枝会剪掉唯一的解，而失败之后不可分辨。 剪掉一批分支之后搜索失败了，你无法从失败本身判断是“本来就无解”还是“我剪掉了那条唯一的路”。据此预测：一个人在采用某个方案之后仍然睡不着，这件事本身不含任何关于该方案是否封掉了他的路的信息——两种情形在结果上同码。这正是本章第八节第二条签名的形式来源。

十一·临床试验方法学：重复交叉设计。 运动生理学近十年为了判定“个体反应差异是真的还是测量噪声”，发展出重复交叉设计——同一个人多次接受干预与对照，用他自己的重复性给出底。这个方法是现成的、便宜的、已经在别的领域跑熟了。睡眠干预试验极少这样做。第十七节把这一条摆成正面对照。

十六·三家反过来削掉本章的话

三家不是各供一块材料。它们各削掉了本章本来要写的一句更强的话。

气候研究削掉的那句：“落在自噪底以内的干预效应可以忽略。”

浮现时刻这个概念的全部要点是，一个受迫信号在浮现之前已经存在了几十年，而且它最终是决定性的。低于内部变率不等于不存在，只等于不可分辨。因此本章不得从第十四节的结果推出任何“这个药没用”“这个方案可以停”的结论，也不得据此建议任何人改变治疗。本章能说的只有一句：在这份记录里，它不可判读。

这一削动的是本章的价值排序——本章原本打算主张“不可判读的效应不值得追求”，现在这句话被取消了。

交通网络削掉的那句：“助眠办法用得越多越互相拥堵。”

布雷斯效应的成立条件是多个自利使用者竞争同一段有限容量，均衡才会移到更差的点上。一个人独自入睡，没有竞争者，也没有均衡。因此本章只能主张“可用通路数可以减少”，不能主张通路之间存在拥堵式的互相取消，更不能推出“方案越简单越好”。

这一削动的是本章的适用范围——本章原本打算把结论推到“多成分方案的成分数与效果负相关”，现在这一整条被切掉了。

搜索理论削掉的那句：“睡不着就该起床。”

重启优于继续搜索，前提是运行时间分布重尾。分布若不重尾，重启严格变差——你丢掉的进度大于你换来的新起点。因此“睡不着就起床”不是一条普适建议，它是一条**有前提**的建议，而那个前提可测。本章因此不得把刺激控制当作本章说法的证明；相反，本章必须承认：**如果某个人的入睡时间分布不重尾，本章对他的处方是错的。**

这一削是三处里最有用的一处，因为它把一条模糊的适用范围变成了一个可以事先测的量。

十七·与最近的几种说法的分界

有几种现成的说法，读者会正当地怀疑本章只是给它们换了个名字。逐条写，把最危险的排在前面，并且先说清楚：**第一位与第二位，本章初稿都没有想到，是查出来才发现的，而它们比本章说得更早、更成熟。**

（一）个体反应差异的真伪之争。阿特金森与巴特勒姆二〇一五年在《实验生理学》上系统地指出：一次干预前后观察到的个体差异，可能全部由被试内随机变异造成；要判定“真实的个体反应差异”，**必须有对照臂，或者有一项事先的重复性研究。**这一支后来发展出重复交叉设计、个体反应标准差及其批评，至今仍在争论。

*它说到哪一步：*它已经说出了本章第十四节的全部结论，而且早了十一年，说得更严格。

*分离线：*它解决的是**样本层**的问题——这批人里到底有没有真实的个体反应变异，用两个臂的变化量标准差之差来估。它**不给任何一个具体的人一个底**，也不改变主结局是一个数这件事。本章要的不是“这批人里有真实差异”，是“这个人今晚这条读数超出他自己的抖动了吗”，以及“超出的是哪一条”。

*判定形式：*取一批睡眠干预试验，按那一支的方法算出真实个体反应变异。**若结果为零**，则本章的第一层与第二层被推翻，暗封无处安放；**若不为零**，两家一致，而本章多要求的只有一件——把它逐人给出，并且不止给主结局那一条读数。

（二）多成分方案的成分网络分析。近年的大型合并分析在两百余项随机试验、三万余名参与者上把行为疗法拆成成分，判定认知重建、睡眠限制、刺激控制有效，睡眠卫生教育非必需，**放松训练可能起反作用。**

*它说到哪一步：*它已经在群体层证明了“某个被广泛推荐的成分会把整个方案往下拉”。这是本章所说那件事的群体层显影，本章完全接受。它自己也写明，未检出的成分间交互可能动摇结论。

*分离线：*它读的是“这个成分平均而言是不是拖后腿”，需要很多试验臂；本章问的是“这个成分对这个人封掉了哪一条”，需要同一个人身上的两条读数。前者的分辨率再高，也永远是一个成分一个系数；后者要的是一张逐人的通路表。

*判定形式：*若某个成分的“反作用”完全由一类可用现行基线变量识别的人群解释，则本章的“不可判读”这一层被削弱——那时它是可判读的，只是没人去读。若反作用在所有可识别分层上都均匀分布，则本章的说法得到支持。

（三）睡眠限制疗法的急性期损害。已有研究报告，睡眠限制疗法在急性期伴随客观总睡眠时长下降、日间嗜睡上升与警觉性客观受损。

*它说到哪一步：*本领域自己已经报告过“一条读数改善的同时另一条读数变差”，而且给出了客观测量。

*分离线：*它报告的是**已命名的不良反应**——日间嗜睡与警觉性受损，有量表、有报告表、有分级，属于安全性。本章说的是在**入睡这件事本身**上少了一条可走的路，它与主结局在同一维度上，因此永远不会被记为不良反应，只会被记为“这个人效果不好”。

*判定形式：*若把“日间功能”作为共同主结局与入睡潜伏期并列报告之后，个体层的方向不一致消失了，则本章的说法多余；若仍然存在（有人两条都好、有人一好一坏、有人两条都坏），则需要的正是本章说的那张表。

（四）简并与冗余。埃德尔曼与加利二〇〇一年在《美国科学院院刊》上论述：结构不同的元件产生相同的功能输出，这是生物系统鲁棒性的来源。

它说到哪一步：它正是本章“多条并行通路”的生物学正主，而且比本章一般得多。

分离线：它讲的是系统具有多路径这个性质，并把它当作好事；本章讲的是一次干预会减少路径数，而减少这件事不产生读数。前者是状态描述，后者是关于干预与记录的。

判定形式：给系统施加一个提升总功能的干预，简并理论预测鲁棒性不因此下降（多路径是内在的）；本章预测某些个体的可达性下降。取同一批人，测干预前后的“可用通路数”，两家预测分岔。

（五）鲁棒而脆弱。为一类扰动优化的系统，对另一类扰动极度脆弱；这是复杂工程系统的一条一般命题。

它说到哪一步：它给出了“优化会制造脆弱”的一般形式。

分离线：它的脆弱是面对未曾优化的扰动；本章的封闭是在同一类扰动上少了一条路。更要紧的是，那一支的核心是脆弱性的分布，本章的核心是读不出来。

判定形式：若把干预设计成对所有已知扰动都稳健（多成分、覆盖面广），鲁棒—脆弱说预测残余风险移到未知扰动上；本章预测覆盖面越广、封掉的路越同质，因而失败者的类型会更集中而不是更分散。这两条预测方向相反，可用多中心数据判。

（六）布雷斯悖论本身。本章第五节正是用它推翻那条“进度可加总”的假定，因此必须交代本章与它的距离。

它说到哪一步：它证明了在自利选路的均衡下，增加容量可以使所有人变慢。

分离线：它的机制是均衡移动，需要多个使用者竞争同一资源。一个人入睡没有竞争者。因此本章从它那里取的只有一条：可加总的进度量不存在；本章没有、也不能主张入睡通路之间有拥堵式的互相取消。

判定形式：若有人证明入睡通路之间存在类似均衡的相互取消（用一条通路会削弱另一条的容量），那是比本章更强的主张，本章的说法将成为它的一个特例；若不存在，本章仍然成立，而布雷斯只作为形式上的来源。

（七）刺激控制与重启。一九七二年提出的刺激控制疗法要求：躺下若干分钟仍未入睡就起床离开卧室。它是本领域最老也最有效的成分之一。

它说到哪一步：它已经在做本章第十五节第九条所说的那件事，而且早了三十年。

分离线：它的理由是条件反射——把床与清醒解绑；搜索理论给出的理由是重尾分布下重启的期望时间更短。两个理由预测不同的适用范围：条件反射说预测它对“床与清醒已建立联结”的人有效；重启说预测它对“入睡时间分布重尾”的人有效。

判定形式：取一批人，同时测他们的入睡时间分布形状与床—清醒联结强度，看刺激控制的效果与哪一个更相关。两家在此分岔，一次前瞻队列即可判。

（八）未完成的任务与反刍。一九二七年的经典实验表明未完成的任务比完成的任务被记得更牢；后续的睡前反刍与认知唤醒研究是它的延长线。

它说到哪一步：它解释了为什么“没做完的事”会妨碍入睡，处方是把它写下来、清空。

分离线：那一支谈的是内容——什么念头占住了工作记忆；本章谈的是通路数，与念头的内容无关。一个心里毫无挂念的人，同样可能被一个方案封掉他唯一那条路。

判定形式：控制睡前认知唤醒量表得分之后，若个体层的方向不一致完全消失，则本章的说法可被反刍解释吸收；若仍然存在，两者是两件事。

（九）心理治疗的负面效应报告规范。已有框架主张系统地定义、发现并分类心理治疗的副作用，并给出核查表。

它说到哪一步：它已经在推动“心理治疗也要报告不良事件”。

分离线：这类框架处理的是可命名的、与主结局在不同维度上的不良事件。本章说的损失与主结局在同一维度上——入睡这件事本身少了一条路——因此它在结构上不会被登记为不良事件，只会被登记为“疗效不足”。

判定形式：把这类核查表原样用到一批睡眠干预试验上，若“入睡通路减少”这一类事件能被它捕获，则本章多余。本章预测捕获不到——核查表按症状与领域分类，没有“小结同维度上的路径减少”这一栏——因而需要新增一栏。

（十）单元划分与聚合效应。改变计数单元会改变结论，这一支从地理学到时间序列都有成熟文献。

它说到哪一步：它管的是同一份材料在不同划分下给出不同的数。

分离线：本章与它正交。把判读窗、分期数、记录时长全部固定，暗封照样存在——它不是划分造成的假象，是通路数的实际变化。本章明确把这一片地划出去，不主张它。

至此本章自己的那一格只剩下：**同一次干预在同一个人身上同时开与封，而现行装置把两者折叠成同一个数，且没有为个体设底以便把它们分开。**判定不确定性归第一位，成分效应归第二位，安全性归第三位与第九位，系统鲁棒性归第四、五位，单元划分归第十位。

十八·与同一批文章的分界

本章与另外几篇出自同一个人、同一个题域，读者会把它们读成一套。它们不是一套，而且有几处彼此对立。

与《那三十秒算不算一段》（第一章）。那篇说的是判读格子把一段睡眠切成可清点的段落，而有一类段落未被追认因而不构成一样东西。它谈的是**同一份记录内部**的清点方式；本章谈的是**两份记录之间**（干预夜与自噪底夜）的可比性。判定形式：若把判读窗改细，那篇的未追认段比例会变；本章的自噪底不会因判读窗变细而消失——它是夜与夜之间的差，不是窗与窗之间的分类。

与《住在线上》（第二章）。那篇说睡与醒之间有一片有厚度、承担功能而被记为分界的区域。它谈的是一条**读数内部的结构**；本章谈的是**多条读数之间的折叠**。两篇有一处真对立：那篇主张界带自身有产出，因此应当另账登记；本章主张多记一夜基线。若资源只够做一件，两篇给的建议不同，而且本章认为自己那一件更便宜、更先做。

与《必须回答，所以从不犹豫》（第三章）。那篇说判读格式强迫每一窗给出唯一归属，因而名次间的分差被丢掉。它谈的是**判读时**丢掉的东西；本章谈的是**干预时**封掉的东西。两者的共同点只是都在说“账上没有那一栏”，而这个共同点不构成同一个判断——判余在任何分辨率下都不出现，暗封只要多记一夜就能露出一部分。

与《谁替睡者接住下一拍》（第十二章）。那篇说完整的睡眠需要外部接手关键回应，交不出去就是病。本章与它方向相反的一处：那篇讲的是把一样东西交出去（回应位置），交出去是睡眠的条件；本章讲的是一样东西被拿走（一条通路），拿走不产生信号。两篇合起来可推出一条新预测：**一个把回应位置成功托管出去的人，其可用通路数应当增加**，因而他对换轨型干预的耐受性更强。这条预测两篇单独都推不出来。

与《失眠：为什么越想睡越睡不着》。那篇是本题域上最险的一处，因为它已经占满了努力悖论整条：睡眠是撤退不是动作，监督入睡需要意识在场而入睡要求意识退场。**它能反过来削弱本章：**如果一个人在换轨之后睡得更差，最省事的解释是他因为方案而更用力了，而不是他的某条通路被封了。分离线：那篇的机制是**内部的**——用力这个动作本身；本章的机制是**结构的**——可用路数减少，与用不用力无关。判定形式：取一个完全不知道自己在被干预的人，例如仅改变了室温或床垫。若他仍出现方向不一致，则努力悖论解释不到，本章的说法成立；若方向不一致只出现在知道自己在被治疗的人身上，本章应当退回成努力悖论的一个变体。**这是本章最容易输掉的一条，本章明写在此。**

与同一批的另外两篇“怎么办”。同一道题（怎样才能快速入睡、睡得深）另有两篇用不同领域做成，三篇合起来是一组而不是三个重复。其中一篇与本章用的是**同一份二十二人交叉记录**，这一点必须先说清楚。

·一篇问的是**刻度**：从入睡点到首次深睡那一段用了多久，这个量算得出来却从未被裁定过；它的缺口在**仪器**——三十秒的判读窗太粗，而一套九级的细刻度一九九四年就有，至今没进临床报告。

·一篇问的是**验收**：使用者索求“快、深、高效”三样，而他手上的验收工具只给一个分数；它的缺口在**验收方**——实验室分得开，要它的人分不开。它给的读数是索求间距与验收分辨率之比。

·本章问的是**噪声底**：缺口在**设计**——没有为个体设一个比较的底，因而开与封无法分开。

三者的分离线是干净的：**仪器/验收方/设计**，各占一处，且互不覆盖。判定形式：把刻度做细而不给个体噪声底，第一篇的缺口补上而本章的没补；把验收工具做细而不改分析口径，第二篇的缺口补上而本章的没补；只多记一夜基线，本章的补上而另两篇的都没补。三条互不蕴含。

与本章共用同一份记录的那一篇，还有一处必须交代。它报告睡眠效率与总睡眠时长在服药夜显著改善（分别约 $p=0.03$ 与 $p=0.05$ ），而本章第十二节报告这两条不显著。本章第十二之二节已经查明差别的全部来源是分析窗，并复算出两种口径下的两套数字。**这不是两篇之间的分歧，是同一件事的两次显影**：那一篇据此说“动的是没被索求的那一端”，本章据此说“显著与否由口径决定”。两篇不冲突，但任何引用其中一篇数字的人，必须一并知道另一种口径给出相反的判读。

与《下一步该做什么，有几个答案》。那篇同为一道“怎么办”的题，也走三条不同终点的路径，因而形式上最容易与本章混。它问的是：把一个人当下的状态与目标交给几位有经验的人，“下一步”是同一件事、几件不同的事、还是都说不出来——它谈的是**倒推这个动作有没有解**，是知识分布的性质。本章谈的是**并行通路数目在干预后变了没有**，是系统本身的性质。两者的读数形状也不同：那篇是唯一/多解/无解三档，本章是一个连续的底。判定形式：一个下一步唯一确定（倒推有唯一解）而通路数只有一条的人，那篇判为好处境，本章判为最危险的处境——两篇对同一个人给出相反的评价。这是最干净的一条分离线。

十九·不适用的边界，与证伪条件

本章不适用于：单夜记录（没有底就没有本章）；跨设备或跨记录方案的比较（底不可比）；主观日记与客观记录混用（两者的抖动不是同一个量）。也不适用于急性情境性失眠（通路结构在几天内变动太大），以及儿童与倒班人群（这两类的夜间变异结构与本章用的样本不同）。**本章也不主张任何人据此改变正在进行治疗**——第十六节第一削已经取消了这个推论。

证伪条件（其中第四条低成本可实做，且所需数据在公开档案里已经存在）：

·若在一批睡眠干预试验里加设两个可比的基线夜，按本章口径算出每人的自噪底，而**九成以上的人**的干预前后差大于自己的自噪底，则第八节第四条签名不成立——现行装置其实读得出来，本章的“不可判读”是假的。

·若按重复交叉设计测得的真实个体反应变异为零，则本章的第一层与第二层被推翻。

·若把入睡通路操作化为若干条可独立测量的读数，逐条测，而任何一次干预对它们的改变**方向全部一致**（要么全开要么全封，从无一开一封），则“换轨型”那一格是空的，三分类型学退化成两分。

·在任一公开的睡眠干预交叉数据集上，把干预效应与同一批口径下无干预者的两夜差**并排**。本章已经做过一次，结果对本章不利（第十四节）。任何人可以用另一份数据集重做；若多份数据集上干预效应稳定地大于无干预两夜差，则本章第十四节的结论只是这一份数据的局限。

·若成分网络分析在纳入个体层数据后，发现“某成分起反作用”完全由现行基线变量可识别的人群解释，则本章的“不可判读”被削弱。

·若“被封的通路”给出的预测与“总量下降”给出的预测**从不分歧**，则暗封是多余的名字，应当删掉。

·**正向读数（顺序，不是相关）**：若本章成立，则在同一个人身上，**可用通路数的减少应当早于主结局的恶化出现**——先是备用路没了，然后碰上一次扰动才睡不着。若两者同时出现，或主结局先动，则本章的时间结构是错的。顺序比相关难被巧合满足。

证伪主件是一次真跑，不是一个日期。本章不押某年某月某个指南会不会改写；本章押的是一次可以立刻重做的检验，并且已经先做了一次、报告了失败。补充一条**可当场执行的最小实验**：任何一

个人，在开始任何助眠方案之前，先用同一台设备连记两夜、什么都不改，把四条读数各自的差记下来。这就是他的自噪底。此后任何方案的效果，先与这两个数比。做这件事的成本是两夜。

一条不写日期而写条件的赌注：若有一份被广泛采用的睡眠干预指南，要求在基线期采集**同一个人的两个可比夜**并把逐人的差报告出来，则本章的第三层主张命中。**以下四种不算命中：**（一）要求一周睡眠日记均值的；（二）只在方法学建议里提及而不进入报告条目的；（三）报告的是组内标准差而不是逐人的差的；（四）由本章作者或本章推动写入的。

二十·反噬、分层引用与伦理边界

反噬。假定本章这条主张被制度采纳，最省事的实现方式是什么？是把“自噪底”做成一个新指标，进考核，进设备的报告页。一旦它进了考核，最省事的达标办法不是给每个人多记一夜，而是**把底做小**——用更稳定的记录条件、更宽的判读容差、更粗的读数粒度，让两夜之间的差看上去很小，于是任何干预都“超出了自噪底”，于是一切都有效。本章因此会摧毁它要保护的那个东西：一个用来防止把噪声读成效果的量，会被改造成一个把噪声读成效果的量。

我看不到出口。任何“规定底的测量必须标准化”的条款，本身就是把底做小的条款。

分层引用。本章的主张分三层，各自可以被单独引用，也各自有不同的倒塌方式：

·**第一层（最强、最易倒）：**入睡由若干条并行通路承担，一次干预会封掉其中一些。这一层是本体主张，本章没有为它提供任何经验支持——第十二至十四节的检验对它不利。引用它的人必须一并引用这句话。

·**第二层（分析工具）：**干预分净开型、净封型、换轨型；换轨型有四条签名。这一层是分类，它的价值在于组织观察，不在于它已被证实。它可以在第一层被推翻之后仍然有用（作为一张检查表），但那时它只是启发式。

·**第三层（最稳、也最便宜）：**现行睡眠干预试验没有为个体设噪声底，因此个体层面的“这次让他变差了”在原理上不可判读。这一层不依赖通路本体，可以由任何人在任何一份试验方案上核对，也可以由第十四节那样的一次计算复核。临床研究若只采用自噪底作为流程变量，不得顺带声称发现了新的睡眠通路。

伦理与证据边界。自噪底是一个可以被拿去考核人的量，因此写死三条：（一）它指的是记录条件与设计，不是这个人的品质；一个自噪底大的人不是“不配合”“不稳定”，只是他的夜与夜之间本来就差得多。（二）不得为了把这个数做好看，而在临床或安全关键情境里安排额外的记录夜、限制活动或延迟必要的治疗。（三）已经据此对某个人做出不利安排的（例如判定为治疗无反应而停止提供服务），必须公开编码规则并给出复核通道。三条缺一，这个读数不应当被使用。

二十一·本章自己落在哪里

用本章自己的判据检本章自己，检的不是这一篇，是产出这一批文章的那套做法。

这一批关于睡眠的文章是一篇一篇写出来的。每写一篇，它就在这个题域里占住一块地：一个说法、一个名字、一条读数。而占住一块地，**同时就封掉了后面几篇的若干条路**——同一个题域里，一个说法一旦被前一篇用过，后一篇就只能绕行，哪怕那条路本来是最好走的。本章第十八节列了六条绕行的痕迹，每一条都是这样来的。

关键在于：**哪一条被封掉了，不在任何记录里。**已经写出来的那几篇是可见的；因为它们已经写出来而没能写成的那些，不留任何痕迹。这一批文章的“成果清单”上，只有开，没有封。

这正是本章说的那件事，落在本章自己的生产过程中。而且它满足第八节的四条签名。短期指标（篇数、字数、被引）表现为改善；封闭不产生信号；越是形成了一套稳定的做法，封掉的就越是同一条路。第四条尤其贴：**没有任何噪声底可以用来判断“这一篇是不是让下一篇更差了”**——因为不存在一个“不写这一篇”的对照。

还有一层要认账。本章第十二节那条塌掉的预测（以为服药夜慢波必然更少），是我照着教科书写的，没有先去看数据。一个“书上写着会发生”的判断，**本身就是一条被封掉的路**——它封掉的是“先去看一眼”这条路，而封掉这件事在我写下那句话的时候没有产生任何信号。第十三节那个百

分之九十点九也是一样：它好看，所以我差一点就用了它；把它拆掉靠的正是本章推荐的那个动作——先算噪声底，再看你那个数有没有大过它。

还有第三层。本章第十二节原本写的是“四条读数在群体上全部不显著”，这句话我当时以为是一个事实，其实它是一个口径下的事实。发现它需要换一把同样正当的尺子重算一遍——而这正是本章给所有人的那条建议。我在写一篇讲“别把口径当事实”的文章时，自己把一个口径当成了事实。第十二之二节是补正，不是补充。

三个交出去的洞。一，本章说不出一个人到底有几条入睡通路，也说不出怎么把它们逐条测出来；没有这一步，第一层永远无法检验。二，本章的自噪底是用两夜算的，而两夜给不出一个分布，只给出一个点；正确的做法需要多少夜，本章不知道。三，本章用来做对照的两个子集是两批不同的人、不同的年代与设备，严格的对照需要同一批人的两个无干预夜，公开数据里没有，本章也无力组织——这个洞是本章最大的一个，第十四节末已经写明。

注释与材料

关于三家的取用。本章取的是三个领域各自的一条核心判断，均可引到原文，未由本章代拟。交通网络那一条是布雷斯一九六八年的结论及其后续的普遍性分析；气候研究那一条是浮现时刻的定义与算法；搜索理论那一条是重尾运行时间分布与随机重启。三者在各自领域里回答的都不是关于睡眠的问题，本章取用的方式在正文第五、十、十五节逐处写明。

关于数据。全部计算使用公开的扩展版睡眠脑电数据库，只用判读标注文件，不用原始信号。遥测子集为二十二人的交叉设计（每人一夜服替马西洋、一夜服安慰剂，条件与熄灯时刻见随库元数据）；盒子子集为无干预的连续两夜记录，取两夜齐全且可解析者七十五人。判读为三十秒窗的六类标签。熄灯时刻由记录起始时间戳与元数据换算；有三个夜的换算结果早于记录起点，按正文第十节的规则从记录起点起算，并在此报告。

关于口径的一致。自噪底的定义、粒度与编码规则见正文第十节的清点手册；正文第十二、十四节与证伪条件第一、四条引用的是同一个定义、同一个粒度、同一份编码规则。

关于预注册。四条主预测在下载并解析数据之后、计算任何检验统计量之前写定存档；补充预注册（噪声底对照）在该项计算运行之前写定。第十三节的探索性分析在主预测结果产出之后进行，正文已标明并要求读者据此折扣。

关于口径。第十二之二节与第十四节的稳健性复核为事后补跑，非预注册，正文已逐处标明。两处补跑的起因是发现另一种同样正当的分析窗给出相反的显著性判读，这一发现本身也已写入正文。

关于不利结果。第三条预测（服药夜慢波更少）方向相反且无差异，本章未将其移入脚注，而是写在正文第十二节，并在第二十一节把它读作本章命题的一个实例。第十三节自行拆掉了本章得到的最好看的那个数。第十四节据补充预注册的结果撤回了本章原本要立的经验主张。

关于本章没有做的事。本章没有做同一批人的“无干预—无干预”两夜对照（正文第十四节末已写明这是本章最大的洞）；没有把入睡通路操作化为可逐条测量的读数；没有在多份数据集上重复第十四节的对照。

参考文献

- Braess, D. (1968). Über ein Paradoxon aus der Verkehrsplanung. *Unternehmensforschung*, 12(1), 258–268. doi:10.1007/BF01918335. 英译本：Braess, D., Nagurney, A., & Wakolbinger, T. (2005). On a paradox of traffic planning. *Transportation Science*, 39(4), 446–450.
- Youn, H., Gastner, M. T., & Jeong, H. (2008). Price of anarchy in transportation networks: Efficiency and optimality control. *Physical Review Letters*, 101(12), 128701. doi:10.1103/PhysRevLett.101.128701 (勘误：*Phys. Rev. Lett.* 102, 049905, 2009)
- Furukawa, Y., 等 (2024). Components and delivery formats of cognitive behavioral therapy for chronic insomnia in adults: A systematic review and component network meta-analysis. *JAMA Psychiatry*. PMID: 38231522.

·Goldberger, A. L., 等 (2000). PhysioBank, PhysioToolkit, and PhysioNet: Components of a new research resource for complex physiologic signals. *Circulation*, 101(23), e215–e220. doi:10.1161/01.CIR.101.23.e215

第七章 夜里那一端，一直没有动

夜里那一端，一直没有动

为何今天连年轻人也很少有高质量睡眠：恢复的兼任者、独任化，与一个三分之二来自白天的夜间读数

摘要

今天的年轻人睡不好，通常被归给三种原因：睡得太少、睡得太晚、或者恢复能力提前衰退。本章提出第四种，并给出一份可复算的检验。恢复在很长一段时间里并不由任何一件以恢复为名的事情完成，而是由一批主要功能不是恢复的活动顺带完成——走路去办事、做半小时家务、和人面对面说一段废话、通勤路上坐着。这批活动在所有账目上都以别的名义存在，因此它们逐条退出时没有任何一栏数字下降，退出反而以“省下时间”的名义被记为进步。恢复由此集中到睡眠这一条专任通道。本章把前一类称为恢复的**兼任者**，把这一集中称为**独任化**，并给出可清点的读数**弧覆盖率**。用二十二名健康年轻人二十四小时连续记录（MMASH 公开数据集，预注册在先）复算：弧覆盖率中位 10.6%，静止时间中位 35.1%，十九人中十人整天只用到三类身体活动中的一类。随后是四处对本章不利的结果。其一，被广泛用作夜间恢复指标的“下潜幅度”，其变异 66.7% 由白天那一端解释，0.2% 由夜间那一端解释；但这两个数由三个矩完全决定，是算术不是发现，Oldham 一九六二年已给出一般形式。其二，两端方差 3:1 有一半是统计量选择的产物。其三，也是最要紧的一处：同一批人同一夜，换用心搏间期变异构造，两端权重由 (0.667, 0.002) 翻成 (0.045, 0.651)。哪一端在说话，是构造者在减号两边各挑了什么统计量的属性，而这个选择从不公布。判据是一句不含程度词的问话：在这份记录里，恢复是一个类别，还是一个时段？

关键词：兼任恢复者 独任化 弧覆盖率 端权对 数学耦合

The Other End Never Moved

一·三个人的一天

先看三个具体的日子，都是真的，都来自本章所用的那份公开记录。

第一个人二十五岁，那一天他清醒了十七点八小时。在他自己填的活动日志里，他有两小时的“轻度活动”，没有中等强度的活动，没有重度活动。他睡了三百四十八分钟，睡眠效率在这批人里属于上游。任何一份健康报告都会说他一切正常。

第二个人二十四岁，清醒二十一点一小时——这批人里最长的一天。他有三小时落在离开静止形态的类别里，是这批人里最高的几个之一。他的总活动量也是这批人里最高的。他睡了二百三十六分钟，是这批人里最短的几个之一。

第三个人的活动日志里，一整天离开静止形态的时间加起来是四十分钟，占他清醒时间的 3.7%。这是这批人里最低的一个。他的其余时间分布在“坐着”“躺着”“看小屏幕”“吃饭”这几类里。他不是病人，不是老人，也不是加班到崩溃的人。他是一名健康的在校学生。

这三个人有一个共同点：他们中没有一个人，在这一整天里做过任何一件被记为“恢复”的事情——除了睡觉。

这不是他们的疏忽。他们所用的那份记录格式，一共提供了十二个活动类别：睡眠、躺着、坐着、轻度活动、中等活动、重度活动、进食、小屏幕、大屏幕、含咖啡因饮品、吸烟、饮酒。这十二类里没有一类叫恢复。有三类是给加速消耗的物质留的，一类也没有留给恢复。

本章要问的问题是：为什么今天连年轻人也很少有高质量的睡眠。本章给出的回答从上面那件小事开始——恢复曾经由一批不叫恢复的事情完成，而当它们退出时，没有任何一本账上的数字下降过。

二·三种回答，每一种都只承认自己

对同一个问题，有三门学问给出了三种回答。三种都写得很硬，三种都主张只有自己那一条是决定性的。

第一种来自老年学。 它说：决定一个人能睡多好的，是他的恢复能力上界，而这个上界随年龄单向下降。这条主张最干净的证据来自 Klerman 与 Dijk 二〇〇八年那项实验。他们把十八名老年人和三十五名年轻人放进实验室，撤掉闹钟、撤掉日程、撤掉社会安排。每天给出覆盖昼夜周期三分之二的睡眠机会——夜里十二小时，下午再加四小时——连续三到七天。结果是，老年人的总睡眠时长渐近值比年轻人短一点五小时。作者的结论写得毫不含糊：在没有社会与昼夜约束的情况下，老年人的白天睡眠倾向和睡眠的最大容量都降低了。

按这条主张，年轻人的上界还在，所以”年轻人睡不好”必定是别的问题——机会不足，或者行为不当。上界是驱动者，别的都是被驱动的结果。

这一家自己知道一处对它不利的事实，但不往那个方向用：在同一项实验里，受试者的总睡眠时长起初长于他们在家里的习惯时长，然后逐日下降到渐近值。也就是说，那个”上界”不是一个一测就得的常数，它取决于机会是怎么给的、给了多久、分成几段。

第二种来自舞蹈学。 它说：决定一个身体明天还能不能动的，是它今天那一段用力有没有把自己的后半段走完。这条主张的经典表述来自多丽丝·汉弗莱。她把一切动作放在”两个死之间的弧”上：一端是纹丝不动的平衡，另一端是再也回不来的倾倒。**两端都是死。** 动作只存在于弧上，恢复也只发生在弧上——它不是停下来，它是同一条弧的后半段。汉弗莱还说了一句在本章里格外要紧的话：受控的倾倒越极端，恢复就必须越有力。

按这条主张，恢复不是外面送进来的补给，是前一段用力自己生成的。今天的年轻人用力的形态单一而低幅——没有落，就没有复。用力的形态是驱动者。

这一家自己也知道一处不利的事实：**代偿。**当一条弧被外部日程切断——演出季、排练表、装台时间——舞者仍然会把动作做完，只是换一组肌肉、换一条链条去做。而代偿在外观上与正常动作不可分辨。这是舞蹈医学最难办的一件事，也意味着”用力的形态”这条驱动在日程固定时反而变成被驱动的。

第三种来自结构学。 它说：决定一处局部失效会不会扩散成整体倒塌的，不是构件的强度，是有没有替代的传力路径。这条主张已经写进了设计规范。美国国防部的《抗连续倒塌设计》与总务管理局的《替代路径分析与设计指南》都把”替代路径法”列为直接设计途径：把某一个承重构件整个拿掉，要求结构能够跨越那个缺口。规范同时写明，连续倒塌需要两个条件同时成立——一个反常荷载引发局部损伤，**加上一个在连续性、延性与冗余上不足的结构。**

按这条主张，强度不是驱动者，路径配置才是。你可以把每一根柱子都做得很粗，只要它们串成一条链，拿掉一根就全完。

这一家自己知道的那件不利的事实，是本章后面要用的：**塑性铰与内力重分布。**一处形成塑性铰之后，结构不再是原来那个结构，荷载会沿新的路径传递到别处。混凝土设计规范甚至允许设计者据此重新分配弯矩——这等于承认，**边界条件不是给定的外生参数，是被上一轮响应改写的状态变量。**

三·三家为什么撞得起来

这三家不是互补的。它们各自说的是”只有我这条是决定性的”，而它们指的方向互不相同。

	它说什么在推动	被推动的是什么	它给的时序读数
老年学	已经长成的身体，与它可用的代谢条件不容	睡眠的组织方式改变（变浅、变碎、提前）	上界先降，行为后跟；行为怎么改都追不回上界
舞蹈学	做出来的动作，与编排要求的路径不容	身体所处的条件改变（代偿、损伤、训练制度）	弧先断，代偿后起，条件最后变
结构学	加载的次序，与约束的配置不容	结构当下的状态改变（内力去哪里、哪儿先屈服）	损伤先累积，裂纹后可见，破坏最后发生

三家分别把驱动放在三个不同的位置上：一家放在”已成的形态与条件之间”，一家放在”形态与路径之间”，一家放在”路径与条件之间”。这不是三种角度，是三种指向不同的箭头。它们不能同时为真——如果上界是决定性的，那么用力的形态只是上界的表现；如果用力的形态是决定性的，那么路径配置只是形态的容器；如果路径配置是决定性的，那么前两样都只是路径上的读数。

正因为如此，这三家的分歧无法靠“判谁对”化解。你不能把它们排在一根轴上比大小，因为它们根本不在同一个语域里作答。

四·三家都不说的那一句

把三家的争论压到最紧，它们争的是同一件事：**恢复该由哪一样东西推动。**

老年学说由内生的上界推动，舞蹈学说由前一段用力推动，结构学说由约束的配置推动。三家吵得很凶，而它们有一句话从来没有人说出口，因为三家都觉得那还用说吗：

在一轮之内，谁推动谁是给定的。

这句话的后果是具体的。它意味着，“这一轮恢复得不好”总是可以归到那个推动者身上——上界不够、用力不对、约束太紧。它意味着，你只要找到那个推动者，就找到了原因。它还意味着一件更隐蔽的事：**没有人需要问“这一轮的发起处是不是换了人”。**

把这句话念给三家听，三家都会说这还用说吗。会引起争论的那种句子不是这一句，那是第四家的立场。这一句不会引起任何争论——这正是它值得被单独拎出来的理由。

五·推翻它的那件材料，来自结构学自己

要推翻上面那句话，从外面搬一个理由来是没有用的。从外面搬来的理由只是第四种主张，三家可以各自退回自己的阵地不理它。

所幸不必从外面搬。**结构学自己手上就有那件材料，而且用了七十年。**

一个超静定结构在某处形成塑性铰之后，它的刚度分布变了，内力沿新的路径重新分配，下一处屈服会出现在别的地方。塑性设计的整套方法建立在这件事上；混凝土规范允许在一定幅度内重新分配弯矩，正是因为结构自己会重分配。这不是新闻，也不是争议，是每一本结构设计教材都写的常识。

而这件常识说的正是：**约束不是外生的参数，是被上一轮响应改写的状态变量。**

它为什么能推翻那句“谁推动谁是给定的”？因为它在结构学自己的领域里回答的是**另一个问题**——**极限承载力有多大、会不会连续倒塌。没有人把它拿来当作“驱动方向可以在一轮之内换向”的证据。**它一直摆在那里，一直被引用，只是被引用来回答别的事情。

于是三家无法退回阵地：那件材料是结构学自己的，而结构学恰恰是三家里把“约束”说得最硬的那一家。

六·本章的判断

有了那件材料，可以写出本章的判断。

今天年轻人睡不好，不是因为恢复的总量不足，不是因为恢复的时机错位，也不是因为恢复的能力提前下降。

>

而是因为：恢复此前一直由一批不以恢复为名的活动兼任，它们的退出以“效率”之名被记录。恢复因此集中到睡眠这一条专任通道上。而判断这条通道好不好的读数，其变异有三分之二来自另一端——于是“通道数归一”这件事，在读数上表现为读数变好。

三个被否掉的说法，各对应一家：总量不足是老年学之外的常识版本，时机错位是节律那一支，能力提前下降是老年学本身。本章不否认这三样都会发生，也不主张它们是错的。本章主张的是，它们各自都能被单独诊断出来并各有处方，而下面这样至今没有诊断口。

兼任者是本章提出的那一类东西：主要功能不是恢复，而实际承担着恢复的活动。它的定义特征只有一条——**它在任何一本账上都以别的名义登记。**走路去办事登记为交通，家务登记为家务，面对面的闲谈登记为社交，通勤路上那四十分分钟登记为通勤时间。因此当它们逐条退出的时候：

- 交通时间下降，被记为**通勤成本降低**；
- 家务时间下降，被记为**家电普及**；

- 面对面的闲谈被线上消息取代，被记为**沟通效率提升**；
- 站着办事变成坐着点几下，被记为**流程数字化**。

每一条都是真的进步。每一条都没有一个抵扣项。

独任化是本章给这个过程起的名字：恢复由多条兼任通道分担，逐步集中到睡眠这一条专任通道上。它有一个结构学里现成的对照——一个超静定结构逐步退化成静定结构。静定结构在正常使用荷载下的读数完全正常，它只是失去了一件事：**任一处失效不再有别处兜底**。

这个过程不是一次完成的，它有轮次，而且中途换过一次发起处：

轮次	谁先动	另两样多久跟上	回写了什么
第一轮	外部条件先动：通勤消失、体力劳动外包、办事线上化	身体的日常形态在数月内跟上（清醒时间大部分停在静止端）；恢复的组织方式随后跟上（能承担恢复的只剩睡眠）	睡眠在这个人的生活里变得空前重要
第二轮	发起处换成了恢复策略自己 ：为了保护睡眠而进一步削减清醒时段的活动	早睡、睡前不运动、不做耗神的事、减少晚间社交	弧被压得更平，兼任者进一步减少

第二轮是本章最要紧的一段。**它不再需要外部条件推动**。一个人只要相信”睡眠最重要”，他就会开始削减清醒时段里那些看上去与睡眠无关、实则一直在兼任恢复的活动。他做的每一件事都合乎现行的睡眠卫生建议，而他做的每一件事都在减少兼任者。

这就是本章所说的”驱动方向在一轮之内换过向”：起初是条件在推动恢复的重分布，之后是重分布的结果在推动条件继续变化。

七·弧覆盖率

要让上面那段话可以被检验，需要一个可以清点的读数。本章给出的是**弧覆盖率**。

汉弗莱那条弧的两端都是死：一端是纹丝不动，一端是再也回不来。恢复只在弧上发生。把这句话翻成一个可清点的量，就是：

弧覆盖率 A = 清醒时间里身体离开静止形态的时间 ÷ 清醒总时间。

清点手册

一个读数如果不写死粒度和编码规则，它就不是读数。以下六条是本章全篇唯一的口径，正文、证伪条款与赌注三处引用的是同一份。

项	内容
读数名	弧覆盖率，记作 A
定义	清醒时段中”弧上类别”的总分钟数，除以清醒时段总分钟数
弧上类别	在本章所用的记录格式里，是编码 4（轻度活动：慢走中速走、家务、劳作）、5（中等：快走、骑车）、6（重度：健身、跑步）
静止端	编码 2（躺着）与 3（坐着，说明书原文含”驾驶”）
不计	编码 7（进食）、8/9（小屏大屏）、10/11/12（咖啡因、烟、酒）——它们不指定身体形态；另有说明书未定义的编码 0（全体出现 32 次），一并不计并单独报数
粒度	一段活动记录以起止分钟为界；区间重叠先取并集再计时长；同一分钟不重复计入
边界例	编码 3 的说明含”驾驶”。驾驶四十分钟与坐着看书四十分钟，在本章口径里 同判为静止端 ——身体形态没有变。这一判定是本章最容易被质疑的一处，写在这里供人反对
不适用	卧床者、行动受限者、以身体活动为职业者（后者的 A 天然高而与恢复无关）

配套还有一个整数读数：**弧类别数 k** = 编码 4、5、6 三类中，当天累计不少于五分钟类别数，取值零到三。它数的是通道有几条，不是通道有多宽。

四条性质

一个跨领域的读数要能用，得同时满足四条。前三条容易：**无量纲**，所以三个领域可以并排比；**可事后清点**，任何一份带时间戳的活动记录都能算；**把一个印象变成一个数**——“这个人一天到晚坐着”从此有了一个数字。

第四条最容易漏也最要命：**它必须与既有的主要指标正交**。如果 A 与总活动量高度相关，那它就只是总活动量的一个代理，白造。

检验这一条的办法是举一个反例：**既有指标最好看、而这个数最难看的一个真实例子**。本章数据里现成有一对。第二十二号受试者的总活动量是 2,638,424（加速度向量幅值之和），在这批人里属于前几名；他的弧覆盖率是 7.6%，低于中位数。第五号受试者的总活动量是 1,687,318，比前者低 36%；他的弧覆盖率是 11.2%，高于中位数。**总活动量高 56% 的那个人，弧覆盖率低 32%**。

原因不难懂：总活动量把一整天的加速度加起来，一个人可以靠长时间的中等强度单调运动把这个和推得很高，而整段时间里他的身体形态一直没变——他停在弧上的一个点，不是走过那条弧。这真是两个量分开的地方。

在本数据的全部十九人里，A 与总活动量的相关系数是 +0.424（p=0.070）。有关联，不是同一件事。

七之二·一句可以拿去问记录的话

一个判断如果只能拿去问人的感觉，它就不是判据。判据要能拿去问一份文件。本章的那一句不含任何程度词：

在这份记录里，恢复是一个类别，还是一个时段？

这句话在六个具体场合的答案各不相同，其中两个是查出来才知道的，而且都反过来改了本章。

拿去问谁	答案
本章所用那份活动记录格式	是一个 时段 。恢复只作为睡眠表里的若干指标存在；十二个活动类别里有三类留给咖啡因、烟和酒，零类留给恢复
主流可穿戴设备的恢复分数	既不是类别也不是时段，是一个 差值 ，主体由夜间数据算出
职业健康里的“恢复需要”量表	是一个人的 状态 （我还需要多少恢复），不是一件事
运动科学的训练单调性	恢复不单列，它是 负荷序列的一个统计量 （均值除以标准差）
舞蹈的排练表	是 编排里的一个位置 ：互补的用力质地被排进序列，它有名字、有时长、有人负责
结构设计的替代路径法	是一个 被逐条立项的类别 ：每一条备用传力路径要被单独列举、单独验算

两处是查出来的，而且两处都改了本章。其一，本章原本打算把”没有恢复这一类”写成一处显而易见的缺失；查完那份说明书才发现，缺的不是”有一类没写”，而是**写了三类给加速消耗的物质、零类给恢复**。那一格因此从”漏了一栏”改成”两种东西一栏有一栏没有”，锋利得多。其二，本章原本以为”备用通道要单独立项”只是一个比方；查规范才知道那是一条强制的验算程序——**结构工程把本章要的那件事做了几十年，逐条列举、逐条移除、逐条验算**。

最后两行答”类别”，前四行不答。别处做了，这里没做。

七之三·十二处兑现

跨领域的判断不能停在”像”。以下十二处，每一处都能据本章的读法在那个领域里预测一件具体的事。

#	领域	兑现
1	结构工程	一个逐步失去冗余的框架，在使用荷

		载下的挠度读数与新建时几乎一样；预测：变化出现在 移除任一构件后的剩余承载上 ，而这一项不在日常检测项目里
2	结构工程	局部加固会把内力吸引到加固处；预测： 只加固一处的改造，最弱点会搬家，而不是消失
3	舞蹈训练	演出季日程固定时，舞者仍能把动作做完；预测： 代偿出现在肌群分配上而不是外观上 ，因此录像评估读不出来
4	舞蹈训练	汉弗莱那条弧的两端都是死；预测： 要求”多休息”的干预与要求”多训练”的干预可以同时失败，二者不构成一根轴
5	老年医学	储备耗竭前静息指标正常；预测： 同年龄段内的静息读数不能区分储备高低，只有应激试验能
6	老年医学	上界取决于机会如何分配；预测：把同样时长的睡眠机会 分成两段给 ，渐近值会与一次性给不同
7	职业健康	若恢复靠”需求停止”，则静止时长应与夜间恢复正相关；本章实测-0.085，预测该关系在更大同质样本里 仍然近零
8	运动科学	训练单调性与弧类别数在”每天做同样负荷、不同形态”的方案上 给出相反的风险预测 ，这一对可以直接分开检验
9	交通与通勤	通勤时间下降被记为成本降低；预测： 通勤取消前后，个人活动类别数的下降幅度大于活动总量的下降幅度
10	家务与家电	家电普及被记为时间节省；预测： 节省下来的时间流向静止端而非弧上 ，因此总活动量下降幅度小于类别数下降幅度
11	可穿戴产品	恢复分数是差值；预测： 同一个人在加大训练量的周里，恢复分数会上升，而其夜间心率底几乎不动
12	临床试验报告	凡以”某端读数改善”为主要终点的试验；预测： 把该终点拆成两端重报，多数试验的效应会落在非目标那一端

第十一与第十二两处是本章最愿意被人拿去检验的。它们不需要新的被试，只需要把已有的数字分开印。

八·两端归属表

本章的辨别装置不是一张分类格，是一张**归属账**。

理由是这样的。睡眠领域里被广泛使用的恢复指标，很多是**差值**：白天减夜里、清醒减睡眠、基线减当下。可穿戴设备给出的”恢复分数”“身体电量”“就绪度”，其主体成分是这一类。差值有一个性质，日常使用时几乎不会被想起来：

一个差值的**变异，可以几乎全部来自它的某一端，而使用它的人以为自己**在读另一端。

所以本章提出，凡是一个被当作”某一端的成绩单”的差值读数，都应该配一张表，逐行写清：这一端的方差多大、那一端的方差多大、两端相关多少、差值的方差里各端单独能解释多少。

这张表的空表头如下，本章第十二节会把它填满。

栏	含义	为什么必须单列
---	----	---------

端一方差	被减数在人群中的离散度	两端方差相差越大，差值越像大方差那一端
端二方差	减数在人群中的离散度	同上
两端相关	被减数与减数之间的相关系数	相关越高，差值抵消掉的共同成分越多
端一单独解释	差值对端一做单变量回归的决定系数	这是“差值其实在测哪一端”的直接读数
端二单独解释	差值对端二做单变量回归的决定系数	同上
端权对	上两栏并列写成一对数	两个份额之和不等于一，缺口就是两端的共同成分
白天变量→端一	每一个白天变量与被减数的相关	用来看干预动的是哪一端
白天变量→端二	每一个白天变量与减数的相关	同上，这一栏是全表最要紧的一栏

最后一栏为什么最要紧：一个人做了运动、控制了屏幕、按时上床，第二天他的”恢复分数”上升了。这句话有两种完全不同的解释。一种是他夜里确实沉得更深；另一种是他白天更累了。在只报一个差值的报告格式里，这两种解释在数字上完全一样。

⚠️ 这张表在本章初稿里是一个空壳，本章当时以为把它填满就会得到一个经验发现。第十二之第二节会说明它并不是——**这张表的前五栏由三个数完全决定，它不是发现的工具，是翻译的工具**：它把一个已经被报出来的差值，翻成”这个数在说哪一端”。最后两栏才需要看数据。

九·数据与预注册

数据

本章用的是 MMASH (Multilevel Monitoring of Activity and Sleep in Healthy people) v1.0.0, PhysioNet 公开，开放数据库许可，任何人可下载复算。受试者是二十二名健康年轻男性，意大利比萨大学学生，每人二十四小时连续记录。所含数据有：逐秒心率与三轴加速度、逐拍心搏间期、按十二类编码的活动日志（起止到分钟）、睡眠指标（在床、入睡、效率、总睡眠时长、觉醒次数），另有唾液皮质醇与褪黑素以及若干问卷。

选它的理由有三条。第一，它是本章所知唯一同时具备**分类活动日志**与**逐秒生理读数**的公开二十四小时数据集。第二，它记录于二〇二〇年前后，受试者是**今天的年轻人**——本章问的正是”今天”。第三，它的受试者在关键混淆变量上高度同质：同一所大学、同一批年龄、同一性别、同一套记录方案。这一点很重要：跨国家、跨地区的两两对比在方法上撑不住任何”控制掉某变量”的说法，而这里是同一制度环境内的多个单元。

预注册

统计之前先写死，文件 PREREG.md, SHA-256:

8df05c6173d0838efa4ffbc8f9a5f243724c656da4e8f52a2e51d851790c277

预注册内容包括第七节那份清点手册、下面五条假设与判定线、以及一条撤销规则。

- **A1 (描述性，不做检验)**：报告弧覆盖率的中位与四分位。预设中位低于 0.15。
- **A2 (主检验)**：控制总活动量与总睡眠时长后，弧覆盖率与夜间下潜的偏相关，放行线 $\geq +0.25$ 。
- **A3**：弧类别数与夜间下潜的等级相关，放行线 $\geq +0.25$ 。
- **A4**：按弧类别数分两组 (≤ 1 与 ≥ 2)，比较组内”总活动量与夜间下潜”的相关强度，预设通道少的一组耦合更强，差值 ≥ 0.20 。
- **A5**：同龄段内年龄与夜间下潜的等级相关，预设绝对值 < 0.25 。
- **撤销规则第一条**：若 A2 的偏相关小于等于零，或绝对值小于 0.10，则”弧覆盖率驱动夜间恢复”这条主张在本数据上整体撤回，原文照录在正文供核对，且不许换一个读数重试。

夜间下潜的定义也在预注册里写死：清醒时段心率中位数，减去主睡眠段心率的第十百分位。主睡眠段取总睡眠时长最长的那一段。

代理变量声明

这几条写在正文而不是脚注，因为它们改变了本章能主张什么。

活动日志是**自报的**，粒度到分钟但由受试者事后填写，不是客观测量。心率由腕表推得，不是心电。夜间下潜是”夜间自主神经下潜”的一个代理，**不是恢复本身**，本章不主张二者等同。n=19、全部男性、全部同一所大学的学生、每人只有一天——**这不构成任何人群推论**。

剔除与数据缺陷

二十二人中三人被剔除，逐个写明理由：两人的主睡眠段心率完全缺失（睡眠段有效心率采样点为零），一人无睡眠记录行。剩余十九人全部进入分析，无二次剔除。

另有两处数据缺陷必须公开：其一，第十八号受试者的年龄字段记为 0（其余十八人为 22 至 34 岁）；其二，活动编码 0 在全体中出现三十二次，而说明书未定义这一编码。前者的敏感性分析见十一节末，后者按清点手册不计入任何类别。

十·结果一：一天停在哪里

先看那个不需要任何统计检验的描述。

读数	中位	四分位	最小	最大
弧覆盖率	10.6%	7.8% / 13.3%	3.7%	18.2%
静止端覆盖率（躺+坐）	35.1%	—	0.0%	80.4%
屏幕覆盖率（小屏+大屏）	5.8%	—	—	—

十九个健康的年轻人，清醒时间的中位数是十七点八小时。他们身体离开静止形态的时间，中位数占清醒时间的**十分之一稍多一点**。最低的那个人是 3.7%——四十分钟。

弧类别数的分布更直接：

弧类别数 k	人数	该组夜间下潜中位	该组弧覆盖率中位
0（三类都没用到）	0	—	—
1	10	22.5 bpm	8.7%
2	7	28.0 bpm	11.5%
3	2	30.5 bpm	13.5%

十九人里有十人，一整天只用到轻度、中等、重度三类身体活动中的一类。两人用满三类。

这三个数——10.6%、35.1%、十九分之十——是本章全部主张里最不依赖统计的一层。它们不需要相关系数，不需要 p 值，不需要任何因果假设。它们只是清点。

而且它们不是关于老人的，不是关于病人的，不是关于 99% 的。这是一批二十来岁、健康、在读大学的年轻人的普通一天。

十一·结果二：主检验命中，而它的意思与本章预期的相反

按预注册的判定线，五条里四条命中。

项	放行线	实测	判定
A1 弧覆盖率中位	< 0.15	0.1063	命中
A2 控制后 弧覆盖率 ~ 夜间下潜 偏相关	≥ +0.25	+0.338 (p=0.157)	字面命中
A3 弧类别数 ~ 夜间下潜 等级相关	≥ +0.25	+0.439 (p=0.060)	命中
A4 通道少的一组耦合更强	差 ≥ 0.20	+0.165	未命中
A5 同年龄段内年龄 ~ 夜间下潜	绝对值 < 0.25	-0.130	命中

A3 的稳健性另做了两项：五千次置换检验的单尾 p 为 0.030；留一法下等级相关落在 +0.354 到 +0.569 之间，没有一个人单独撑起这个结果。A2 的留一区间是 +0.255 到 +0.395。

如果本章就此打住，它会是一篇看上去很顺的文章。弧覆盖率高的人夜间下潜更深，通道数多的人夜间下潜更深，年龄在同龄段内不解释任何东西——三条连起来正好是本章第六节那段话的经验版本。

本章没有打住，因为夜间下潜是一个差值，而第八节那张表还没有填。

填的办法很简单：把差值拆成两端，分别看每一个白天变量动了哪一端。结果如下。

白天变量	- 白天那一端（清醒心率中位）	- 夜间那一端（睡眠心率第十百分位）
弧覆盖率	+0.608 ($p=0.006$)	+0.413 ($p=0.079$)
弧类别数	+0.504 ($p=0.028$)	+0.238 ($p=0.327$)
总活动量	+0.400 ($p=0.090$)	+0.058 ($p=0.812$)
总睡眠时长	-0.300 ($p=0.212$)	-0.186 ($p=0.446$)

请注意第二列的符号。它们全是正的。

弧覆盖率高的人，夜里的心率不但没有更低，反而更高（+0.413）。他们之所以“下潜幅度”更大，是因为白天那一端被抬得更高（+0.608），减数几乎没动。

再控制一次：把白天心率、总活动量、总睡眠时长一起控制掉之后，弧覆盖率与夜间心率底的偏相关是 **+0.162** ($p=0.508$)——方向仍然是错的，而且不显著。

换一个更干净的恢复读数，效应干脆消失。用逐拍心搏间期算出的夜间与日间 RMSSD 之比（十九人全部可算，中位 1.396），与弧覆盖率的相关是 **+0.116** ($p=0.636$)，控制总活动量与总睡眠时长后是 +0.090；与弧类别数是 +0.159 ($p=0.515$)，控制后 +0.139。

换三种夜间读数，结论一样：弧覆盖率与“睡眠心率第十百分位”的相关是 +0.413，与“全夜最低心率”是 +0.424，与“全夜心率中位”是 +0.341。三个都是正的。

于是必须写下这一句：

本章原先设想的因果链——“多用几条恢复通道，夜里就沉得更深”——在这份数据上没有得到支持，而且方向相反。A2 按预注册的字面标准命中了，但它命中的是一个我事先没有想清楚的量。

撤销规则第一条的字面条件（偏相关小于等于零，或绝对值小于 0.10）没有触发，所以按纪律本章不能事后改阈值把它判成触发。本章能做也必须做的是另一件事：承认预注册里那个读数是**我自己设计的一处缺陷**。我把“夜间恢复”操作化成一个差值，却没有在预注册里要求把两端分开报。这不是运气不好，是设计阶段就该看出来的——一个减法里有两个数会动，而我只准备读其中一个。

还有一个方向相反的旁证。静止端覆盖率与夜间下潜的相关是 **-0.085** ($p=0.731$)，基本是零。一个人一天里躺着坐着的时间长短，与他夜里下潜多深没有关系。“停下来就能恢复”这句话，在这份数据上一点支持也没有。

最后是数据缺陷的敏感性：把年龄记为 0 的那一位剔除后重算，A2 的偏相关是 +0.335，A3 的等级相关是 +0.414，年龄与夜间下潜的等级相关变为 -0.279。结论不因这一位而改变。

十一之二·十二个白天变量，一个也没通过多重比较校正

第十一节那张表只有四个白天变量。既然要判“白天有没有动到夜间那一端”，就该把手上能算的都算了，而不是只算对本章方便的那几个。这一节是打磨时补的，非预注册。

白天变量	- 白天那一端	- 夜间那一端	- 差值
弧覆盖率	+0.608 ($p=0.006$)	+0.413 ($p=0.079$)	+0.466 ($p=0.044$)
弧类别数	+0.504 ($p=0.028$)	+0.238 ($p=0.327$)	+0.463 ($p=0.046$)
总活动量	+0.400 ($p=0.090$)	+0.058 ($p=0.812$)	+0.462 ($p=0.046$)
静止端覆盖率	+0.031 ($p=0.899$)	+0.170 ($p=0.486$)	-0.085 ($p=0.731$)
屏幕覆盖率	+0.232 ($p=0.339$)	+0.026 ($p=0.914$)	+0.274 ($p=0.257$)
重度活动覆盖率	+0.192 ($p=0.431$)	+0.292 ($p=0.226$)	+0.030 ($p=0.903$)
进食覆盖率	-0.470 ($p=0.042$)	-0.071 ($p=0.774$)	-0.541 ($p=0.017$)
咖啡因覆盖率	+0.295 ($p=0.220$)	+0.299 ($p=0.214$)	+0.155 ($p=0.527$)

饮酒覆盖率	-0.076 (p=0.759)	-0.412 (p=0.080)	+0.204 (p=0.401)
清醒总时长	+0.313 (p=0.192)	+0.086 (p=0.727)	+0.332 (p=0.164)
总睡眠时长	-0.300 (p=0.212)	-0.186 (p=0.446)	-0.243 (p=0.316)
入睡时刻	+0.076 (p=0.756)	+0.235 (p=0.333)	-0.075 (p=0.761)

十二次检验里最小的 p 值是 0.079，而 Bonferroni 校正后的阈值是 0.0042。没有一个通过。

这一表逼出两条更正，两条都对本章不利。

第一条：本章的标题比数据强。“夜里那一端一直没有动”是一个断言，而这十二次检验支持的只是一句更弱的话——这份数据分辨不出白天的任何一个变量有没有动到夜间那一端。分辨不出与没有动，是两件事。标题保留，因为它是这道题的形状；但读者应当按这一节读，不按标题读。

第二条：那个唯一看上去有点意思的负号也不能用。饮酒覆盖率与夜间心率底的相关是 -0.412 (p=0.080)。方向与常识相反（酒精通常抬高夜间心率），样本里只有少数人有饮酒记录，而且这一条在十二次检验里排第二，本身就在多重比较的靶心上。本章不据它立论，只把它写在这里供人去更大的样本里查。

十二·结果三：66.7% 对 0.2%

把第八节那张空表填满。

栏	值
端一（白天心率中位）方差	56.32 (均值 79.9 bpm, 标准差 7.50, 区间 69-93)
端二（夜间心率第十百分位）方差	18.77 (均值 54.1 bpm, 标准差 4.33, 区间 47-64)
两端相关	+0.610 (p=0.006)
差值方差	$35.40 = 56.32 + 18.77 - 2 \times 19.85$
端一单独解释	66.7%
端二单独解释	0.2%
四分位间距	白天端 7.0 bpm 夜间端 4.5 bpm 差值 7.5 bpm

这一行是本章全部经验内容里最硬的一行：

一个被广泛用作”夜间恢复”的读数，它在人群里的变异，三分之二可由白天那一端单独解释，只有千分之二由夜间那一端单独解释。

原因是算术的，不需要任何生理学。两端相关 +0.610，说明有一个共同的个体因素同时抬高两端——很可能就是这个人的基础心率水平。相减把这个共同因素抵消掉了大半。抵消之后剩下的，是两端各自的”额外变异”；而白天那一端的方差是夜间那一端的三倍。差值于是几乎就是大方差那一端。

这条算术有一个直接后果，写在这里让每个人自己去核对自己的设备：

如果一个人今天做了更多的运动，他的”恢复分数”会上升。上升的原因可以是他夜里沉得更深，也可以只是他白天更累了。在只给一个差值的报告里，这两种情形的数字完全一样。

而本章第十一节那张表说的正是：在这十九个人身上，动的是前一种解释的反面——白天变量把白天那一端抬上去了，一个也没有把夜间那一端压下去。

这就补上了第六节那个回路里最难解释的一环。独任化为什么不产生信号？为什么恢复通道从多条掉到一条，没有任何一个数字下降？

因为信号本身来自另一端。一个人减少了兼任通道（少走路、少做家务、少出门办事），把剩下的力气集中到一次高强度训练上，他的白天心率中位会上升，他的”恢复分数”会上升，他的设备会告诉他今天恢复得不错。通道数从三条掉到一条这件事，在他能看到的每一个数字上都表现为改善。

这就是本章要打的那一格。它通过一切形式审查：他确实更活跃了，数据确实是真的，设备确实没有说谎，医生看到的每一项指标确实在合格范围内。三条签名齐全——

- 1. 短期指标表现为改善：通道集中之后，被留下的那条通道读数上升；
- 2. 失效不产生任何信号：兼任者在所有账目上以别的名义存在，退出时无一栏下降；
- 3. 自我加固：分数变好 → 更相信这条路对 → 继续把力气集中到这条路 → 兼任者更少。

十二之二·这两个数是算术，不是发现

上一节那两个数是本章初稿里最得意的一处。打磨时我把它们放回算式里，它们塌了一半。

差值 $R = \text{端一} - \text{端二}$ 。只要知道两端的方差与它们的协方差，两端各自对差值的解释份额就已经完全确定，不需要看任何一个人的数据：

- $\text{Cov}(\text{端一}, R) = \text{Var}(\text{端一}) - \text{Cov} = 56.32 - 19.85 = 36.47$
- $\text{Cov}(\text{端二}, R) = \text{Cov} - \text{Var}(\text{端二}) = 19.85 - 18.77 = \mathbf{1.08}$
- 由此 $R^2(\text{端一}, R) = 0.6674$, $R^2(\text{端二}, R) = 0.0018$

与实测逐位相同。也就是说，66.7% 与 0.2% 不含任何超出这三个矩的经验内容。把它们说成”发现”是不对的。

这件事一九六二年就被 Oldham 说清楚过：取两串互相独立、标准差相同的随机数 x 与 y ， x 与 $x-y$ 的相关约为 0.707，也就是解释份额 0.5。一个差值天生就与它自己的某一端相关，这不需要任何机制。后来这件事在临床医学里有了一个名字，叫**数学耦合**（Archie 一九八一年），并且成为一条方法学禁令：不要把一个差值拿去与它自己的组成部分做相关或回归。

把 Oldham 的零情形按本章实测的方差比推一遍：

情形	端一份额	端二份额
两端独立、等方差（Oldham 一九六二）	0.500	0.500
两端独立、方差比 3.00:1	0.750	0.250
本章实测（方差比 3.00 、两端相关 +0.610 ）	0.667	0.002

于是本章能主张的经验内容只剩两条，而且都很朴素：**两端的方差比是 3.00；两端的相关是 +0.610**。剩下的都是算术。

★但那个 1.08 值得单独看一眼。它是端二与差值的协方差，几乎正好是零——因为两端的协方差（19.85）与端二的方差（18.77）几乎相等。这意味着夜间那一端与”下潜幅度”在这批人里近乎正交：知道一个人下潜多少，你对他夜里心率底一无所知。这一条本章认为是真的，但它是三个矩的一次巧合，换一批人就不一定还在小数点后那么近。本章不把它当作规律，只把它当作一个提醒：正交不是设计出来的，是碰上的。

十二之三·3:1 的方差比，有一半是统计量选的

还有一处更难看的。本章的”端一”取的是清醒心率的中位数，”端二”取的是睡眠心率的第十百分位。一个是中间的统计量，一个是下尾的统计量。下尾统计量天然被压缩，方差本来就该小。

把两端放在**配对的分位**上重算：

分位	白天端方差	夜间端方差	比
0.05	21.65	16.62	1.30
0.10	23.43	18.77	1.25
0.25	35.60	21.33	1.67
0.50	56.32	25.51	2.21
0.75	77.94	30.23	2.58
0.90	137.18	53.51	2.56
0.95	222.54	83.34	2.67

白天端在**每一个**配对分位上方差都更大，所以”白天端更散”这件事本身不是分位选择造出来的。但比值不是 3.00，是 1.25 到 2.67；本章报的 3.00 落在这个范围的上沿，而它之所以在上沿，正是因为把中位数配了第十百分位。

再换一把尺子。白天心率的水平本来就高（均值 79.9），夜间低（均值 54.1），绝对方差大一点是意料之中的。按**变异系数**（标准差除以均值）比：

分位	白天端	夜间端	比
0.10	0.0735	0.0801	0.92

0.50	0.0939	0.0868	1.08
0.90	0.1198	0.1095	1.09

在相对尺度上，两端的离散程度几乎一样。

于是本章必须再收窄一次：“白天那一端更散”只在绝对单位上成立，而且幅度比本章初稿报的小；在相对单位上它不成立。这一处收窄不是修辞，它改变了本章能说的话——本章不能再说”白天那一端天生更有信息”，只能说”在这个具体的构造下，白天那一端的绝对离散度更大，因此差值随了它”。

十二之四·换一个量，说话的那一端会翻

前两节把本章的数字削掉了一大半。这一节是打磨轮里唯一一处新东西，而它比被削掉的那些硬。

同一批人、同一批夜、同一份记录，换一个生理量：不用心率，用心搏间期的逐拍变异（RMSSD，十九人全部可算）。同样取白天与睡眠两端，同样做归属分解。

	白天端	夜间端	两端相关	端一份额	端二份额
心率（下潜幅度 = 白天 - 夜间）	方差 56.32	方差 18.77	+0.610	0.667	0.002
心搏间期变异（夜间 - 白天）	方差 61.3	方差 167.9	+0.407	0.045	0.651

方向翻了。

用心率构造，差值几乎完全随白天那一端；用心搏间期变异构造，差值几乎完全随夜间那一端。两者用的是同一批人、同一批夜、同一段睡眠、同一段清醒。变的只有一样：放在减号两边的是哪一个统计量。

这一处把本章的主张从”白天那一端在说话”改成一句更一般、也更难反驳的话：

一个差值读数里，说话的永远是绝对离散度大的那一端；而哪一端更散，由构造者挑了什么量、在两边各取什么统计量决定。它既不是这个人的属性，也不是这一夜的属性。

由此可以给出本章真正要的那个东西，它是一对数而不是一个数：

端权对：一个差值读数的两端，各自对该差值的解释份额。写成两个数并列。心率下潜是（0.667，0.002），心搏间期变异比是（0.045，0.651）。

端权对有三条性质，值得写在使用它的地方。其一，它不需要新数据——两端的方差与协方差三个数就能算出来，而这三个数在任何一份原始记录里都是现成的。其二，两个份额之和不等于一（本章两例分别是 0.669 与 0.696），因为两端不独立，剩下的那部分落在两端的共同成分里。其三，它可以事先算，不必等结果——一份研究方案在收集数据之前就能报出它打算用的那个差值的端权对，只要它给出两端的预期方差。

以及一条使用纪律，它是本章最不肯让步的一条：报告一个差值读数而不报端权对，等于把一个数交给读者，而不告诉他这个数在说哪一端。

十三·最强的三个竞争解释

一个判断如果只挑好打的靶子，它的证伪就是装饰。以下三条是本章能想到的最难反驳的三种”其实不就是……吗”。

其一：其实不就是体能吗。 体能好的人白天敢动、夜里静息心率低，两端一起变，于是”下潜幅度”大。这条解释确实能说明两端相关 +0.610 的来路，本章接受它。但它说明不了第十一节那张表：如果体能是主因，弧覆盖率高的人夜间心率底应该更低，而实测是 +0.413，方向相反。分岔点是明确的：体能解释预测弧覆盖率与夜间心率底负相关；本章的观察是正相关。若在更大样本里这个符号翻成负的，本章第十一、十二两节的读法应当被推翻。

其二：其实不就是自报日志不准吗。受试者事后填写活动日志，粒度粗、有回忆偏差，所以弧覆盖率是噪声。这条解释有力，本章无法排除。但它有一个可检的推论：如果弧覆盖率主要是噪声，它不应该与任何客观读数有强关联。实测它与逐秒加速度算出的总活动量相关 +0.424，与白天心率中位相关 +0.608——噪声不会有这两条。所以更可能的情形是：弧覆盖率含有真实信号，只是它测的东西与本章原先想的不一樣。

其三：其实不就是样本太小吗。十九个人，一天，一个性别，一所大学。这条完全成立，本章不辩解。但要注意它的方向：小样本会让真效应显不出来，不会让效应方向翻转。第十一节那四个正号（+0.413、+0.238、+0.058、+0.424）是四个独立读数上的同向结果，样本量小不解释这个一致性。至于第十二节那两个数（66.7% 与 0.2%），它们是方差的分解，不是显著性检验——方差多大就是多大， n 小只影响估计的精度，不改变”白天那一端方差是夜间那一端三倍”这个事实的方向。

按机制证伪的格式写：

控制个体基础心率水平之后，若”弧覆盖率”与”夜间心率底”之间不存在负向的剂量—反应关系，则本章第六节所说的”通道分担改善夜间恢复”这一环为伪。

这一条本章自己已经跑过一次，结果是本章输了：控制白天心率、总活动量与总睡眠时长后，偏相关 +0.162。本章因此不再主张这一环，只保留读数归属那一段。

十四·三家反过来给本章加的三条限制

一个跨领域的判断，如果三个来源只贡献材料而不砍掉它任何一句话，那多半是它没有真正进过那三个领域。以下三条各砍掉了本章原本打算说的一句更强的话。

老年学砍掉的：本章不得主张”恢复能力不下降，全是通道的问题”。Klerman 与 Dijk 的实验把社会与昼夜约束整个撤掉，老年人的睡眠上界仍然更低。这不是通道能解释的。本章的适用范围因此缩到：上界尚未成为约束的年龄段。若没有这一削，本章原本会写下这样一句更响亮的话——“睡眠能力的年龄差异是通道差异的副产品”。那句话是错的。

结构学砍掉的：本章不得把弧类别数当成一个可加的量。替代路径法要求逐条移除、逐条验算，正是因为路径之间不可相加——拿掉不同的构件，剩下结构的行为完全不同。而且规范只做一次移除，因为第一次重分布之后结构已经不是原来那个结构。弧类别数因此只计通道存在与否，不度量它们的可替代性。没有这一削，本章原本会写下”通道数越多越好，处方是把 k 做大”。那正是本章不能给的处方。

舞蹈学砍掉的，是三条里最狠的一条：本章不得主张”弧覆盖率越高越好”。汉弗莱那条弧有两个端点，两端都是死。纹丝不动是一种死，再也回不来是另一种死。把一个人从静止端推向弧上是一件事，把他推过弧的另一端是另一件事，而后者也是死。

这一削的分量在于，数据正好在同一个地方打了本章的脸：弧覆盖率高的人夜间心率反而更高（+0.413）。学科的限制与数据的打脸指向同一处。本章因此把主张收窄为一句更弱、也更站得住的话：

弧覆盖率高与夜间恢复深浅之间的关系，在本章这份数据上没有被判定；本章只主张这个量目前根本没有被记录，因此这个关系至今没有人有资格判定。

十五·与最近的几种说法的分界

以下十二条，每一条都写明在什么观察下本章对而该说法错、在什么观察下相反。凡是只能写出”本章更强调”“本章更系统”“视角不同”的，本章视同承认被那一条覆盖。

第一，数学耦合与初值定律（Oldham 一九六二；Archie 一九八一，《外科学年鉴》193:296-303；上溯 Wilder 一九三一年的初值定律）。这是本章第十二节那一击的正主，比本章早六十四年，而本章初稿一个字也没提。写在第一位，是因为漏掉它这件事本身就是本章第八节要讲的那个毛病：它不冷门，它是临床方法学的标准警告，只是登记在”变化的测量”名下。

Oldham 证明：两串互相独立、标准差相同的随机数 x 与 y ， x 与 $x-y$ 的相关约为 0.707。Archie 把这一类问题命名为数学耦合——一个变量直接或间接包含了另一个的全部或一部分，然后二者被拿去做相关。Wilder 的初值定律（一九三一）则是同一件事在生理学里的老名字：反应的大小依赖于初始值；后世反复论证，那个负相关里有一大截是 $x(y-x)$ 的算术效应而非生理效应。

本章接受这一整支，并且承认第十二节那两个数就是这个效应本身。分离线不在数字上，在对象上，而且是可判定的：

数学耦合是一条给分析者的禁令：不要把差值拿去与它自己的一端做相关。本章说的是一条给报告者的要求：既然差值被印出来并被人依据它行动，就必须同时印出它两端的方差与协方差。

判定形式：一个研究者可以完全遵守数学耦合的禁令——从不做那种相关——而本章所说的问题原样存在（他仍然只印一个差值）；反过来，公布端权对不需要做任何一次耦合回归（它由三个矩算出）。两件事互不蕴含。若能找到一份文献，在提出数学耦合警告的同时要求报告两端的离散度，则本章这一格不成立。

第二，血压的夜间下降率（一九八八年至今，四十年文献）。这是本章最近的应用领域正主，本章初稿同样零命中。下降率的定义是（白天平均 - 夜间平均）÷ 白天平均，白天那一端既在分子里又在分母里，整个领域对此心知肚明。该领域已经反复报告：夜间下降的分类（下降者与非下降者）重复性很差，三五月内三成到四成的人会翻类。SAMPLE 研究更进一步——治疗引起的左心室质量变化与二十四小时、白天、夜间三个平均值的变化都显著相关，相关系数都在 0.33 以上；唯独与昼夜差值的变化不相关（收缩压 -0.03，舒张压 -0.008）。

这一条既是本章最危险的祖宗，也是本章第四个“别处做了、这里没做”的正面对照，而且是最强的一个。高血压科的动态血压报告单上，白天平均与夜间平均是分开印的，下降率只是第三行；医生看得到两端。睡眠与恢复类产品的报告只印一个分数。判定形式：若找到一款面向公众的恢复追踪产品，其报告界面像动态血压报告单那样把两端分印，则本章第三层主张不成立。

第三，差值分数问题（Cronbach 与 Furby，一九七〇年，《心理学公报》74:68-80；上溯 Lord 一九五六、一九五八）。该文论证的是：当两端高度相关且方差相近时，差值的信度低于两端各自的信度，因为减法削掉了个体间的真实变异而保留了测量误差。

分离线：他们说的是信度，本章说的是归属。而本章这一处恰恰不满足他们所描述的典型情形——两端方差不是相近的（56.32 对 18.77）。但必须同时承认第十二之三节那一处：按变异系数比，两端几乎一样散，所以“方差不相近”这句话只在绝对单位上成立。判定形式：若在更大样本里两端的绝对方差趋于相等，则本章这一处退化为经典差值分数问题、无独立贡献。同族还有 Pearson 一八九七年那篇关于比值伪相关的经典。

第四，训练单调性（Foster，一九九八年，《运动医学与科学》30:1164-1168）。这是本章最危险的一位——它已经在实务里用了二十七年，说的也是“不是量的问题，是形态的问题”：单调性等于是周内日负荷的均值除以标准差，负荷高而单调性高的时段与过度训练、疾病显著相关，独立于总负荷。

分离线有现成的文献背书：单调性度量的是负荷大小在日与日之间的变异，按后续研究者的说法，它“并不反映训练内容的多样性”。本章的弧覆盖率与弧类别数度量的是同一天之内身体形态的类别数。判定形式：取一批人，让他们每天做同样时长、同样强度、但形态不同的活动（周一游泳、周二爬楼、周三搬东西）。单调性预测这批人风险高（日负荷完全一样，标准差为零），本章的读数预测风险低（ $k=3$ ）。若这批人的结局与单调性一致，本章的读数被它吸收；若与本章一致，两者是两个量。

第五，努力—恢复模型（Meijman 与 Mulder，一九九八年）。职业健康心理学的正主，二十七年来来的框架底座。它的核心假设是：工作引起的负荷反应是可逆的，只要需求停止，生理系统就回到基线，这个回归过程就叫恢复。

这一条与本章正面对立，而且本数据构成一个直接的反例：静止端覆盖率与夜间下潜的相关是 -0.085。一个人一天里躺着坐着的时间越长，他夜里并不沉得更深。判定形式：若在更大样本里”需

求停止的时长”与夜间恢复读数呈正向剂量—反应，则该模型对而本章错；若这个关系持续为零或为负，则“恢复等于需求停止”这一条至少不完整。这也正是汉弗莱那句“纹丝不动是两种死之一”在数据上的位置。

第六，恢复体验四因子（Sonnentag 与 Fritz，二〇〇七年，《职业健康心理学杂志》12(3):204）。心理脱离、放松、掌控、控制感，其中“掌控”指的正是“以另一种方式用力也能恢复”——这几乎就是汉弗莱那条弧的问卷版。

分离线在对象上，不在程度上：**该量表处理的是闲暇时间里的体验**，它的整个框架把恢复安放在“下班之后”。而本章所说的兼任者**既不是工作也不是闲暇**——走路去办一件工作上的事、做家务、上班路上发呆，在该量表的四个因子中一个也放不进去。判定形式：给一批人做该量表，同时清点他们的兼任通道数。若量表得分能预测通道数（相关 > 0.6 ），本章这一格被它覆盖；若两者近乎正交，则量表照不到这一类活动。

第七，必要多样性定律（艾什比，一九五六年，《控制论导论》第十一章）。“只有多样性才能消解多样性”。本章“通道数”这个想法的结构祖宗，比本章早七十年。

本章不主张对这条定律有任何首创权，也不主张自己是它的一个应用。分离线在账目而非结构：艾什比给出的是调节者与被动调节者之间的一条不等式，它默认两边的多样性都是可以清点的。本章说的恰恰是**一边压根没有被清点**——兼任者不在任何一张清单上，因此这条不等式的一边在实践中从来没有被填过。判定形式：若能找到一份现行的健康或劳动记录，其中列举了“承担恢复功能的活动类别”，则本章这一格不成立。本章找了一份最完整的公开数据集，它有十二个类别，其中三个留给咖啡因、烟和酒，零个留给恢复。

第八，一般适应综合征（塞利耶，一九三六、一九五〇）与生理安全余量（坎农，一九三二年，《身体的智慧》）。两位都在一九八〇年前把“储备耗尽之前读数正常”这件事说清楚了：抵抗期的各项指标看上去是好的，直到衰竭期。

分离线：他们说的是一个**储备量在被耗尽**，本章说的是**分担这项功能的通道数在减少**。判定形式：储备说预测干预方向是“减少总消耗”；本章预测“总消耗不变而通道数增加也应有效”。这两条给出的处方不同，而且可以在同一批人身上分开检验——本章承认自己这一半至今没有被检验过。

第九，全天候资本主义（克拉里，二〇一三年，《24/7：晚期资本主义与睡眠的终结》）与倦怠社会（韩炳哲，二〇一〇年）。这两本书在人文一侧把本章第六节那个回路说得比本章更早也更彻底：睡眠是最后一块未被殖民的领地；深度无聊的消失使注意力扁平化。

本章与它们的关系不是竞争，是分工，而且本章欠它们的比它们欠本章的多。分离线在**可裁决性**上：它们给出的是一个时代诊断，本章给出的是一个可以在任何一份带时间戳的活动记录上清点的数。判定形式：本章的读数可以在个体层面上被推翻（例如第十一节那一处），它们的诊断不能。这不是本章的优点，是两种工作的不同职责。

第十，社会时差（Roenneberg 一路，二〇〇六年起）。工作日与休息日睡眠中点之差，是“时机错位”这条解释的正主，也是本章第六节三个被否掉的说法里的第二个。

分离线：社会时差度量的是**睡眠区间的位置**，本章度量的是**清醒时间**的内容。判定形式：取两个社会时差完全相同的人，一个 $k=1$ 、一个 $k=3$ 。若他们的夜间恢复读数无差别，则本章的量被社会时差覆盖；若有差别，则两者正交。本数据无法做这个检验（缺工作日/休息日的配对），本章如实记为未判。

第十一，久坐行为研究（sedentary behaviour 一支，以及“运动也救不了久坐”那一系列）。这一支已经把“静止时间长本身有害，独立于运动量”说了十几年，本章的静止端覆盖率与它高度重叠。

本章不主张这一格。**明确划出去：**静止时间的健康后果归这一支。本章留下的只有两句。一句是这一支度量的是**总时长而非通道数**：两个人同样静止 35%，可以一个 $k=3$ 、一个 $k=1$ 。另一句是本章第十一节那个零相关（静止端与夜间下潜 -0.085 ），它说明**这一支的读数与夜间恢复读数之间的关系至今没有被判定**。判定形式：若久坐时长能预测夜间恢复读数（相关 > 0.3 ），本章这两句都不必说。

第十二，可穿戴设备的恢复分数效度研究。这一支比对设备读数与实验室金标准，问的是“准不准”。

分离线是最干净的一条：**准确性不等于归属**。一个完全准确的差值读数，其变异仍然可以几乎全部来自它不打算测的那一端。判定形式：找一份已发表的恢复分数验证研究，看它有没有把两端的方差分别报出来。若有，则本章这一节是重复劳动；若没有，则本章要求的只是一个不需要任何新数据的重算。

一次没做成的外部核对，以及它自己说明了什么

本章第十八节第二条要求：拿一批已发表的研究，把两端的方差分别算出来，看方差比是不是普遍地不等。打磨时试着做了这件事，**没有做成**，检索词与结果如实记在这里。

检索走的是动态血压这一支，因为它是唯一一个几十年来同时报告两端平均值的领域。检索词：白天与夜间收缩压均值及标准差、昼夜下降率的重复性、两端相关系数、人群研究。检索到的文献确实成对报告两端的**平均值**，而报告标准差时，报的多半是**每个人在该时段内部的血压波动**（血压变异性，是一个个体内的量），而不是**各人该时段平均值之间的离散**（个体间的量）。端权对需要的是后者，而后者在我检索到的表里没有成对出现。

这一次失败本身与本章的第三层主张同形：一个领域可以把两端的位置都印出来，却仍然不印两端的离散度——于是端权对在已发表的表上算不出来。但这只是一次未完成的检索，不是一个结论。任何持有原始动态血压数据的人，用三行代码就能算出端权对；本章把它列为第一顺位的待办，且不预测结果。

十六·与同一批另外几篇的分界

本章属于同一批以睡眠为题的系列，作者相同。同一批里的兄弟篇是彼此最近的邻居，而它们通常互不引用。下面逐篇给出判定形式。

《只要有一条路是通的》（第六章）（交通学、气候学、逻辑学）——本批里最险的一篇，因为它也讲“并行通路”。那篇的主张是：快速而深入地入睡，靠的是在若干条并行的入睡通路里保住至少一条可通行；每种助眠办法在开一条的同时可能封掉另一条，而封这件事不产生任何读数，它把这叫做暗封。

两篇的分野在三处，任一处都足以分开。其一，**通路的时间尺度不同**：那篇说的是一夜之内、入睡这一件事上的并行通路；本章说的是清醒时段里承担恢复的通道，尺度是一天到数年。其二，**退出的原因不同**：那篇的通路是被某个干预顺手封掉的，封的动作有一个明确的施动者；本章的兼任者不是被谁封掉的，是以效率之名被逐条移除的，而移除方并不知道自己在移除什么。其三，**缺口的位置不同**：那篇的缺口在设计——试验不为个体设噪声底，所以“这次让他变差了”原理上不可判读；本章的缺口在归属——两端都记了，而报告只给它们的差。

判定形式有两条，互不蕴含。给每个人在基线期多记一夜、做出个体噪声底 ⇒ 补那一篇，不补本章：噪声底做出来之后，恢复读数仍然是一个差值，两端仍然不分报。把每个差值读数的两端分别公布、并报各自的方差份额 ⇒ 补本章，不补那一篇：两端分报之后，个体仍然没有噪声底。

顺带说一处，那篇引了布雷斯悖论——加一条路可以让所有人变慢。本章第十四节被舞蹈学砍掉的那一句（不得主张弧覆盖率越高越好），方向与布雷斯悖论一致。两篇在这一点上互相加固，而不是互相覆盖。

《三样东西，一个分数》（第五章）（工程学、戏剧学、信息学）。那篇讲的是三个索求端被压成一个分数，缺口在使用者的验收分辨率——实验室分得开，要它的人分不开。本章讲的是两个来源不同的量被相减成一个差，而差的变异三分之二来自其中一端。判定形式：把分数拆成分项公布 ⇒ 补那一篇，不补本章（每个分项仍然可能是一个差值）；把每个差值的两端分报 ⇒ 补本章，不补那一篇（两端分报之后，多个索求端仍然被压成一个分数）。两篇处理的是同一段路上的两次不同合并。

《下去的那一段，没有人记》（第八章）（生理学、物理学、哲学）。那篇的缺口在仪器——三十秒的判读窗太粗，一九九四年就有九级刻度却没进报告。本章的缺口不在刻度：本章所用的读数逐秒可得，精度绰绰有余。判定形式：把刻度做细⇒补那一篇，不补本章。

《那三十秒算不算一段》（第一章）《住在线上》（第二章）《必须回答，所以从不犹豫》（第三章）《何谓睡眠：谁替睡者接住下一拍？》（本批的四篇”是什么”）。这四篇的缺口都落在同一份记录内部：一个段落有没有被追认、一条分界有没有厚度、一次判定的第二名去了哪里、一份记录里有没有”谁接住下一拍”这个字段。本章的缺口不在记录内部，在报告的一次算术上——被减数和减数都被完整记录了，只是相减之后只公布结果。判定形式：给记录增设新字段⇒补那四篇，不补本章（本章要的两个数早就在记录里）；在报告里把已有的两个数分开印⇒补本章，不补那四篇。

同批的另一篇”为什么”（心理学、文学、材料学）。那篇问的是为何年老之后难以深睡，落点是”进得去待不住”——进入深睡的次数几乎不随年龄下降，而每次停留的时长掉了三分之二。本章问的是今天的年轻人。判定形式：那篇的读数是同一段睡眠内部的段数与段长分列；本章的读数是清醒时段的一个覆盖率加一个差值的两端归属。两篇甚至不在同一个时段里取数。此外那篇给了本章一个现成的对照：在一九八〇年代末采集的那批健康人里，二十五至三十九岁组的深睡占比是17.19%、每段平均长度6.26个判读窗、零深睡者为0.0%——那一代的年轻人睡得很好。本章所问的”今天”，正需要这样一个基线，而本章自己的数据没有它。

《失眠：为什么越想睡越睡不着》（同作者，站内早前一篇）。那篇讲努力悖论：你派出一个守夜人去监督入睡，而入睡要求意识退场。本章与它方向相反：那篇的守夜人是内部的，监督的是入睡本身；本章的削减是外部的，削的是清醒时段里与入睡无关的活动。判定形式：若一个人完全不监督自己入睡、也不焦虑，而他的兼任通道数为一，本章预测他仍然睡不好，那篇预测他睡得好。这是一条可以在同一批人身上分开的预测。

不过必须承认一处紧张：本章第二十节给出的处方要求人清点自己一天的活动类别，而那篇会说，任何形式的自我监测可能都是在派守夜人。本章接受这个风险，并在处方里写死了限制——这项清点针对清醒时段，不针对睡眠，且不在睡前进行。

十七·不适用的边界

以下六种情形，本章的判断不适用，写在这里是为了让它可以被反对得更准。

其一，以身体活动为职业的人。建筑工、护工、快递员、舞者的弧覆盖率天然很高，而这不说明他们恢复得好——他们更可能落在弧的另一端。第十四节那一削在这里是硬的。

其二，行动受限者与卧床者。弧覆盖率对他们没有意义，而他们同样需要恢复。本章不主张这个读数适用于所有人。

其三，睡眠障碍患者。本章的数据是健康人。呼吸暂停、不宁腿、失眠障碍各有自己的机制与治疗，本章不涉及。

其四，上界已经成为约束的年龄段。见第十四节老年学那一削。在上界本身已经很低的人身上，通道的事情可能完全不重要。

其五，只有一天记录的任何个体判断。本章的读数在个体层面上不稳：每人只有一天，无法区分”这个人一贯如此”与”这个人今天如此”。本章的所有数字都是人群层面的描述。

其六，本章不主张兼任者的减少是坏事。通勤时间下降、家务被机器接手、办事不必跑腿，这些是真实的好处，而且很多人为了它们付出了很大代价才得到。本章主张的只有一件事：它们的恢复功能从来没有被记过账，因此这笔账在退出时也没有被结过。

十八·证伪条件

以下八条，任一成立本章即为伪或需大幅修订。第一条已经跑过一次，且判本章错，仍然留在表上。

1. 已跑，本章输。控制个体基础心率水平后，若弧覆盖率与夜间心率底之间不存在负向的剂量—反应关系，则”通道分担改善夜间恢复”这一环为伪。实测偏相关+0.162，方向相反。该环已撤回。

2. 已尝试，未完成（见第十五节末）。取任一批已发表的恢复读数验证研究，把两端方差分别算出。若在多数研究里两端的绝对方差相近（比值落在 0.7 到 1.4 之间），则本章第十二节所依赖的“方差不等”这一前提不成立，本章那一节退化为经典差值分数问题的重述。本章检索到的表里两端标准差没有成对出现，这一条至今没有被判定。2a. 新增（打磨时补）。在另一份含两端连续读数的公开数据里重算端权对。若心率类差值与心搏间期变异类差值的端权对不再指向相反的两端，则第十二之四节那次方向翻转是本数据的偶然，本章最要紧的那一节应当撤回。2b. 新增。把两端放在配对分位上重算方差比。若在多数数据里配对分位下的方差比落在 0.8 到 1.25 之间，则“两端离散度不等”整条不成立，端权对退化为一个恒等于 0.5 附近的常数、不携带信息。本章自己这一次的结果是 1.25 到 2.67，已经贴着这条线的上沿。
3. 若能找到任何一份被广泛使用的健康、劳动或运动记录格式，其中列有“承担恢复功能的活动类别”这样一个字段，则本章的第三层主张不成立。
4. 在同一批人身上同时清点兼任通道数与恢复体验四因子量表得分，若两者相关超过 0.6，则本章这一格被该量表覆盖。
5. 让一批人保持每天总负荷与日间变异完全不变，只改变身体形态的类别数。若结局与训练单调性一致而与弧类别数无关，则本章的读数被单调性吸收。
6. 取社会时差完全相同的两组人，弧类别数分别为一与三。若两组的夜间恢复读数无差别，则本章的量被社会时差覆盖。
7. 若在一份 n 大于两百、含多日重复测量的公开数据里，弧覆盖率与夜间心率底的相关变为负号且显著，则本章第十一节那次打脸是小样本的偶然，本章应当撤回第十一、十四两节的收窄，恢复原来的主张。本章欢迎这个结果。
8. 若在多日重复测量里，弧覆盖率的个体内变异远大于个体间变异（组内相关系数小于 0.2），则这个量不是一个人的属性，而是一天的属性，本章所有涉及“某人通道少”的表述都要改写。

样本纪律的自我限制：以上第二、五、六条都涉及跨案例比较。本章明确要求这些检验在同一制度环境内的多个同质单元上做——同一所学校的多个院系、同一家公司的多个班组、同一个平台的多个社群。用两个国家或两个地区的对比来做这些检验，本章不承认其结论，无论结果对本章有利还是不利。

一条赌注

赌注写条件不写日期，因为本章要的那件事随时可以做，不需要等任何人。

本章赌：在二〇三〇年十二月三十一日之前，至少有一款面向公众的睡眠或恢复追踪产品，在它的报告界面上把“恢复分数”的两端分别显示，并注明各端在人群中的方差份额。

四条不算命中：

- 只在开发者文档或白皮书里说明公式的，不算——本章要的是报告界面。
- 把两端分开显示但不给方差份额的，不算——分母不写，使用者仍然无法判断自己在读哪一端。
- 由本章作者或其所在机构推动写入的，不算。
- 仅在科研版或研究者接口中提供的，不算——本章的那一格在普通使用者那一侧。

十九·反噬预言

一个判断如果被制度采用，最省事的实现方式往往正好摧毁它要保护的东西。本章的反噬有两层。

第一层：弧覆盖率一旦进入考核，最省事的达标法是在文书上做手脚。让员工每两小时站起来走三分钟并打卡，弧覆盖率立刻上升，而那三分钟与恢复没有关系。这不是假想——凡是可以清点的行为读数，都会走上这条路。

第二层，比第一层严重得多：本章的判断一旦被采纳，最自然的处方是“多安排几种活动”，而这可能恰好是错的。第十四节那一削说得很清楚——弧有两个端点，两端都是死。一个已经很疲惫的

人被要求增加活动类别数，他去的方向可能是另一个端点。而本章的数据在这一点上正好站在反面：弧覆盖率高的人夜里心率更高。

我找不到出口。本章提出的这个量，其唯一诚实的用法是描述与研究，而它的形状恰恰是最容易被拿去当指标的那种形状——无量纲、可清点、听起来像个好东西。我能做的只有把它写在这里，并在下一节把限制写死。

还有一处更小的反噬要说：本章第十八节那条赌注要求把两端分报。真做了之后，会出现一个新的**问题——使用者会去优化那个夜间端**。而夜间心率底很大程度上是一个人的基础属性，短期内改不动；把一个改不动的数摆在人面前，会制造一种新的挫败。这一处本章也没有解法。

二十·自反：本章自己的读数，承重也落在它不打算测的那一端

本章的判据必须能用来判本章自己。用一次，结论是不好看的。

第一件。本章造了一个读数叫弧覆盖率，并声称它测的是“恢复通道的分担情况”。而第十一节那张表说：这个读数与白天心率中位的相关是 +0.608，与夜间心率底的相关是 +0.413，与总活动量的相关是 +0.424。**它与白天那一端的关联，明显强于它与它声称要说明的那一端的关联。**也就是说，本章自己造的读数，其承重同样落在它不打算测的那一端——这正是本章第十二节指控整个行业的那件事。

第二件。本章在预注册里把“夜间恢复”操作化为一个差值，却没有要求把两端分开报。**我写了一篇关于“差值读数的两端必须分报”的文章，而我自己的预注册里没有这一条。**两端分报是在数据跑完之后、在探索性分析里才做的。按本章第十八节第二条的标准，我自己那份预注册不合格。

第三件，也是最要紧的一件。本章的核心主张是：兼任者之所以退出而无人察觉，是因为它们在账目上以别的名义存在。而**这篇文章本身就是一个实例——第十五节那位最危险的祖宗（Cronbach 与 Furby，一九七〇年）在本章的第一版设计里完全没有出现。**它不冷门，它是心理测量学教科书里的标准章节，被引用了五十六年。它躲过了检索，不是因为它藏得深，是因为我在“睡眠”“恢复”“活动”这几个词里找它，而它登记在“变化的测量”这个名下。

一件东西不在你的清单上，可以只是因为它登记在别的名下。这句话是本章的主张，而它第一个应验在本章身上。而且不止一次——数学耦合、血压夜间下降率这两条更近、更致命的祖宗，本章初稿同样一个字没提，它们分别登记在“外科统计方法”与“高血压”名下。

第四件：本章的标题比本章的数据强。“夜里那一端一直没有动”是一个断言；第十一之二节那十二次检验支持的只是“这份数据分辨不出”。标题保留，因为它是这道题的形状，但它属于本章自己第十九节所说的那种最省事的实现方式——一个听上去像结论的说法，比一个诚实的说法好记。

第五件，也是最难写的一件：本章最得意的那个数，是一九六二年的算术。66.7%与0.2%由三个矩完全决定。我把一次减法的必然后果当成了一个经验发现，写了整整一节，还给它起了名字。发现这件事的办法不是新的检索技巧，是把自己的数字放回算式里推一遍——而这件事本可以在写下它的当天就做。本章第十八节要求别人做的事情，本章自己晚做了一轮。

二十一·如果要用这个读数，先接受三条限制

弧覆盖率和弧类别数是可以拿去考核人的读数。凡是这种形状的东西，都必须配限制，否则不该发表。

- 1. 这个读数指的是位置，不是人。**一个岗位的弧覆盖率低，说的是这个岗位的活动结构，不是坐在那里的人懒。任何把它写进个人绩效的用法，都不在本章的授权范围内。
- 2. 不得为了把这个数做好看，而在安全关键的岗位上安排活动轮换。**手术台、驾驶舱、值班室、流水线上的关键工位，其安排应当由安全考虑决定，不由任何恢复读数决定。
- 3. 已经按这个读数对任何人作出不利安排的，必须公开编码规则，并给出复核通道。**编码规则指第七节那份清点手册的完整版本——哪些类别算弧上、粒度如何、边界例怎么判。不公开这三样，读数就不可复核，不可复核的读数不得用于对人的判定。

三条缺一，本章认为这份稿子不该被用于任何考核场景。

二十二·三层可以分开引用的主张

本章的三层主张强度不同，倒下的顺序也不同。引用者可以只取其中一层。

第一层（最强，最易倒）：恢复此前主要由兼任者完成，它们的退出以效率之名被记录，恢复因此独任化。这一层至今没有直接证据。本章的数据支持不了它，第十一节那次打脸打的正是它的因果环节。它现在是一个待检验的假说，不是一个结论。

第二层（中等）：弧覆盖率与弧类别数是两个可清点、与总活动量不重合的量，而现行的健康与活动记录不产生它们。这一层由第十节那三个数（10.6%、35.1%、十九分之十）与第七节那对反例（第二十二号与第五号受试者）支撑，不依赖任何因果主张。

第三层（最稳，也最便宜）：凡是被当作“某一端的成绩单”的差值读数，都必须同时公布端权对——两端各自的方差、两端的协方差，以及由此算出的两个解释份额。本章给了两个实例，而这两个实例指向相反的两端：心率下潜是（0.667, 0.002），心搏间期变异比是（0.045, 0.651）。同一批人、同一夜，换一个量，说话的那一端就翻了。

这一层不需要新数据、不需要新被试、不需要任何新设备——三个矩在任何一份原始记录里都是现成的，一份研究方案甚至可以在收集数据之前就报出它的端权对。第一层与第二层全部倒掉，这一层照样成立；第十二之二节证明前一版的两个数是算术之后，这一层反而更硬了——因为它要求的正是那个算术所需的三个输入，而那三个数至今不被公布。

引用纪律两条。其一，若只采用弧覆盖率作为流程变量，不得顺带声称本章发现了恢复的新机制。第一层与第二层是可以分开取的。其二，若采用端权对，不得把它读成“这个读数是错的”。端权对不判对错，它只说这个数在说哪一端；一个端权对为（0.9, 0.0）的读数，在只关心那一端时是完全合格的读数。

二十三·注释与材料

数据。MMASH (Multilevel Monitoring of Activity and Sleep in Healthy people) v1.0.0, Rossi A., Da Pozzo E., Menicagli D., Tremolanti C., Priami C., Sirbu A., Clifton D., Martini C., Morelli D. (2020), PhysioNet, 开放数据库许可 v1.0。配套数据论文见 *Data* 5(4):91 (2020)。下载与目录结构见 PhysioNet 项目页；本章所用文件为每位受试者目录下的 Activity.csv、Actigraph.csv、sleep.csv、RR.csv、user_info.csv。

预注册。文件 PREREG.md, SHA-256 为 8df05c6173d0838efa4ffbc8f9a5f243724c656da4e8f52a2e51d851790c277，写于任何统计量计算之前。

脚本。run.py 执行预注册的五条；run2.py、run3.py、run4.py 为探索性补跑，三者都在预注册之外，本章所有标注“探索性”或“打磨时补”的结果均出自这三个脚本。四个脚本与预注册文件一并留档，可按上述哈希核对。

口径一致性自查。弧覆盖率的定义（第七节）、证伪条款（第十八节）、赌注（第十八节末）三处引用的是同一份清点手册，阈值与粒度一致。夜间下潜的定义在第九节、第十一节、第十二节三处一致。

探索性与预注册的分界。第十节与第十一节前半（五条判定线）为预注册内容；其余全部为探索性分析——第十一节后半、第十一之二节的十二变量表、第十二节的归属分解、第十二之二至十二之四三节（矩推导、配对分位、心搏间期变异）。本章的探索性分析全部用于拆本章自己，没有一处用于加固本章原来的主张；第十二之二与十二之三两节削掉的正是本章自己最想要的那个数。

参考文献

2. Humphrey, D. (1959). *The Art of Making Dances*. New York: Rinehart. (Grove Weidenfeld 1991 重印本, ISBN 0-8021-3073-9)
3. Laban, R., & Lawrence, F. C. (1947). *Effort*. London: Macdonald & Evans.

18. Ashby, W. R. (1956). *An Introduction to Cybernetics*. London: Chapman & Hall. (第十一章“必要多样性”)

20. Cannon, W. B. (1932). *The Wisdom of the Body*. New York: W. W. Norton.

22. Han, B.-C. (2010). *Müdigkeitsgesellschaft*. Berlin: Matthes & Seitz.

字数与版本。本稿 23,900 汉字，三十二节，真表十六张。版本二（第二轮修订），二〇二六年八月九日。第一版 19,100 汉字。

人机分工声明。本章的问题、三个领域的选定与全部判断由作者提出并负责。人工智能参与了文献检索的辅助、公开数据的下载与复算、脚本编写、结构检查与语言编辑。第十一节与第十二节的全部数字由脚本计算，任何人可按上述哈希与文件名复核。本章的所有错误由作者承担。

二十四·第一轮修订说明（相对初始设计的六处）

1. **主张整体收窄。**初始设计的承重命题是”多用几条恢复通道，夜里就沉得更深”。第十一节的探索性拆分推翻了这一环，第十四节的舞蹈学限制指向同一处。本稿把第一层主张降为待检验假说，并在第二十二节写明它至今没有直接证据。
2. **新增第八节与第十二节（两端归属）。**这两节在初始设计里不存在。它们是探索性分析逼出来的，也是本稿唯一有把握的一层。
3. **新增第十五节第一条（差值分数问题）并公开承认初稿漏了它，**同时在第二十节把这次遗漏读作本章命题的一个实例。
4. **第十八节第一条改为”已跑过且判本章错”，**保留在证伪表上而不是删掉。
5. **新增第九节的剔除与数据缺陷一段，**公开三人剔除的理由、年龄字段记为 0 的那一位、以及说明书未定义的编码 0。
6. **新增第十七节第六条：**本章不主张兼任者的减少是坏事。初始设计里这一句没有写，容易被读成一种怀旧。

二十五·第二轮修订说明（相对第一版的七处）

第一版交出去之后又做了一轮，主要动作是用一个新的补跑拆自己。七处改动如下。

1. **新增第十二之二节：**证明 66.7% 与 0.2% 由三个矩完全决定，是算术不是发现；给出 Oldham 一九六二年的零情形与按实测方差比的推导。**本章第一版最得意的一节，在这一节里塌了一半。**
2. **新增第十二之三节：**配对分位下方差比是 1.25 至 2.67 而非 3.00，按变异系数两端几乎相同。主张由”白天那一端天生更有信息”收窄为”在这个具体构造下白天那一端绝对离散度更大”。
3. **★新增第十二之四节（本轮唯一的新东西）：**换成心搏间期变异，端权对由 (0.667, 0.002) 翻成 (0.045, 0.651)。据此提出**端权对**这个读数，并把主张改成一句更一般的话——**哪一端在说话由构造决定，不由人或夜决定。**
4. **新增第十一之二节：**十二个白天变量的完整表与 Bonferroni 校正，没有一个通过。据此更正标题的强度。
5. **第十五节重排，**前两位换成本章初稿完全没提的两支：**数学耦合与初值定律**（Oldham 一九六二／Archie 一九八一／Wilder 一九三一）与**血压夜间下降率**（四十年文献，且是本章第四个正面对照——动态血压报告单把两端分印）。原第一位的差值分数问题降为第三位。
6. **第十五节末新增一次未做成的外部核对，**附检索走向与失败原因：文献报的是各时段内部的血压波动（个体内），不是各时段平均值之间的离散（个体间），端权对因此在已发表的表上算不出来。
7. **第二十节自反新增三笔：**漏掉的不止一条祖宗；标题比数据强；**本章最得意的那个数是一九六二年的算术，**而发现这件事只需要把自己的数字放回算式里推一遍。

第八章 下去的那一段，没有人记

下去的那一段，没有人记

论“快速入睡”与“睡得深”能否同时达成至今未被判定，以及判定它所缺的两样东西都不是技术难题

摘要

“怎样才能快速入睡？怎样才能睡得深？”三门学问给出三种回答：生理学说由睡眠压力与昼夜相位的当前值决定；物理学说不够——同一份熔体慢冷得晶体、快冷得玻璃，差别是残余应力，它在静态测量上为零、只在下一次受扰时兑现，故终态由**路径速率**决定；哲学说有一类状态只能作为别的行动的副产品出现。三家共同踩着一句没说出口的话：**下降过程是一条中性通道，怎么下去不影响下去之后是什么**。推翻它的材料来自物理学自己。

本章把这条否认搬到睡眠上，用四十夜公开判读记录构造最粗代理**下降时长 T**，做了预注册检验。**四条预测**：一条命中但不稳，一条未命中，一条表面命中而被本章自己的稳健性检验拆掉，一条**指向竞争解释**。本稿补了原设计缺的自助区间：**六项里五项跨零**，唯一不跨零的正是随后被拆掉的那一条；放大显示**约需八十夜区间才离开零**。在四十夜这个样本上，本章的检验原理上不可能成功。

因此本章**不主张存在一种被快速冻住的睡眠**，只主张一件更小也更难反驳的事：这个问题至今没有被判定过，而判定它所缺的两样都不是技术难题——**一把细于三十秒的下降刻度**（Hori 九级一九九四年即有）与**一次标准化扰动**。这类量本章命名为**未判速率**（可算、无栏、未裁），其所在那一格命名为**算熟未裁**。文末给出十九条判决性分界（第一条是初稿漏掉的研究缺口那一支）、四处反向约束与四条使“未裁”为假的条件。

关键词：下降时长 | 未判速率 | 算熟未裁 | 成就动词 | 退火与淬火 | 缺件清单

一·一个每天被回答、而从来没有被判定过的问题

搜索“怎样快速入睡”，回答以百万计。搜索“怎样睡得深”，回答同样以百万计。两组回答常常出自同一篇文章、同一个人、同一段视频，中间用一个顿号连起来，仿佛这是一件事的两半。

本章要问的不是哪一条建议管用。本章要问的是一件更靠前的事：**这两个目标能不能同时达成，有没有被判定过？**

答案是没有。而没有的原因不是没人关心——关心的人太多了；也不是缺少数据——公开的整夜记录数以万计；是因为判定它所需要的那个自变量，**没有出现在任何一份记录里**。

一份标准睡眠报告会告诉你：躺下多久睡着（入睡潜伏期，一个时刻差）、各期各睡了多少分钟（若干存量）、醒了几次、效率多少。它不会告诉你**你是怎么下去的**——是一路滑下去，还是先掉到很深又浮上来再掉下去，还是磨了很久才开始掉。这一整段过程的形状，在报告上没有一栏。

本章一开始想做的，是证明这一栏的缺失代价高昂。本章做了检验，**检验失败了**。第六、七两节完整报告这次失败，包括本章如何用自己前一篇文章的方法拆掉了自己剩下的那一条读数。第八节写失败之后剩下什么——剩下的比原来想主张的小得多，但它站得住。

先从最锐的那一刀开始，因为本章全部的力气都从它来。

二·物理学那句“不许”

有一句话，任何一个做过材料的人都会当作常识：

同一份熔体，慢慢冷下来得到晶体，快速冷下来得到玻璃。

两者都是固体，都不流动，都有确定的形状。密度可以差得很少，硬度可以差得很少，用宏观手段量，你甚至可能分不出哪块是哪块。**差别在结构，以及结构里冻着的残余应力**。而残余应力有一个要命的性质：**它在静态测量上接近于零**。你把玻璃放在那里，它不动，什么也不发生；直到你敲它一下、或者温度变一下，它就沿着那些冻住的应力裂开——而晶体不会。

也就是说，物理学自己承认存在这样一类差别：它由路径速率造成，在任何终态读数上都读不出来，必须等下一次扰动才兑现。

这件事在流变学里被压成了一个无量纲数。物质有它自己的弛豫时标 τ ——受扰之后回到新的平衡要多久；过程有它的时标 T ——你多快去改变条件。两者的比 $De = \tau/T$ 被称作 Deborah 数。 De 很大，材料表现得像固体； De 很小，同一份材料表现得像流体。同样东西是什么，取决于你多快去动它。

把这句话对准睡眠，得到的追问是干净的：入睡这件事，有没有它自己的弛豫时标？如果有，下降得比它快，会不会留下什么？

这不是一个比喻。它可以被写成一条会输的预测，第五节会写。而先要说明，为什么另外两家不会问出这个问题。

三·另外两家：一家给了处方，一家不许开处方

3.1 生理学：由两个状态量的当前值决定

睡眠调节最成功的框架，是把入睡的时机与深度交给两个过程：一个随清醒时间累积、随睡眠消散的稳态过程，与一个不管你睡不睡都在走的昼夜过程。三十多年后作者们自己做了一次重估，这个双过程结构仍然是领域的骨架（Borbély 1982; Borbély, Daan, Wirz-Justice & Deboer, *Journal of Sleep Research* 25: 131–143, 2016）。

这一家最硬的一条落地读数不在脑电上，在皮肤温度上：远端（手脚）与近端（躯干）皮温之差，是入睡潜伏期最强的生理预测量之一，而暖脚促进快速入睡（Kräuchi, Cajochen, Werth & Wirz-Justice, *Nature* 401: 36–37, 1999）。

这一家的单因主张是：入睡的快慢与深浅，由若干可测状态量的当前值决定。把睡眠压力抬高（少睡、别打盹、缩短卧床）、把相位对齐（光照、固定起床时间）、把远端血管扩开（暖脚、温水浴），就会又快又深。

它自曝的地方也很清楚：双过程模型是一个关于水平的模型。它预测某一时刻入睡的倾向有多大，它不预测你以什么速率下去，也不主张速率本身会留下什么。速率在这套框架里不是一个变量，是一个结果。

3.2 哲学：有一类状态不能被当作目标

第三家不回答“怎样才能”，它质疑这个问句。

有一批状态似乎具有这样的性质：它们只能作为为别的目的而采取的行动的副产品而出现；它们永远不能被有智慧地或有意图地带来，因为试图这么做这件事本身就排除了那个人正试图带来的状态。（Elster, *Sour Grapes*, Cambridge University Press, 1983, ch. II）Elster 把这类状态叫作“本质上的副产品”，并把“试图带来它们”称作副产品的谬误。

这一家的可贵之处在于，它自己先把最难的反驳摆出来了：即便承认这些状态无法仅凭意志的命令而被带来，难道就不可能存在某些技术，让我们经由间接手段把它们带来吗？这一问正是失眠行为治疗存在的理由——刺激控制、睡眠限制、矛盾意向，它们没有一条是在教人“睡”，它们都在改别的东西。

这一家的单因主张是：能不能把一件事当作目标去追求，这件事本身是有条件的；而对于落在这一类里的状态，追求它的动作与达成它的条件互斥。

它自曝的地方：Elster 没有给出一条判据，用来判断某个具体状态到底落不落在这一类里。“间接技术”这条出口一旦打开，边界就模糊了——刺激控制算间接吗？暖脚算间接吗？定一个固定起床时间算间接吗？这条线他没有划。

而这条线在更早的地方有一个现成的划法，Elster 自己没有用。Ryle 区分过成就动词与任务动词：“找”是任务动词，“找到”是成就动词；任务动词描述一段可以被执行、被中断、被做得更卖力的活动，成就动词只标记一个结果的到达，它没有对应的执行过程（Ryle, *The Concept of Mind*, Hutchinson, 1949）。判据因此可以写得很机械：一个状态落在“本质上的副产品”这一类里，当且仅当指向它的动词是成就动词而没有配套的任务动词。

拿它去量：“睡着”是成就动词，而汉语与英语里都没有一个描述“正在执行睡着这件事”的任务动词。你可以躺着、可以放松、可以数羊——这些都不是“睡”这个动作的执行，它们是别的动作。**这正是本章全部论证的语言学地基**：如果连描述那段过程的动词都不存在，那么那段过程在记录上没有一栏，就不是疏忽，是同一件事在两处显影。

这条判据同时约束本章：它说明“慢慢地睡下去”这个指令在语法上就不成立——它要给一个成就动词配一个执行副词。第十节第三小节那一削，根子在这里。

3.3 三家两两之间是真打架

生理学与物理学打的是状态函数与路径函数。生理学说初末条件定了，终态就定了；物理学说同样的初末条件、不同的速率、不同的终态——而这两个词在热力学里正是一对反义词。

物理学与哲学打的是可工程化。物理学说只要控制冷却速率就能拿到想要的终态，这是冶金车间每天在做的事；哲学说你一旦把那个速率当成要达成的目标，自败的动作就已经开始了。冷却曲线可以被设定，入睡曲线不能——如果哲学那条对的话。

哲学与生理学打的是处方的合法性。哲学说睡眠不能被意愿；而睡眠限制疗法是失眠的一线治疗，确实有效。Elster 自己把这个反驳写在书里，却没有回答它。

还要干掉一个总想挤进来的第四家。临床指南、睡眠卫生手册、手环上那串建议，都不构成一种立场——它们是上面三种主张各自落地的场所。本章后面反复引用判读手册条文与公开记录，那些是兑现材料，不是第四个说法。

四·三家共同踩着的那块地

把三家的争论压成一句：三家争的是**入睡该由什么来加速**——由当前值、由路径速率、还是由别的事。

而填在横线上的那样东西——“下降过程”——三家都默认它是这样一种东西：

下降过程本身是一条中性通道；怎么下去，不影响下去之后是什么。

念给生理学听，它会说这还用说吗：它测的是入睡潜伏期与各期时长，全是端点与存量，中间那一段只是从一个端点走到另一个端点的路。念给哲学听，它也会说这还用说吗：它关心的是你能不能到达那个状态，不关心你以什么速率到达。

念给物理学听，它会说：**这句话在我这里是错的。**

这就是本章的入口。推翻三家共同踩着的那句话的材料，必须是三家之一自己早已掌握、早已印出来、只是从没往这个方向使的事实。玻璃与晶体的那一条正是：它是教科书内容，不是本章的推想；两侧的极端也是物理学自己给的。

推翻之后，问题就换了：不再是“怎样下去更快”，而是——

下去的那一段，是什么样子的？以及，有没有人在记它？

答案是：没有人在记。这就是第五节要构造的东西。

五·我打算怎么判它

5.1 一次判定需要哪些件

本章不给分类格，给一张**缺件清单**：要判定“快速下降与缓慢下降是否导向不同的终态”，需要下面五件，缺一件这个问题就悬着。

① **一个下降速率的定义**——它必须能从记录里算出来，而不是靠自报。② **一把刻度**——它的分辨率必须细于下降过程本身的时间尺度。③ **一个终态读数**——它必须能分辨结构而不只是总量。④ **一次标准化的扰动**——因为待检验的差别按假设在静态读数上为零。⑤ **一个能排除睡眠压力的设计**——否则“下降快”与“睡眠压力高”不可分。

对现有的公开整夜记录逐项核对：① 可以构造（见 5.2）；② **没有**——最细粒度是三十秒判读窗；③ 部分有（段数、段长）；④ **没有**——公开数据里没有受控扰动；⑤ **没有**——观察性数据，两者共变。

五件里缺三件。本章不假装这三件不重要。本章要做的，是用剩下的两件半，把这个问题往前推到它能被推到的地方，然后**如实报告它停在哪里**。

5.2 下降时长 T

在剪裁后的判读序列里：

- **入睡点 O** = 第一个非清醒的判读窗的位置；
- **首达深睡 D** = 第一个 N3 判读窗的位置；
- **下降时长 T = D - O**，单位是三十秒的判读窗。

T 是本章的自变量。它粗，它只用了两个端点，它完全不描述中间那一段的形状——这些都是它的缺点，且都由第②件的缺失造成。但它有一个好处：**它可以从任何一份现成的判读记录里算出来，一行代码，任何人可复算。**

5.3 一句不含程度词的判据

本章给自己定的判据是：

同一份记录里，“怎么下去的”这件事，有没有一栏是它自己的？如果没有，判定它需要的哪几件是缺的？

这句话拿去问报告模板与字段表，不用问任何人的判断。在五个场合各跑一遍：

场合	答案	来源
临床睡眠报告	没有。只有入睡潜伏期（一个时刻差）与各期时长（存量）。	查出来的（报告模板通行字段）
Hori 九级脑电分期	有。入睡这一段被切成九级，从 α 波列一路到纺锤波。	查出来的（Hori et al. 1994）
冶金热处理规程	有。冷却速率是工艺参数本身，写在规程第一行。	查出来的（退火/淬火工艺）
消费级睡眠手环	没有。给出“入睡时间”与“深睡时长”，不给下降曲线。	从命题可推
麻醉诱导记录	有一半。诱导速度被记录，但记的是给药速率，不是被诱导者的下降形状。	从命题可推

第二、三条反过来改了本章的定位：把线打开成一段有形状的过程，这个动作在别处已经被完成了——脑电研究里做了三十年，冶金车间里做了几百年。本章因此不主张一个新原理，只指出一个已被别处完成的动作在这里缺席。这一条第九节还要用。

5.4 预注册

以下文本在跑任何统计之前封存，其后一字未动。数据为 Sleep-EDF Expanded 1.0.0 的 sleep-cassette 部分，按文件名排序的前四十份判读标注，与本书前四章同一批，未作挑选。

- **剪裁：**首尾各保留最多三十分钟离床清醒。
- **快降组 / 慢降组：**按 T 的中位数二分。
- **H1（本章承重的“读不出来”）：**两组在睡眠效率、深睡百分比、快速眼动百分比、觉醒次数四项上差异很小。判定用效应量而非 p 值：Mann-Whitney 的 rank-biserial r 0.30 记为“读不出来”；四项中至少三项达标则 H1 命中。
- **H2（分歧）：**首个深睡段结束后重新回到深睡所需的窗数，快降组更长（r 0.20）。
- **H3：**整夜深睡段数，快降组更多（r 0.20）。
- **H4（最强竞争解释的正面检验）：**若 T 与深睡百分比的等级相关 ρ 0.50，则下降速率只是睡眠压力的代理，本构念退回为双过程模型的一个侧面，本章须据此改写，不得保留原表述。
- **撤销规则：**若 H1 落空（四项中两项及以上 $r \geq 0.30$ ），则“差别在当晚读数上读不出来”直接判错，本章不得改口说“至少某几项读不出来”。若 H2 方向相反，则残余应力的类比在本数据上不成立。

— 写在跑之前的四条局限：本数据无标准化扰动，H2用的是自发跌出，只能算相关；剪裁把首段清醒截到三十分钟，故常规入睡潜伏期被削顶，不用作分组变量；三十秒是最细粒度，T是粗刻度；每人一夜，个体差异与夜间差异不可分。

六·四条预测的下场

四十夜，四三三二九个判读窗，三六一小时。四十夜全部可用，无一夜整夜无深睡。

T的分布很偏：最小17窗（8.5分钟），四分位42 / 54 / 97窗，最大**752窗（6.3小时）**。中位数53.5窗，合26.8分钟。快降组 n=20，T中位39窗（19.5分钟）；慢降组 n=20，T中位93窗（46.5分钟）。

6.1 H1 命中

读数	快降组中位	慢降组中位	rank-biserial r	判定
睡眠效率	0.847	0.818	+0.285	读不出来
深睡百分比	0.173	0.156	+0.265	读不出来
快速眼动百分比	0.221	0.230	-0.140	读不出来
觉醒次数	11.5	12.5	-0.010	读不出来

四项全部低于0.30，H1命中。但**前两项贴着线（0.285、0.265）**，这一点第六节第五小节与第七节会先后要了它的命。

⚠️ 顺带记一件本章自己身上发生的事。本系列第三篇讲的是：一次险胜与一次压胜，在只保留名次的记录上完全相同。而**H1的结算正是如此**——0.285与0.010在预注册表上是同一个记号“读不出来”，判余0.015的那一项与判余0.29的那一项不可分辨。本章用四条预测的“命中／未命中”做结算，就是在给自己造一份独判格式的记录。这一处不是文字游戏：如果不是第七节另外去拆，光看结算表，本章会得出一个比实情强得多的结论。

6.2 H2 未命中

首个深睡段结束后重新回到深睡：快降组中位4.0窗，慢降组3.0窗，方向与预测一致（快降组更慢），但**r=+0.090，远低于预注册的0.20**。按预注册规则，H2判未命中，原表述不保留。无删失（四十夜全都回到过深睡）。

6.3 H3 命中，而且看上去是全章最有力的一条

整夜深睡段数：快降组中位**23.5**，慢降组**14.5**，r=+0.360。

这个数当时看着很好：两组的深睡百分比几乎一样（0.173对0.156），而快降组的深睡被切成的段数多出六成。同样多的深睡，一组是整块的，一组是碎的——正是玻璃与晶体的形状。

第七节会说明为什么这个读法是错的。

6.4 H4 未触发撤销线，但指向了竞争解释

T与深睡百分比的等级相关 **$\rho = -0.361$** ，未超过预注册的0.50撤销线。附带的三条：T与睡眠效率 $\rho = -0.258$ ，T与记录总长 $\rho = +0.209$ ，T与觉醒次数 $\rho = +0.035$ 。

-0.361不算大，但它的方向完全符合生理学的标准解释：**睡眠压力越高，下降越快，深睡越多**。撤销线没被触发，不等于这条解释被排除了。

6.5 四十夜根本裁不了：区间

预注册用效应量而非p值作判定，这是为了避免小样本上的p值游戏。但它留下一个空档：**效应量本身也有抽样误差，而原设计没有报告它**。本节补上——用五千次自助重抽给每一项算95%区间（随机种子写死）。

读数	快降中位	慢降中位	r	95% 自助区间	
睡眠效率	0.847	0.820	+0.285	[-0.075, +0.625]	跨零
深睡百分比	0.173	0.156	+0.265	[-0.090, +0.615]	跨零
快速眼动百分比	0.221	0.230	-0.140	[-0.485, +0.235]	跨零
觉醒次数	12.0	12.5	-0.018	[-0.375, +0.350]	跨零
深睡段数	23.5	14.5	+0.365	[+0.005, +0.673]	不跨零
碎裂密度	0.1531	0.1195	+0.085	[-0.300, +0.450]	跨零

六项里五项的区间跨零。唯一不跨零的那一项，正是第七节第二小节要拆掉的那一条——而且它的下界只有 +0.005，贴着零。

这一节的后果比它看上去大。它说明的不是“本章的检验失败了”，而是更硬的一句：在四十夜这个样本上，本章的检验原理上不可能成功。无论结果朝哪个方向，都读不出可靠的东西。

那么要多少夜？以实测分布为总体做自助放大（假定实测效应为真，这是一次外推而非功效计算）：

每组夜数	深睡百分比的区间	睡眠效率	深睡段数
20（本章）	[-0.10, +0.60]	[-0.07, +0.62]	[+0.03, +0.68]
40	[+0.01, +0.51]	[+0.04, +0.53]	[+0.14, +0.60]
80	[+0.09, +0.43]	[+0.11, +0.46]	[+0.20, +0.52]
200	[+0.16, +0.37]	[+0.17, +0.39]	[+0.26, +0.47]

约八十夜，深睡百分比那一项的区间才第一次离开零。

这个数把第八节那句“未裁”从一个否定性断言变成了一个有标价的断言：判定这件事不是“没人想做”，是它需要八十夜以上的样本、一把细于三十秒的刻度、一次标准化扰动——而三样里只有第一样是钱的问题。

七·我怎么拆掉了自己剩下的那一条

一条预测命中之后，最省事的做法是把它写进摘要。本章的做法相反：用本书前一章文章教出来的方法，先去拆它。

前一篇《那三十秒算不算一段》（第一章）报告过一个数：把两个相邻分期并类之后，四十夜的段数从 4477 掉到 2907——三成五的段落是判读规则制造的，不是生理事件制造的。那一篇的结论直接对着本章的 H3：段数这个量本身就不可靠，因为它对“怎样算断开”这条规则极其敏感。

于是本章对 H3 做了四道检验。它没有全部通过。

7.1 段数的断开阈值：这一关过了

把“间隔多少窗以内不算断开”从零放宽到四：

断开阈值	快降组中位	慢降组中位	r
严格（任何一窗中断即断开）	23.5	14.5	+0.360
容许 1 窗中断	14.5	9.0	+0.360
容许 2 窗中断	12.5	7.5	+0.377
容许 4 窗中断	10.0	5.5	+0.495

四种口径下 r 全部大于 0.20，方向一致，甚至更强。H3 不是判读规则制造的。

7.2 把深睡总量除掉：这一关没过

但还有一个更朴素的解释没被排除：快降组的段数多，会不会只是因为它的深睡本来就多？

于是把总量除掉，算碎裂密度 = 深睡段数 ÷ 深睡窗数：

断开阈值	快降组中位	慢降组中位	r	T 与碎裂密度的 ρ
0 窗	0.1531 段/窗	0.1195	+0.090	+0.022
2 窗	0.0776 段/窗	0.0580	+0.130	-0.037

差异消失。两种口径下 r 都低于 0.20 的阈值，而 T 与碎裂密度的等级相关基本为零。

也就是说：H3 那个“多出六成的段数”，几乎全部由“快降组的深睡总量多”解释掉了。快降组的深睡不是被切得更碎，只是更多，所以段数更多。玻璃与晶体那个形状，在这份数据上没有出现。

7.3 去掉四个极端夜：H1 也塌了一半

T 的分布极偏，最长的一夜从入睡到首次深睡用了 6.3 小时。去掉最长的四夜重做：

读数	全样本 r	去掉四夜后 r	
睡眠效率	+0.285	+0.105	读不出来

深睡百分比	+0.265	+0.481	★读得出来
快速眼动百分比	-0.140	+0.025	读不出来
觉醒次数	-0.010	-0.204	读不出来
深睡段数	+0.360	+0.472	
碎裂密度	+0.090	+0.012	消失得更干净

H1 的第二项翻掉了。在去掉四个极端夜的样本上，深睡百分比读得出来，而且效应不小。也就是说：H1 的“读不出来”依赖于那四个极端夜的存在。

按预注册的撤销规则，H1 是在全样本上判定的、且四项全部达标，所以形式上仍算命。但本章不打算靠这个形式。一条结论如果去掉四个观测就翻掉，它不配被写进摘要当作发现。

7.4 三分位上那个诱人的形状，是噪声

把四十夜按 T 分成三份，出现了一个很吸引人的东西：

组	T 中位	深睡百分比	碎裂密度中位
最快 1/3	33 窗	0.260	0.1409
中间 1/3	53 窗	0.161	0.1017
最慢 1/3	111 窗	0.114	0.1805

深睡百分比单调下降——这是睡眠压力那条解释，方向一致。而碎裂密度是一个 **U 型**：太快和太慢，深睡都更碎；中间那一段最不碎。**如果这个 U 型是真的，它就是本章要的那个东西——单调的睡眠压力解释预测不出非单调，而“存在一个最优下降速率”正是物理学那一边会预测的形状。**

本章把 T 打乱做了五千次置换检验：中间三分之一对两端合并，实测 $r=+0.080$ ，置换检验 $p=0.3457$ 。

是噪声。再看一眼中间那组的均值（0.2252）与中位数（0.1017）相差一倍多，说明那个低中位数是被少数几夜拉出来的。作为对照，改按深睡百分比分三份，碎裂密度单调下降（0.2500 / 0.1515 / 0.1004），说明碎裂密度主要由深睡总量决定，与 T 无关。

7.5 顺带一条：T 本身是稳的

有一种为失败开脱的说法是“刻度太粗，所以测不出来”。本章自己检验了这条借口。把判读窗从三十秒粗化到六十、九十、一百二十秒，重算 T 并重新二分：

窗宽	快降组与三十秒时重合	Jaccard	换组夜数	T 的等级相关
60 秒	17/20	0.77	5	+0.987
90 秒	17/20	0.74	6	+0.867
120 秒	15/20	0.71	6	+0.902

T 在这个范围内相当稳定。所以不能说“因为刻度粗所以什么都测不到”——在三十秒到两分钟这个区间里，T 是一个稳定的量。它只是**不携带超出睡眠压力的信息。**

7.6 这次检验的总账

预测	下场
H1 两组在常规读数上读不出来	形式上命中，但去掉四个极端夜后深睡百分比读得出来（ $r=+0.481$ ）—— 不稳，本章不据此立论
H2 扰动后恢复更慢	未命中 （ $r=+0.090$ 0.20），原表述撤回
H3 深睡段数更多	表面命中（ $r=+0.360$ ）， 被 7.2 拆掉 ：除掉深睡总量后差异消失
H4 是否被睡眠压力吸收	未触发撤销线（ $\rho=-0.361$ ），但方向与睡眠压力解释完全一致
三分位 U 型（非预注册）	置换检验判为噪声（ $p=0.35$ ）

结论：在判读级三十秒粒度、无受控扰动的观察数据上，下降速率不携带任何超出睡眠压力的信息。

本章原本想主张的那件事——存在一种被快速冻住的睡眠，它在当晚读数上看不出来、只在下一次扰动时兑现——在这份数据上没有得到任何支持。本章因此不主张它。

八·失败之后剩下什么

一次检验失败，可以有两种读法：

(a) 那件事不存在；(b) 那件事存在，但这份记录读不出它。

而本章无法区分这两者。要区分，需要第五节缺件清单里缺的那三件：细于三十秒的下降刻度、一次标准化的扰动、一个能排除睡眠压力的设计。本章一件也没有。

这个“无法区分”本身，就是本章剩下的那一格，而且它比原命题窄得多、也硬得多。

承重命题（改写后）：“快速入睡”与“睡得深”能否同时达成，至今没有被判定过。不是因为无人关心，也不是因为缺少数据，是因为判定它所需的自变量从未被记录、所需的扰动从未被施加。本章构造的下降时长 T 是一个稳定、可从任何现成记录算出、却既不在任何报告栏里、也从未被任何研究设计单独裁决过的量。这一类量，本章称作未判速率。

一个未判速率有三个特征，缺一不可（这一格的三个签名与另外两个实例见 8.1、8.2，它怎么会错见 8.3）：

1. **可算**——从现成记录一行代码算得出（ T 满足；7.5 证明它还是稳的）；

2. **无栏**——不出现在任何常规报告字段里（ T 满足）；

3. **未裁**——从未有一个研究设计以它为自变量、并配备足以裁决它的刻度与扰动（ T 满足，这正是本章的失败所证明的）。

第三条是最要紧的。一个量可以被算了很多次而从未被裁决过——本章自己就算了一次，然后如实报告它裁不了。“算过”不等于“判过”。

而这三条里，第二、三条都不是技术难题：

— **刻度早已存在。**Ogilvie 综述了三十年的入睡研究后写道：所回顾的研究并不支持“睡眠在某一瞬间开始”这一想法，而是表明进入睡眠是一系列连续、交织的变化，始于放松的困倦，穿过第一期，常常延续到第二期的最初几分钟；从清醒的脑到睡眠的脑的转变，被 Hori 的九级脑电系统准确地追踪（*Sleep Medicine Reviews* 5: 247–270, 2001）。Hori 的九级把入睡这一段切成： α 波列、 α 间断（ $\geq 50\%$ ）、 α 间断（50%）、脑电平坦、涟波、孤立顶尖波、顶尖波簇、顶尖波与不完整纺锤波、纺锤波（Hori, Hayashi & Morikawa, 1994）。一九九四年就有了九级刻度，而临床报告里这一整段至今是一个数字。

— **扰动只是一次实验设计。**给两组下降速率不同的人施加同一个标准化扰动，看恢复曲线，这不需要任何新技术。

这就是本章全部的主张：不是“快与深互斥”（数据不支持），不是“存在淬眠”（数据不支持），而是——有一个可算、无栏、未裁的量，横在一个每天被回答一百万次的问题中间；而补齐它需要的两样东西，一样已经存在了三十年，另一样是一次实验设计。

8.1 这一格为什么长期没人发现：三个签名

一个未判速率所在的那一格，有三个可以逐条核对的签名。三条齐了，它才是本章说的那一格，而不是普通的“研究还不够多”。

1. **短期表现为改善。**这个量被算的次数越多，这个领域看起来越像已经研究透了。入睡过程被研究了三十年，文献以千计——从外面看，这是一个成熟领域。

2. **失效不产生任何信号。**从来没有人裁决过它，这件事不会在任何地方报错：没有一份综述会写“本领域至今未裁决 X ”，因为综述的框架是“已有证据说什么”，不是“哪个变量从未做过自变量”。

3. **越正规化越自我加固。**一个量被算得越多、被引用得越多，后来者越不会去问“有人以它为自变量裁决过吗”，而是直接引用前人的算法。算的次数本身，被读成了裁决的证据。

本章把满足这三条的那一格命名为**算熟未裁**：算得很熟，从未被裁。它与“研究空白”的分别见第九节第一小节；与“阴性结果”的分别在于，阴性结果是裁决过的结果，算熟未裁根本没进过裁决。

8.2 这个类不是为睡眠定制的：另外两个实例

一个只有一个成员类，是一个改了名的个例。本章举两个别处的成员，读者可以自己核对那三条签名。

— **药物滴定速率**。同一个终剂量，快滴定与慢滴定的耐受性与停药率不同，这在若干药物上是明文写进说明书的。这是本章的第三个正面对照：那些药物把速率写进了规程，因而不属于算熟未裁；而在另一些药物上，滴定速率只留在处方者的经验里，病历上只记终剂量与起止日期——那一部分就落在这一格。同一个领域内部两种情形并存，这比跨领域对照更能说明问题不在技术。

— **康复训练的负荷上升曲线**。病历记起始功能与出院功能（两个端点），不记中间那条上升曲线的形状。曲线可以从每次治疗记录里算出来（可算），不在任何评估量表里（无栏），也几乎没有以它为自变量、配备随机分配的试验（未裁）。

两个实例都不在睡眠里，且都不依赖本章的任何一个数。这说明未判速率是一个可以被别人拿去用的类，而不是本章为自己的失败造的一个词。

8.3 “未裁”这一条怎么会错

本章的承重命题里，“未裁”是唯一一条否定性断言，因此必须写清它怎么被推翻：

1. 若有人出示一项以下降速率为自变量、且配备了细于三十秒的刻度与标准化扰动的研究，“未裁”立刻为假，本章的承重命题作废。这一条不需要争论，只需要一条文献。

2. 若在任何一份睡眠研究优先级清单里出现等价条目（见 9.1），则“未被识别”这一层不成立。

3. 若某份通行的临床报告模板里已有等价于 Hori 九级的下降分级字段，“无栏”为假。

4. 若在八十夜以上的样本上，T 与终态结构的关系仍全部落在零附近，则这一格从“未裁”变成“已裁而为负”——本章这一格随之消失，而这是本章最希望发生的结局。

第四条要特别说明：本章并不主张下降速率一定重要。本章主张的是这件事至今没有被判定过。判定的结果若是“不重要”，本章的目的同样达到——一个空档被填上，与它被填上什么内容无关。

九·与最近的十九种说法的分界

逐条列三样：它走到哪一步为止、线划在什么地方、以及一个能把“本章对”与“它对”分开的具体判法。哪一条只写得出程度上的比较，就算这一条被它覆盖，不留。

排序按危险程度。第一条排在最前，因为初稿把它整支漏了——而它几乎是本章落点的同形物。

9.1 可避免的研究浪费与研究缺口——初稿漏掉的那一支

说到哪一步：有一整支元研究在做本章第八节做的那件事。Chalmers 与 Glasziou 论证，医学研究中大量投入被浪费掉，而问题的选择本身就是浪费的第一个来源（*The Lancet* 374: 86–89, 2009）。Cochrane 系统综述的标准结尾是一节“对研究的启示”，专门写“这个问题至今没有被足够的证据回答，回答它需要什么样的试验”。James Lind Alliance 的优先级设定伙伴关系更直接：把患者、照护者与临床医生凑在一起，专门去找那些重要而至今无人回答的问题，产出一份清单。

本章初稿一句也没提。而第八节那句“判定它所缺的两样东西都不是技术难题”，几乎是这一支的标准句式。这是初稿最大的一处缺席，写在这里而不是删掉。

分离线（两条，而且它们反过来给本章加力）：

1. 处方的时间顺序相反。研究缺口那一支的处方是“去做那个研究”——问题已经能被提出、变量已经能被测量，缺的是有人出钱出人。本章的处方是“先加一个字段，否则那个研究做不成”：下降速率这个自变量至今不在任何记录格式里，所以不是没人做，是做的人得先自己造一把刻度。一个是资源问题，一个是格式问题。

2. 它们各自看不见对方那一格。优先级设定问的是“哪些问题重要而未被回答”，它的候选池由已经被问出口的问题构成；一个连变量都没有字段的问题进不了候选池，因此永远不会出现在任何一份优先清单上。本章所指的那一类，恰恰是从优先级设定的筛子里漏下去的。

判定形式：若在任何一份已发表的睡眠研究优先级清单里出现等价于“入睡下降速率是否影响终态结构”的条目，则本章所指的缺席不成立，本章这一格并入研究缺口那一支。就本章所查而言，没有。

检索词（补上，因为“未裁”是一个否定性断言，不出示检索就不算检索）：sleep onset descent rate / sleep onset speed outcome / rate of falling asleep sleep quality / hypnagogic transition duration outcome / sleep onset trajectory predictor / 入睡速度 睡眠质量。未命中不等于未被占位，本章只能标〔已检索，未命中〕，不能据此宣称第一。

9.2 入睡过程研究与 Hori 九级分期——本章的正面对照

说到哪一步：见第八节。这一支从上世纪八十年代起就论证入睡不是一个瞬间而是一段过程，并造出了追踪它的细刻度。

分离线在被追问的对象不同：这一支追问的是入睡过程本身是什么样的（脑电如何演变、主观体验如何变化、反应时如何延长）；本章追问的是这段过程有没有进入任何一份用于判断“睡得好不好”的记录。前者是一个现象学与电生理问题，后者是一个记录格式问题。三十年的细刻度研究与临床报告里那一个数字，可以毫无矛盾地并存——这个并存本身正是本章的对象。

判定形式：若在任何一份通行的临床睡眠报告模板中，存在等价于 Hori 九级的下降分级字段（有独立名称、独立时长统计），本章所指认的缺席不成立。就本章所查的模板而言，没有。

9.3 双过程模型与睡眠压力——本章这次检验的实际赢家

说到哪一步：入睡与深睡由稳态压力与昼夜相位的当前值决定。

分离线：双过程模型是关于水平的，本章问的是速率。二者本可独立——直到本章自己的数据把它们连了起来：T 与深睡百分比 $\rho = -0.361$ ，三分位上深睡百分比随 T 单调下降。

本章必须承认的：在这份数据上，速率携带的一切信息都可以由压力解释。这一条是本章这次检验的实际赢家，本章输了。

判定形式：设计一次睡眠压力被固定的实验——同一批人、相同清醒时长，只操作下降速率（例如操作卧室温度或远端加温的时机）。若两组的终态结构仍无差异，则速率确实不承载信息；本章这一格连“未裁”都算不上，只剩“已裁而为负”。这是本章第一顺位的待办。

9.4 注意—意图—努力路径，以及努力悖论

说到哪一步：Espie 等人提出，从注意到睡眠、到形成明确的入睡意图、到在入睡过程中使劲，这条路径抑制了正常睡眠—觉醒的自动性（*Sleep Medicine Reviews* 10: 215–245, 2006）。这是失眠领域对“越用力越睡不着”最完整的学术表述，也是哲学那一家在临床上的对应物。

分离线：这一支处理的是睡不着——意图与努力如何阻止入睡发生。本章的全部对象是已经睡着的人：本章的四十夜里，下降时长最短的一夜只用了 8.5 分钟。本章问的不是能不能睡着，是睡着这件事以什么速率发生、以及那个速率有没有被记录。

判定形式：若把样本限制在入睡潜伏期正常（无入睡困难）的人群，本章所指认的记录缺口消失，则本章属于失眠研究的一个侧面；若缺口照旧（本数据即是如此——判读报告对好睡者与差睡者用同一套字段），则两者不同。

9.5 睡眠反应性与 FIRST 量表

说到哪一步：这一支用问卷评估个体在常见应激情境下的睡眠受扰程度，并在实验室首夜做多导睡眠图与随后的生理唤醒评估，以确定是否存在一种对睡眠受扰与高唤醒的特质易感性。（Drake, Richardson, Roehrs, Scofield & Roth, *Sleep* 27: 285–291, 2004）这是“扰动响应”这个想法最近的祖宗。

分离线：FIRST 测的是人的一个特质，靠自报，方向是应激→睡眠受扰；本章原想测的是一夜之内的路径量，从记录算出，方向是下降速率→下次扰动的恢复。两者的载体不同（人 vs 一夜）、来源不同（问卷 vs 记录）、方向相反。

判定形式：若个体的 T 在多夜之间高度稳定（组内相关系数 0.7），则 T 是一个特质而非夜间量，应并入睡眠反应性这一支；本数据每人只有一夜，无法判定，如实记为未判。

9.6 玻璃转变与残余应力（物理学，本章的工具来源）

说到哪一步：过冷液体与玻璃转变是凝聚态物理的成熟领域；冷却速率决定终态结构，残余应力在受扰时释放。

分离线：这一支处理的是分子构型的冻结，其弛豫时标可以直接测量；本章借用的只是它的形状——“路径速率决定终态、而差别在静态读数上为零”。本章并不主张睡眠中存在任何对应于分子构型的東西。

本章必须自曝的：借来的形状在本数据上没有兑现（第七节）。因此本章对这个类比的使用，现在只剩一个功能——它提供了那条被推翻的三家共同踩着的那句话，不再提供任何机制。

判定形式：若在配备了细刻度与标准化扰动的设计里，下降速率仍不影响扰动后的恢复，则这个类比整个失效，本章第二节应当删去。

9.7 矛盾意向

说到哪一步：让失眠者尽力保持清醒，反而缩短入睡潜伏期（Ascher & Efran, 1978）。这是哲学那条“副产品”在临床上最直接的兑现，也是 Elster“间接技术”那个出口的一个实例。

分离线：矛盾意向改变的是意图的对象，本章关心的是速率有没有被记录。一个用矛盾意向成功入睡的人，他的下降时长照样没有一栏。

判定形式：若矛盾意向组与常规组的 T 分布无差异，则意图操作不影响速率，两条正交；若有差异，则速率是意图的函数，本章须把哲学那一条从条件场提到机制位。本章未做。

9.8 冶金热处理规程——第二个正面对照

说到哪一步：退火、正火、淬火、回火，每一种的核心参数都是冷却速率，写在工艺规程第一行。

分离线：这一支的对象是可以被外部设定的冷却曲线；睡眠的下降曲线不能被外部设定——而按哲学那一家，它也不能被内部设定。本章因此不能把冶金规程当处方模板，只能把它当作“速率作为一个被记录的参数”的存在证明。

判定形式：若有任何一种干预被证明能在不改变睡眠压力的前提下系统性地改变 T，则速率是可设定的，冶金类比可以进一步使用；否则它只是一个记录格式的对照。

9.9 早期预警信号与临界减速

说到哪一步：系统接近临界转变时，受扰后回到平衡变慢，这个变慢可以在转变之前被测到（Scheffer et al., *Nature* 461: 53–59, 2009）。

分离线：临界减速处理的是转变之前的可预测性；本章处理的是转变之后留下了什么。两者时间方向相反。

判定形式：若入睡前的自相关升高能预测 T，则两者相连，本章的自变量可以被提前预测；本数据没有原始信号，未做。

9.10 局部睡眠

说到哪一步：睡眠不是全脑同步的开关，局部脑区可以先睡（Vyazovskiy et al., *Nature* 472: 443–447, 2011）。

分离线：局部睡眠说的是空间上的不同步；本章说的是时间上的速率。一个全脑高度同步入睡的人，他的下降速率照样可快可慢。

判定形式：若 T 与入睡时的空间不同步程度高度相关，两条应合并；本数据只有单导判读结果，未做。

9.11 自家五处（此前四篇未做够的动作，本章提前做）

（一）/column/insomnia/《失眠：为什么越想睡越睡不着》（同一作者）——最险的一处。那一篇已经占满努力悖论整条：睡眠是撤退不是动作、守夜人、油门与刹车、不是学会入睡而是停止阻止入睡。它回答的正是本章标题里那半句“怎样才能快速入睡”。

分离线，而且这条线反过来给本章加力：那一篇的全部对象是睡不着的人；本章的全部对象是睡得着的人——四十夜里下降最快的一夜只用了 8.5 分钟。那一篇说别使劲、放手；本章原想说放手之后下得太快也要付账——而本章这次没能证明这一点。因此本章与那一篇的关系现在是：本章尝试给它加一条限定，尝试失败了；那一篇的结论不受本章影响。

判定形式：若在入睡潜伏期正常者中，下降速率与任何终态结构有关，则那一篇需要一条限定；本章的检验说没有。

（二）本系列第一篇《那三十秒算不算一段》。那一篇讲段落账与时长账在同一份记录上分岔。本章原想给出一条使它们分岔的原因（下降速率），而第七节第二小节用那一篇的方法拆掉了自己——这是本系列内部第一次“后一篇用前一篇的方法否认自己”。分离线：那一篇的对象是账本单位，本章的对象是一个路径量；那一篇成立与否与本章无关。

（三）本系列第二篇《住在线上》（第二章）。那一篇讲睡与醒之间那片被记成线的区域。词面很近，必须分开：那一篇问的是边界那块地有没有厚度（空间性的，问的是“在哪里”）；本章问的是穿过它用了多久（时间性的，问的是“多快”）。一个厚度为零的边界照样可以被快穿或慢穿；一个厚度很大的边界照样可以没有速率字段。

（四）本系列第三、四篇（判余、托拍事件）：那两篇的缺口分别是一个名次差与一个接手人，都是当晚账上的存量；本章的缺口是一个跨时段的比值。四篇加本章都在说“账上少一栏”——这是本系列的共同全书赌的那句话，必须承认；本章与它们的分别，在于少的那一栏的类型。

（五）/frontier/sleep-medicine/ 睡眠医学面板（同一作者，两幕二十条）：那份面板讨论的是“哪个读数有资格代表睡得好”；本章问的是更靠前的一步——一个从未上过任何读数表的量。

9.12 这些邻居共同站着的那块地

逐个问“它结构性地看不见什么”：

— **双过程模型把水平当作给定的解释单位**，因此看不见速率；一旦承认速率独立承载信息，它就要多一个自变量——而本章这次没能证明它需要。

— **入睡过程研究把过程本身当作研究对象**，因此看不见“这个过程有没有进入判断好坏的记录”；一旦承认，它三十年的细刻度就要面对一个问题：为什么它没能进入报告。

— **临床报告模板把端点与存量当作给定的字段集**，因此看不见路径；一旦承认，它就要多一栏。

— **努力悖论把睡不着当作给定的问题域**，因此看不见睡得着的人身上也可能有账；一旦承认，它的适用范围就要写明。

背后是同样被当作给定的东西：**要判断一件事做得好不好，看它的结果就够了。**

⚠ 这句话本章不主张自己是第一个看见它的。它在别处有过很多主人。本章只报告一件具体的事：在睡眠这一处，看结果确实不够——因为结果读不出路径；而本章尝试证明“路径确实携带了结果读不出的东西”，没有成功。

十·四家反过来给本章加的约束

以下四处，每一处都写明：本章原想写哪句更强的话，哪一家不许，削掉之后剩什么。

10.1 生理学削掉的：速率不是一个独立的自变量

原想写：下降速率是一个独立于睡眠压力的路径变量。

生理学不许，而且是用本章自己的数据不许的：T 与深睡百分比 $\rho = -0.361$ ，三分位上深睡百分比随 T 单调下降。碎裂密度的差异在除掉深睡总量后完全消失。

削掉之后剩：T 是一个可算但在本数据上不独立的量。本章只能主张它未被记录、未被裁决，不能主张它承载独立信息。这一处动的是本章的核心主张，是本稿最大的一次退让。

10.2 物理学削掉的：类比不能替代机制

原想写：睡眠中存在对应于残余应力的东西。

物理学不许。残余应力之所以是一个物理量，是因为它有明确的定义、有测量方法、有释放时的能量核算。本章没有任何对应物——本章有的只是一个形状上的相似。而且物理学自己规定：要证明残

余应力存在，必须施加一次受控扰动并测量释放。本章的 H2 用的是自发跌出，不是受控扰动，因此即使 H2 命中也不足以证明。

削掉之后剩：物理学在本章里只提供那条被推翻的三家共同踩着的那句话（路径速率可以决定终态且差别在静态读数上为零），不提供任何机制。第二节因此是一节关于“问题该怎么问”的论证，不是一节关于睡眠的论证。

10.3 哲学削掉的：本章不能给出“控制你的下降速率”这条处方

原想写：既然速率要紧，那就应当学着慢慢地睡下去。

哲学不许，而且这一条比前两条更根本。按 Elster，入睡若落在“本质上的副产品”这一类里，那么“慢慢地睡下去”这个指令是自败的：执行它需要你下降过程中持续监控自己的下降速率，而这种监控正是那条注意—意图—努力路径要说的东西。

削掉之后剩：任何基于下降速率的建议，都必须作用在别的东西上——环境温度、光照时刻、卧床时间、远端加温的时机——而不能作用在“你要注意自己下得多快”这个指令上。这一处动的是价值排序，也把本章的处方形态整个定死了（见十一）。

10.4 睡眠医学自己削掉的：本章没有发现任何新现象

原想写：下降过程的形状是一个从未被人看见的对象。

睡眠医学自己不许。它看了三十年，造了九级刻度，写了综述，明确论证过入睡不是一个瞬间。本章之所以在初稿里能写下那句话，是因为本章一开始只查了临床报告模板，没查入睡研究的文献。

削掉之后剩：本章只能主张一句更窄的话——下降过程被研究了三十年，而它至今没有进入任何一份用于判断睡得好不好的记录。这句话不是关于知识的，是关于记录格式的。

十一·在一个不可裁决的空档里，还能说什么

用户的原问题是：怎样才能快速入睡？怎样才能睡得深？本章有责任回答它，而不是只回答“这个问题没被判定过”。本章能说的分三层，一层比一层弱，也一层比一层稳。

第一层（稳）：把两个问题拆开问。“快速入睡”与“睡得深”在现有证据下没有被证明冲突，也没有被证明相容。任何把它们连着承诺的建议，都超出了现有证据。这是本章唯一可以确定说的话。

第二层（有证据，但不是本章的）：按状态量给的处方是有效的。抬高睡眠压力（固定起床时间、缩短卧床、不补觉）、对齐相位（早晨见光）、扩开远端血管（睡前温水浴、暖脚）——这些有随机对照证据，且与本章无关。本章的失败不影响它们。

第三层（形态约束，来自哲学）：无论将来发现什么，处方都不能作用在“你要注意自己怎么下去”上。这是十的第三小节那条削的直接后果。如果有一天证明下降速率要紧，那么可用的干预只能是改变条件（温度曲线、光照时刻、卧床时刻），而不能是改变注意力的指向。一条要求人监控自己入睡过程的建议，在原理上就是坏建议——不管那个被监控的量多么正确。

这三层里，第三层是本章自己的，而且它不依赖第六、七节的任何一个数。

十二·用本章的判据检本章自己

本章的判据是：这件事有没有一栏是它自己的；如果没有，判定它缺哪几件。把它转过来对准本章。

先对准这次检验。本章的四条预测，两条落空、一条被自己拆掉、一条指向对手。按第五节那张缺件清单，本章自己缺三件——刻度、扰动、控制设计——而本章明知缺三件仍然跑了这次检验。这不是错误：把一个缺件的检验跑完并如实报告它停在哪里，正是让第八节那个“未判”变成一个可指认对象的唯一办法。但这也意味着，本章的失败不构成任何证据，只构成一次刻度的示范。

再对准本章的自变量。T 是用两个端点算出来的——入睡点与首达深睡。它对中间那一段的形状一无所知。一个先掉到很深再浮上来再掉下去的人，与一个匀速滑下去的人，只要用的窗数相同，在本章眼里就是同一夜。本章批评临床报告只记端点，而本章自己的自变量也只用了两个端点。这不是疏忽，是第②件缺失的直接后果；但它必须被写在正文里，而不是写在局限一节的第四条。

最后对准这一篇文章的位置。本章是本书九章里的第五篇，也是第一篇属于“怎样”题的。前四篇每一篇都报告了一个成立的判断；本章报告的是一次失败。按学术生产的常规结算方式，一篇没有阳性结果的文章记不上什么账——它在引用上、在评价上、在读者的记忆里，都近乎不存在。而这正是本章所描述的那个结构在本章自己身上的一次运行：一件确实发生了、确实占用了时间、也确实排除了一条错误路线的工作，因为账本只记阳性结果，所以在账上是零。

还有一层是这一稿才看见的。初稿漏掉了研究浪费与研究缺口那一整支元研究——一支专门做“哪些重要问题至今没被回答”的学问，而本章第八节那句话几乎是它的标准句式。漏掉它的那份初稿，与没有漏掉它的稿子，在纸面上一模一样：同样有三家、有前提、有预注册、有失败报告。这正是算熟未裁的第二个签名在本章自己身上发作——没查到那一支，不会在任何地方报错。

最后一句留给那件本章没能做成的事。要判定“又快又深”能不能同时达成，只需要两样东西：一把已经存在了三十年的刻度，与一次实验设计。本章可以被引用的最稳的一层，不是那个新名字，也不是第七节那五张表——那五张表里已经有四张是用来拆本章自己的。最稳的是这一句：这两样东西加起来，比大多数人以为的便宜得多。

注释与材料

公开数据为 Sleep-EDF Expanded 1.0.0 版 sleep-cassette 部分，按文件名排序的前四十份判读标注，与本书前四章同一批，未作任何挑选。第五节第四小节的预注册文本在取数之后、跑任何统计之前写定，其后一字未动；四条预测的不利下场按封存时写下的规则原样写入第六、七两节，未作补救性重述。第七节第一至五小节的稳健性检验与第四小节的五千次置换检验为事后（非预注册）分析，本章不把它们伪装成预注册的一部分——它们全部是用来拆本章自己的，其中第二小节直接取消了本章唯一一条命中的实质预测。解析脚本、预注册文本与四份分析脚本一并提供，可从原始 EDF+ 标注起逐步复算；置换检验的随机种子已写死在脚本中。

关于本稿新增的分析：第六节第五小节的自助区间与样本量外推、第八节第一至三小节、第九节第一小节，均为初稿之后新增。自助区间使用与原稿同一批四十夜、同一套读数定义，随机种子写死于脚本；样本量外推假定实测效应为真，是一次外推而非功效计算，不得当作功效分析引用。

关于文献核验：本稿写作过程中逐条核对了 Elster (1983, 章节与原句)、Ogilvie (2001, 摘要原句)、Hori 等 (1994, 九级的具体分级)、Espie 等 (2006, 出处与核心表述)、Drake 等 (2004, 出处与设计) 五条的来源页面。其余各条依通行出处引用，未在本轮逐页核验，读者引用前请自行复核——本系列第一篇曾因引注挂错过一次正文更正，此处预先声明。

人机分工声明：本章的三家选源由人指定；文献检索与核对、检验设计、脚本、真跑与初稿由人机协作产出。四条先验预测在拿到任何结果之前即已封存；H2 未命中、H3 被稳健性检验取消、H1 在去掉四夜后翻掉、三分位形状被判为噪声——四件不利结果全部按封存时写下的规则原样报告，未作任何补救性重述。

参考文献

4. Elster, J. (1983). *Sour Grapes: Studies in the Subversion of Rationality*. Cambridge University Press. (第二章“本质上的副产品”)
12. Reiner, M. (1964). The Deborah Number. *Physics Today*, 17(1), 62.
16. Ryle, G. (1949). *The Concept of Mind*. Hutchinson. (成就动词与任务动词之别，第五章)
18. Berry, R. B., et al. *The AASM Manual for the Scoring of Sleep and Associated Events*. American Academy of Sleep Medicine. (判读窗与占优判定条文)
22. Higgins, J. P. T., Thomas, J., et al. (Eds.). *Cochrane Handbook for Systematic Reviews of Interventions*. (“对研究的启示”一节的写法)
19. Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberyé, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194. (Sleep-EDF 数据库来源)

修订说明（本稿相对初稿的六处改动）

初稿完成后经一次外部评审，指出一处整支缺席与三处可补。逐条列明，供比对：

1. **新增 9.1“可避免的研究浪费与研究缺口”**，排在分界节第一位，并承认初稿把这一整支漏了。那一支（Chalmers & Glasziou 2009／Cochrane“对研究的启示”／James Lind Alliance 优先级设定）专门做“哪些重要问题至今无人回答”，而本章第八节那句话几乎是它的标准句式。两条分离线（处方的时间顺序相反；连变量都没有字段的问题进不了优先级候选池）反过来给本章加力。并附检索词表——因为“未裁”是一个否定性断言，不出示检索不算检索。

2. **新增 6.5 自助区间**：六项里五项跨零，唯一不跨零的正是第七节第二小节要拆掉的那一条，且下界只有 +0.005。据此把结论从“本章的检验失败了”升级为“**在四十夜这个样本上，本章的检验原理上不可能成功**”；并给出样本量外推——**约需八十夜**，深睡百分比那一项的区间才第一次离开零。这个数把“未裁”从否定性断言变成有标价的断言。

3. **第三节第二小节补 Ryle 的成就动词／任务动词**（原稿只挂在参考文献第 16 条，正文未用）：一个状态落在“本质上的副产品”这一类里，当且仅当指向它的动词是成就动词而没有配套的任务动词。这条判据补上了 Elster 自曝的那个缺口，也解释了“慢慢地睡下去”为什么在语法上就不成立。

4. **第六节第一小节补一段自诊**：H1 的两项贴着阈值（0.285、0.265）与贴着零的那一项（0.010），在预注册结算表上是同一个记号——**这正是本系列第三篇的那一格在本章自己的结算表上发作**。本章用“命中／未命中”结算，就是在给自己造一份只保留名次的记录。

5. **新增 8.1-8.3**：那一格的三个签名（短期表现为改善／失效不产生任何信号／越正规化越自我加固），命名为**算熟未裁**；两个不在睡眠里的实例（药物滴定速率——同时是第三个正面对照，因为同一个领域内部两种情形并存；康复训练的负荷上升曲线）；以及四条使“未裁”为假的条件。

6. **第十二节补一段自反**：初稿漏掉研究缺口那一整支，这件事本身是算熟未裁第二个签名在本章自己身上的一次发作——没查到那一支，不会在任何地方报错。

仍未做的三件（与初稿相同，未因本次修订减少）：八十夜以上的样本；细于三十秒的下降刻度；一次标准化扰动。

编三·每个人都得归进一类

个体层

这一编从三十秒和一个分数，走到一个人。

一个人躺进诊室，最后被归进一类：失眠，或者不是。这一层的强制作答发生在归属那一步——不管两套证据说的是不是同一件事，最后必须给出一个类别。

第九章指出：主诉给一套划分，客观界线给另一套，两个方向都大量漏出，而没有第三套划分能同时细化二者；制度用同一个词称呼两边。第十章指出：一次事件被读成了一个人身上的一种能力——同一条三十分钟的线，在“这一夜发生了什么”和“这个人是什么样的人”两个位置上，量到的是两样东西。第十一章指出：这个系统里能重复的那个量没有位置，而所有的解释都在找位置；于是清单越长，能找到的原因越多，“还有一个原因没找到”这个信念被越加固。

三章全部处理失眠，全部用同一批一百五十三夜的不同切法。**其中前两章是两条互不知情的产线，碰到了同一个方向**——自述健康的人身上，客观界线大面积命中，且随年龄陡升。它们互为对方的独立复现，这是全书最有力的一处，也是本书唯一一次不靠论证、只靠两条独立路径撞在一起就成立的地方。

第九章 两张名单

两张名单

论“失眠”至少指着两套互不细化的划分，以及为什么不存在一个更细的第三方去裁决它们

Two Rosters: On Insomnia as Two Mutually Non-Refining Partitions

提 要

“失眠是什么？”这个问题今天有三种互不相容的答法。群论说，一样东西是什么，由**在哪些变换下它保持不变**决定——先指定群，才有“同一个”；离开群谈同一性没有意义。量子力学说，一个态是什么，由**可观测量代数**决定；而它自己有一条定理：同一个密度矩阵可以由无穷多种不同的系综实现，这些系综在该系统的一切可观测量上完全等价。佛学说，睡眠（middha）在唯识百法里列在**四不定之中**——它本身不定善恶，是什么取决于与它相应的是什么。

三家争的是失眠的身份由什么给定，而三家共同踩着一句没说出口的话：“这个人现在是不是在失眠”这个问句有一个答案。推翻它的材料来自量子力学自己：存在一整类关于“它实际上是哪一个”的问题，在原理上没有事实可依——不是我们测不准，是那个问句没有事实与之对应。

本章用两份公开数据做了一次预注册检验，问的是一件很朴素的事：主诉给出的划分与客观界线给出的划分，有没有一套细化另一套？若客观细化主诉，则主诉无问题的人应当全部落在客观正常一侧。四条先验预测全部命中：七十八名库说明书载明为健康、未服睡眠药物的受试者中，**百分之八十五点九在某一夜命中至少一条常用临床量化界线**（六十岁以上两组均为百分之百）；反方向，二十二名载明为“轻度入睡困难”者的安慰剂夜里，**入睡潜伏期超过三十分钟的只有一人（百分之四点五）**，中位数十点五分钟，而他们命中最多的是一条他们并未主诉的维度。

两个方向都大量漏出，因此两套划分互不细化。既然互不细化，就不存在一个更细的第三方去裁决谁对——这不是测量精度问题，是等价关系的结构问题。本章把“同一个人在两套互不细化的划分下归属不同、且不存在细化二者的第三套”这一格命名为**异归**，给出一张可以拿去核对任何一个临床类别的**归属四问**，并报告一处对本章自己不利的稳健性结果：换一个同样正当的分析窗，甲组的命中率从百分之八十五点九变成百分之百，而睡眠效率那条界线本身从百分之七十五点六跳到百分之百——**这条界线的取值依赖一个从未被声明的口径**，而这一点本身就是本章的论点在本章数据上的一次发作。

⚠ 本章**不主张**这些自述健康的人“其实有失眠”，也**不主张**客观界线是错的。本章主张的是更窄的一句：它们与主诉不构成同一个划分，而现行制度用同一个词称呼两者。

关键词：异归 | 互不细化 | 归属四问 | 双向漏出率 | 等价关系 | 系综分解

一·两张名单

先摆两张名单，本章全部的力气从它们来。

第一张：七十八个人，年龄从二十五岁到一百零一岁，来自一个公开的睡眠数据库。这批人的入组条件写在库的说明书里：健康，不服用任何与睡眠有关的药物。**没有一个人说自己有睡眠问题。**

第二张：二十二个人，来自同一个库的另一个子集。他们的入组条件同样写在说明书里：**轻度入睡困难**，其余健康。**每一个人都说自己有睡眠问题，而且说的是同一种。**

现在把三条在睡眠研究里被反复使用的量化界线，同时压到这两张名单上——入睡潜伏期超过三十分钟、入睡后清醒超过三十分钟、睡眠效率低于百分之八十五。

结果在第九节。这里先说它的形状：**两张名单都大面积地不服从。第一张名单上的人大批命中界线，第二张名单上的人大批一条也不命中。**

本章要问的不是“谁错了”。本章要问的是一个更靠前的问题：**当两套划分互相不服从时，凭什么认为其中一套是另一套的细化？**

而这个问题，需要一门专门研究“什么叫同一个”的学问才能问清楚。

二·群论：先指定群，才有“同一个”

数学里处理“两样东西算不算同一样”的标准办法，是先指定一个作用在集合上的群，然后把“同一”定义为**同轨道**：两个元素算同一个，当且仅当存在群里的某个元素把一个送到另一个。

这套装置的核心不在于它给出了什么答案，而在于它规定了**问题的形式**：

在没有指定群之前，“这两样东西是不是同一样”不是一个有意义的问句。

同一副扑克牌，在“任意置换”这个群下，所有排列都同轨道——只有一个类；在“循环移位”这个群下，类多得多；在平凡群下，每一张排列自成一类。**牌没有变，变的是你允许哪些操作把差别抹掉。**

配套的还有一条计数关系：一个元素所在轨道的大小，等于群的大小除以它的稳定子的大小——**保持它不变的操作越多，与它同类的东西就越少。**身份的“牢固”与它所在类的“宽窄”是同一件事的两面。

这一家的单因主张是：一样东西是什么，由**在哪些变换下它保持不变**决定；先有群，才有类，才有身份。

它自曝的地方非常干脆：群论本身不告诉你该用哪个群。哪些差别算“同一件事的不同写法”、哪些差别算“两件事”，是从外面输入的。数学只保证：一旦你指定了群，等价类就唯一确定；至于该指定哪个群，它一个字也不说。

这一处自曝后面要用两次。

三·量子力学那条不许

第二家给出的东西比第一家重。它不只是说“身份依赖于你允许什么”，它给出了一个**硬定理**，说明有一类关于身份的问题在原理上没有答案。

一个量子系统的混合态由密度矩阵描述。密度矩阵可以由一个“系综”实现——一堆纯态，各自带一个概率。自然的想法是：这个混合态**实际上是**那些纯态之一，只是我们不知道是哪一个。

这个想法是错的，而且错得很彻底。

一九三六年薛定谔已经处理过这件事；一九八四年吉辛、一九九三年休斯顿、约萨与伍特斯给出了完整的分类（Hughston, Jozsa & Wootters, *Physics Letters A* 183: 14–18, 1993）。结论可以引到很直白的形式：由于辅助空间与基底可以任意选取，一个混合态的纯化不是唯一的；事实上，给定的混合态有无穷多个纯化。而这些不同纯化对应的不同系综，只需选择去测量同一个给定纯化的不同可观测量，就能被实现出来。

把这句话翻成本章要用的形状：

同一个密度矩阵可以由无穷多种彼此不同的系综实现，而这些系综在该系统的一切可观测量上完全等价。“它实际上是哪一个系综”这个问句，没有事实与之对应。

请注意这不是认识论上的谦虚——不是“我们暂时测不出来”。它是**可观测量代数的结构性结论**：能被测的东西已经被密度矩阵穷尽了，剩下的差别不在物理里。

这一家的单因主张是：一个态是什么，由**能对它做的测量**决定；测不出差别的两样东西，不是“差别很小”，是**没有差别**。

它自曝的地方同样要紧：这条结论的适用范围由“可观测量代数”划定，而代数是相对于一个系统与一组可及操作说的。**换一组可及操作，同一批系综可能就分得开了**——纠缠的另一端做一次测量，本地的系综就被挑选了出来。所以“没有事实可依”这句话，本身也是**相对于一组操作**说的。

第二家与第一家在这里第一次正面撞上：第一家说身份由不变性（内禀结构）给定，第二家说由可测性（外部操作）给定。同一句“什么叫同一个”，一个答在结构那边，一个答在操作这边。

四·佛学：睡眠是不定

第三家把前两家的对象整个换了一个位置。

唯识学把心所法分类，其中有一组叫**四不定**——悔、眠、寻、伺。“不定”这个名字说的正是本章要的那件事：这四个心所本身**不属于善，也不属于不善，也不是决定性地属于无记**；它们是善是不善，取决于与它们同时生起的是什么。同一个“眠”，随善心而起则善，随烦恼而起则不善。

而对“眠”（middha）本身的界定同样值得注意：它的作用被描述为令心昧略、不得自在——**不是一种内容，是一种“主宰权减弱”的模式**。

这一家的单因主张是：一样东西是什么，由与它相应的是什么决定；离开相应者去问它自身的性质，问的是一个不存在的东西。这条主张有一个更一般的名字——缘起：不是先有一个自性，再进入关系；是关系成立了，那个东西才作为这样一个东西成立。

它自曝的地方是本章第十三节要用的：如果一切身份都是相待而有的，那么**“两套划分互不细化”这件事本身也不构成一个特别的困难**——它只是缘起的一个普通实例，没什么可惊讶的。这条会削掉本章一句更响亮的话。

三家两两之间是真打架

群论与量子力学打的是身份归内还是归外。群论说身份由不变性给定——**有一个东西在那儿不变**；量子力学说测不出的差别不是差别——**没有一个“在那儿”的东西等着被测出来**。前者是结构实在论的形状，后者是操作主义的形状，两者不能同真。

量子力学与佛学打的是“没有事实可依”的性质。量子力学说那是**可观测量代数的结构性结论**，是一条定理，有明确的适用边界，换一组操作就可能改变；佛学说无自性是**普遍的**，不待某个特定代数而成立。一个是局部定理，一个是一般命题，二者的证成方式互不承认。

佛学与群论打得最直接。群论要求先指定群，然后有一个在群下**不变**的东西——不变性正是“自性”的数学形态；佛学说没有任何东西在一切关系下不变。群论会说：那你至少得承认“在这个群下不变”；佛学会说：那个群本身也是缘起的。这一处**两家谁也说服不了谁，而正是这一处，把本章的问题逼了出来**。

顺带挡掉第四家。临床指南、保险核定、可穿戴设备的评分算法都不是一家——它们是这些主张落地的地方。本章后面大量使用公开数据与量化界线，那是兑现材料，不是第四种立场。

五·三家共同踩着的那块地

把三家的争论压成一句：三家争的是**失眠的身份由什么给定**——由不变性、由可测性、还是由相应者。

而“失眠”这样东西，三家都默认它是这样一种东西：

“这个人现在是不是在失眠”这个问句，有一个答案。

群论默认它——它只是要求你先说清楚在哪个群下；量子力学默认它——它只是要求那个答案必须能被某个测量给出；佛学默认它——它只是说那个答案由相应者决定。三家都在争**答案由什么给定，没有人问过这个问句是不是有答案**。

念给谁听都会说这还用说吗。这正是三家共同踩着的那句话的验收标志。

推翻它的材料，来自量子力学自己

借一条外来的理由去推翻它，得到的不过是第四种立场。要真正取消它，用的东西必须是这三家中某一家自己手上早有的——已经写下、已经发表，只是从来没有朝这个方向使过。

它就是第三节那条定理。休斯顿等人在量子信息论里回答的是“给定密度矩阵的系综有哪些”，这是一个技术问题；**没有人把它拿来当作关于“一类问句有没有答案”的一般证据**。而一旦这样用，三家共同踩着的那句话就塌了一半：

存在一整类形如“它实际上是哪一个”的问句，在原理上没有事实与之对应——不是因为难测，是因为可测的东西已经穷尽了，剩下的差别不在那里。

另一半由群论自己拆掉：群论承认它**不指定群**。于是“失眠”这个类要成立，必须有人从外面指定：哪些差别算同一件事的不同写法。而临床上从来没有人指定过这件事。

两半都塌之后，问题就换了：

“失眠”这个词，指着的是哪一个划分？而如果不止一个，它们之间是什么关系？

六·异归，与归属四问

6.1 承重命题

“失眠”至少指着两套划分：主诉给出的一套，客观量化界线给出的另一套。两套之间没有一套细化另一套——两个方向都有大比例的成员漏到对面。既然互不细化，就不存在一个更细的第三方去裁决它们；而现行制度用同一个词称呼两者，并在不同场合默默地切换用哪一套。

本章把“同一个人在两套互不细化的划分下归属不同、且不存在细化二者的第三套”这一格命名为异归。

“细化”这个词在这里是精确的：划分甲细化划分乙，当且仅当甲的每一个类都完整地落在乙的某一个类里面。若客观划分细化主诉划分，那么“主诉无问题”这个类里的人，应当全部落在客观正常一侧——命中率应为零。零假设因此不需要模拟，它由定义给出。

6.2 归属四问

要判断某个类别有没有异归，本章不给分类格，给四个可以拿去问文件的问题：

① 这个类由哪些变换下的不变性定义？（即：哪些差别被判为“同一件事的不同写法”？写不出这份变换清单，这个类就没有指定群。）

② 有没有第二套划分同时在用？（主诉／客观读数／处方记录／保险编码，任意两套即可）

③ 两套之间有没有一套细化另一套？（做一张交叉表，看有没有空格）

④ 若没有，制度上由谁的划分作准，依据是什么？（若答案是“看场合”，则异归成立且未被声明）

对“失眠”逐项核对：① 无——没有任何文件写下哪些差别算同一件事；② 有，至少两套；③ 见第九节；④ 诊断用主诉，研究纳入用客观界线，设备评分用第三套，而没有一份文件写明为什么这样切换。

这张清单不是睡眠专用的。它可以拿去核对疼痛、疲劳、注意力缺陷、轻度认知障碍——任何一个“主诉与客观读数并存”的类别。

七·可判定：一次预注册的双向检验

预注册（跑统计之前封存，其后一字未动）

数据：（甲）Sleep-EDF Expanded 1.0.0 的 sleep-cassette 部分，七十八名二十五至一百零一岁受试者、一百五十三夜，库说明书载明为健康、未服任何睡眠相关药物者——记作主诉无睡眠问题组。

（乙）同库 sleep-telemetry 部分，二十二人的安慰剂夜，库说明书载明为“轻度入睡困难但其余健康”——记作主诉有入睡困难组。

⚠ 甲乙是两批人、两种记录方案、不同年代设备。本章不做两组之间的读数比较，只让每一组各自与同一组固定的量化界线比对。这个用法比组间比较稳健得多，但仍须声明。

口径：分析窗=熄灯时刻起八小时（熄灯时刻取自随库发布的受试者表）。入睡潜伏期=窗起点到窗内第一个非清醒判读窗的分钟数；入睡后清醒=窗内首个与末个睡眠窗之间的清醒分钟数；睡眠效率=总睡眠÷四百八十分钟。三条界线：入睡潜伏期 > 30 分、入睡后清醒 > 30 分、睡眠效率 < 85%。

先验预测：

- B1 甲组中命中至少一条界线者（两夜取“任一夜命中”） $\geq 30\%$ 。
- B2 该比例随年龄上升，最高年龄组比最低组高出至少 20 个百分点。
- B3 乙组中入睡潜伏期 > 30 分者 < 30%。

·B4 三条界线本身也互不细化：任取两条，都存在只命中其一的人（双向都有）。

撤销规则：若 $B1 < 10\%$ ，则“客观划分不细化主诉划分”在本数据上不成立，本章的经验支点作废，不得改口为“至少某一条界线上不成立”。若 $B3 \geq 50\%$ ，反方向不成立，本章只能主张单向失配。

写在跑之前的四条局限：“主诉无睡眠问题”是库说明书的入组条件，不是逐人施测的量表，是一个粗代理；分析窗定死为八小时，而本系列已记录过“换窗会翻面”的事故，故所有比例须在第二种窗下复算并一并报告；临床诊断从不只用客观界线，本章并不主张这些人“其实有失眠”；一夜一测，无重复测量的信度估计。

八·两个方向的漏出

8.1 甲组：主诉无问题的人，大批命中界线

七十八人全部可用，一百五十三夜。

命中至少一条界线者：67 / 78 = 85.9%。

界线	命中人数	比例
入睡潜伏期 > 30 分	21 / 78	26.9%
入睡后清醒 > 30 分	57 / 78	73.1%
睡眠效率 < 85%	59 / 78	75.6%

B1 命中（预测线 30%，撤销线 10%）。

按年龄分：

年龄段	n	至少一条	SOL>30	WASO>30	SE<85%
25-39	20	60.0%	20.0%	35.0%	60.0%
40-59	21	85.7%	28.6%	71.4%	61.9%
60-79	21	100.0%	14.3%	90.5%	85.7%
80-104	16	100.0%	50.0%	100.0%	100.0%

最高减最低 40.0 个百分点，**B2 命中**（预测线 20）。六十岁以上的两组，没有一个人能在这三条界线下保持干净。

B4 也命中：只命中入睡潜伏期而不命中入睡后清醒的有 3 人，反向 39 人；只命中入睡后清醒而不命中睡眠效率的 7 人，反向 9 人。三条界线之间同样互不细化——它们不是一件事的三种测法。

8.2 乙组：主诉有入睡困难的人，大批一条也不命中

二十二人安慰剂夜，年龄十八至七十九岁。

界线	命中人数	比例
入睡潜伏期 > 30 分	1 / 22	4.5%
入睡后清醒 > 30 分	10 / 22	45.5%
睡眠效率 < 85%	6 / 22	27.3%

B3 命中（预测线 30%，实测 4.5%）。入睡潜伏期中位数 **10.5 分钟**，四分位 5.5 与 20.0 分钟。

探索性（非预注册，标明如此）：三条界线一条也不命中者 $12 / 22 = 54.5\%$ 。

8.3 这两张表放在一起是什么形状

主诉无问题的人里，**85.9% 命中**；主诉有困难的人里，**54.5% 一条不命中**。若客观划分细化主诉划分，前一个数应当是零；若主诉划分细化客观划分，后一个数也应当是零。两个都不是零，而且都很大。两套划分互不细化。

还有一处更尖的细节。乙组这批人主诉的是**入睡困难**——而恰恰是入睡潜伏期这条界线，他们只有一个人命中（4.5%）。他们命中最多的是**入睡后清醒**（45.5%），一个他们并没有主诉的维度。主诉指着一个地方，客观读数亮在另一个地方。

九·换一个窗，本章自己的界线就动了

本系列有一条铁律：凡用这批数据报比例，必须在第二种同样正当的分析窗下复算。照做，结果对本章不利，也对本章有利，两面都得报。

把甲组的分析窗从“熄灯起八小时”换成“熄灯到记录结束”：

读数	八小时窗	全窗
至少一条	85.9%	100.0%
入睡潜伏期 > 30 分	26.9%	26.9%
入睡后清醒 > 30 分	73.1%	76.9%
睡眠效率 < 85%	75.6%	100.0%

三件事要说清楚。

其一，本章的结论方向不变，而且更强——两种窗下 B1 都远超撤销线，全窗下七十八人无一幸免。

其二，睡眠效率那条界线从 **75.6%** 跳到 **100%**。原因不神秘：全窗更长，分母更大，效率必然更低。但这正说明——“睡眠效率低于百分之八十五”这句话，在没有声明分母是什么之前，不是一个确定的判据。而三条界线里，只有入睡潜伏期在两种窗下完全不动（26.9%/26.9%）。

其三，这件事本身就是本章的论点在本章数据上的一次发作。本章说“失眠”这个词指着两套没有指定群的划分；而本章自己用来做检验的那三条界线，其中两条的取值依赖一个从未被任何文件声明的口径。第六节那个第①问——这个类由哪些变换下的不变性定义——本章自己也答不上来。

十·与最近的十二种说法的分界

逐条列三样：它走到哪一步、线划在哪里、一个能把“本章对”与“它对”分开的具体判法。只写得程度比较的，算被它覆盖，不留。前三条离本章最近。

10.1 矛盾性失眠与睡眠状态错觉——本章最危险的一位

它走到哪一步：睡眠医学早就注意到主诉与客观记录不一致，并且为它起了名字：有一类患者主诉严重失眠而多导记录基本正常。这一支已经把本章第八节乙组那种现象命名、分型、写进了分类手册。

分离线在谁被判为错的那一方。这一支把不一致读成病人的知觉出了问题——名字里的“错觉”“误知”已经把裁决做完了：客观是准的，主观偏了。本章说的是另一件事：**两套划分互不细化，因此不存在一个更细的第三方去裁决谁偏了**。而本章第八节甲组那个方向，这一支处理不了——没有名字叫“客观错觉”，尽管 85.9% 的自述健康者命中界线。

判定形式：若把客观界线当作准绳，则甲组那六十岁以上的百分之百应当被读成“大量隐性失眠”；若把主诉当作准绳，则乙组那 54.5% 应当被读成“主观错误”。**两种读法都必须放弃另一半数据**。若有人给出一套同时容纳两个方向的裁决规则，本章这一格作废。

10.2 客观短睡表型

它走到哪一步：一支重要的工作用客观睡眠时长把失眠再分亚型，并报告不同亚型在心血管与代谢结局上的差别。这是“用客观读数把主诉切细”这个动作做得最成功的一次。

分离线在它预设了什么：亚型化预设主诉划分是母类、客观读数在其内部切分——即客观细化主诉。本章第八节甲组那个数直接反对这个预设：客观界线大量落在主诉划分之外。

判定形式：把同一套亚型规则施加到主诉无问题的人群上。若几乎没有人被切进任何亚型，则客观确实只在主诉内部起作用，本章错；若大批人被切进去（本章数据即是如此），则那不是母类内部的细分。

10.3 分类系统与残余范畴

它走到哪一步：科技与社会研究里有一整支处理分类系统如何造成后果，包括落在所有类别之外的东西如何被处置。

分离线：那一支处理的是一套分类系统内部的边缘与残余；本章处理的是两套分类系统之间的关系，而且是一种特定关系——互不细化。一套系统可以毫无残余而与另一套完全不可比。

判定形式：若把两套划分合并成一套更细的公共加细（数学上总是存在），临床上是否可用？若可用，则问题只是“还没人合并”；若合并后的类数多到无法承载（本章数据：三条界线加主诉，最多十六类，其中大部分只有个位数的人），则不可比是实质的。

10.4 群论本身

它走到哪一步：见第二节。它给出了“同一个”的完整形式，也给出了自曝——不指定群。

分离线：群论说指定群之后等价类唯一；本章说的是在没有指定群的地方，实际上有两套等价类在同时运行。前者是数学定理，后者是一个制度事实。

判定形式：若临床上能找到一份文件，写明哪些差别算“同一种失眠的不同写法”，则群已被指定，本章第六节第①问有答案，本章这一格作废。

10.5 量子力学那条定理

它走到哪一步：见第三节。它给出本章的推翻材料。

分离线，也是本章必须自曝的一处：那条定理是关于可观测量代数的数学结论，有严格的前提；本章借用的只是它的形状——“有一类关于身份的问句没有事实可依”。本章并不主张睡眠里存在任何对应于密度矩阵的东西，也不主张两套划分之间有任何量子性质。这是一个结构类比，不是一次物理推导。

判定形式：若两套划分之间能找到一个可测的量，其取值决定哪一套作准，则“没有事实可依”不成立，本章的类比失效。

10.6 佛学的不定与缘起

分离线：唯识说“眠”是不定，是一条关于心所本性的教义主张；本章说的是两套具体划分在一批具体数据上互不细化，是一条可被数据推翻的经验主张。教义为真时本章可以为假（两套划分完全可以恰好互相细化），反之亦然。

10.7 诊断标准本来就不用客观界线——本章最应正面接的一条反驳

它说到哪一步：这是对本章最省事、也最难反驳的一条批评。现行的失眠诊断标准以主诉与日间功能受损为核心，并不要求多导睡眠记录；那三条量化界线主要用于研究纳入与描述，不是诊断门槛。因此“自述健康者命中客观界线”根本不构成矛盾——没有人拿那些界线去诊断他们。

本章接受这条批评，并且认为它加强了本章。如果那些界线不用于诊断，那么它们在做什么？它们大量出现在研究的纳入与排除标准里、出现在可穿戴设备给出的分数里、出现在与劳动能力和保险有关的评估里。一套不用于诊断的划分，仍然在决定谁进入研究、谁拿到什么分数、谁被认为“有问题”。本章第六节第④问要的正是这个：由谁的划分作准，依据是什么。批评者的回答恰好是“看场合”——而这正是异归成立且未被声明的定义。

判定形式：若能找到一份文件，写明在哪些场合用主诉、哪些场合用客观界线、以及为什么，则异归被声明，本章的主张退化为“已知问题”。

10.8—10.12 自家五处

(一) 站上《失眠：为什么越想睡越睡不着》(同一作者)——自家这一侧最险的一处。那篇处理的是失眠为什么发生：意志监督入睡，而入睡要求意志退场。本章完全不碰这条，通篇不使用意志、用力、松弛一族的任何说法。

分离线在两篇各自预设了什么：那一篇从一个已经被确定为失眠的人出发，问他身上发生了什么；本章问的正是“确定”这一步是怎么做的。两篇可以同时为真，而且互补——但顺序上本章在前：那一篇讲的机制，究竟发生在两套划分的哪一套所指的人身上，那一篇没有说，也不需要说。判定形式：若那一篇描述的机制在甲组（主诉无问题、客观命中）的人身上同样出现，则两套划分指着同一批人，本章错。

(二) 本系列《三样东西，一个分数》(第五章)(索验比)。那篇讲诉求端与验收端的分辨率不匹配；本章第八节乙组那个“主诉入睡困难而命中的是入睡后清醒”与它形状相近，必须切开：那篇是一维尺子的粗细问题(分不开)，本章是等价关系的结构问题(两套划分互不细化)。一把无限精细的尺子解决那篇的问题，解决不了本章的问题——两套划分可以各自精确到任意程度而仍然互不细化。

(三) 本系列前几篇的记录制度全书赌的那句话(账上少一栏/边界没有地/速率未裁)：那几篇讲一套记录里缺什么，本章讲两套划分之间是什么关系。缺项与不可比是两件事：一套毫无缺项的记录，照样可以与另一套互不细化。

(四) 本系列《办法都在夜里》(第十四章)：那篇讲干预点的配置，本章讲类别的归属。

(五) 站上睡眠医学面板：那份整理的是“哪个读数有资格代表睡得好”，本章问的是更靠前一步——这些读数所划出的类，与主诉划出的类是什么关系。

十一·三家反过来给本章加的约束

11.1 群论削掉的：本章不能说哪一套划分更对

原想写：客观界线划出的类比主诉划出的类更基本。

群论不许。等价类的粗细由所选的群决定，而群是从外面输入的；没有哪个群天然更基本。说客观更基本，等于偷偷指定了一个群还不声明——正是本章批评的动作。

削掉之后剩：本章只主张两套互不细化，不主张任何一套更对。这一处动的是价值排序。

11.2 量子力学削掉的：本章不能说“没有事实可依”

原想写：“这个人是不是在失眠”这个问句没有事实与之对应。

量子力学不许。那条定理的“没有事实可依”是相对于一个明确的可观测量代数说的，而且它自曝过：换一组可及操作，同一批系综就分得开了。睡眠这边没有任何东西对应于可观测量代数，本章因此没有资格作这个断言。

削掉之后剩：本章只能主张——在现有的两套划分与现有的记录上，不存在一个细化二者的第三套。这是一句关于当下可用工具的话，不是一句关于事实结构的话。这一处动的是本章的适用范围，也是最大的一次让步。

11.3 佛学削掉的：本章不能把这件事说成一个反常

原想写：两套划分互不细化，这是一件反常的、需要被修理的事。

佛学不许。若一切身份都是相待而有，那么两套划分互不细化只是缘起的一个普通实例，谈不上反常，也未必需要修理。真正值得说的不是“它们不一致”，而是“制度在不同场合默默切换而不声明”。

削掉之后剩：本章的处方不是“统一两套划分”——那可能既做不到也不必做——而是一条纯程序性的要求：凡使用某一套划分的地方，须声明用的是哪一套。这一处也动了价值排序，且它把本章从一篇批判文章变成了一篇登记文章。

11.4 数据自己削掉的：本章的两组不可直接比

原想写：甲组与乙组的对照证明了双向失配。

数据不许。甲乙是两批人、两种记录方案、不同年代设备。本章因此**不做组间比较**，只让每组各自与同一组固定界线比对。双向失配这个结论是两次**独立的单组结论**并列，不是一次对照实验。若将来有一份同一批人身上同时有细致主诉量表与多导记录的公开数据，本章的检验应当在那上面重跑——**那才是这个问题该有的数据。**

十二·这不是说客观界线错了

本章有责任说清楚它没有主张什么，因为这一篇最容易被误读成一个廉价的结论。

本章不主张那七十八个人“其实有失眠”。他们说自己没有睡眠问题，本章没有任何理由推翻这一点。

本章不主张那三条界线是错的。它们各自被造出来的用途上工作得很好。

本章不主张主观比客观可靠，或者反过来。见第十一节第一小节。

本章主张的只有一句：这两套划分不是同一个划分，两个方向都有大比例的成员漏到对面，而没有一个更细的第三方能裁决它们；因此当一个词同时指着两套划分、并在不同场合默默切换时，**至少应当声明这一次用的是哪一套。**

这条处方很小，小到几乎不像一个结论。但它有一个好处：**它不需要任何新的研究经费，也不需要等任何人同意。**

十三·用本章的判据检本章自己

第六节那四问，转过来对准本章。

第①问：本章自己的类由哪些变换下的不变性定义？答不上来。本章把“主诉无睡眠问题”这个类，整个建立在一份公开库的入组条件上——**那是一句说明书里的话，不是逐人施测的量表。**本章一边批评临床没有指定群，一边自己用了一个连边界都是转述的类。

第②问：有没有第二套划分？有，而且本章自己就制造了一套——第九节那两种分析窗，给出的是两套不同的命中名单。

第③问：有没有一套细化另一套？八小时窗的命中者是全窗命中者的子集（85.9% $\subset$ 100%），这一次恰好是细化关系。但那只是因为全窗更宽松；**若换一条更严的界线，两套窗给出的名单是否仍然嵌套，本章没有检验。**

第④问：由谁作准？本章选了八小时窗作主结果，选择的理由是“与本系列前几篇一致”——一个关于连贯性的理由，不是一个关于正确性的理由。这一条必须写在这里，而不是写在局限一节。

还有一层。本章这一篇属于一个九篇的系列，而**这个系列本身也在做一次划分：**把睡眠这件事切成“是什么／怎样／为何”三组。这个划分同样没有指定群——没有任何文件写明哪些问题算同一个问题的不同问法。**按本章自己的判据，这个系列的目录也处在异归里。**

最后一句留给那件本章没能做成的事。要真正判定这两套划分的关系，需要的是同一批人身上同时有细致主诉量表与多导记录的数据。这样的数据存在，只是不在本章手上。**本章可以被引用的最稳的一层，不是“异归”这个名字，也不是第八节那两张表——那两张表来自两批不能直接比较的人。最稳的是那四个问题：它们不需要数据就能问，而且问的对象是文件，不是人。**

注释与材料

本章第八、九节的全部数字来自 Sleep-EDF Expanded 数据库 1.0.0 版：sleep-cassette 部分一百五十三份判读标注（七十八名受试者），sleep-telemetry 部分的安慰剂夜二十二份；年龄与熄灯时刻取自随库发布的两份受试者表。第七节的预注册文本在取数之后、跑任何统计之前写定，其后一字未动；四条先验预测的结果与两种分析窗下的复算按封存时写下的规则原样报告。解析与统计脚本随文提供，可从原始 EDF+ 标注起逐步复算。

关于哪些是本轮核实的：量子系综分解那条定理的出处与表述（Hughston, Jozsa & Wootters, 1993；上溯 Schrödinger, 1936 与 Gisin, 1984）在本轮逐条核对了来源页面。**唯识四不定的分类、群论的轨道—**

稳定子关系、以及三条量化界线的常用取值，依通行出处引用，本轮未逐页核验，读者引用前请自行复核——本系列第一篇曾因引注挂错做过一次正文更正，此处预先声明。

一处必须声明的用法限制：本章的两组来自不同子集、不同记录方案、不同年代设备，本章不做组间读数比较，双向失配是两次独立的单组结论并列。凡引用本章数字者，请一并引用所用的分析窗——第九节已说明为何这一点不可省。

人机分工声明：本章的三家选源由人指定；文献检索与核对、口径设计、预注册文本、脚本、真跑与初稿由人机协作产出。四条先验预测与撤销规则在拿到任何结果之前即已封存；第八节第二小节那个 54.5% 为探索性分析，已在正文标明；第九节对本章不利的一半（睡眠效率界线随窗跳到百分之百）按封存时写下的规则原样报告。

参考文献

4. Nielsen, M. A., & Chuang, I. L. (2010). *Quantum Computation and Quantum Information* (10th anniv. ed.). Cambridge University Press. (密度矩阵与系综分解的标准处理)
5. Artin, M. (2011). *Algebra* (2nd ed.). Pearson. (群作用、轨道与稳定子)
7. 世亲《大乘百法明门论》；护法等《成唯识论》卷七（四不定：悔、眠、寻、伺）。
8. 玄奘译《瑜伽师地论》本地分（睡眠心所的界说）。
9. Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberyé, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194. (Sleep-EDF 数据库来源)
10. Berry, R. B., et al. *The AASM Manual for the Scoring of Sleep and Associated Events*. American Academy of Sleep Medicine. (判读窗与分期定义)
11. American Academy of Sleep Medicine (2014). *International Classification of Sleep Disorders* (3rd ed.). (失眠障碍的诊断要件以主诉与日间功能为核心)
12. Edinger, J. D., et al. (2004). Derivation of research diagnostic criteria for insomnia: report of an American Academy of Sleep Medicine Work Group. *Sleep*, 27(8), 1567–1596. (研究用量化界线的来源)

第十章 那一次，还是那个人

那一次，还是那个人

论失眠的诊断把一夜里的一次事件读成了一个人身上的一种能力，以及为什么改一个定义就能让四分之一的夜换类

摘要

关于“失眠是什么”，三门互不往来的学问给出三种判法。西医学说看主诉与日间后果就够了——现行的慢性失眠障碍诊断要求入睡或维持困难、伴日间损害、每周至少三次、持续至少三个月，且在睡眠机会与环境充足的条件下；而权威学会明确建议不要为慢性失眠常规做多导睡眠图。易经学说，一件事算什么由它所处的时与位决定，同一个动作时不同则吉凶相反。戏剧学说，出没出错不重要，重要的是错之后那几个动作——救场是一门有工序的技艺。

三家共同踩着一句没说出口的话：失眠是一个可以在事后被判定的属性。推翻它的材料来自易经自己：《系辞》给“无咎”下的定义是“善补过也”——最常见的判词不是“没有过错”，是“出了过错而补上了”；吉凶悔吝四判词全部预设已经出了事，这套体系里没有“正常”那一格，而且它判的是动作的走向，不是物的属性。

本章用一百五十三夜公开判读记录做了预注册检验，四条预测三条落空。唯一命中的那条很硬：清醒总量主要由最长的那一次决定（ $\rho=+0.877$ ），远强于由中断次数决定（ $\rho=+0.631$ ）；而与同样次数的独立随机过程相比，实测最长那一次长出 1.67 倍（153 夜里 123 夜高于零模型， $p=6.6\times 10^{-16}$ ）；而前后两次之间几乎没有惯性（ $\rho=+0.121$ ）。

由此提出：“维持困难”是一个能力名词，而承重的是一次事件。本章把那一次命名为长断，把这个动作命名为能力化误读。文中另报告作者自己的一处计算错误：把最后一次醒来之后的时间误计入清醒总量，使中位数由 38 分钟变成 67.5 分钟、超过常用阈值的夜由 58.8% 变成 82.4%——同一批人、同一份记录，一个定义之差，超过四分之一的夜换了类别。这不是插曲，是本章命题的实例，已写入正文。文末给出十四条判决性分界、三处反向约束与三层可分别引用的主张。

关键词：长断 | 能力化误读 | 无咎与补过 | 时位 | 维持困难 | 定义换类

一·一个不用仪器的诊断

先摆一件在睡眠医学内部是常识、而在外面几乎没人知道的事。

诊断慢性失眠，不需要做睡眠监测。

这不是说做不起，也不是说做了没用。这是一条明确的建议：美国睡眠医学会在“明智选择”清单里写着，除非症状提示合并了别的睡眠疾病（如呼吸暂停、不宁腿、周期性肢体运动），否则不建议为慢性失眠患者做多导睡眠图。诊断依据的是临床病史——你说了什么，白天怎么样。

现行的诊断标准是这样构成的：入睡困难、睡眠维持困难或早醒；伴有日间损害；每周至少三次；持续至少三个月；并且是在睡眠机会与环境充足的条件下发生的。最后那一条是排除条件——如果你是因为要照顾婴儿、上夜班、或者住在铁路边上而睡不着，那不算慢性失眠障碍。

把这几条并排看，会看见一个很奇怪的形状：在这个定义里，夜本身不承担判定作用。判定由三样东西完成——你的主诉、你白天的状态、你的环境是否充足。夜里到底发生了什么，不进入判定。

而同一门学问，在做药物试验的时候，入组标准是这样的：连续两夜的多导睡眠图，入睡潜伏期均值不低于三十分钟、或入睡后清醒时间均值不低于三十分钟、或睡眠效率不高于百分之八十五。这里夜说了算，而且精确到分钟。

同一个病，诊断时说仪器不必要，试验时又必须用仪器把它定出来。两套定义并存，而没有人要求它们一致。

本章从这里开始。要问的不是哪一套对，是一个更前面的问题：这两套定义在判什么？

二·三种判法

2.1 西医学：看眼前这些就够了

第一家的立场最清楚：**看眼前摆着的这些就够了**——病人报告什么，白天表现如何。来路不必追，机制不必知，仪器不必用。

这不是懒惰，是有理由的。现行分类明确取消了此前把慢性失眠分成“原发”与“共病”的做法，理由写得很直白：那种细分**并没有提高诊断准确性，也没有区分出不同的治疗选择**；而且把失眠称作“共病的”，会误导人以为它是次要的、随着原发病治好就会消失。于是一切慢性失眠归入一个诊断。

这一家的单因主张是：**失眠是什么，由眼前这份主诉与它的日间后果决定**；夜里的客观结构、白天的成因、身处的环境，要么不必要，要么只是排除项。

它自曝的地方，就是第一节那件事：**它自己的疗效试验用一套完全不同的、以夜为准的定义**。这一处第五节还要用。

2.2 易经学：看时与位就够了

第二家来自一套两千多年的判断体系，而本章只取它的结构，不取它的任何占断用法。

《易》判一件事，靠的不是这件事本身的性质，是它所处的**位与时**。同一根阳爻，在不同的爻位上，判词可以完全相反；《彖传》反复出现的那句“时义大矣哉”说的正是这件事——该不该动、这么动算好算坏，看时，不看动作本身。

这一家的单因主张是：**一件事是什么、算好算坏，由它在整个格局中的时与位决定**，不由它自身的性质决定。同一个“夜里醒着”，在不同的时位上是完全不同的两样东西。

它自曝的地方是本章全部论证的入口，第五节整节处理。

2.3 戏剧学：看错之后那几个动作就够了

第三家来自剧场。

剧场里有一条不成文而极硬的规矩：**演出中出了错，不停、不解释、不道歉**。台词忘了，就用下一句把它接过去；道具倒了，就让角色把它扶起来；搭档跳了一段，就顺着他跳过的地方往下演。这门技艺有个名字叫救场，而它的全部要点是：**错本身不重要，重要的是紧接着的那两三个动作有没有把它纳进这出戏**。

这与排练里的做法正相反：排练可以停、可以重来、可以讨论。演出不能。同一个错，在排练里是一次修改的机会，在演出里是一次必须被接住的事件。

这一家的单因主张是：**一段过程的完整性，不由中间出没出错决定，由错之后那几个动作的组织决定**。

它自曝的地方很关键：救场技艺预设了**还有下一个动作可用**。如果戏本身停了——停电、演员倒地、观众席起火——救场无从谈起。**它只处理可继续的错**。这一处第十节要用。

三·为什么这三种判法不能同时对

	单独够用的那一样	它据此把另两样看成
西医学	眼前的主诉与日间后果	时位是排除条件；错后的动作是治疗内容，不是诊断内容
易经学	所处的时与位	主诉只是表层；具体动作只在时位允许时才有意义
戏剧学	错之后那几个动作的组织	主诉是观众的事；时位是外部条件，不改变技艺本身

三处分歧不在同一根轴上，因此不可裁决。而两两之间是真打架：

西医学与易经学打的是标准化本身。西医学要一个跨时不变的判据——每周三次、持续三个月，对二十五岁和八十五岁用同一条线；易经学说同一个现象在不同时位上是不同的东西，用一条线去量它，量到的是线不是东西。一家的立身之本，是另一家**诊断为病根的那件事**。

西医学与戏剧学打的是判定的对象。西医学问“这一夜出了什么事”，并把它记成症状；戏剧学问“出事之后那几个动作是什么”，并认为那才决定了这一段算不算完整。而现行的判定恰恰记不住后者——病人只记得“我醒了好几次”，记不住每一次是怎么回去的。一家的日课是另一家的盲区。

易经学与戏剧学打的是人能不能做。戏剧学说救场是技艺，可训练、可执行、有工序；易经学说吉凶由时位定，人能做的是趋时，不是改命。一个说能做，一个说只能顺。

顺带挡掉第四家。诊室、药厂、手环、睡眠卫生手册都不构成一种立场——它们是上面三种主张各自落地的场所。本章后面反复使用公开判读记录，那是兑现材料。

四·三家共同踩着的那块地

把三家的争论压成一句：三家争的是**该用什么来判定失眠**——用眼前的主诉、用时位、还是用错后的动作。

而填在横线上的那样东西，三家都默认它是这样一种东西：

失眠是一个可以在事后被判定的属性；判定它，就是判定这个人有没有那个属性。

西医学默认它，所以它的全部工作是找一组够好的判据；易经学默认它，所以它主张换一套更细的判法（按时位）；戏剧学默认它，所以它主张把判定的落点挪到错后的动作上。**三家争的是判法，没有一家问过被判的那样东西是不是一个属性。**

念给谁听都会说这还用说吗。而这正是三家共同踩着的那句话的验收标志。

五·易经自己拆掉了它

一条从别处借来的反驳只能立成第四种立场，那是入局不是散局。要让这条前提失效，用的东西必须是三家自己早已印出来、只是一直没往这个方向使的。

它在《易》那里，而且是这套体系最基本的一层：**判词**。

《易》的判词有四个常用等级：吉、凶、悔、吝。而《系辞》给出的定义，几乎每一条都在说同一件事：

“无咎者，善补过也。”

这一句要慢慢念。**无咎不是没有过错，是善于补过。**在这套判断体系里，最常出现、也最接近“常态良好”的那个判词，它的定义**内建了一次已经发生的过错**。

同一段还有：“**吉凶悔吝者，生乎动者也。**”判词生于动——不是物的属性，是动作的后果。而“悔”与“吝”是两个中间判词，《系辞》称之为“忧虞之象”：同样是有过，一个向吉，一个向凶。

把这三条并起来看，会看见一个与三家共同踩着的那句话正面冲突的结构：

1. 这套体系里没有“正常”那一格。四个判词全部预设已经出了事，区别只在往哪儿走。
2. 判词判的是动作，不是属性。“吉”不是“这东西好”，是“在这个时位上这么动会好”。
3. 最好的常态判定是“补上了”，不是“没出事”。

于是三家共同踩着的那句话塌了一半：**被判定的那样东西，未必是一个属性。**它可以是一次动作的走向。

另一半由西医学自己拆掉。第一节那两套并存的定义说明：同一个“失眠”，在诊室里是一个属性（这个人有没有），在试验里是一个夜间量（这两夜的数字是多少）。如果它真是一个属性，这两套定义不会各自独立地用了几十年而没有人要求它们对齐。

两半都塌之后，问题换了：

现行判定里那个被称作“失眠”的东西，究竟是一个人身上的什么，还是一夜里的什么？

六·长断

6.1 承重命题

失眠不是一组可判定的夜间症状（西医学），不是补过程序的缺席（戏剧学），也不是时位不当（易经学），而是——把一夜里的一次事件，读成了一个人身上的一种能力。

“睡眠维持困难”这个词组里，“维持”是一个能力名词。它指的是一种持续的、属于这个人的、可以在别的夜里再次显形的东西。而第七节的数据说：一夜的清醒时长，主要由最长的那一次中断决定，而不是由中断的次数决定；那一次，比同样次数的随机过程应有的长得多。

6.2 三个用词，定死

— 长断：一夜之内最长的那一次中断（从入睡后第一次醒来算起，到最后一次入睡为止，其间最长的一段清醒）。它可以从任何一份判读记录里算出来，不需要新设备。

— 能力化误读：用一个能力名词去指一次事件。它的特征不是错，是语法上无法被察觉——“他维持不住”与“他那一次没回去”在汉语与英语里都读得通，而处置方向相反。

— 补程：一次中断之后回到睡眠所走的那一段。戏剧学的救场、《易》的补过，指的都是它。它有没有走成，决定了那一次是短断还是长断。

6.3 辨别装置：一次事件与一种能力，怎么分

不给二维格，给一条五问的核对序列。对任何一条以能力名词命名的临床判据，逐项问：

- ① 这条判据实际操作时用的是哪一个数？（写下量与单位）
- ② 那个数在一次记录里，主要由几个事件贡献？（一个 → 事件；多数均摊 → 能力）
- ③ 把最大的那一个事件去掉，这个数还剩多少？
- ④ 这个数在同一个人的不同记录之间稳不稳？（稳 → 偏能力；不稳 → 偏事件）
- ⑤ 判据的名字是能力名词还是事件名词？与②③的答案一致吗？

第五问是这套装置的全部要点：名字与承重之间可以长期不一致，而不产生任何报错。

对“睡眠维持困难”逐项核对：① 入睡后清醒总时长（分钟）或觉醒次数；② 见第七节，主要由一个事件贡献；③ 见第七节；④ 本章数据每人两夜，不足以判定，如实记为未判；⑤ 名字是能力名词，而②③的答案是事件——不一致。

6.4 一句不含程度词的判据

这条判据用的那个数，去掉最大的那一次事件之后还剩多少？如果剩得很少，那么这条判据的名字里那个能力名词指的是什么？

这句话可以拿去问诊断标准、问报告模板、问试验入组表，不用问任何人的判断。

在五个场合各跑一遍，答案互不相同，其中至少两个是查出来的：

场合	答案	来源
睡眠维持困难	去掉最长一次，清醒总量掉三分之一；而该量与最长一次的相关（+0.877）远高于与次数的相关（+0.631）	查出来的（第七节）
药物试验入组	用的是两夜均值，即已经在向能力靠；但阈值仍是总量，不区分一次长断与多次短断	查出来的（公开入组标准）
演出成败	去掉最严重的一次意外，这场戏通常仍然成立——因为救场把它纳进去了	从命题可推
驾驶资格评估	去掉最严重的一次违规，多数人的记录变成清白；而资格判定用的是“是否有过”	从命题可推
学期成绩	去掉最低的一次，均分大幅回升；而“这个学生学习能力如何”用的是均分	从命题可推

第二条反过来改了本章：试验入组用两夜均值，说明这一行早已知道单夜不足以定人。本章因此不能说“临床完全没有意识到事件与能力之别”，只能说更弱的一句——它在入组时用均值处理了这件事，而在诊断标准的措辞里没有。

七·真跑：三条落空，一条命中

7.1 预注册（跑任何统计之前封存）

数据：Sleep-EDF Expanded 1.0.0 的 sleep-cassette 全部一百五十三份判读标注，七十八名二十五至一百零一岁受试者；年龄与熄灯时刻取自随库发布的受试者表。分析窗=熄灯时刻起八小时——这一声明是必需项，同一批数据换窗会改变显著性判读。判读窗三十秒。

对象：入睡点为分析窗内第一个非清醒的分期窗；此后每一段连续清醒记为一次中断，其长度记为该次中断的回归时间。

先验预测：

- G1（重尾）：每夜最长的一次中断，占该夜清醒总量的比例，中位数大于 40%。
- G2（惯性）：同一夜内，第 n 次中断的回归时间与第 n-1 次的等级相关大于 0.20。
- G3（总量由个别决定）：跨夜比较，清醒总量与最长一次的相关，大于它与中断次数的相关。
- G4（年龄）：按年龄三分位内标准化后，G2 的相关下降。

撤销规则：若 G3 反向（次数比最长一次更能解释清醒总量），则“一夜的清醒主要由个别中断决定”不成立，本章第六节须改写，不得改口为“两者都重要”。

写在跑之前的三条局限：三十秒是本库最细粒度，短于三十秒的觉醒不可见，故中断次数是下界；本库受试者是健康人群，不是失眠患者，本检验只能说明中断的时间结构，不能说明失眠患者身上是否相同；判读依旧体例，其觉醒规则与现行手册不同。

7.2 结果

一百五十三夜全部纳入，中断对 2566 组。每夜中断次数中位 16 次（四分位 10-24），清醒总量中位 38.0 分钟。

- G1 未命中。最长一次占清醒总量的比例，中位 32.6%（四分位 21.8-43.8%），未达预设的 40%。一夜的清醒不是被一次吃掉的。
- G2 未命中。前后两次中断的回归时间相关 $\rho = +0.121$ ($p = 8 \times 10^{-10}$, $n = 2566$)，未达预设的 0.20。样本大到什么都显著，而效应量很小——上一次醒了多久，几乎不预测下一次醒多久。
- G3 命中，幅度很大。清醒总量与最长一次的相关 $\rho = +0.877$ ；与中断次数的相关 $\rho = +0.631$ 。撤销规则未触发。
- G4 未命中，且方向相反。年龄三分位内标准化之后，G2 的相关从 +0.121 升到 +0.163。年龄不是这条惯性的混杂来源；去掉年龄，惯性反而更清楚一点。

7.3 那个 32.6%，与它的零分布

按本系列上一篇留下的规矩：报一个比例并据此立论之前，先给出它在零假设下该是多少。

若一夜的清醒是 k 次独立同分布的中断，最长的一次占总量的比例，纯粹由 k 决定（k 越多，占比越小）。用每一夜各自实际的 k 做同分布模拟：

	实测	同 k 的独立指数零模型
最长一次占清醒总量（中位）	32.6%	19.5%
实测高于零模型的夜数	123 / 153 (80%)	随机期望 50%
配对检验	—	$p = 6.6 \times 10^{-16}$

实测是零模型的 1.67 倍。G1 之所以未命中，不是因为最长那次不特别——是因为我把预设的线画得太高（40%），而没有先算过在这个 k 下它本来该是多少（19.5%）。用零模型作标尺，32.6% 是一个很大的偏离；用直觉作标尺，它看起来平平无奇。这一条是本章这次检验里最要紧的一处方法教训，写在正文而不是附注。

同时报告一个与零模型一致的方向：最长一次占比与中断次数的相关是 $\rho = -0.521$ ，负号与零模型预测一致（k 越多占比越小）。结构不在“占比与次数的关系”上，在“给定次数之后那一次有多长”上。

7.4 合起来：一夜的清醒是一次事件

把三条放在一起：

清醒总量主要由最长的那一次决定 ($\rho = +0.877$ 对 $+0.631$)；那一次比同样次数的随机过程长出 **1.67 倍**；而前后两次之间几乎没有惯性 ($\rho = +0.121$)。

没有惯性这一条特别要紧，它正是“能力”这个读法最需要的东西而数据没有给：如果失眠是一种维持能力的缺陷，那么同一夜里、同一个人身上，前一次醒得久应当预示后一次也醒得久。实测几乎没有。

一夜的清醒时长，主要是“有没有摊上那一次长的”，而不是“这个人维持得住维持不住”。

7.5 年龄那三个数，正好是易经那一家要说的

	ρ (与年龄)
清醒总量	+0.574
最长一次	+0.486
中断次数	+0.395

按正确口径分层：二十五至五十岁那一层，清醒总量中位 **17.5 分钟**，超过三十分钟的夜只有 25.6%；六十九至一百〇一岁那一层，中位 **68.5 分钟**，超过三十分钟的夜 **82.4%**。

同一条三十分钟的线，在年轻那一层筛掉四分之三的夜晚，在年长那一层留下五分之四。**这不是两个人群病得不一样，这是一条线在两个时位上量到了两样东西。**易经那一家的主张在这里得到一次经验兑现，而且它不需要接受《易》的任何一条占断用法。

八·我把清醒时长算错了，而那正是本章要说的事

这一节本该在附录里，写在正文是因为它是本章命题的一次实例。

第一次跑这份数据时，我把最后一次醒来之后到分析窗结束的那段时间，也计入了清醒总量。临床上的入睡后清醒时长不是这么算的——它只算入睡点到最后一次入睡之间的清醒，早醒之后那段属于另一个指标。

改正之后：

	错的口径	正确口径
清醒总量中位	67.5 分	38.0 分
超过三十分钟的夜	82.4%	58.8%
最长一次占比中位	20.9%	32.6%
与同 k 零模型比	0.98 倍 (无法区分)	1.67 倍 ($p=6.6\times10^{-16}$)

三件事同时发生：超过四分之一的夜换了类别；核心比例从“与随机无法区分”变成“显著高于随机”；而**两套数字都可以从同一份记录里算出来，都没有造假，都有人这么用过**（早醒后清醒时长在某些研究里确实被并入清醒总量）。

这正是第五节那个结构：判定的不是一个属性，是一次动作——而这里的动作是“我把边界画在哪里”。一个定义上的选择，决定了这些人里有多少“维持困难”。

而它与本章的第二个连接更直接：如果一个定义之差能让四分之一的夜换类，那么“每周至少三次、持续至少三个月”这条频次门槛数的到底是什么？它数的是症状夜的个数，而症状夜由一条阈值线切出来，而那条线的位置可以合法地移动。本章不据此主张现行门槛是错的——本章只主张，这条门槛所依赖的那个“症状夜”，其边界的稳健性从来没有被要求报告过。

九·与最近的十四种说法的分界

逐条写四件：它走到哪一步、分离线、以及一个能分出谁对的判定形式。只写得出程度比较的，算被它覆盖，不留。排序按危险程度。

9.1 极值统计与“重尾”那一整支——本章最近的一位

它走到哪一步：一个总量主要由其中最大的一项决定，这在统计学里有专门的分支（极值理论），在别的领域也早有说法——少数几件事占掉大部分后果。凡是报告“总量由个别事件贡献”的工作，都在这一支里。

分离线：那一支处理的是分布的形状，处方是换一套统计量（用极值而不是均值）。本章处理的是名字与承重之间的错位——“维持困难”这个能力名词与它实际所指的一次事件之间的不一致，而这一层不是统计问题：换成极值统计，判据的名字仍然是“维持困难”，处置方向仍然是训练维持能力。

判定形式：若把诊断判据的操作定义换成“最长一次中断的长度”而措辞不改，临床处置随之改变，则本章这一层多余、极值统计足够；若措辞不改则处置不改，本章对。这一条可以在指南修订时直接观察，本章未做。

9.2 反常性失眠与主客观不一致

它走到哪一步：有一类患者主诉严重失眠，而客观记录显示睡眠效率很高。睡眠医学早已把它单列，并把主客观分离当作一个待解释的现象。

分离线：那一支承认主客观可以不一致，并把客观那一侧当作真值（所以叫“反常”“错觉”）。本章不作此假定：第一节那两套并存的定义说明，这一行自己也没有把哪一侧定为真值——诊断用主观，试验用客观。本章处理的是两套定义并存这件事，不是某一类患者的知觉偏差。

判定形式：若主客观不一致只出现在被诊断为反常性失眠的少数人身上，本章应收缩到那一群；若在任何睡眠障碍的人身上，客观判据同样普遍越线（本章数据：年长层 82.4% 的夜超过三十分），则这不是一小群人的知觉问题。本章数据支持后者。

9.3 三因素模型（易感—诱发—维持）

它走到哪一步：慢性失眠的经典解释把成因分成易感、诱发、维持三层，并指出真正让失眠慢性化的是维持因素（卧床时间延长、日间补觉、对睡眠的过度关注）。这是行为治疗的整个地基。

分离线：这一支已经把“维持”这个词用掉了，但它指的是跨夜的维持（是什么让这件事持续了三个月）；诊断标准里的“睡眠维持困难”指的是夜内的维持（这一夜醒了之后回不去）。两个“维持”是两个东西，而它们共用一个词。本章只处理后。

判定形式：若把夜内的长断次数作为自变量，能预测跨夜的慢性化，则两个“维持”有实质联系，本章应把它们并起来讨论；若不能，则共用一个词是纯粹的词面巧合，且是一个应当被改掉的措辞。本章数据每人只有两夜，做不了，如实记为未做。

9.4 觉醒指数与睡眠碎片化

它走到哪一步：临床报告里有觉醒指数（每小时觉醒次数），它已经在数中断。

分离线：它数次数，本章关心长度——而本章数据说，清醒总量与最长一次的相关（+0.877）远高于与次数的相关（+0.631）。更要紧的是：报告里同时有清醒总量与觉醒次数，恰恰没有最长那一次的长度，而后者可以从同一份记录里零成本算出。

判定形式：若在一份常规睡眠报告模板里已有“最长清醒段”这一栏，本章的处方多余。就本章所查的常规模板而言，没有。

9.5 过度觉醒模型

它走到哪一步：一支很有分量的解释把失眠看作二十四小时的过度觉醒状态，而不只是夜里的睡眠不足。

分离线：这是一个不折不扣的能力/状态读法，而且它明确把落点放在跨时段的持续状态上。本章与它正面不同：本章说夜内的承重是事件而非状态。两者可以并存——一个人可以有持续的高觉醒，而他那一夜的清醒总量仍然主要由一次长断决定。

判定形式：若过度觉醒的指标能预测长断的长度（而不只是中断的次数），则状态读法覆盖了事件；若它只预测次数不预测长度，则两层各自成立。这一条可做，本章未做。

9.6 救场技艺（戏剧学，本章的工具来源）

分离线：救场处理的是有观众在场、且戏可以继续的错。夜里没有观众，也没有导演；而更要紧的是，救场是一个可执行的动作，而回到睡眠不是——这一点由本系列已发的另一篇处理（睡着是成就动词，没有配套的任务动词）。因此本章不从戏剧学取任何处方，只取它那条判定：完整性由错后的动作决定。

9.7 《易》的占断用法

分离线：本章只取判词的结构（判动作不判属性、没有正常那一格、无咎即善补过），不取任何一条占断用法，不主张任何卦爻与睡眠有对应关系，也不主张时位可以被算出来。凡把本章读作支持以《易》占断睡眠的，都超出本章。

判定形式：把本章第五节里“《易》”三处换成任何一套“判动作不判属性、且无正常态”的判断体系，全章的论证与预测一字不变——若不然，说明本章暗中依赖了占断，那是本章的错。

9.8—9.11 自家四处

（一）站上的失眠专文（/column/insomnia/，同作者）——最险的一处。那一篇讲的是用力的方向：意志一介入就锁死，处方是停止阻止入睡。本章讲的是判定的对象：一次事件被读成一种能力。同案分离：一个完全不用力、上床即睡、却在凌晨三点摊上一次九十分钟长断的人，在那一篇的框架里没有问题，在本章的框架里仍然会被现行判据记成“维持困难”。并且必须承认一处紧张：本章第十一节建议报告“最长一次”，而那一篇会说，多给一个可看的数字就是多给意志一个盯着看的东西。本章接受这条批评，处置见第十一节第三层。

（二）睡眠专栏的判定本体一篇。那一篇讲一个时间单元只能有一个标签、险胜与压胜在账上相同。本章讲一个判据的名字与它的承重不一致。**分离线：**那一篇的问题在标签的分辨率（同一层里分不出强弱），本章的问题在词类（能力名词指着事件）。二者独立：把判余记下来，“维持困难”这个名字照样指错东西。

（三）睡眠专栏的边界本体一篇。那一篇讲睡与醒之间是一片有厚度的地带。本章的长断是入睡之后的整段清醒，与那条边界无关：把界带完全测清楚，长断照样存在。

（四）本系列另几篇的共同全书赌的那句话（记录里少一栏）。本章的第一层主张确实是“报告里没有最长一次这一栏”，属于同一族，本章不主张对这条家族命题有首创权。但本章的第二、三层不在这一族里：第二层是词类错配，第三层是边界定义的稳健性，两者都与“少一栏”正交——把那一栏加上，名字仍然是能力名词，阈值仍然可以移动。

9.12—9.14 三条更远的

（一）个别事件主导的总量在别的领域早有名字（少数客户贡献大部分收入、少数缺陷造成大部分故障）。**分离线：**那些说法处理的是资源分配（把力气放在关键少数上），本章处理的是命名与判定。**判定形式：**若把长断当作“关键少数”去优先干预就能解决问题，本章这一层多余；而第十节第二处说明这条处方本章开不出来。

（二）诊断阈值的任意性这一支（凡有连续量被切成二分的地方都有人批评过）。**分离线：**那一支批评的是阈值的位置（该定在哪儿），本章第八节指出的是被切的那个量本身的定义可以移动——不是线画在哪儿，是量算到哪儿为止。两者独立：把阈值论证得再充分，只要清醒总量的算法有两种合法版本，四分之一的夜仍然会换类。

（三）年龄相关的睡眠变化这一支已有大量文献。**分离线：**那一支的处方通常是“按年龄给不同的参考值”；本章第七节第五小节的用法不同——它是易经那一家的一次兑现，说明同一条线在两个时位上量到的是两样东西，而这与“给不同参考值”不是一回事：给不同参考值仍然假定被量的是同一样东西。

9.15 这些邻居共同站着的那块地

逐个问“它结构性地看不见什么”：

— 诊断标准把判据的名字当作给定，因此看不见名字与承重的错位；一旦承认，措辞就要重写，而重写措辞会牵动处置。

— 觉醒指数把次数当作给定，因此看不见长度；一旦承认，它的分母就不再中立。

— 极值统计把统计量的选择当作问题所在，因此看不见词类；一旦承认，它就不是一个统计问题。

— 阈值批评把被量的量当作给定，因此看不见量的边界可移；一旦承认，所有阈值论证都要先附上一份边界敏感性。

背后是同样被当作给定的东西：一个判据的名字，忠实地指着它所判的东西。

十·三家反过来给本章加的约束

10.1 西医学削掉的：本章不能说现行诊断错了

原想写：诊断不用仪器、只用主诉，是一个错误。

西医学不许，而且理由很硬：取消原发/共病细分，是因为那种细分没有提高诊断准确性，也没有区分出不同的治疗选择。一个不能改变处置的区分，加进诊断标准只会增加成本与误分。本章若要求把长断写进诊断，必须先证明它改变处置，而本章证明不了。

削掉之后剩：本章不主张改诊断标准，只主张在报告里加一栏（第十一节第一层），并把“维持困难”这个措辞的错配指出来。这一处动的是本章的实用价值，是最大的一次损失。

10.2 戏剧学削掉的：本章开不出处方

原想写：既然承重的是那一次，就应当针对那一次做干预——训练“补程”。

戏剧学不许。救场之所以可训练，是因为它是一个可执行的动作，而且台上那个人是醒着的。回到睡眠不是可执行的动作，而长断发生时那个人正处在最不可能执行任何技艺的状态。把救场搬过来会得到一条形式正确而实际有害的建议——那正是本站失眠专文警告过的方向。

削掉之后剩：戏剧学在本章里只提供一条判定（完整性由错后的动作决定），不提供任何处方。本章因此只有一个诊断，没有治疗建议。

10.3 易经学削掉的：本章不能主张时位可以被算出来

原想写：既然同一条线在不同时位上量到两样东西，就应当按时位给出不同的判据。

易经学不许——恰恰是它不许。《易》的整个要点是时位不能被化约成一张对照表；“变通者，趣时者也”说的是趋时，不是查表。把“按年龄分层给参考值”当作对时位的实现，是把一个不可穷举的东西换成了一个可穷举的表格，而那正是《易》所反对的判法。

削掉之后剩：本章只主张一件更弱的事——在报告里同时给出该数与该人群的分布位置，让读的人自己看见这条线在这个位置上切掉了多少。不主张任何分层参考值。这一处动的是价值排序。

十一·能说什么：三层，一层比一层弱，也一层比一层稳

第一层（最稳，不依赖第七节任何一个数）：在报告里加一栏“最长一次清醒段”。这不需要新设备、新算法、新研究，只需要在已有的判读结果上取一个最大值。加完之后，“这一夜醒了九十分钟”这句话就分成了两句：一次九十分钟，还是十八次五分钟。这是本章唯一一条不会塌的建议。

第二层（有证据，但证据是本章的，样本很小）：一夜的清醒主要是一次事件。清醒总量与最长一次的相关+0.877、与次数+0.631；那一次比同k随机过程长1.67倍；前后两次几乎无惯性。这一层依赖一百五十三夜健康人的记录，在失眠患者身上是否成立，本章答不了。

第三层（本章自己的，且是一条警告）：不要把一个能力名词当作它所指之物的描述。“维持困难”听起来像是这个人缺一东西，而它实际操作时用的那个数，主要由一次事件决定。第八节给了一个更近的例子：我自己在算这个数的时候，一个边界之差就让四分之一的夜换了类别。名字很稳，量很不稳。

十二·用本章的判据检本章自己

本章的判据是第六节那五问，其中第五问是：这条判据的名字是能力名词还是事件名词，与它的承重一致吗？把它对准本章。

“长断”是一个事件名词，而本章用它做的事，正在把它变成一个能力名词。只要有人开始比较“谁的长断更长”，长断就成了一个属人的量；再往下一步就是“长断倾向”，然后是“长断易感人群”。本章提出的那一栏，会在被使用的过程中变回它所批评的那种东西，而本章没有任何办法阻止这件事——第一层建议一旦被采纳，第三层警告就开始失效。

第二笔账在第八节那个错误上。我在同一天里，一边指出别人把定义当作事实，一边自己把最后一次醒来之后的时间算进了清醒总量，并据此写下过“绝大多数健康人满足失眠的客观判据”这样的

句子。那句话是错的，它建立在一个我没有核对过的边界上。它被发现，靠的正是本章推荐的那个动作——换一个合法的口径重算一遍，看什么变了。

第三笔账在样本上。一百五十三夜，一个公开数据库，全部是能到实验室过夜的健康人。本章关于失眠说了这么多，而本章一个失眠患者的记录也没有看过。这不是一处局限，这是一处结构：本章所批评的那个判据，其判定对象恰恰是本章取不到的那群人；而本章能取到的这群人，按同一条客观线，年长那一层有 82.4% 的夜越线。两件事合起来，正好是本章第五节那个结论的形状——被判定的东西和判定所用的东西，不是同一样东西。

最后一句留给那件本章没能做成的事。要判定长断到底是事件还是能力，只需要同一个人的多夜记录，看长断的长度在夜与夜之间稳不稳。这样的数据集公开存在。本章可以被引用的最稳的一层，不是那个新名字，也不是第七节那些相关系数，是这一句：在报告里取一个最大值，成本为零。

注释与材料

公开数据为 Sleep-EDF Expanded 数据库 1.0.0 版 sleep-cassette 部分全部一百五十三份判读标注，未做挑选；年龄与熄灯时刻取自随库发布的受试者表。分析窗为熄灯时刻起八小时；入睡后清醒时长按临床口径计算，即入睡点到最后一次入睡之间的清醒，不含最后一次醒来之后的时间——第八节报告了不按此口径计算时的全部差异。预注册文本在取数之后、跑任何统计之前写定，四条预测与撤销规则其后一字未动，三条落空按封存时写下的规则原样写入第七节。解析与统计脚本随文提供，可从原始判读标注起逐步复算。

关于第七节第三小节的零模型：同 k 独立指数模拟，每夜按其实际中断次数各模拟四百次取中位，随机种子写死于脚本。该零模型不是对生理过程的建模，只用于回答一个算术问题——给定这么多次中断，最长的一次本来该占多大比例。

关于三家的引述：西医学部分依据现行国际睡眠障碍分类第三版关于慢性失眠障碍的诊断要件与该学会“明智选择”清单中关于多导睡眠图的建议，以及公开的药物试验入组标准；易经部分依据《系辞》“无咎者，善补过也”“吉凶悔吝者，生乎动者也”“悔吝者，忧虞之象也”与《彖传》论时诸条；戏剧学部分依据剧场救场惯例，未引具体理论著作，如实记为行业惯例而非文献主张。

人机分工声明：三门学科由人指定；文献核对、检验设计、脚本、真跑与初稿由人机协作产出。四条先验预测与撤销规则在拿到任何结果之前封存，其后未动；第八节那处计算错误由作者自查发现，改正前后两套数字一并报告，未作补救性重述。

第十一章 每次都不是同一个地方

每次都不是同一个地方

为何会失眠：一个有大小、而没有位置的量

摘要

关于失眠的解释与办法，几乎全部建立在同一个假定上：有一个位置。是那杯咖啡、是那件心事、是十一点后的屏幕、是凌晨三点的那次醒来。本章用三个领域的最硬主张互相顶撞，指出这个假定本身可能是错的，并给出一份可复算的检验。用 Sleep-EDF 公开数据的七十八名健康受试者、一百五十三夜（每人两夜，预注册在先）复算得到：同一个人两夜之间，深睡百分比的相关是 $+0.82$ ；而“这一夜里最长的那段没有深睡的时间落在哪儿”，两夜之间的相关是 -0.08 。一个人有多少深睡，是他的属性；这些深睡在夜里怎么排，是这一夜的偶然。本章把这类“有大小而没有位置”的量称为**无位置量**，把位置在同一人不同夜之间的可重复程度称为**位重复度**。预注册的五条里，一条触发了撤销规则——把一夜切成十等份逐份检验，没有任何一份比别的份更重要（最强与中位之比 1.21）；“一夜由某个关键点决定”这条因此整体撤回。本章另公开三处自己设计上的缺陷：一处数学耦合、一处零模型选错、一处预注册条款的风险度只有 0.284。判据是一句不含程度词的问话：上一次你说找到原因的那一夜，那个位置和上上次是同一个吗？

关键词：无位置 位重复度 窄口 夹点 跨夜可重复性

It Was Never the Same Place Twice

Why Insomnia Happens: A Quantity with Magnitude but No Location

一·一张清单

一个睡不好的人，手上通常有一张清单。这张清单是真实的，我把它按常见顺序排出来：

咖啡不能过午。晚上不看屏幕。房间要暗、要凉、要安静。不要在床上做别的事。睡前一小时不谈话。别吃太饱，也别饿着。白天要运动，但不能太晚。心事要在睡前写下来。睡不着就起来，别躺着。别看时钟。褪黑素。呼吸法。白噪音。重力被。换床垫。

这张清单有一个共同点，而且它一眼看不出来：**每一条都在指一个位置**。咖啡指向下午三点，屏幕指向十一点，床垫指向身体下面，那次醒来指向凌晨三点。清单不是关于“睡眠”的，是关于“哪儿出了问题”的。

于是也就有了一种很熟悉的经验：每一条都试过，每一条都有**时候管用**。而“有时候管用”这件事，从来没有被当成一个需要解释的现象——它被读成“这条对我不完全适用，还得再找”。

本章要问的是：**如果那个位置每次都不是同一个呢？**

二·三种回答，每一种都只承认自己

对“为何会失眠”，有三个领域各给了一个硬回答。三个都主张只有自己那一条是决定性的，而三条指向的方向互不相同。

认知学：只有预测误差是驱动

按预测加工与主动推断这一路的说法，大脑不是被动接收信号的器官，是一台不断生成预测的机器；感知是预测与输入之间的误差被消化的过程。而降低预测误差只有两条路：**改模型，或者改世界让它符合预测**。行动的唯一理由，就是后者。

按这条主张，睡不着是因为误差降不下来：某个通道上的输入没有被预测到，系统就不能停止采样。**驱动是误差，别的都是误差的表现**。

这一家自己知道一件对它非常不利的事，而且是他们自己提出来的，叫**黑暗室问题**。Friston、Thornton 与 Clark 在二〇一二年的那篇文章里把它摆到桌面上：如果最小化感觉意外就是

全部的驱动力，那么最优策略应当是**坐进一间全黑的屋子里不动，直到死掉**——因为那里的输入完全可预测。生物显然不这么做。

这个自曝对本题格外要紧，因为**睡眠就是那间黑屋子**：一个把外部预测误差降到最低的状态。如果最小化误差是唯一驱动，人应当随时都能睡着，而且睡着之后应当不愿意醒。**黑暗室问题正是”为何会失眠”的镜像。**

化学工程：整个系统由一个点单独决定

化工里有一套叫夹点分析的方法，一九七八年成形，一九八三年由 Linnhoff 与 Hindmarsh 在《化学工程科学》上写成可以徒手操作的设计法。它的核心主张硬得没有余地：

把一个工厂里所有需要加热的物流合成一条曲线、所有需要冷却的物流合成另一条曲线，把两条曲线沿热量轴平移到最小允许温差处，**两条曲线贴得最近的那一点，就是夹点**。夹点决定了这个系统理论上最少需要多少外部加热和多少外部冷却。而后是一条设计禁令：**不许跨夹点传热**。只要不跨，就保证能达到那个理论目标；一旦跨了，加热与冷却的需求会同时上升——你在一个局部看起来完全合理的地方做了一次换热，两头都变差。

按这条主张，失眠是某处的余量最小、而那一处被跨越了。驱动是那个点，别的地方做什么都不改变结果。

这一家的自曝就在方法的地基上：**决定夹点在哪儿的那个最小温差，是一个经济权衡的选择，不是系统里的物理量**。把它从十度改到二十度，夹点会跳到另一条物流上，“离目标还有多远”这个读数整张重画。也就是说，在化工自己最成熟的方法里，**位置是被规定出来的，不是被测出来的**。

体育学：干扰的量是决定性的，而过程中的表现是反指标

运动学习里有一个一九七九年的经典发现，Shea 与 Morgan 做的：让人学三个动作，一组按块练（先把第一个练完再练第二个），一组随机穿插练。按块练的那组在**练习期**表现明显更好；而在隔一段时间的**保持与迁移**测试里，随机练的那组明显更好。后来这被叫做情境干扰效应，说法是：高干扰损害获得、改善保持，低干扰反过来。

按这条主张，决定学没学会的是练习期的干扰量；而**练习期的表现是最终结果的反指标**——过程中越顺，最后越差。

这一家的自曝也很明确：**这个效应在实验室的简单任务上稳，在真实运动技能与初学者身上经常不复现**，其可推广性在这个领域内部被反复讨论了三十多年。

三·三家为什么撞得起来

这三家不是互补的。它们各自说的是”只有我这条是决定性的”，而它们指的方向互不相同。

	它说什么在推动	被推动的是什么	它给的时序读数
认知学	收到的输入与生成的预测长期不容	智能体去改变它所处的环境	误差先升，行动后起，环境最后变
化学工程	换热的匹配次序与温位的约束不容	系统当下的能耗状态改变	夹点先定，网络后设计，公用工程量最后算出来
体育学	练习期做出来的表现与安排的练习顺序不容	学习者可动用的条件改变	干扰先加，表现后降，保持最后升

三家分别把驱动放在三个不同的位置：一家放在”输入与模型之间”，一家放在”次序与约束之间”，一家放在”表现与安排之间”。这不是三种角度，是三支指向不同的箭头。

如果误差是决定性的，那么夹点只是误差最大的地方的一个名字；如果夹点是决定性的，那么干扰只是跨夹点操作的一种；如果干扰量是决定性的，那么误差只是干扰的读数。三条不能同时为真，而它们的分歧也没法靠”判谁对”化解——它们根本不在同一个语域里作答。

四·三家都不说的那一句

把三家的争论压到最紧，它们争的是同一件事：**失眠该由哪一样东西推动**。

认知学说由预测误差推动，化工说由那个余量最小的点推动，体育学说由干扰量推动。三家吵得很凶，而它们有一句话从来没有人说出口，因为三家都觉得那还用说：

决定性的那一样东西，既有大小，也有位置。

误差在某个通道上、某个时刻最大；夹点在某个温度上；干扰在练习序列的某一段里。三家都在找”哪里”。

这句话的后果非常具体。它意味着”找到原因”和”找到位置”是同一件事。它意味着一个正确的解释必须能指着某处说”就是这儿”。它还意味着一件更隐蔽的事：**如果一个东西没有位置，它就不像一个原因。**

把这句话念给三家听，三家都会说这还用说吗。会引起争论的句子不是这一句——这一句不会引起任何争论，而这正是它值得被单独拎出来的理由。

五·推翻它的那件材料，来自化学工程自己

要推翻上面那句话，从外面搬一个理由没有用。从外面搬来的只是第四种主张，三家可以各自退回阵地不理它。

所幸不必从外面搬。**化工自己手上就有那件材料，而且每天在用。**

夹点的位置由最小允许温差决定，而最小允许温差是设计者选的一个数——它由投资与能耗的经济权衡定，不由物料定。把它从十度改到二十度，夹点会跳到另一条物流上；换一个数，“离理论最优还差多少”这个读数整张重画。

也就是说：**在化工自己最成熟、最可靠的那套方法里，那个决定一切的位置，不是系统的属性，是分析者选的一个参数的函数。**

这件材料在化工里回答的是另一个问题——“最小温差该取多大”。它是一个工程经济学问题，出现在每一本教材里。**没有人把它拿来当作”位置可以不是系统的属性”的证据。**

于是三家无法退回阵地：那件材料是化工自己的，而化工恰恰是三家里把”位置”说得最硬的那一家。

六·本章的判断

有了那件材料，可以写出本章的判断。

失眠不是唤醒太高，不是某处的余量被跨越了，也不是干扰的量不对。

>

而是：这个系统里真正稳定的那个量没有位置。一个人有多少深睡，在他自己的两夜之间高度可重复；而这些深睡在夜里排在哪儿，两夜之间几乎毫无关系。大小是他的，位置是这一夜的。于是每一次”我找到原因了”，都会找到一个不同的位置；而每一次之后的改善，几乎全部是减法本身。

三个被否掉的说法，各对应一家。本章不主张它们是错的，也不主张它们不重要。本章主张的是：**它们共享一个形式上的要求——指出位置——而这个要求可能在这件事上无法被满足。**

无量量是本章提出的那一类东西：在个体之间高度稳定、而在同一个体的不同次之间没有固定位置的量。它的定义特征只有一条——它的大小可重复，它的位置不可重复。

这类东西为什么难办？因为一切”找原因”的方法，本质上都是**定位方法**：把时间轴切开，看哪一段异常；把条件列开，看哪一条相关。定位方法在无量量上不是效率低，是**每次都会给出一个答案，而每次的答案不一样。**

这个过程有轮次，而且中途换过一次发起处：

轮次	谁先动	另两样多久跟上	回写了什么
第一轮	一夜睡得差 → 人去定位	总能找到一个”合理的位置”（那杯咖啡、那件事、那个时刻）；下一夜常常真的好一点	“这次找对了”这个经验被记住
第二轮	发起处换成了那条清单自	每一条都”有时管用”，	“还有一个原因没找到”这

	己：清单越长，能找到的位置越多	因为下一夜的改善本来就会发生	个信念被加固
--	-----------------	----------------	--------

第二轮是本章最要紧的一段。它不再需要任何一次真正的失眠来推动。一条足够长的原因清单，加上一个天然存在的”下一夜通常好一点”，就足以自我维持。

而第一轮里那个”下一夜常常真的好一点”，本章第十二节会把它算出来：在这份数据上，它几乎全部是减法本身，不是补偿。

七·位重复度与窄口

要让上面那段话可以被检验，需要两样东西：一个可以定位的对象，和一把量”位置可不可重复”的尺。

窄口

化工的夹点是”两条曲线贴得最近处”。要把它搬到一夜睡眠上，需要一个能从判读记录直接算出来的对应物。本章取的是：

窄口：一夜的睡眠期之内，最长的一段连续“既没有深睡也没有快速眼动”的区间。

理由是这样的：深睡与快速眼动是两种离清醒最远的状态。一段只由清醒、浅睡一期、浅睡二期构成的时间，是这一夜里离清醒最近、余量最小的一段。它是这份记录上最接近”夹点”的东西。

- 窄口长度：**该段的窗数乘以零点五分钟。
- 窄口位置：**该段中点在睡眠期内的相对位置，零表示刚入睡，一表示最后一个睡眠窗。

位重复度

位重复度：同一个人不同两夜之间，某个结构量的位置的相关系数。

它要与另一个数并排着看，才有意义：

量重复度：同一个人不同两夜之间，同一个量的大小的相关系数。

两个数并列写。若大小可重复而位置不可重复，这个量就是本章所说的无位置量。

清点手册

一个读数不写死粒度和编码规则，它就不是读数。以下是本章全篇唯一的口径，正文、证伪条款与赌注三处引用的是同一份。

项	内容
睡眠期	从第一个属于浅睡一期／二期／深睡／快速眼动的窗，到最后一个这样的窗。首尾的清醒与非五分期标注（体动、未定）不计入
窄口	睡眠期内最长的连续“无深睡、无快速眼动”区间； 并列最长时取最早的那一段 （写死，防事后挑）
粒度	判读窗三十秒，故一切长度的分辨率下界是零点五分钟
边界例	若整夜没有任何深睡与快速眼动，窄口即整个睡眠期，位置记为零点五，并单独报数。本数据中这样的夜有 0 个
位重复度	只在同一受试者的两夜之间算，跨人不算；用等级相关，并同时报皮尔逊相关
不适用	只有一夜记录的个体；分期判读体例不同的两夜之间

一句关于命名的话

“余量最小”是本章对窄口的**解释**，不是测量。本章实际测的只是”无深睡无快速眼动的最长连续段”。这两者的距离必须留在读者眼前，因为本章接下来要用这个量得出一个否定性的结论——而一个否定性的结论最容易被”你测的不是那个东西”驳倒。本章承认这一条随时可以被驳，并在第十九节把它写成了一条证伪条件。

八·数据与预注册

数据

Sleep-EDF Expanded 的 sleep-cassette 子集：一百五十三份判读标注，七十八名受试者，每人两夜，年龄二十五至一百零一岁，健康、未服用任何睡眠相关药物。配套的受试者表提供年龄与每夜熄灯时刻。数据在 PhysioNet 公开，任何人可下载复算。

选它的理由有两条。其一，**它是一项老化研究，每人两夜——而”两夜”**正是本章核心检验所需的最小结构；一个人只有一夜，位重复度在原理上算不出来。其二，它的受试者是**健康志愿者，不是失眠患者**。这一点必须写在最前面：**本章因此不能主张关于失眠人群的任何事，只能主张关于这个结构量本身的事**。如果连健康人的位置都不可重复，那么”位置”这个维度本身就需要重新考虑；但这不等于说失眠患者的位置也不可重复。

预注册

统计之前先写死，文件 PREREG.md, SHA-256:

b671800c039b04ff359be951172278a991dedd2a8f6f02b44d446bb1ab23998e

五条判定线与一条撤销规则：

- B1（瓶颈不在两端）：**窄口位置的中位落在 0.25 到 0.75 之间；且位置小于 0.15 或大于 0.85 的夜合计少于 30%。
- B2（单点主导还是均摊）：**把睡眠期切成十等份，逐份算该份内”最长的无深睡无快速眼动段”，各自与整夜深睡百分比求等级相关。计**最强一份的绝对值 ÷ 十份绝对值的中位。大于等于 1.8 判”单点主导”；小于等于 1.3 判”均摊”**。
- B3（跨瓶颈操作）：**药物子集配对，预测两端改善而窄口不改善。
- B4（非单调）：**窄口长度三分位与整夜深睡百分比是否非单调，**且必须先给零分布**。
- B5（两端与瓶颈不相关）：**入睡潜伏期与总睡眠时长同窄口长度的相关，绝对值均小于 0.30。

撤销规则第二条：若 B2 判为”均摊”，则”一夜由某个关键点决定”这条**整体撤回**，本章此后只保留窄口位置的描述，不得主张它决定整夜。

预测风险度

上一次做这类工作时学到一件事：一条预注册的条款如果在零假设下也几乎必然命中，那么它命中时不提供任何证据。所以这一次把风险度也写进了预注册，并在跑真实数据之前算出来：

条款	零假设下的命中率	风险度
B1（打乱同一夜的分期顺序后，位置仍落在非两端）	71.5%	0.284
B2（打乱结局后，比值仍大于等于 1.8）	80.0%	0.201

这两个数都很低，尤其 B1。也就是说：**B1 即使命中，它提供的证据也很有限**。这一条写在结果之前，不写在结果之后。

代理变量声明

判读窗三十秒。本数据采集于一九八七至一九九一年。深睡的判读依赖幅度阈值，而慢波幅度随年龄下降——这意味着老年受试者的深睡量与年轻人的不完全可比，本章在第十七节把它列为不适用边界之一。

九·结果一：瓶颈不在两端，但它也不在中间

一百五十三夜全部可用，无一剔除；整夜没有深睡也没有快速眼动的夜有零个。
窄口位置的分布如下：

读数	值
中位	0.542
四分位	0.250 / 0.789
均值	0.530
最小 / 最大	0.039 / 0.954
位置小于 0.15 或大于 0.85 的夜	26.1% （判定线：小于 30%）

按预注册，B1 命中：中位落在区间内，两端夜低于三成。撤销规则第一条未触发。

但这一条几乎不值钱，而且我事先就知道它不值钱。它的风险度是 0.284——把同一夜的分期顺序完全打乱、只保留各期的构成比例，位置仍有 71.5% 的机会落在非两端。一个七成会自动命中的条款，命中时说明不了什么。

比这条判定更有信息的是分布的形状。把位置切成十等分数一下：

位置区间	0-1	1-2	2-3	3-4	4-5	5-6	6-7	7-8	8-9	9-1
夜数	7	16	25	8	15	11	13	24	20	14

这不是一个集中在中段的分布，也不是一个集中在两端的分布。它铺得相当开，而且有两个鼓包，一个在前三成处，一个在后三成处。

于是第一条结果的正确读法是：窄口不在两端，也不在中间，它到处都是。而”到处都是”这句话，用夹点那套语言说，就是这个系统没有一个稳定的夹点位置。

十·结果二：撤销规则被触发

第二条是本章原本最想要的那一条。把睡眠期切成十等份，逐份算该份内”最长的无深睡无快速眼动段”，看哪一份对整夜深睡百分比的预测最强。若一夜真由某个关键点决定，应当有一份显著强于其余。

第几份	1	2	3	4	5	6	7	8	9	10
等级相关绝对值	.441	.506	.511	.454	.475	.453	.548	.396	.419	.319
该份长度中位（分）	40.0	29.0	32.0	28.0	28.5	30.0	28.5	33.5	33.0	40.0

十份全部显著相关，方向全为负（p 从一千三百亿分之一到十万分之六）。最强的是第七份，0.548；十份的中位是 0.453。

比值 **1.21**。判定线是：大于等于 **1.8** 判单点主导，小于等于 **1.3** 判均摊。

判为均摊。撤销规则第二条触发。

本章原打算主张的”一夜由某个关键点单独决定”这一条，整体撤回。原文照录在此供核对：预注册第 B2 条写的是”若一夜真由某个关键点决定，应当有一份显著强于其余”。实测没有。本章此后不得改用别的位置定义重试，也不得把 **1.21** 说成”接近 **1.3** 所以还有希望”。

夹点这个形状，在这份睡眠数据上不成立。不是没测出来，是十份等价。

十一·结果三：0.82 对 -0.08

撤销之后，本章手上还剩什么？剩下一个问题：如果位置不决定整夜，那么位置本身是个什么东西？

七十八人里有七十五人两夜齐全。对每一个人，把第一夜的读数与第二夜的同一读数放在一起求相关。这是一个跨夜的检验，两个数来自两份不同的记录，不共享任何算术。

读数	同一人两夜之间的等级相关	p
深睡百分比	+0.823	1.4×10^{-19}
觉醒总量	+0.529	1.1×10^{-6}

窄口长度	+0.471	2.0×10^{-5}
总睡眠时长	+0.209	0.072
窄口位置	-0.079	0.502

把最上面一行和最下面一行并排读一遍：

一个人有多少深睡，在他自己的两夜之间相关 **+0.82**——这是一个很强的个人特征，强过觉醒量，强过总睡眠时长。

>

而这些深睡在夜里排在哪儿——用”最长的那段没有深睡的时间落在何处”来度量——两夜之间的相关是 **-0.08**，**p** 为 **0.50**。完全没有。

大小是他的，位置是这一夜的。

这就是本章所说的无量纲：量重复度 **+0.82**，位重复度 **-0.08**。两个数并列写，是本章要的那个读数。

而它顺带解释了第十节那个撤销：十份之所以等价，是因为位置本身在人身上没有固定值。一个每夜都把窄口摆在同一处的人群，会给出一个尖锐的十份剖面；一个位置随机的人群，只能给出一条平的。第十节那条平线，不是本章的测量失败，是第十一节这个 **-0.08** 的必然后果。两条结果互相解释。

十二·结果四：那个”第二夜就好了”，里面有多少是减法

本章第六节说过，那条自我强化回路的第一轮靠一件事维持：去掉一个”原因”之后，下一夜常常真的好一点。这一节把它算出来。

第一夜与第二夜的深睡百分比：

量	值
第一夜方差	116.56
第二夜方差	78.83
两夜协方差	76.38
两夜相关	+0.797
第一夜与”第二夜减第一夜”的相关（按三个矩推导）	-0.570
同上，实测	-0.570

推导值与实测值逐位相同——这不是巧合，一个差值与它自己一端的相关，由两端的方差与协方差三个数完全决定。这件事一九六二年 Oldham 就写清楚了：两串互相独立、标准差相同的随机数，前一串与”后减前”的相关约为 -0.707。

于是可以这样读：

“第一夜差的人，第二夜会好一些”——这句话在这份数据上成立，相关是 **-0.570**。而如果两夜之间毫无关系、纯粹是掷骰子，这个数应当是 **-0.707**。实测的 **-0.570** 比纯随机的情形还要弱。

也就是说：那个”下一夜通常好一点”的经验，在这份数据上几乎全部是减法本身的性质，而不是身体在补偿。两夜之间的正相关高达 +0.797，恰恰说明补偿很弱——一个人差就一直差，好就一直好。

这一条对本章第六节那条回路是致命的支持，也是对”我找到原因了”这句话最直接的一次拆解。你去掉了一个位置上的东西，第二夜好了一点。这件事在一个位置随机、大小稳定的系统里，本来就会发生。

十三·我自己设计里的三处缺陷

上面四节里有三处是我自己的问题，写在这里，不写在脚注。

其一，B2 与 B4 是耦合的。窄口长度这个自变量，与整夜深睡百分比这个结局，来自同一份分期序列——一段”没有深睡”的时间越长，整夜深睡当然越少。这不是发现，是定义蕴含的一部分。具体有多耦合，可以量出来：

相关	值
窄口长度 与 睡眠期总长	+0.757
窄口长度 与 深睡总分钟	-0.421
窄口占睡眠期的比例（中位）	0.217（四分位 0.165 / 0.386，最大 0.762）
窄口占比 与 整夜深睡百分比	-0.424

窄口长度有四分之三以上是”这一夜有多长”。用它去预测整夜深睡，本来就该显著。B2 那张十份表里”十份全部显著”这件事，因此不构成任何发现——它只说明我的自变量和结局是同一件事的两种写法。

要紧的是：这一处耦合不影响第十节的结论。因为十份共享同一种耦合，它们之间的相对强度仍然可比，而比较正是 B2 做的事。耦合抬高了十个数的绝对水平，没有理由抬高其中某一个。

其二，我给 B2 选的零模型是错的。我预注册时打算用置换检验：把结局随机打乱，看比值分布。跑出来是这样的——

	值
实测比值	1.208
零分布中位	2.411
零分布 90% 区间	1.485 – 5.212
实测所在百分位	0.2%

看上去很漂亮：实测比零分布低得多，好像”十份显著地均匀”。但这个对照不成立。打乱结局之后，十个相关系数变成十团纯噪声，噪声之间的最大与中位之比自然很大；而实测的十个数都真实地相关、都远离零，它们之间的比值自然小。这个检验比较的是”有信号”与”没信号”，不是”集中”与”均摊”。

所以那个 0.2% 不能用。第十节的结论只能靠一件朴素的事情支撑：最强的一份（0.548）只比最弱的一份（0.319）高 72%，而”由一个点决定”要求的是量级上的差别。这句话没有 p 值，但它是诚实的。

其三，B5 我第一次算错了口径。我最初用”记录起点到第一个睡眠窗”当入睡潜伏期，得到与窄口长度 -0.526 的强相关。而这份数据的记录从白天就开始，那个量主要反映”记录开了多久才上床”，不是入睡潜伏期。按受试者表里的熄灯时刻重算（一百零四夜可算），结果完全不同：

相关	错误口径	熄灯口径
入睡潜伏期 与 窄口长度	-0.526	+0.237 (p=0.016)
总睡眠时长 与 窄口长度	—	+0.006 (p=0.948)

熄灯口径下的入睡潜伏期中位是 10.2 分钟（四分位 5.0 / 22.1）。B5 在正确口径下命中：两端与窄口长度的相关绝对值都小于 0.30，其中总睡眠时长那一条几乎正好是零。

这一处更正逼出一个附带的结论，比 B5 本身有用：总睡眠时长与”这一夜最长的那段没有深睡的时间”相关 +0.006。睡得久不代表那段短，睡得短不代表那段长。报告上最常用的那个数，与夜里这段结构毫无关系。

十四·结果五：非单调不成立

体育学那一家给了本章一条最想要的预测：过程中的表现是最终结果的反指标，因此窄口长度与整夜质量的关系可能是非单调的——太短不好，太长也不好。

按窄口长度分三组：

组	窄口长度	n	整夜深睡百分比中位
短	< 83.2 分	51	15.53
中	83.2 – 151.2 分	51	6.68

长	> 151.2 分	51	1.81
---	-----------	----	------

单调下降，而且幅度很大——从 15.53% 掉到 1.81%，将近八倍。没有任何非单调的迹象。

按预注册，本条还要报零分布：把结局随机打乱五千次，“非单调”在零假设下的出现率是 66.0%。预注册写的是“若零假设下出现率大于 60%，本条不作证据”。这里不必动用这条豁免，因为实测根本不是非单调。

情境干扰那条在这里没有兑现。★但要说清楚它为什么没兑现，而不是简单地把它划掉：这三组的差别，有很大一部分正是第十三节那处耦合——窄口长而深睡少，本来就是同一句话。这份数据在原理上无法检验体育学那条预测，它需要的是“同一夜的深睡总量固定、只改变它的排布”这样一个对照，而分期记录给不出。这一条列进第十九节的证伪表。

十五·最强的三个竞争解释

其一：其实不就是深睡少的人窄口自然就长吗。这是最省事也最难反驳的一条，第十三节已经承认它对 B2 与 B4 成立。但它说明不了第十一节：位置的跨夜相关是 -0.08，而深睡量的跨夜相关是 +0.82。如果位置只是深睡量的影子，位置的跨夜相关应当跟着深睡量走，至少是正的、可观的。它不是。分岔点很清楚：若在更大样本里位置的跨夜相关升到 +0.3 以上，本章的核心读法被推翻。

其二：其实不就是位置这个量测得太糙吗。窄口位置是一个由“最长段”定义的量，最长段一旦有并列，位置会跳。这条也有力。但它有一个可检的推论：如果位置只是噪声，那么它不应该与任何东西相关。实测中它与入睡潜伏期的相关是 -0.107，与整夜深睡量的相关也不强——这一点确实与“位置基本是噪声”相容。本章接受这条解释与本章的解释在当前数据上不可分，并把区分它们的设计写进第十九节：需要同一个人的多夜记录（三夜以上）与一个更稳的位置定义（例如深睡的时间重心，而不是最长空段的中点）。

其三：其实不就是这批人是健康人吗。完全成立，本章不辩解。健康人的睡眠结构本来就更规整，位置的随机性可能正是“没有病”的表现；而失眠患者也许恰恰有一个稳定的位置（例如每夜三点）。这是本章最愿意被推翻的一条，也是最容易做的一次检验——任何一份带两夜以上记录的失眠队列都能算。本章的预测是：失眠患者的位重复度也不会显著高于健康人。若显著更高，本章的主张应当缩到“健康人”这个范围之内。

按机制证伪的格式写：

控制个体的深睡总量之后，若窄口位置在同一人多夜之间的相关显著大于零（大于 +0.3），则“位置不是这个人的属性”这一条为伪。

十六·三家反过来给本章加三条限制

认知学砍掉的：本章不得主张“没有位置就没有机制”。黑暗室问题的解决方案是：生物最小化的不是任意模型下的意外，而是它自己那个生成模型下的意外，而那个模型里写着“要吃饭、要见人、要动”。这意味着位置可以不在时间轴上，而在模型里。本章测的是时间轴上的位置，测不到模型里的。没有这一削，本章原本会写下“失眠没有原因”这句更响亮的话。那句话是错的。

化学工程砍掉的：本章不得把“没有夹点”读成“没有约束”。夹点分析里还有一件东西，本章一直没提：热复合曲线与冷复合曲线的形状本身，也就是“这个系统里各种温位上有多少热量”。夹点只是两条曲线的最小间距处；曲线本身携带了全部约束。对应到睡眠：位置不可重复，不意味着没有结构；本章的 +0.82 恰恰说明结构非常稳。本章因此只主张“位置不是约束的所在”，不主张“约束不存在”。

体育学砍掉的，是三条里最狠的一条：本章不得主张位置的随机性是好事。情境干扰效应的要点是高干扰改善保持——一个自然的推论是“位置随机也许正是有用的”。但那条效应在真实技能与初学者身上经常不复现，这个领域自己承认可推广性有限。而本章第十四节的数据不但没有支持非单调，还给出了一条陡峭的单调下降。学科的限制与数据的结果指向同一处：本章对“位置随机是好是坏”这个问题，没有任何发言权。

十七·不适用的边界

其一，本章的数据是健康人。本章不能主张关于失眠患者的任何事。第十五节第三条已经把这写成了一条待检条件。

其二，深睡的判读依赖幅度阈值，而慢波幅度随年龄下降。这批人年龄跨二十五到一百零一岁，老年受试者的深睡量与年轻人不完全可比。第十一节那个 +0.82 是跨全年龄段算的，因此它有一部分来自年龄这个稳定的个体差异——这不削弱“大小可重复”，但它提醒：可重复不等于内生。

其三，两夜不足以估计位置的可重复性。两夜之间的相关是位重复度最粗的一种估计。正确的做法是多夜的组内相关系数。本章的 -0.08 应当被读作“没有看到”，不是“证明为零”。

其四，本章只测了一种位置。“最长的无深睡无快速眼动段落在哪儿”是位置的一种写法，不是唯一一种。第十五节第二条已经承认，一个更稳的定义（例如深睡的时间重心）可能给出不同的结果。

其五，本章不主张那张清单上的条目是错的。咖啡、屏幕、床上不做别的事，都有各自的证据基础，其中一部分证据相当硬。本章主张的只有一件事：“每一条都有时管用”这个经验，需要另一种解释，而现有的解释是“还没找到那一条”。

十八·与最近的几种说法的分界

以下十二条，每一条都写明在什么观察下本章对而该说法错、在什么观察下相反。只能写出“本章更强调”“视角不同”的，本章视同承认被那一条覆盖。

第一，特质与状态的区分（Cattell 与 Scheier 一九六一年；Spielberger 的状态—特质焦虑量表，一九七〇年）。这是本章“大小可重复、位置不可重复”这句话最近的祖宗，比本章早五十五年，而本章初稿完全没有想到它。写在第一位，是因为漏掉它这件事本身就是本章第六节讲的那个毛病：它不冷门，它是人格心理学的标准章节，只是登记在“焦虑的测量”名下。

分离线是可判定的：特质—状态区分的是同一个量的两个成分——你的焦虑有多少是稳定的、多少是此刻的，两者相加等于观察值。本章说的是同一个现象的两个不同属性——大小与位置，它们不相加，而且位置这一侧没有特质成分（-0.08）。判定形式：若能把窄口位置分解成“稳定成分+当下成分”并测出一个非零的稳定成分，则本章这一格退化为状态—特质模型的一个应用；若稳定成分反复测出为零，则那套分解在这里不适用。

第二，失眠的三因素模型（Spielman 一九八七年：易感因素、诱发因素、维持因素）。这是失眠领域四十年的标准框架，也是那张清单的理论来源。它把原因分成三层，每一层都是一个可以指认的东西。

分离线：三因素模型处理的是因素的种类，本章处理的是位置的可重复性，二者正交而不是竞争。但有一处真正的紧张，值得写出来：三因素模型的临床用法是“找出你的维持因素并去掉它”，而本章第十二节说明，去掉任何一个维持因素之后的改善，与减法本身难以分开。判定形式：若某个维持因素被去掉之后的改善，显著大于同一人两夜之间的自然回落（本数据里那个 -0.570 与 -0.707 之间的差），则该因素是真的；若不能分开，则那次改善不构成证据。这一条不否定三因素模型，它给它加了一个对照组要求。

第三，回归到均值（高尔顿一八八六年）与数学耦合（Oldham 一九六二年；Archie 一九八一年）。本章第十二节那一击的正主。高尔顿那篇比本章早一百四十年。

分离线：他们说的是一条统计学禁令（不要把改变量与初始值做相关，那个负相关有一大截是算术）；本章说的是一条临床归因的后果（一个人根据“第二夜好了”得出“我找对了原因”，而那个改善在算术上本来就会出现）。判定形式：完全遵守那条统计学禁令的人——从不做那种相关——本章所说的归因错误仍然原样存在，因为它发生在个人经验里而不是论文里。反过来，本章要求的只是一个对照：把“去掉某因素后的改善”与“什么都不做的两夜之间的自然变化”并排放。

第四，约束理论（Goldratt 一九八四年）与最小因子律（李比希一八四〇年）。“系统的产出由瓶颈单独决定”“作物由最缺的那种养分决定”——这是“由一个点决定”这个形状最有名的两位，李比希在一八四〇年。

分离线：这两位说的瓶颈是可指认、可稳定定位的——那台机器、那种养分。本章的第十节恰恰是这个形状没有兑现的一次记录。判定形式：若在任何一份睡眠数据上，十等份检验给出一份显著强于其余（比值大于等于 1.8），则约束理论的形状在睡眠上成立而本章这一节错。本章自己那次是 1.21。★顺带说一句：夹点与约束理论的瓶颈不是一回事——约束理论的瓶颈是产能最低的资源，夹点是推动力最小的位置，而那里的设备可能完全正常。这个区别本章用了很久才想明白，写在这里省得读者也绕一圈。

第五，第一个睡眠周期的特殊地位（慢波活动在第一个周期最高，为整夜之冠；纺锤波密度在第一个周期预测失眠症状的增加）。这是睡眠领域内部最接近“夜里有一个关键位置”的一支，而且它有实证。

分离线，而且它反过来加固本章：那一支指的位置是”第一个”——也就是夜的开头，仍然是一端。本章第一节说清单上的每一条都指向两端之一，这一支是研究者版本的同一件事。判定形式：若把”第一个周期”之外的某个中段周期也做成同样强的预测量，本章这一格弱化；若几十年里能被做成预测量的始终是第一个周期与晨醒前，则”读数都在两端”这条描述成立。

第六，睡眠碎片化与觉醒指数。这一支度量”一夜被打断了多少次”，是现行报告里最接近”结构”的读数。

分离线：它数次数，不问位置。一个人有十次觉醒，这十次是挤在一段里还是均匀散开，觉醒指数完全一样。判定形式：取两夜觉醒次数相同而分布不同的记录，按觉醒指数判为等同，按本章的位置读数判为不同。这条差别可以在任何一份现有数据上当场演示。

第七，睡眠状态错觉与主客观不一致。那一支处理”病人报告的与仪器测的不一样”，把不一致当作一种症状。分离线：那一支比较的是两个观察者，本章比较的是同一个观察者的两个夜晚。判定形式：一个主客观完全一致的人，其位重复度仍然可以是零；反之亦然。二者正交。

第八，过度唤醒模型。失眠的主流生理解释：这些人的中枢与自主神经在二十四小时里都处于更高的激活水平。分离线：过度唤醒是一个大小，不是一个位置——这一支恰恰是本章所说的”有量无位”的一个例子，而它没有把这一点说出来。判定形式：若过度唤醒指标能预测窄口位置，则本章这一格被它吸收；若只能预测大小，则它与本章说的是同一件事的两半。

第九，刺激控制与睡眠限制。这两套是失眠认知行为治疗里证据最硬的成分，而且它们有一个共同点：**都不指位置——刺激控制说的是”床与清醒别绑在一起”，睡眠限制说的是”把在床时间压到接近实际睡眠时间”。**分离线：它们改的是规则，不是某个时刻。★这一条对本章是正面的：**证据最硬的两个成分，恰恰是清单上唯二不指向位置的两条。**判定形式：若把这两套改写成指向具体时刻的版本（“十一点必须上床”）而效果不变，则本章这条观察没有内容。

第十，个体内变异与组内相关系数（费雪一九二五年一路）。位重复度这个读数在方法学上就是一个跨次的相关，它的正主是组内相关系数。本章不主张对它有任何首创权。分离线只在应用上：这类系数通常被用来评估测量工具不可靠（低了就说仪器不行），本章用它来问被测的那样东西有没有位置（低了可能是那个东西本来就没位置）。判定形式：若能证明窄口位置的低相关全部来自测量误差，则本章的读法错，正确的读法是”这个量测不准”。这两种读法在当前数据上不可分，第十五节第二条已承认。

第十一，前一晚睡眠预测后一晚（睡眠的夜间变异研究）。那一支已经在做跨夜相关，而且报过总睡眠时长的跨夜相关不高。分离线：那一支比较的是同一个读数在两夜之间稳不稳，目的是决定”研究要记几夜才够”；本章比较的是两个不同性质的读数（大小与位置）的稳定性差异。判定形式：那一支的结论是”要多记几夜”；本章的结论是”再多记几夜，位置也不会变稳”。这两条可以在多夜数据上直接分开。

第十二，第一夜效应。陌生环境的第一夜睡眠更差，这是睡眠实验室四十年的常识，而本数据每人正好两夜。分离线与一处必须承认的风险：**本章第十一节的跨夜相关，可能被第一夜效应污染——**如果第一夜系统性地更差，两夜的差别里有一部分不是自然波动。本章查过：第一夜与第二夜的深睡百分比方差分别是 116.56 与 78.83，第二夜确实更集中。**这一点削弱第十二节那个 -0.570 的解释**（两夜不同方差，减法的性质随之改变），但它不影响第十一节那个 -0.08，因为一个系统性的第一夜效应会同时移动两夜的位置，不会把位置的相关压到零。

十九·与同一批另外几篇的分界

本章属于同一批以睡眠为题的系列，作者相同。同一批里的兄弟篇是彼此最近的邻居，而它们通常互不引用。下面逐篇给出判定形式。

《失眠：为什么越想睡越睡不着》（同作者，站内早前一篇）——**本章最险的一处。**那篇讲努力悖论：你派出一个守夜人去监督入睡，而入睡要求意识退场；四步锁死；不是学会入睡，是停止阻止入睡。

两篇的分野是干净的，而且**方向不同**。那篇诊断的是**动作的性质**——监督这个动作本身与入睡不相容；本章诊断的是**归因的位置**——每次找到的位置都不一样。判定形式：**一个完全不监督自己入睡、也不焦虑的人**，按那篇预测他睡得好；本章对他不作预测，但本章预测他的**窄口位置在两夜之间同样不可重复**。这条可以在同一批人身上分开。

必须承认一处紧张，而且它对本章不利：**本章的处方（别追位置）与那篇的处方（别用力）在实践上会被读成同一句“别管它”**。两篇给的理由完全不同，而读者只会记住结论。本章第二十一节把它写进了反噬。

《只要有一条路是通的》（第六章）（交通学×气候学×逻辑学）。那篇讲一夜之内的并行入睡通路，与暗封——某种助眠办法在开一条的同时封另一条，而封这件事不产生读数；它要的是给每个人一个噪声底。

分离线在**维度上**：那篇处理的是**有没有通路（存在）**，本章处理的是**位置可不可重复（定位）**。判定形式有两条，互不蕴含。**给每个人在基线期多记一夜、做出个体噪声底 ⇒ 补那一篇，不补本章**：有了噪声底，位置这个维度仍然没有被记录，而本章的 -0.08 说明它在原理上不可预测。**在报告里加一栏“这一夜的窄口位置” ⇒ 补本章，不补那一篇**：记了位置，个体仍然没有噪声底。

★顺带一句，那篇引的布雷斯悖论与本章第十节那次撤销方向一致：**加一条路可以让所有人变慢，与“没有一个点是关键点”，说的都是“局部改善与整体后果之间没有简单的传导”**。两篇在这一点上互相加固，不互相覆盖。

《夜里那一端，一直没有动》（第七章）（年老学×舞蹈学×结构学）。那篇讲一个差值读数的变异归哪一端，并提出端权对。本章提出的是量重复度与位重复度，也是一对数。**两对数必须切开**。端权对说的是一个**差值的两端**——它由两端的方差与协方差三个矩算出，是一次算术；量位对说的是**同一个量的两个属性在不同次之间的稳定性**——它需要至少两次观察，不是算术而是重复测量。判定形式：端权对可以在单次横断数据上算出，位重复度不能；位重复度可以在没有任何差值读数的地方算出，端权对不能。

★而本章第十二节直接用了那篇建立的那套算术（差值与它自己一端的相关由三个矩决定），并且用它拆了本章自己的一条支撑。这是本系列里第一次出现一篇的工具被下一篇拿去拆自己。

《下去的那一段，没有人记》（第八章）《三样东西，一个分数》（第五章）《办法都在夜里》（第十四章）。三篇分别落在刻度、验收分辨率、干预点的配置。判定形式：**把判读刻度做细 ⇒ 补第一篇，不补本章**（刻度再细，位置照样不可重复）；**把验收工具做细 ⇒ 补第二篇，不补本章**；**给每条建议标上滞后距 ⇒ 补第三篇，不补本章**（标了滞后距，位置仍然没有一栏）。

本批的四篇”是什么”与另两篇”为什么”。前者的缺口都落在同一份记录内部的某个字段；后者分别落在维持与进入的分别（留驻）、以及转换的分拣。本章的缺口既不在字段也不在转换，在**位置这个维度整个不存在**。判定形式：给记录增设任何一个新字段 ⇒ 补那几篇；而本章要的不是一个字段，是**同一个人的第二夜**。

二十·证伪条件

以下八条，任一成立本章即为伪或需大幅修订。第二条已经跑过一次，且判本章错，仍留在表上。

1. 控制个体的深睡总量之后，若窄口位置在同一人多夜之间的相关显著大于 $+0.3$ ，则”位置不是这个人的属性”为伪。本章这次两夜的实测是 -0.079 ($p=0.502$)，但两夜是最粗的估计。
2. 已跑，本章输。若把睡眠期切成十等份，某一份对整夜的预测显著强于其余（最强与中位之比大于等于 1.8 ），则”没有关键点”为伪。实测 1.21 ，判为均摊，“一夜由某个关键点决定”已按撤销规则整体撤回。
3. 在失眠患者的多夜记录上重算位重复度。若显著高于健康人，本章的适用范围缩到健康人。本章预测不会显著更高——这是本章最愿意被推翻的一条，也是最便宜的一次检验。
4. 换一个位置定义重测。用深睡的时间重心、或第一次出现深睡的时刻，替代”最长空段的中点”。若任一更稳的定义给出跨夜相关大于 $+0.3$ ，则本章测的是一个坏定义，而不是一个无位置。
5. 若能证明窄口位置的低相关全部来自测量误差（例如把判读窗细化到五秒后相关显著上升），则本章的读法错，正确的读法是”这个量测不准”。
6. 本章第十四节那条无法检验的预测：需要一个”深睡总量固定、只改变排布”的对照。若在这样的设计下窄口长度与整夜质量呈非单调，则体育学那条在睡眠上成立，本章第十四节的否定结论作废。
7. 若某个被去掉的维持因素带来的改善，显著大于同一人两夜之间的自然回落，则本章第十二节对”我找到原因了”的拆解不适用于该因素。注意这一条是本章最希望有人做成的：它会把清单上的某几条从”有时管用”变成”真的管用”。
8. 若在一份 n 大于两百、每人三夜以上的公开数据里，位重复度的组内相关系数显著大于零，本章的核心读数应当撤回。

样本纪律：第三、七、八条都涉及跨案例比较。本章要求这些检验在同一制度环境内的同质单元上做——同一家睡眠中心的多批患者、同一套判读体例。用两个国家或两个队列的对比来做这些检验，本章不承认其结论，无论结果对本章有利还是不利。

一条赌注

赌注写条件不写日期，因为本章要的那件事随时可以做。

本章赌：在二〇三〇年十二月三十一日之前，至少有一份面向临床的睡眠报告模板，在结构读数里加入一个位置类的字段，并要求至少两夜。这里说的位置不是”第几点醒的”这类事件时刻，而是一个描述”某个结构量落在夜的哪个相对位置”的量。

四条不算命中：

- 只报告事件时刻（几点入睡、几点醒）的，不算——那是时钟，不是相对位置。
- 只在研究用的分析软件里能算而不进报告模板的，不算。
- 由本章作者或其所在机构推动写入的，不算。
- 只记一夜的，不算——位置的可重复性在一夜上没有定义。

二十一·反噬预言

第一层：本章最容易被读成”别管了”。一个人读完本章，最省事的实现方式是停止一切尝试。而本章第十八节第九条恰恰指出，证据最硬的那两套办法（刺激控制与睡眠限制）都还在，它们只是不指位置。把”别追位置”读成”别做任何事”，是本章能造成的最直接的伤害。

第二层，更麻烦：位重复度一旦进入评估，最省事的做法是把它做高。一个诊所若被要求报告位重复度，它会发现”让位置看起来更可重复”的最简单办法是改定义——换一个粒度更粗、更容易稳定的位置定义。位置这个量对定义极其敏感，第二十条第四款说的就是这件事。一个对定义敏感的量进了考核，被优化的一定是定义。

第三层：本章自己就是那张清单的第十七条。一个失眠的人读完本章，很可能把“我在追一个随机的位置”当成新的那一条——它照样是一个解释，照样会在下一夜带来一点改善（第十二节那个数说明那点改善本来就会来），照样会加强“我找到了”的感觉。**本章因此有可能延长它所描述的那条回路。**

我看不到出口。一个关于“别把改善当证据”的说法，本身会因为读者感到的改善而被相信。本章能做的只有把这句话写在这里，并在第二十节第七条里把那个对照要求写死。

二十二·自反

本章的判据必须能用来判本章自己。用一次，结论不好看。

第一件。本章提出“位置不可重复”，而本章自己只用了**两夜**。两夜是位置可重复性最粗的估计，一个真正的位重复度需要多夜与组内相关系数。**我用一个我自己承认不够的方法，得出了一个否定性的结论——而否定性结论恰恰是最需要功效的那一种。**

第二件。本章第十节撤回了“关键点”，靠的是一个比值 1.21；而我为这个比值准备的零模型是错的（第十三节第二条）。**撤回的结论我保留，因为撤回是往严的一侧走；但支撑它的那个统计学装置我没有。**剩下的只是一句朴素的话：最强的一份只比最弱的一份高七成。

第三件，最要紧的一件。本章第十八节第一条那位祖宗——特质与状态的区分，一九六一年——本章初稿一个字也没提。它不冷门，它是人格心理学的标准章节。它躲过检索，不是因为它藏得深，是因为我在“睡眠”“位置”“可重复性”这几个词里找它，而它登记在“焦虑的测量”名下。

而这正是本章自己描述的那个毛病的一个实例，只是换了个方向。本章说：人们在找位置，而每次找到的位置都不一样。我说的是：我在找祖宗，而我只在自己熟悉的那几个位置找。**两件事的形状是一样的——定位方法总会给出一个答案，而答案取决于你从哪儿开始找。**

第四件。本章的标题是一个断言。“每次都不是同一个地方”——严格地说，这份数据支持的是“在七十五个健康人的两夜之间，没有看到位置的重复”。**没有看到，与每次都不同，是两件事。**标题保留，因为它是这道题的形状；但读者应当按第十一节读，不按标题读。

二十三·如果要用这两个读数，先接受三条限制

量重复度与位重复度是可以拿去评估人和机构的读数。凡是这种形状的东西，都必须配限制，否则不该发表。

- 1. 这两个读数指的是量，不是人。**位重复度低说的是“这个结构量在这个人身上没有稳定位置”，不是“这个人的睡眠不规律”，更不是“这个人不配合”。任何把它写进依从性评估的用法，都不在本章的授权范围内。
- 2. 不得为了把位重复度做好看而更换定义。**任何一次定义更换，必须与旧定义的结果并排公布。第二十一节第二层说的正是这条被违反之后会发生什么。
- 3. 已经按这两个读数对任何人作出不利安排的，必须公开完整的清点手册并给出复核通道。**清点手册指第七节那份的完整版本——睡眠期怎么切、窄口怎么定、并列怎么办、粒度多少。不公开这几样，读数就不可复核；不可复核的读数不得用于对人的判定。

三条缺一，本章认为这份稿子不该被用于任何评估场景。

二十四·三层可以分开引用的主张

第一层（最强，最易倒）：失眠的解释与办法建立在“有一个位置”这个假定上，而这个假定在睡眠的结构量上可能无法满足。这一层是一个概念主张，本章的数据只能对它提供一个方向上的支持，不能证实它。

第二层（中等）：量重复度与位重复度是一对可以分开算的量，而它们可以差得很远。本章这一次是 +0.82 与 -0.08。这一层由第十一节那张跨夜表支撑，不依赖任何因果主张，也不受第十三节那处耦合影响——跨夜相关的两个数来自两份不同的记录，不共享任何算术。

第三层（最稳，也最便宜）：现行睡眠报告的全部结构性读数——入睡潜伏期、总时长、效率、各期百分比、觉醒指数——没有一个指向一夜之内的某个相对位置。这是一条可核的格式事实。前两层全部倒掉，这一层照样成立。而它要的东西极便宜：在已有的判读记录上加一列，不需要新设备、新被试、新导联。

引用纪律两条。其一，若只采用位重复度作为流程变量，不得顺带声称本章发现了失眠的机制。第一层与第二层可以分开取。其二，不得把低位重复度读成“这个人的睡眠有问题”。本章这批人是健康志愿者，他们的位重复度就是 -0.08。

二十五·注释与材料

数据。 Sleep-EDF Database Expanded, PhysioNet, sleep-cassette 子集全部一百五十三份判读标注，与 SC-subjects.xls。Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberyé, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194. 数据集经 PhysioNet 分发 (Goldberger et al., 2000, *Circulation* 101(23):e215–e220)。

预注册。 文件 PREREG.md, SHA-256 为 b671800c039b04ff359be951172278a991dedd2a8f6f02b44d446bb1ab23998e，写于任何统计量被计算之前，含五条判定线、一条撤销规则、以及跑数据前算出的预测风险度。

脚本。 run.py 执行预注册的 B1、B2、B4、B5 与风险度；run2.py 为补跑，含 B5 的口径更正、跨夜检验、耦合摊开与置换零分布。两个脚本与预注册文件一并留档。

预注册与探索性的分界。 第九、十、十四节为预注册内容；**第十一节（跨夜检验）、第十二节（回归到均值的对照）、第十三节（三处缺陷）全部是探索性分析。**本章的探索性分析里，第十二、十三两节全部用于拆本章自己；第十一节是唯一一处用于建立正面主张的探索性结果，因此它必须被当作待检验的观察，而不是结论——第二十节第一、四、八条都是为它准备的。

B3 未执行。 预注册的 B3 要用药物子集做配对检验（两端改善而窄口不改善）。本章没有做，理由写在这里：该子集的记录窗口口径会改变若干读数的显著性判读，而本章在预注册里只写了“熄灯起八小时”这一种窗，在只有一种窗的情况下报告一个配对结果，会重复本系列此前踩过的一次口径事故。B3 因此列为待办，不作为本章的证据，也不计入命中率。

口径一致性自查。 窄口与位置的定义（第七节）、证伪条款（第二十节）、赌注（第二十节末）三处引用的是同一份清点手册。跨夜相关的算法在第十一、十二、二十节三处一致。

参考文献

字数与版本。 本稿约一万八千汉字，二十五节，真表十四张。版本一，二〇二六年八月九日。

人机分工声明。 本章的问题、三个领域的选定与全部判断由作者提出并负责。人工智能参与了文献检索的辅助、公开数据的下载与复算、脚本编写、结构检查与语言编辑。第九至十四节的全部数字由脚本计算，任何人可按上述哈希与文件名复核。本章的所有错误由作者承担。

二十六·这一稿撤回了什么

一篇文章应当把自己撤掉的东西列在最后，供人核对。

1. “一夜由某个关键点单独决定”——整体撤回。依据是预注册第 B2 条与它的撤销规则，实测比值 1.21，判为均摊。见第十节。
2. “窄口长度与整夜质量呈非单调关系”——不成立。实测单调下降，15.53% / 6.68% / 1.81%。见第十四节。而本章同时承认，这份数据在原理上无法检验这条预测。
3. B2 与 B4 的置换零分布——撤回，不作证据。那个 0.2% 的百分位比较的是“有信号”与“没信号”，不是“集中”与“均摊”。见第十三节第二条。
4. B5 的第一次计算——作废，口径错误。错误口径给出 -0.526，正确口径给出 +0.237 与 +0.006。见第十三节第三条。
5. B3——未执行，理由见第二十五节，不计入命中率。

6. “失眠没有原因”这句更响亮的话——不写。认知学那一削（位置可以在模型里而不在时间轴上）挡住了它。见第十六节第一条。

编四·每笔账都得记在一个人名下

主体层

这一编处理最粗的那一层，也是最容易被当成“不属于医学”的那一层：账记在谁名下。

前三编处理的都是一个人的记录内部。这一编问的是记录的边界画在哪里——为什么一份睡眠记录只关于躺在床上的那一个人。

第十二章指出：睡着的时候，“下一拍由谁接住”这件事从没有进过任何字段；于是身体已经内转、而回应仍由睡者承担的那一类夜，与真正被接住的夜，在报告上完全一样。第十三章指出：一个人的睡眠可以长期依赖另一个人的例行安排，两侧都不知道；那个人换了班、搬了家、店关了门，支承撤走，两侧都不产生任何记录——**没有人违约**。第十四章指出：所有办法都压在睡前那两三个小时，而这条分布从来没有被要求给出理由；它可能只反映症状可见的位置，不反映因果强度的位置。

这一编在第十七章的验证表上最吃亏：三章的读数有两个在睡眠判读记录上根本算不出来。而算不出来的理由，逐条就是这两章的论点——记录里没有那一栏。这不是验证的失败，是命题在验证的记录上原样发作了一次。

第十二章 谁替睡者接住下一拍

何谓睡眠：谁替睡者接住下一拍？

社会学、化学与音乐学共同揭示的“托拍事件”

作者：王德生 | v2.0（打磨稿） | 2026年8月9日

睡眠不是世界没有下一拍，而是下一拍在睡者不回答时仍被可靠地接住。

提要

睡眠通常被还原为三种东西：行为停止、神经化学修复，或获得暂不参与社会的豁免。三种解释各自抓住真实，却都把睡眠安放在一个单独账本里。本章以社会学、睡眠化学与音乐学为材料，提出：睡眠不是行为静止，不是某一种分子清除，也不是一纸社会许可，而是一种“托拍事件”——外界对下一次回应的要求被可靠地转交或延期，身体由对外应答轨道转入可逆的内稳态重组，同时保留一条由高显著信号触发的返还通道。三环闭合，睡眠才作为完整事件发生。由此可辨认一类现有睡眠账本没有字段的存在：“悬答眠”，即身体已经入睡，下一次必须作出的回应却无人接手。它可以拥有正常时长和分期，却在关系结构上仍未脱离值守。本章复算一个含130,150首独立曲目的公开睡眠歌单数据集：仅30.37%的曲目同时满足宽松的慢速、低能量、低响度与高器乐性条件；81组阈值敏感性分析的覆盖率为20.82%—37.63%。这一不利结果排除了“睡眠音乐具有足以定义睡眠的单一镇静声学指纹”，迫使音乐维度从固定音形退回到“交拍—托拍—返拍”的关系位置。全章给出二维辨别格、零情态判据、下一拍托管率的清点手册、跨物种边界、判决性分界与机制证伪条件。本稿另作三项更正。其一，把职业健康领域的值守睡眠研究列为本章最近的一位，并公开承认初稿漏掉了它。其二，把本章主张收窄为一句：回应去向与身体内转是两笔各自成立而互不登记的账。其三，声明音乐维度在本稿只承担时间语法的命名，不承担经验证据。

关键词：睡眠本体论；托拍事件；悬答眠；回应托管；选择性唤回；睡眠音乐

一·凌晨三点：三个看起来都在睡的人

凌晨三点，三个人都闭着眼。第一位是值夜班后躺进值班室的医生，呼吸已经变慢，脑电也可能进入非快速眼动睡眠，但传呼机仍只认他一个人。第二位是母亲，孩子在隔壁，伴侣明确接管凌晨两点到五点的照护；她听不见普通街声，却会被事先约定的紧急信号唤醒。第三位独居在安全住所，手机进入自动分流，火警与门禁仍能越过静音，第二天清晨熟悉的音乐把他带回清醒。三个人都“睡了”，但睡眠并不是同一个存在。

睡眠科学很擅长回答身体发生了什么：肌张力下降，反应阈值升高，脑电结构改变，腺苷、去甲肾上腺素、食欲素、胶质细胞钙信号、线粒体氧化与脂质代谢进入不同于清醒期的组织。社会学很擅长回答谁被允许睡、谁必须警醒、谁的睡眠不断替别人的需求让路。音乐学则告诉我们，休止并不等于时间消失；一个无声位置仍被节拍、期待与下一次进入所规定。问题在于，这三类知识长期分别记账：身体账本记录睡眠阶段，制度账本记录工时与照护，声音账本记录速度、响度与节奏。没有一个账本询问：睡者停止即时回应之后，世界的下一次请求由谁接住，又以什么信号把回应权交还给他。

这不是给睡眠增加一个“社会因素”，也不是把音乐当作修辞。若睡眠只是身体内部的状态，那么同一段脑电、同一段时长、同一种姿势就应属于同一种睡眠；外界由谁值守，只会影响睡得好不好，不会改变睡眠是什么。本章将论证相反结论：睡眠的最小存在单位不是孤立身体里的静止状态，而是一段跨越身体、关系和时间边界的完整事件。身体的内转、外界的接手和选择性的返还，缺一环都只得到睡眠的一个局部。

因此，本章的核心判断写成三重否定：睡眠不是行为静止，不是分子修复，也不是社会豁免；睡眠是下一次回应被外界托住、身体进入可逆内稳态重组、并由选择性信号返还回应权的“托拍事件”。这一定义的锋刃不在于给已有睡眠换名字，而在于让一类过去没有本体位置的事件显形：身体已经睡着，下一次回应却无人接手的“悬答眠”。

二·三个账本：每一种定义为什么都对，却都不够

行为学与生理学的传统定义通常从可观察特征开始：姿势相对固定、活动减少、对刺激的反应阈值升高、状态可迅速逆转，并在剥夺后出现补偿性反弹。Campbell 和 Tobler 以这类标准跨物种识别睡眠，Borbély 的双过程模型则用睡眠稳态压力与昼夜节律的交互解释何时入睡以及睡眠强度如何变化。这一账本非常重要，因为它把睡眠从昏迷、麻醉、冬眠和单纯休息中区分出来。但它把判定点放在有机体内部：只要这个身体呈现出一组状态特征，睡眠就被认为已经完整存在。

睡眠社会学从相反方向打开问题。Aubert 与 White 早在 1959 年就把睡眠理解作为一种社会撤退；Schwartz 把它描述为从社会要求中获得的周期性豁免；Williams 进一步讨论“做睡眠”的礼仪、权利、义务以及“社会上保持注意的睡者”；Meadows 则通过“协商的夜晚”展示伴侣如何安排声音、空间、照护与身体。这个账本说明，睡眠从来不是关门之后的纯私人事件。一个人何时能失去响应能力，往往依赖另一个人仍然醒着，或依赖制度承认这段不回应具有正当性。

音乐学提供第三种、常被低估的知识。音乐中的休止不是没有东西；它有时值、有位置、有前后关系。Hasty、London 以及 Lerdahl 与 Jackendoff 关于节拍、层级与期待的研究都表明，听者不是逐个接收孤立声响，而是在时间结构中预测下一次进入。休止之所以是休止，不是因为声压降为零，而是因为下一拍仍在场。音乐也不只以“慢、轻、无歌词”帮助入睡。大规模睡眠歌单研究发现，睡眠音乐内部存在显著异质性，其中若干簇甚至比平均睡眠音乐更快、更响、更有能量。声音在这里不是睡眠的固定外形，而可能是进入、维持或返回某种时间关系的标记。

三个账本分别把睡眠放在三个位置：身体账本问“内部状态是否改变”，社会账本问“缺席是否被允许”，音乐账本问“时间中的空位如何被保持”。它们的共同盲点并不是没有考虑另两类因素，而是都默认睡眠可以由自己那一维单独读出。身体状态仿佛足以定义睡眠；社会许可仿佛足以保障睡眠；声学形态仿佛足以表现睡眠。恰恰是各领域自己积累的材料，使这一默认前提无法继续成立。

表1 三种睡眠账本的能力与盲点

账本	它可靠回答的问题	它单独回答不了的问题
社会学	谁被允许暂退、谁仍承担照护与值守、夜晚如何被协商	身体是否进入睡眠轨道；返还信号如何选择性生效
睡眠化学	内部状态如何切换，睡眠压力、神经调制与代谢如何重组	外部请求由谁处理；同样内转为何有不同关系后果
音乐学	缺席如何在共同时间中被保持，何时预期下一次进入	身体是否睡着；何种制度把声音赋予交拍或返拍位置

三·社会学的发现：睡眠首先是一种不再立即回答的资格

一个社会如何看待睡者，首先体现在它如何处理“不回答”。课堂上闭眼、会议中打盹、照护中熟睡、宗教仪式里的昏沉与卧室里的夜眠，生理上可能相近，社会判定却完全不同。Williams 所谓睡眠礼仪，正是围绕何时可以睡、在哪里可以睡、睡给谁看、是否打扰他人以及醒来后应承担什么而展开。睡眠者从互动现场退开，但这种退开必须被共同体解释：是合法休息、懒惰、疾病、失职，还是危险。

Schwartz 的“周期性豁免”已经逼近本章要处理的关键：人不能在每一时刻都承担社会角色，睡眠让角色义务暂时松开。然而，“被允许不回答”与“下一次回答已经有人接手”不是同一件事。值班制度可以在文字上给医生安排八小时休息，却仍把传呼的最终责任留给他；家庭可以承认母亲需要睡觉，却默认孩子一哭仍由她先起身；企业可以规定下班，却让平台通知、客户升级和安全事故直达个人手机。形式豁免存在，回应托管却没有发生。

Meadows 的“协商的夜晚”与 Hislop、Arber 关于女性睡眠中断的研究揭示，睡眠不是个人单方面完成的动作，而是伴侣、孩子、家务、空间和身体相互安排的结果。Venn 等人把夜间照护称为“第四班”，更直接指出：某些人的睡眠之所以完整，是因为另一些人在通常应睡的时段继续工作。这里真正稀缺的并不只是时间，而是让一个人暂时不再成为最后回应者的关系资源。

因此，社会学给出的最低条件不是“社会支持”这种宽泛变量，而是可核查的回应转交：从计划入睡到计划醒来的区间里，哪些事件仍可以合法要求此人行动；每一类事件由谁接手；没有人接手时是否存在可执行的延期规则；什么信号结束转交并把回应权归还给睡者。若这些问题没有答案，社会从睡者身上撤走的只是一部分可见任务，潜在值守仍留在身体上。

这也解释了为何两个人拥有相同的床、相同的八小时、相同的安静房间，却可能拥有不同种类的睡眠。一个人的缺席被系统吸收，另一个人的缺席随时可能制造损失。前者可以让外部世界真正“继续”，后者只能把所有未发生的事件悬在自己身上。传统睡眠社会学已经看见睡眠的权利与协商，却很少把“下一次回应的持有人”写成一个独立字段。本章的第一步，就是把这个空字段填出来。

四·化学的发现：身体不是停机，而是在可逆地换轨

如果只从外观判断，睡眠容易被误认为活动减少。化学与神经科学却显示，睡眠是高度组织化的状态转换。Borbély 的稳态过程提示，清醒时累积的睡眠压力会在睡眠中释放；Porkka-Heiskanen 等人的研究把腺苷与 prolonged wakefulness 后的睡眠诱发联系起来；食欲素系统的缺失会破坏清醒稳定，形成发作性睡病样表型；下丘脑、脑干和基底前脑的多条回路共同调节睡眠—觉醒切换。不存在一枚“睡眠分子”能够独占整个过程。

近年的材料进一步把睡眠压力与代谢、免疫和线粒体联系起来。Haynes 等人在果蝇中发现，清醒期间的神经元—胶质细胞脂质转移与胶质线粒体氧化形成日周期，完整一夜睡眠与脂质清除、线粒体氧化恢复及神经元线粒体自噬相关。Ma 等人在小鼠中显示，小胶质细胞可通过 P2Y₁₂ 相关信号、胞内钙变化与去甲肾上腺素传递调节睡眠。Sarnataro 等人又把睡眠剥夺后的转录变化、线粒体呼吸和 ATP 合成连接到睡眠压力。Rorsman 等人则提出电压门控的脂质过氧化“记忆”参与睡眠压力积累。无论这些机制未来如何修订，它们共同排除了“睡眠等于消极静止”。

同时，化学材料也反过来限制任何单功能定义。Xie 等人的工作曾强化“睡眠促进脑代谢物清除”的解释，Miao 等人却在另一套测量设计中得到睡眠和麻醉时脑清除降低的结果，并明确挑战“睡眠的核心功能是清除毒素”这一说法。科学上的争论不是麻烦，而是重要约束：本章不能把睡眠定义成修复、排毒、记忆巩固或某一条代谢任务。可保留的最低判断只是，睡眠包含一条可逆的内部重组路径，它重新分配清醒期的能量、氧化、神经调制与信息处理负荷，而具体完成什么不能由单一机制预先写死。

化学账本能告诉我们身体是否换轨，却不能告诉我们外部世界有没有停止向这个身体索要即时回应。一个传呼机旁的医生可以进入非快速眼动睡眠；一位独自照护婴儿的母亲也可以出现完整睡眠周期；野外动物甚至可能在保持部分警戒时出现局部或单半球睡眠。内部重组与外部脱责并不总是同步。化学上的“已睡”因此不是完整睡眠的终点，而是托拍事件的内环。

本章把这一路径称为“内转”，并给它三个最低约束：第一，它必须可逆，不能把昏迷和不可逆损伤纳入；第二，它必须改变对外即时响应与内部资源组织之间的比例，不能把安静清醒等同于睡眠；第三，它必须具有历史依赖，剥夺会改变后续进入与强度，不能把一次药理抑制自动视为自然睡眠。内转是睡眠成立的必要条件，却不是充分条件。

五·音乐学的发现：休止不是空白，下一拍仍然存在

音乐把一个极难说清的事实变得可听：不发声不等于不在时间里。谱面上的休止符不产生音高，却占据确定时值；演奏者在休止期间并未退出作品，而是继续数拍、保持相位、预备下一次进入。一个休止若无人保持，后面的音即使准确，也会失去共同时间。睡眠与休止的相似之处不在“安静”，而在于两者都要求某种缺席被持续托住。

这种结构并非纯粹比喻。动态注意理论和节拍知觉研究表明，听者的注意会围绕预期时点周期性聚集；音乐中的重拍、切分、延宕和终止式改变的，是下一事件何时值得准备。睡眠中的感觉通道也不是彻底关闭。Kouider 等人发现，睡眠者仍能对任务相关语音形成可检测反应；Koroma 等人显示，快速眼动睡眠中有意义语音既可能被放大，也会在眼动时段受到选择性抑制；Konkoly 等人甚至在清醒

梦条件下实现了实验者与睡眠者的双向交流。睡眠并非世界被切断，而是信息进入的权重和时机被重新编排。

这使“返拍”成为睡眠定义的一部分。睡者不能对所有信号保持同等开放，否则内转不断被打断；也不能对所有信号完全关闭，否则危险、照护与环境突变无法被处理。完整睡眠保留一条选择性返还通道：闹钟、孩子哭声、火警、熟悉呼名、昼夜信号或内部生理变化，可以在满足条件时把即时回应权交还给睡者。这个通道不由响度单独决定，而由信号在关系中的意义、学习史与责任结构共同决定。

睡眠音乐的全球数字痕迹进一步否定了单一音形。Scarratt 等人分析大规模 Spotify 睡眠歌单，发现睡眠音乐总体上较慢、较轻、较多器乐性，但内部形成六个有意义的亚群，其中若干群包含更快、更响、更有能量的流行曲目。Trahan 等人的调查也显示，人们用音乐助眠的理由、曲目与风格高度多样。音乐若只靠低能量声学直接“压低”神经系统，这种异质性很难解释；若音乐承担的是边界标记、注意转移、熟悉性保护或交还时点的组织，异质性反而成为预期。

所以，音乐学不是来证明“听音乐能睡”，而是为睡眠提供一种时间语法：交拍、托拍、返拍。交拍是外界不再把每个请求直接压到睡者身上；托拍是内部重组占据一段被保护的时值；返拍是某个被认可的信号恢复回应。音乐把睡眠从“无动作”改写成“有结构的缺席”。

六·三种解释为何不能互相吞并

若社会许可足以定义睡眠，那么一名被正式批准休息、却整夜守着唯一报警电话的人应当已经拥有完整睡眠；化学会反驳，因为他的身体可能根本没有稳定内转。若内部状态足以定义睡眠，那么一名在危险环境里短暂进入非快速眼动睡眠的人与一名在可靠照护系统中睡眠的人属于同一类；社会学反驳，因为两人的缺席由不同结构承担。若慢速、低响度与重复节奏足以定义睡眠，那么所有高能量睡眠歌单都是误用；音乐实践和公开数据又会反驳，因为同一功能可以由完全不同的音形实现。

三个领域不是在同一条轴上争“谁更重要”，而是在三个不同位置给睡眠落界。社会学把睡眠边界放在角色义务何时松开；化学把边界放在有机体何时换轨；音乐学把边界放在缺席何时被共同时间保持，并在什么时刻返回。三条边界可以错位：责任先交出而身体仍清醒，是受保护的休息；身体先入睡而责任未交出，是悬答眠；声音与仪式已开始标记夜晚而身体没有换轨，是睡眠准备；内部重组、责任转交与返还通道同时闭合，才是完整睡眠。

这种错位不是异常边角，而是现代睡眠的常态。平台劳动让工作停止与责任停止分离；家庭照护让躺下与脱离值守分离；医院让患者的生理睡眠与报警系统的持续侵入分离；耳机、白噪声与播放列表又把传统由共同体完成的夜间边界部分外包给技术。若仍把睡眠视为一个身体里的单点状态，所有错位都会被压扁成“睡眠质量不好”。而“质量”只是程度词，它不能说明究竟缺了哪一环。

本章因此拒绝把三种解释综合成一张因素清单。因素清单仍假定睡眠本身先已存在，然后被社会、化学和声音分别影响。更准确的说法是：三类关系共同组成睡眠事件。不是先有一个睡眠，再有环境对它好或坏；而是外界接手、内部换轨与选择性返还的闭合，本身让这一段时间成为睡眠。

七·共同前提的坍塌：睡眠并不只存在于睡者体内

三种解释表面上互相排斥，底下却共享同一个前提：睡眠是一种已经属于睡者的东西，研究者只需决定从行为、分子还是制度哪一面把它读出来。这个前提看似不言自明。人是睡者，床在他的身下，脑电贴在他的头上，睡眠当然“在他里面”。然而，三家自己的材料都在瓦解这一点。

社会学知道，睡眠需要别人不来索取回应，需要伴侣调整动作、家庭分配照护、组织安排替班、城市控制噪声与光线。没有这些关系，身体即使闭眼，仍可能处在待命。化学知道，睡眠并非全脑、全身同时切换：局部睡眠、不同睡眠阶段、感觉选择性和代谢过程的异步，都说明“一个身体是否睡着”不是简单的全有或全无。音乐学知道，同一段静音只有嵌入共同节拍才成为休止；脱离前后关系，它只是没有声响。三家共同掌握的事实合起来迫使我们承认：睡眠不是被环境包围的内部物，而是内部与外部共同完成的关系事件。

这一转向取消了一个长期混淆：把“没有发生打扰”当作“打扰已被接住”。某个夜晚没有来电，不等于来电责任已经转交；孩子没有哭，不等于照护者获得了脱离值守的资格；野外没有捕食者出现，不等于动物处在可放心关闭警觉的生态位。偶然没有事件，只产生安静；可靠地安排事件由谁处理，才产生托拍。

于是，睡眠的单位必须从“睡了多少分钟”扩展为“哪一段回应关系完成了交接”。分钟仍然重要，脑电仍然重要，化学机制仍然重要，但它们不再单独拥有定义权。一个睡眠事件的边界从交拍开始，在内部换轨中持续，以返拍结束。它不是把睡者从世界中删除，而是让世界临时改变向谁索要回答。

八·托拍事件：睡眠的新定义

本章把完整睡眠定义为一种托拍事件：在一个有明确起止边界的区间内，有机体把对外部关键请求的即时回应位置暂时移出自身，身体进入可逆且具有历史依赖的内稳态重组，同时保留一条对高显著信号开放的返还通道；外托、内转与返拍三者闭合，睡眠才作为完整事件发生。

第一环是“外托”。它不要求一定有另一个人坐在门外。替代持有人可以是伴侣、同事、护理团队、群体哨兵，也可以是可靠的自动系统、可执行的延期规则、降低事件概率的庇护空间或生态时间窗。关键不在“有人关心”，而在下一次需要回应的事件已经有明确去处。一个只会把报警转发给睡者的设备不算接手；一个能自动处置低级事件、只在预设紧急条件下唤回睡者的系统才构成部分外托。

第二环是“内转”。身体从维持即时外部应答的配置，转入睡眠阶段所特有的神经调制、代谢、免疫、突触与能量重组。本章故意不用“修复”作为定义，因为修复预设了单一正向功能，也无法容纳清除机制争论、梦境活动、局部睡眠与不同阶段的差异。内转只要求状态可逆、具有睡眠稳态历史，并改变外部响应与内部处理的组织关系。

第三环是“返拍”。外托不能无限延伸，否则睡眠与失联、昏迷或被剥夺行动权混在一起。完整睡眠需要一种返还条件：到达约定时点、出现高显著危险信号、内部生理状态完成转换，或替代持有人主动交还。返拍不是单纯醒来，而是即时回应位置重新回到睡者。一个人在闹钟响前自然醒来，也可以是内在昼夜与稳态过程完成了返还；一个被无关通知反复惊醒，则是返还通道被低价值信号劫持。

三环之间有先后，却不要求机械同刻。社会交拍可以发生在入睡前，内部换轨随后展开，返拍条件则在整段睡眠中潜伏。重要的是闭合：谁接手、身体如何转、什么把回应交还。托拍不是把睡眠浪漫化为被照顾，而是把一个原本被省略的系统条件写入定义。

这一判断可以被一句不含“充分、真正、合理、良好”等程度词的问题检验：从这段入睡开始到醒来，哪些事件仍可要求此人回应；这些事件由谁接手；哪一个信号把回应权交还给此人？若第一问能列出事件，第二问没有答案，而第三问仍把所有返回都压给睡者本人，那么我们面对的不是完整托拍，而是悬答。

表2 托拍事件的三环结构

环节	最小问题	成立证据	缺失时的事件
交拍／外托	睡者不回应时，关键事件去哪里？	替代持有人、自动处置或可执行延期	责任仍悬在睡者身上
托拍／内转	身体是否进入可逆、历史依赖的睡眠重组？	行为、生理、脑电或相应物种标准	受保护的清醒或休息
返拍	什么条件把即时回应交还给睡者？	约定时点、高显著信号、内部完成或主动交还	失联、任意打扰或返还通道被劫持

九·悬答眠：一个现有账本没有字段的存在

“悬答眠”指有机体已经进入可识别的睡眠内转，而至少一项具有后果的即时回应义务仍只由睡者本人承担，且没有可靠替代持有人或可执行延期规则。它不是失眠的同义词。失眠通常围绕难以入

睡、难以维持、早醒或日间损害展开；悬答眠则描述睡眠事件的关系结构。一个人可以很快入睡、拥有正常总时长和看似完整的睡眠分期，却仍处在悬答眠中。

它也不等于“主观担心”。有些人清楚自己值守，因此感到焦虑；另一些人已习惯夜间随叫随到，主观上并不紧张。悬答眠的判定不依赖自我报告，而依赖责任配置：若真正发生事件，最后必须行动的人是谁。这个区别使它能够与一般压力、人格焦虑和高唤醒理论分开。心理紧张可能是结果或中介，却不是定义。

悬答眠之所以是一类新存在，而不只是睡眠质量的新读数，有三个理由。第一，它在现行账本中没有稳定字段。多导睡眠图、可穿戴设备、睡眠日记、排班表和家庭分工记录，通常不会同时记录“这段睡眠期间每类事件的最后回应者”。第二，它能在所有常规形式审查中通过：时长够、效率高、阶段齐、房间静，却仍缺一项外部闭合。第三，它的失效往往不产生信号。若整夜没有紧急事件，系统会误以为安排有效；下一夜继续把责任留给睡者，结构因没有暴露事故而自我加固。

最典型的悬答眠不是不断被吵醒，而是“可以随时被叫醒却恰好没有被叫醒”。这句话把潜在值守从实际中断中分离出来。传统指标只记录发生的唤醒，悬答眠记录谁被迫持有所有可能的唤醒。它让我们看见：零次中断可能来自两条完全不同的路径——一条是责任已经转交，另一条是风险没有在当夜兑现。

这一新类也改变干预方向。若问题只是睡眠时长不足，处方是增加时间；若问题只是噪声，处方是降噪；若问题是悬答眠，单纯延长卧床可能延长值守。有效干预必须转交事件、设置替代持有人、建立分级自动处置或明确延期规则。换句话说，悬答眠不是“睡得不够深”，而是“睡者仍是世界的最后接口”。

九·五 本章主张的那一格

初稿把悬答眠说成一类无人处理过的存在。这句话不成立，须当场收窄。

“身体睡着而人仍在值守”这件事，职业健康研究做了三十八年，名字叫值守睡眠（on-call sleep），详见第十八节第一条。“对与自己有关的信号选择性醒来”这件事，睡眠心理生理学做了六十五年，起点是Oswald、Taylor与Treisman一九六〇年在《Brain》上那篇——睡着的人听见自己的名字比听见别的名​​字醒得更快。“陌生环境里身体自己保留一份警戒”这件事，Tamaki等人二〇一六年在《Current Biology》上给出了神经证据，并直接称之为“夜间值守”。这三块地都有主人，本章不再声称对它们有命名权。

划出去之后，本章剩下的那一格是这样一句：**在同一段睡眠上，回应位置的去向与身体的内转是两笔各自成立、而互不登记的账；睡眠记录制度只生产后者，值守记录制度只生产前者，二者没有共同的事件标识。**

这一格之所以还站得住，有三个可核的支点。第一，它不是关于效应有无的主张，而是关于记录制度的主张：值守睡眠研究测的是“on-call与否”这个二分变量，本章要求的是按事件类别枚举的连续读数R；多导睡眠图记录里没有R这一栏，值守研究的问卷里没有睡眠分期与事件时间戳的同轴对齐。第二，它给出一个最锋利的检验位置——实际中断为零的那些夜晚；在那里，时长、效率与分期三项全部正常，两笔账的差别只能由责任去向读出。第三，它可以被一次否定：只要有人拿出一份把职责枚举、告警路由、返还信号与睡眠分期记在同一时间轴上的公开数据集，本章的记录缺口主张当场作废，剩下的只是一个已被别处测过的效应。

后面各节的分界、证伪与赌注，都按这一格重写，不再按初稿那个更大的主张。

十·二维辨别格：身体是否内转，下一拍是否有人接住

为了避免新概念滑回模糊比喻，可以用两条彼此独立的轴定位一段夜间状态。纵轴是身体是否进入睡眠内转；横轴是关键回应是否完成外托，并附有明确返拍条件。两轴组合产生四种事件。那一格不是一眼可见的坏状态，而是“身体已睡、责任未交”的悬答眠：它在短期睡眠指标上可能表现正常，结构失效没有专门报警，并会因为多次平安无事而被继续采用。

这张表也说明，受保护的休息不等于睡眠。一个人下班后阅读，所有工作请求都已转交，身体仍在清醒轨道；他拥有完整外托，却没有内转。相反，麻醉或药理抑制也不能仅凭低反应性自动归入托拍睡眠，因为其历史依赖、阶段组织与返还通道可能不同。二维格不是替代临床判定，而是给临床判定补上关系维度。

对研究而言，两轴必须分别测量，不能用一个指标代理另一个。脑电、眼动、肌电、活动与自我报告可用于判定内转；职责清单、转交规则、告警路由、照护排班和实际事件日志用于判定外托。若研究只问“你是否担心被叫醒”，它仍把关系结构还原成心理变量；若只问“昨晚是否被叫醒”，它又把潜在托管还原成偶然结果。

表3 身体内转 × 下一拍托管

	下一拍已托管（含返还规则）	下一拍未托管
身体已内转	完整托拍眠：外托—内转—返拍闭合	悬答眠：身体已睡，仍是最后回应者
身体未内转	受保护的清醒／休息：责任已交，身体未睡	暴露清醒：责任与应答均留在本人

十一·公开数据真跑：睡眠音乐没有足以定义睡眠的单一声学指纹

理论若只在概念上漂亮，很容易把音乐学写成一层装饰。为使音乐维度接受反证，本章在统计运行前保存了预注册：使用 Scarratt 等人公开的睡眠歌单独立曲目数据，检验“睡眠音乐主要由低速度、低能量、低响度、高器乐性构成”的固定声学命题。主分析把“宽松镇静型”预先定义为速度不高于 90 BPM、能量不高于 0.40、响度不高于 -15 dB、器乐性不低于 0.50，四项必须同时满足。若这一形态足以定义睡眠音乐，至少 70% 的曲目应通过。

先说清这次复算的身份，因为初稿把它说大了。Scarratt 等人自己在原论文中已经报告了六个亚群、且其中若干比平均睡眠音乐更快更响——本章第五节引用的正是这一条。因此，本次复算**是一次干净的复现，不是一次判决性检验**：它以预注册的方式重跑了一个原作者已经给出结论的命题，并把结论量化到一个可以逐位核对的比例上。初稿第十一节末称“这一结果比原论文更不利于本章早期设想”，这句话不准，本稿撤回。复算的全部价值在两点：一，它是本章自己下的注在自己身上兑现——预注册写死了 70% 的放行线，实测 30.37%，本章因此必须撤回一个自己更喜欢的主张；二，它给出了 81 组阈值下的完整覆盖率区间，使“换个阈值就能救回来”这条退路被事先堵死。

数据包含 130,150 首四项特征均不缺失的独立曲目。主分析只有 39,523 首同时满足四条件，覆盖率为 30.37%，远低于预注册的 70% 放行线。为避免某一阈值决定结论，本章又在速度 80/90/100、能量 0.30/0.40/0.50、响度 -18/-15/-12、器乐性 0.40/0.50/0.60 之间运行 81 种组合，覆盖率最低 20.82%，最高 37.63%。即使采用最宽松组合，仍有超过六成曲目不能被单一镇静声学形态覆盖。

按数据集既有六簇复算，只有第 4 簇的速度、能量和响度均未高于全样本中位数；按预注册的相对中位数规则，其余五簇合计占 52.54%。这一结果比原论文“存在若干更快、更响、更有能量的亚群”更不利于本章早期设想。它迫使我们撤回“睡眠在音乐上显露为统一减速”的强主张。睡眠音乐可以总体偏慢偏轻，却不能由慢、轻、无歌词单独定义。

这次复算不证明音乐导致睡眠，也不证明任何特定曲目承担返拍功能。Spotify 音频特征的算法不是公开测量标准，歌单归类受到平台、用户与文化选择影响，数据没有听者人口学、实际播放时点、入睡结果或实验随机化。复算能裁决的只有一件事：固定声学指纹不足以吸收睡眠音乐。音乐维度必须改写为关系命题——一段声音的睡眠功能取决于它在某个听者、某段生活安排和某个交还时点中的位置。

不利结果由此提升了概念的粒度。原先的“托拍”还容易被理解为外部节奏把身体减慢；复算后，托拍只能指一段被共同保持的时间位置。快歌、讲话、宗教声音、熟悉旋律甚至沉默，都可能在特定使用中标记交拍；同一首歌若被设为紧急铃声，又可能承担返拍。决定功能的不是音形本身，而是声音被写入哪一条回应序列。

表4 睡眠音乐公开数据复算结算

项目	预注册口径	复算结果
----	-------	------

样本	四项特征不缺失的独立曲目	130,150 首
宽松镇静型	tempo≤90; energy≤0.40; loudness≤-15; instrumentalness≥0.50	39,523 首; 30.37%
固定声学命题放行线	覆盖率≥70%	未通过
81 组敏感性分析	3×3×3×3 阈值组合	20.82%—37.63%
相对中位数激活型簇	簇均值的速度、能量、响度均高于样本中位数	5 个簇; 合计 52.54%
可支持的结论	单一声学形态是否足以定义睡眠音乐	不足以; 不能据此推断音乐导致睡眠

十二·六个具体场景：同样闭眼，事件结构不同

第一种场景是婴儿照护。父母都躺下并不意味着两人都获得托拍。若凌晨两点到五点由一方明确接手喂养、安抚和异常判断，另一方的关键回应得到转交；若口头上说“谁听见谁起来”，两人都可能保留最后回应位置。零情态问题会得到不同答案：孩子哭由谁先处理，发热由谁判断，何种信号才唤回另一人。这里差别不在爱或责任感，而在交接是否闭合。

第二种场景是医疗值守。两名医生都被安排在值班室休息，一名有分级响应与明确后备，普通事件由团队处理，只有预设严重等级才返还给他；另一名虽名义休息，却是所有电话的唯一接收者。若当夜无事，两人的睡眠时长可以相同。托拍理论预测，后者即使没有实际来电，也更可能保持与责任相关信号的优先加工；传统“是否被叫醒”记录无法区分。

第三种场景是重症监护患者。患者进入睡眠阶段，但监护设备把大量低价值告警直接转成声音，护理流程又要求患者对某些程序反复配合。这里外托可能被技术伪造：系统看似在“替患者监测”，实际却不断把回应退回给患者身体。真正的托管应当由护理团队吸收低级事件，只把与患者行动直接相关且不可延期的信号返还。减少分贝并不能替代告警路由的重写。

第四种场景是平台劳动者。手机静音不等于责任转交。若订单取消、客户升级和安全异常仍在夜间累计，并在醒来后以惩罚方式追索，睡者只是把回应推迟而没有获得可执行豁免。相反，自动分流、时区轮班或明确的服务时段能让请求在睡眠区间找到别的持有人。两种安排都可能没有夜间通知，但只有后一种形成外托。

第五种场景是不安全居所。一个人可以因为极度疲劳而睡着，身体内转已经发生，但门锁、照明、同伴值守和环境风险都不足以托住外部请求。此时“改善睡眠卫生”的建议若只强调作息、咖啡因和屏幕，会把结构问题个体化。托拍视角并不否认个人行为，而是先问：在这段失去响应能力的时间里，什么保护了身体免于必须自己处理每一个危险。

第六种场景来自非人动物。群居动物可能通过哨兵、群体同步或位置安排分配警觉；独居动物则利用洞穴、巢穴、潮汐、昼夜生态位或降低暴露的姿势来减少必须即时回应的事件。这里没有法律意义上的“义务”，但有生态意义上的响应负担。托拍不是人类制度的专属，而是有机体何以暂时降低对环境的即时接口、又不失去关键唤回能力的问题。

表 5 零情态判据在具体场景中的读数

场景	仍可要求的回应	持有人／规则	返拍信号	判定
婴儿照护轮班	持续哭声、发热、呼吸报警	伴侣按时段与事件分级接手	预设紧急条件	R 可达 1；完整托拍
“谁听见谁起”	同上	无明确持有人	任何声音直接返还	R<1；悬答眠
医生分级值守	普通咨询、紧急升级	普通事件由团队处理	严重等级触发	部分或完整托拍
平台劳动者静音	订单升级、客户时限、安全事件	仅把通知延后，无免责	醒后惩罚性追索	静音不等于托管
安全庇护中的独居者	门禁、火警、普通消息	门禁／保安先核验；普通消息延期	确认危险或晨间时点	技术与制度联合托拍

十三·跨物种边界：谁不需要“别人”也能睡

任何把睡眠定义为“社会责任由他人接手”的理论都会在人类中心主义上失败。许多动物独自睡眠，没有伴侣、组织或设备替班。本章所谓外托因此不是道德承诺，而是关键响应的功能性去向。它

可以由其他个体承担，也可以由环境结构承担：巢穴降低攻击路径，夜色改变捕食概率，潮汐形成低风险窗口，身体位置保留一侧感觉输入，群体同步把个体暴露分散。

Rattenborg 等人关于军舰鸟飞行中睡眠的研究表明，睡眠与持续环境任务可以以不完全、侧化的方式共存。局部睡眠研究也显示，清醒动物的部分皮层区域可能出现睡眠样活动。此类材料不是说“任何局部下降都是托拍”，而是说明睡眠的内转本来就可以分层发生。托拍事件也应允许部分闭合：某些回应由外界接住，另一些仍由有机体保留；某些神经网络内转，另一些维持警戒。

跨物种检验使概念必须让步。人类场景中的“职责清单”不能直接搬给动物；只能转写为可观察的事件类别、响应阈值和替代机制。例如，在同一物种、相近能量状态与捕食风险下，若有哨兵或更安全庇护时，个体睡眠的内转更完整、对非关键信号的反应更低，而对真正威胁仍保留选择性，这将支持托拍结构。若没有任何风险转移、庇护或选择性返还，动物仍能出现同等稳定的完整睡眠，外托作为普遍必要条件就会被推翻。

技术设备的“睡眠模式”提供另一条边界。电脑或手机进入低功耗状态、保留唤醒中断，与托拍在形式上相似，但这并不意味着设备具有生物睡眠。设备没有睡眠稳态压力、生命维持与剥夺反弹，因而缺少内转的生物历史。概念的跨域价值在于辨认结构，不在于把一切低活动都宣布为睡眠。

十四·可裁决的时序预测：交拍、内转、返拍的先后

托拍理论最容易被误读成一句“安全和支持有助于睡眠”的常识。要避免这种吸收，必须给出竞争理论预测不出的时序。第一条预测是交拍应当先于内转。控制总工作量、睡眠机会、房间条件、昼夜相位和基线焦虑后，当关键回应应在入睡前被明确转交，睡眠潜伏期、与责任相关刺激的加工或夜间自主神经指标应在随后的时间窗改变；若先出现生理改善，交接记录只是事后补写，理论的方向就错了。

第二条预测是未完成交拍会选择性保留与责任相关的感觉通道，而不是普遍提高所有刺激的反应。一个值守者可能对自己的传呼音、孩子哭声或特定呼名保持更高敏感，却不一定对等响度的陌生声音同样敏感。若控制声学强度、熟悉度与睡眠阶段后，责任相关性没有独立作用，返拍通道与责任结构相连的主张就不成立。

第三条预测涉及回写。某一夜外托成功，不只影响当夜；连续多夜可靠交接会改变下一轮的预期，使身体更早允许内转。反过来，一次替代持有人失约或错误返还，会在后续夜晚提高睡者对相关信号的保留，即使新的安排在纸面上已经修复。若长期记录显示交接可靠性从不影响随后几夜的响应阈值或睡眠组织，托拍的循环性解释将受挫。

第四条预测是返拍的意义依赖位置。同一首声音在睡前作为“现在不再处理请求”的固定标记，可能帮助交拍；被设为紧急报警后，可能立即破坏托拍。随机交叉实验可以在控制音量、速度和曲目不变的条件下，只改变其制度角色与训练关联。若角色改变不影响入睡、唤醒或选择性加工，而声学特征独自解释全部差异，音乐的关系命题就应撤回。

第五条预测是悬答眠与实际中断可分离。选择没有发生夜间事件的夜晚，只比较回应托管完整与不完整的被试；若两组在多夜后的生理、认知或情绪指标仍出现差异，说明潜在值守具有独立作用。若所有差异都由实际来电、哭声和噪声次数解释，悬答眠就退化为普通睡眠中断，没有新本体地位。

十四·六 本章最强的三个竞争解释

上面五条预测若被证实，仍不足以支持托拍事件，因为至少有三种更省事的解释能产生同样的观察。初稿没有单列这一节，本稿补上，并给出各自的岔点。

第一，陌生度而非责任。 Tamaki 等人证明，人在陌生环境里睡第一夜时，默认模式网络出现半球间的睡眠深度不对称，较浅的那一侧对偏差刺激反应更强、唤醒更多、行为反应更快；而这种不对称在随后几夜完全消失。这给出一个不需要任何责任配置的值守机制：只要环境不熟，身体自己留一份警戒，几夜之后自动撤销。若悬答眠只是这一机制的延伸，那么它应当随环境熟悉度衰减，而不随责任路由变化。**岔点：**在同一张床、同一间房、连续多周的记录中，若托管完整与不完整两组的差异

随夜次同步衰减到零，是陌生度解释；若差异随交接安排的变化而变、与夜次无关，才是责任解释。这一条可以在家庭照护样本里直接跑，成本很低。

第二，焦虑中介而非关系结构。本站已发的失眠专文给出过一个更省事的机制：入睡要求意识退场，而“监督入睡”这件事本身恰恰需要意识在场，于是睡者派出的那个守夜人，其存在本身就是睡眠无法发生的原因。若这一机制成立，外托可能根本不是独立维度，它只是消除了“睡者自己派守夜人”的理由，最终仍通过唤醒水平起作用。**分岔点：**本章第八条分界已经写出该条件——在自陈平静、生理唤醒指标正常的被试身上，责任路由若仍有增量作用，两者分离；若效应在低唤醒者身上完全消失，本章应并入唤醒理论，作为它的一个前因变量而非一个本体维度。这是本章最容易输掉的一条，须优先检验。

第三，负荷与预期唤醒时间。值守者往往同时承担更长的工作时间、更不规则的作息与更早的预期起床时刻。预期起床时刻单独就能改变睡眠后段的结构。**分岔点：**把总工作量、上床时刻、预期起床时刻与实际起床时刻全部匹配后比较；若托管效应随之消失，本章应退回为负荷研究的一个描述性分层。

三条都指向同一件事：本章若要成立，必须在“陌生度已控制、唤醒已控制、负荷已控制”之后仍剩下一个由回应去向解释的余量。在拿到这样的余量之前，托拍事件只是一条待检验的记录制度主张。

十五·睡眠不平等的新单位：不是谁有时间，而是谁能退出最后回应位置

睡眠不平等常用时长、作息、住房、轮班与噪声暴露衡量。这些指标不可替代，但托拍理论指出另一个单位：谁能在一段时间内不再成为系统的最后接口。一个人拥有八小时空档，却要独自承担孩子、老人、客户、设备或安全事件；另一个人只有七小时，却有清晰交班、可靠替补和分级返还。两者谁的睡眠更完整，不能只由分钟决定。

这一单位也解释了为何照护劳动对睡眠的侵入常被低估。实际中断容易记录，潜在值守难以记录。负责者可能整夜没有起身，统计上没有失眠，也没有额外家务分钟；但他不能像非负责者一样放弃对某类信号的最终判断。社会把这段无事件的警觉当作零劳动，又把其后果归为个人睡眠问题。

组织层面的审计因此应把“休息安排”与“责任路由”分开。排班表写了休息，不代表告警系统完成转交；手机进入免打扰，不代表错过请求不会被惩罚；家庭成员口头同意照护，不代表每类事件有明确边界。最便宜的制度改进可能不是再增加一项健康教育，而是在流程里增加三个字段：事件类别、替代持有人、返还条件。

不过，本章不把所有社会差异都归入托拍。收入、住房、疾病、歧视、轮班和照护负担有各自机制；托拍只处理它们如何穿过“下一次回应由谁承担”这一接口。若一个变量不改变责任去向、内转或返拍，它不应为了理论宏大而被强行吸收。概念的价值来自切开相似现象，而不是吞并全部睡眠社会学。

十六·化学反过来给理论加的限制：不能把外托写成万能原因

社会结构可以决定一个人是否获得脱离值守的条件，却不能凭空制造睡眠内转。食欲素系统异常、呼吸暂停、疼痛、药物、神经退行性疾病、昼夜错位与许多代谢状态，都可能在外托完备时破坏睡眠。若理论把所有睡眠失败都解释为下一拍无人接住，它就把一个锋利维度膨胀成万能原因。

因此，托拍事件的第一条适用限制是：外托只解释睡眠何以能完整闭合，不单独解释身体是否有能力进入和维持各睡眠阶段。临床诊断仍需多导睡眠、病史、呼吸、药理、精神健康和神经系统评估。下一拍托管率不能替代睡眠效率、觉醒指数、氧饱和度或昼夜相位。

第二条限制来自清除争论。本章不能预设内转的价值方向总是“清除越多越好”。不同测量可能把流入、流出和再分布混在一起；不同脑区、分子大小与睡眠阶段也可能不同。于是“化学重组”必须保持可分解：研究者应明确测量哪条通路、什么时间尺度、与何种竞争机制比较，而不能用“身体在修复”作为不可证伪的总括。

第三条限制是麻醉与镇静。外部责任可以完全由手术团队接手，身体也呈现低响应，但药理状态不一定具有自然睡眠的阶段结构、稳态历史与选择性返还。托拍理论只有在内转被独立确认时才称之为睡眠；否则只是受托的无响应状态。这个让步削弱了“只要有人接住下一拍就能睡”的强说法，却保住了概念的边界。

第四条限制是局部性。睡眠可能不是整个有机体同相发生，因此“身体已经睡着”需要预先指定尺度：全夜临床睡眠、某一阶段、局部皮层活动，或特定动物的行为睡眠。悬答眠也必须与同一尺度配对。若尺度漂移，任何现象都能被塞进理论，概念就失去裁决力。

十七·音乐学反过来给理论加的限制：没有一种声音天然属于睡眠

公开数据已经迫使本章取消“慢、轻、重复就是睡眠声学”的强主张。进一步说，没有一首歌、一个频段、一种节拍天然具有托拍资格。音乐的作用至少依赖听者的熟悉度、文化语境、个人记忆、播放时点、控制感、用途与责任结构。同一段声音可以是安抚、噪声、身份标记、祈祷、工作背景或报警。

这意味着音乐不能成为托拍事件的必要材料。没有音乐的睡眠同样完整；自然醒也不需要外部声音。音乐学贡献的是辨别结构，而不是处方。研究问题不应只问“什么音乐最助眠”，还应问“这段声音在交拍前、托拍中或返拍时承担什么位置”。位置不同，最优声学特征可能相反。

第二个限制是熟悉性与条件化可能解释许多效果。若一首歌长期与入睡共同出现，它可能通过学习降低唤醒；这并不自动证明托拍理论。本章必须给出更强分离：在声学、熟悉性和既往配对相同的情况下，改变它是否代表责任已转交、是否代表紧急返还，若结果随制度角色改变，才支持托拍位置的独立作用。

第三个限制是声音不能替代实际交接。播放白噪声、冥想音频或睡前歌单，若孩子、病人、客户或设备仍只能由睡者回应，最多改变主观体验和感觉输入，不能把悬答眠变成完整托拍。把音乐当作低成本遮蔽物，反而可能让制度更容易把责任继续留给个人。

第四个限制涉及伦理。研究不能为了证明选择性返拍而在安全关键环境中故意撤走后备、制造危险或反复惊醒受试者。可接受的设计应利用既有轮班差异、模拟任务、低风险实验和匿名日志。理论若只能靠制造悬答来证明，就违背了它要揭示的结构。

十七·五 一次降格声明：音乐学在本章承担什么，不承担什么

把上面四条限制合起来看，会得到一个本章必须自己说出口的结论：在本稿中，音乐学不承担经验证据，只承担时间语法的命名。

理由是本章自己的数据。第十一节的复算排除了固定镇静声学形态；第十七节前四条又逐一取消了音乐作为必要材料、作为独立机制、作为交接替代物的资格。剩下的是什么？是“交拍—托拍—返拍”这三个名字，以及“休止有时值、有位置、需要被保持”这一条来自记谱与节拍研究的结构直觉。它确实帮助本章把睡眠从“无动作”改写成“有结构的缺席”，但它不产生任何本章独有的可测预测——把三环改称“移交—内转—归还”，全章的定义、指标、场景、分界与证伪条件一字不必改。

按本章自己在第二十四节立下的证据分层纪律，一个不改变任何预测的维度不能被写成材料位。因此本稿明确：音乐学在此处的位置是命名法，不是三分之一的证据来源。**唯一能让它重新承重的，是第二十二节第二项“同声异职”实验**——在声学参数与熟悉度完全匹配的两段声音之间，只交换它们在责任序列中的制度角色（“普通事件已由系统接手”与“仅高等级事件返还”），若脑电加工、微觉醒或次日记忆随角色交换而改变，音乐的关系命题才第一次获得独立证据。在该实验做出来之前，本章不主张音乐维度具有证据地位。

这一让步使本章的跨领域结构从“三家共同承重”收缩为“社会学供结构、化学供边界、音乐学供语法”。这不是谦辞，是记账：把没有兑现的那一份从资产栏划掉，剩下的才是可以被后来者接手的東西。

十八·与十种最相近解释的判决性分界

新概念最危险的失败不是被反驳，而是被一个早已有名字的概念完全吸收。托拍事件与悬答眠必须在同一案例上给出不同预测，不能只说“更全面”或“更深入”。下面十组分界把最接近的解释逐一请进来。

十八·一 值守睡眠：本章最近的一位，以及初稿漏掉它这件事

初稿把最近的一位漏掉了，而且漏得离谱。本稿把它排在分界节第一位，并先把话说明白：“身体已经睡着、人却仍在值守”这件事，职业健康与睡眠医学已经研究了三十八年，它有名字，叫值守睡眠。

这一支的起点是 Torsvall 与 Åkerstedt 一九八八年发表在《Sleep》上的船舶轮机员脑电研究——不是问卷，是脑电，比本章早三十八年。此后：Ferguson、Paterson、Hall、Jay 与 Aisbett 二〇一六年在《Chronobiology International》上综述指出，值守期间获得的睡眠可能恢复价值大幅下降，原因是**对接到呼叫的预期与对漏掉呼叫的担忧**；Hall 等人二〇一七年在《Sleep Medicine Reviews》上作系统综述，八项研究一致显示居家值守损害睡眠数量、多数情形下也损害质量；Vincent 等人二〇二一年在《Behavioral Sleep Medicine》上报告，**56% 的值守工作者称即使当夜没有接到任何呼叫，睡眠仍然受到影响**；Karan 等人二〇一九年与 Vincent 等人二〇二〇年进一步指出，受影响的不止值守者本人，还包括与之同床的伴侣。

这些结果对本章是三重打击。

第一重：本章的核心现象不是新的。“可以随时被叫醒却恰好没有被叫醒”这句话，本章第九节当作最典型的悬答眠提出；Vincent 等人那 56% 就是它的问卷版本，而 Torsvall 与 Åkerstedt 的脑电版本比它早了近四十年。

第二重：本章第十四节第一条预测已经被人做过。本章预测交拍应先于内转；Wuyts 等人二〇一二年在《Biological Psychology》上直接操纵了这一点——**只在睡前给出模拟的值守指令**，随后的睡眠即发生改变。这是一个更干净的因果设计，方向与本章一致，而本章却把它写成待做检验。

第三重，也是最有意思的一重：这一支里有一个结果反过来剽本章一句话。Jay、Aisbett 与 Ferguson 二〇一六年用一个 105 分贝的闹钟唤醒值守中的被试，并让被试明确知道漏掉这个闹钟的概率极低；结果**睡眠效率、总睡眠时间与各时期长在值守夜与非值守夜之间没有差异**。本章第八节把外托写成“关键回应的即时位置移出自身”，重心全在“有没有人接手”；这个实验说明，在没有任何人接手的条件下，仅仅把**返还通道做得足够可靠**，也可能足以让身体允许内转。于是本章必须让步：三环之中，返拍不是外托的附属条款，它可能是一条与外托部分可替代的独立路径。本稿据此把第八节“三环闭合”的表述限定为“三者共同构成完整形态”，而不再主张每一环都不可或缺；何种组合足以闭合，是一个经验问题，不是定义问题。

那么本章还剩什么？三条分离线，都写成同案不同预测：

其一，**变量的形状不同**。值守睡眠研究用的是二分变量：这一夜 on-call 还是不 on-call。本章用的是按事件类别枚举的连续读数 R。二者在同一批人身上会分开：两名同为“on-call”的医生，一名有分级响应与后备、只有预设严重等级才返还给他，另一名是所有电话的唯一接收者——在二分变量上二人同码，在 R 上一个接近 1、一个等于 0。若在充分样本中 R 相对二分变量不提供任何增量预测，本章这一条应当撤销。

其二，**记录制度的缺口是独立主张**。值守研究的数据里有排班与呼叫记录，却少有同轴对齐的睡眠分期；睡眠实验室的数据里有分期，却没有职责枚举与告警路由。本章主张这两本账至今没有共同的事件标识——这一条不依赖托拍本体成立，也不依赖任何效应量，它可以被一份公开数据集直接推翻。

其三，**适用范围不同**。值守睡眠是一个职业健康范畴，其分析单位是受雇者与轮班制度。本章把同一结构推到家庭照护、重症患者、平台劳动、不安全居所与非人动物，并主张在这些场景里“最后回应者”同样可枚举。这一推广可能失败：若离开雇佣关系后责任路由无法可靠编码，本章应收缩回职业场景，与既有文献合流。

最后一件必须交代的事，直接损害本章的一条制度赌注。Bumpstead 等人二〇二四年在《Sleep Medicine Reviews》上清点 65 份关于值守工作的指导材料，归纳出八类策略，其中第八类就叫**呼叫管理**。也就是说，本章第十九节第八条那个“到 2030 年至少有一份指南加入相关字段”的赌注，写得太松——这类字段在监管与行业指南里已有一整个类别。该综述同时指出，这些策略普遍缺乏支持其有效性的经验证据。本稿因此把赌注整体重写，见第十九节第八条修订版。

第一，行为睡眠定义依据静止、反应阈值、可逆性与补偿反弹判定睡眠。它说到身体是否进入睡眠；托拍理论另判外部回应是否闭合。若两段行为和脑电相同、责任配置不同却所有后续指标完全相同，行为定义足以吸收本章；若责任配置产生独立差异，二者分开。

第二，双过程模型解释睡眠压力与昼夜相位。它预测在相同清醒史和生物钟下睡眠倾向相近；托拍理论预测明确交接仍会改变责任相关信号加工与睡眠闭合。若控制双过程变量后外托无任何增量，本章退回环境修饰。

第三，“睡眠用于修复、清除或突触稳态”把睡眠功能放在内部收益。托拍理论不与这些功能争胜，而判断功能何时有机会不被外部请求打断。若完全相同的内在功能指标足以解释全部结果，外托没有独立位置；若功能指标正常而悬答结构仍预测差异，二者不能互换。

第四，Schwartz 的周期性豁免与 Aubert—White 的社会撤退已经说明睡眠暂停社会参与。它们可能正面占位“睡眠是一种脱责”。本章的分离线是：正式豁免存在但下一次回应无人接手时，前者可判定角色已暂退，本章判定事件未闭合。若这种案例不存在，或不产生任何可测差异，本章被经典社会学吸收。

第五，Meadows 的协商之夜说明伴侣通过谈判共同形成睡眠。托拍理论把协商结果压成可核查的责任路由。两对伴侣可以同样报告“协商良好”，但一对有明确事件持有人，另一对依赖“谁先听见谁处理”。若二者结果无差异，协商概念足够；若责任路由比总体关系质量更有预测力，托拍获得独立性。

第六，睡眠机会或卧床时间衡量可供睡眠的时段。它预测机会相同则潜在睡眠资源相同；托拍理论预测机会相同而最后回应者不同，睡眠事件仍不同。实际中断次数为零的夜晚是最锋利检验，因为它排除了单纯时间损失。

第七，社会时差描述生物时间与社会时间错位。它预测工作日与自由日的睡眠中点差异带来后果；托拍理论预测在相同作息与时差下，回应托管仍有增量。若托管效应只在时差大时出现并被时差完全解释，本章不应独立命名。

第八，高唤醒与警觉理论把睡眠困难归于持续的生理或认知激活。托拍理论允许睡者主观平静、当夜无中断，却因责任未转交而属于悬答眠。若没有主观或生理高唤醒就从不出现任何效应，高唤醒概念可覆盖本章；若责任路由在低焦虑者中仍有作用，二者分离。

第九，Winnicott 式“承托环境”强调安全、照料和被保持的经验。它与“托”最接近，但主要解释心理发展与关系容纳。托拍事件要求列出具体回应、内部睡眠轨道与返还条件。一个人感到被支持，却仍是所有夜间事件的唯一最后回应者，承托环境可判良好，本章仍判悬答。若这类反例不存在，本章只是承托概念的睡眠版。

第十，节律同步、音乐条件化和声音掩蔽解释音乐如何影响入睡。它们通常预测特定声学、重复配对或噪声覆盖产生效果；托拍理论预测在声学 and 熟悉性不变时，声音被赋予“交接完成”或“紧急返还”的不同位置，会产生相反作用。若角色赋值不改变任何结果，音乐维度应由既有同步与条件化理论接管。

表 6 十种相近解释与同案分离线

相近说法	它已经说到哪一步	同案分离线
行为睡眠标准	判定身体是否睡着	相同脑电与时长下，责任配置不同是否仍有差异
双过程模型	解释睡眠压力与昼夜时相	控制清醒史与时相后，交接是否仍改变责任相关加工
修复／清除／突触稳态	解释睡眠的内部功能	功能指标正常时，悬答结构是否仍有独立后果

周期性豁免／社会撤退	说明角色义务暂时松开	正式获准休息但无人接手时，是否已构成完整睡眠
协商的夜晚	说明伴侣共同安排睡眠	总体协商感相同，具体责任路由是否更有预测力
睡眠机会／卧床时间	测量可用于睡眠的时间	机会相同且实际中断为零，托管差异是否仍有效
社会时差	测量社会时间与生物时间错位	作息与时差相同时，R 是否有增量
高唤醒／警觉	解释持续认知或生理激活	低焦虑者中，未托管责任是否仍形成独立类型
承托环境	解释安全与被保持的体验	主观被支持但仍是唯一最后回应者时，两理论判定相反
音乐同步／条件化	解释声学和学习如何影响入睡	声学和熟悉性不变，只改“交拍／返拍”角色是否改变结果

十八·十二 同一作者已发论述的自撞申报

以上分界处理的都是本领域之外的解释。还有一类近邻更危险，因为它们出自同一位作者、发表在同一个站点上：读者不会把它们当竞争理论，只会把本章当重复。初稿一处也没有申报，本稿补齐五处，并逐条给出分离线。

其一，失眠专文。该文主张睡眠是撤退而非动作，并给出“守夜人”这一形象：睡者派出自己的意志去监督入睡，而入睡要求意识退场，监督入睡本身却需要意识在场。它与本章的关系不是相邻，是**对手**：那篇的守夜人是内部的、监督入睡本身的；本章的值守是外部事件的、与入睡过程无关的。分离线是同案不同预测——一位对入睡毫无监督焦虑、上床即睡的照护者，在那篇的框架里没有问题，在本章的框架里若他仍是所有夜间事件的最后回应者，就属于悬答眠。这一条已在第十四节第二个竞争解释里写成可检验形式。

其二，《交出去的不止那件事》。该文的判断是：每一次委托都同时委托掉了对这次委托的监督权，因为验收能力本来就是“自己做”的副产品，而这是唯一一种不会被发现的损失。它与本章**方向相反**：那篇说交出去有代价，本章说交不出去才是病。二者可以同时成立，因为对象不同——那篇交出的是**能力**，本章交出去的是**回应位置**；能力交出去会萎缩，回应位置交出去恰恰是睡眠得以发生的条件。这一对照给出本章一条边界：托拍工程只应转移回应位置，不应连同判断一起转移。第二十三节的第三项原则“持有”因此要加一句——替代持有人须拥有完成一级行动的信息、权限与能力，但事后的复核与规则修订权不随之转移，否则本章的处方会踩进那篇所描述的损失。

其三，《挂着》。该文命名的“挂起”指判决已启动却未落定、判决权交给一个永不到场的“以后”、且没有一个不能再推的到期日。它与“悬答”连词面都接近，必须切开：挂起是**交出去了却回不来**，悬答是**根本没有交出去**；挂起没有到期日，悬答有明确的起止（计划睡眠区间）。两者的读数也相反：挂起的病灶在返还无期，悬答的病灶在移交未发生。若某个案例既无移交又无到期日，那不是本章的对象，应按那篇处理。

其四，《叫不来的那一半》。该文主张有一类过程不认命令只读信号，而信号有两种、方向相反——修复类等的是“安全了”，创造类等的是“没路了”，并且以失眠为三个来源之一。本章的外托几乎就是“安全了”这个信号的一种制度实现。分离线在于：那篇问的是信号内容（等哪一句话），本章问的是信号的来源结构（那句话由谁有资格说、说完之后责任落在谁身上）。一个人可以主观感到“安全了”而实际上仍是最后回应者；本章预测这种情形下内转仍受损，那篇的框架不预测这一点。

其五，本书前三章。前三篇分别处理判读单位的追认、判读窗与状态之间的界带、以及独判格式的判余，三者都在追问睡眠**记录**的结构；本章追问的是睡眠**关系**的结构。共同点是“记录制度决定了什么能被看见”，这一条家族命题四篇共享，本章不主张对它有首创权。差别在于：前三篇的缺口在同一份记录内部，本章的缺口在两份互不登记的记录之间。第一篇与本章使用了同一组三学科，须特别声明：两篇的材料用法不同，且两篇的对象不同（一篇是段落，一篇是回应位置），但读者有权把它们读成同一套方法的两次应用，本章不回避这一判断。

十九·机制证伪：八条会让本章翻车的观察

第一条检验控制总工作量、睡眠机会、昼夜相位、基线焦虑、住房噪声和实际夜间中断。在同一组织或家庭内，若“下一拍托管率”与睡眠潜伏期、夜间自主神经活动、责任相关刺激反应或次日功能之间不存在剂量关系，外托不是独立机制，相关困境应归于负荷或焦虑。

第二条检验信号选择性。控制响度、频谱、熟悉度和睡眠阶段后，若与个人责任相关的声音不比等强度无关声音更易被加工或唤醒，返拍通道不按责任映射，本章关于选择性返还的主张为伪。

第三条检验悬答眠的独立性。只选实际中断为零、总睡眠时长与分期匹配的夜晚，若责任已转交与未转交两组在多夜累积的生理、认知和情绪结果上没有任何差异，悬答眠只是未兑现的风险，不构成独立睡眠类型。

第四条检验先后。若责任转交总是在睡眠指标改善之后发生，或交接时间变化从不早于睡眠变化，本章把因果方向写反；外托可能只是睡得好之后的主观解释或组织记录。

第五条检验跨物种必要性。在同一物种、相近能量与风险条件下，若没有哨兵、庇护、低风险时间窗或部分警戒转移的个体仍表现出同等稳定、同等深度且选择性返还不受损的睡眠，外托不能作为睡眠的一般组成，只能保留为人类社会睡眠变量。

第六条检验概念冗余。若下一拍托管率在多个独立样本中与社会支持、时间在床、实际中断或安全感任一既有指标的相关始终高于 0.90，且不提供任何独特预测，它只是旧指标的重命名，应当取消。

第七条已经用公开数据执行。若至少 70% 的独立睡眠歌单曲目同时满足宽松的慢速、低能量、低响度和高器乐性条件，且激活型亚群低于 10%，本章不得否定固定镇静声学；实际复算得到 30.37% 与规模可观的相对激活簇，因此这一竞争命题未通过。但该结果只排除声学充分性，不验证托拍因果。

第八条是一项写死日期的制度赌注。初稿的版本写得太松，本稿重写并收紧。**旧版**：到 2030 年 12 月 31 日，至少一个公开可核的指南将在休息安排中加入“替代回应者、可延期事件或分级返还”的字段。**为什么必须重写**：Bumpstead 等人二〇二四年清点的 65 份值守指导材料中，已经有一整类策略叫“呼叫管理”，旧版赌注在提出之日很可能已经命中，因而不构成任何风险。

新版赌注：到 2030 年 12 月 31 日，至少一份公开可核的指南或标准，将要求在同一份记录上**同时**登记以下三项，并把三者绑定到同一段计划睡眠区间：（一）该区间内可要求当事人即时回应的**事件类别清单**；（二）每一类别的**替代持有人或可执行延期规则**；（三）该区间内当事人的**睡眠或休息记录**。

不算命中的情形有五条：一，上述 65 份指南里既有的呼叫管理条款，无论其后续修订版本；二，只写“减少呼叫次数”“合理安排轮班”“保证休息时长”“建议关闭手机”一类不指名事件类别的表述；三，三项中只落实其中一项或两项，或三项分存于不同文件而无共同标识；四，仅在研究方案中出现而未进入任何指南、标准或报告格式；五，由本章作者或本站推动写入的条款。

若届时无一命中，说明把回应去向与睡眠记录绑到同一时间轴这件事，在制度上要么不必要，要么不可行，本章的第三层主张（记录缺口）应当撤回。

若这些检验推翻本章，我们仍会学到重要东西：睡眠的外部关系可能只通过实际中断起作用，责任意义可能不进入睡眠感觉加工，或者人类社会睡眠与跨物种睡眠需要两套不同本体。证伪不是把问题关掉，而是把干预重新指向负荷、节律、声学或临床机制。

表 7 机制证伪矩阵

编号	控制后观察	使本章失败的结果
F1	负荷、时相、焦虑、噪声、实际中断	R 与任何睡眠／次日指标均无剂量关系
F2	响度、频谱、熟悉度、睡眠阶段	责任相关声音无独立选择性
F3	实际中断为零、时长与分期匹配	已托管与未托管完全无差异
F4	交接与睡眠变化的时间序列	睡眠总是先改善，交接从不领先
F5	同物种、同风险与能量条件	无任何外部风险转移仍有同等完整睡眠
F6	多样本指标比较	R 与既有单一指标始终 $r>0.90$ 且无增量

F7	公开睡眠歌单真跑	≥70%曲目满足固定镇静形态（实测未发生）
F8	截至 2030-12-31 的公开指南	无指南加入替代回应者／延期／分级返还字段

二十·下一拍托管率：一份可复核的清点手册

为了让外托不退化成“感觉有人照应”，本章提出一个辅助读数：下一拍托管率，记为 R 。它不是睡眠本身，也不用于给个人打分；它只描述某一计划睡眠区间内，关键即时回应义务被系统接住的比例。定义为： $R = \text{已有明确替代持有人、自动处置或可执行延期规则的关键回应类别数} / \text{除以该区间全部预先枚举的关键回应类别数}$ 。 R 取 0 到 1 的连续值； $R = 1$ 表示枚举范围内全部闭合， $0 < R < 1$ 表示部分托管， $R = 0$ 表示无托管。

“关键回应类别”必须同时写明四项：触发事件、期望行动、最迟回应时间和未回应后果。例如“孩子任何声音”太宽，需拆成普通翻身、持续哭声、呼吸报警与发热；“工作消息”也需拆成可次日处理的普通询问、规定时限内的客户升级与安全事故。未写明后果的泛化担忧不进入分母，以防把无穷可能全部算成义务。

编码规则一：替代持有人必须能够完成行动或作出一级判断；只会把通知转给睡者的设备不计入分子。规则二：口头的“有事再说”不算可执行延期，必须有明确时限或组织政策。规则三：覆盖必须贯穿整个计划睡眠区间；只覆盖前半夜按比例拆分为两个时段记录，不得整夜算一次。规则四：同一事件链只计一次，避免报警、电话和处置被重复计数。规则五：返还信号另列字段，不进入 R 的分子；它记录何种条件结束托管。规则六：任何分类争议都保留原始文字与复核人，不得只保留最终数值。

边界例一：自动门锁在异常时只向睡者手机报警，不算托管；若先由保安核验，只有确认入侵才唤回睡者，则“异常门禁”得到托管。边界例二：伴侣说“我今晚照顾孩子”，但没有对发热、喂药与急诊作出分工，可按普通安抚已托管、医疗判断未托管。边界例三：医院告警算法自动静音低优先级提示，但没有人承担复核，不能因为声音消失就算托管。

R 的空表至少包含：日期、计划睡眠区间、事件类别、触发条件、最迟响应、默认责任人、替代持有人或规则、覆盖起止、返还信号、实际是否发生、编码依据、复核结果。研究正文、证伪条件和制度赌注均使用同一 R 定义，不在不同章节更换阈值。悬答眠的操作判定为：内转证据成立且 $R < 1$ ，并至少存在一个未托管类别要求睡者在睡眠区间内即时回应。

这个读数必须与常用指标保持正交。一个人可以拥有八小时、90% 的睡眠效率和 $R = 0$ ；也可以只有短暂午睡、睡眠效率一般，却在该区间 $R = 1$ 。若实证发现 R 总能由时长或实际中断完全推回，读数即失去存在理由。

二十·七 分母怎么定：类别粒度规则

上面六条编码规则管的是分子——什么算已托管。分母同样能被操纵，而初稿只防了一个方向。“未写明后果的泛化担忧不进入分母”防的是分母被吹大；没有防的是分母被压小：把“孩子任何声音”整个算作一类，与拆成普通翻身、持续哭声、呼吸报警、发热四类，会在同一个夜晚给出完全不同的 R 。若伴侣接手前两类，前一种算法得 0 或 1，后一种得 0.5。这不是估计误差，是构念本身可被编码者移动。而 R 直接进入第十九节第一条的剂量关系、第六条的冗余检验与第八条的制度赌注——分母可动，这三条证伪就都松动。本稿因此补一条粒度规则。

成类判据：一个候选事件类别成立，当且仅当它的四项字段——触发事件、期望行动、最迟回应时间、未回应后果——**都能被独立填写，且不能由已列出的任何其他类别蕴含**。四项中任一项只能写成“视情况而定”，该候选不成类，须继续拆分或并入上位类别；若两个候选在四项上完全一致，二者合并为一类。判据只谈可填写性，不谈重要性，以免编码者用“这条更要紧”来调整分母。

双人枚举与两个 R ：分母由两名不知道本研究假设的编码者独立枚举。报告三个数：交集分母下的 R （保守）、并集分母下的 R （宽松）、以及两份清单的 Jaccard 相似度。**分析一律以保守 R 为主，**

宽松 R 作敏感性分析，两者都必须公开。若某一夜的两个 R 落在悬答眠判定线（ $R < 1$ 且存在至少一个未托管的即时回应类别）的两侧，该夜记为不可判定，进入报告但不进入主分析；不可判定夜的比例本身是编码手册质量的读数，超过 15% 即应重写手册而不是继续分析。

这条规则也给出一个新的失败方式，须一并写明：若在多个独立样本中，保守 R 与宽松 R 的差值系统性地大于托管完整组与不完整组之间的差值，那么本读数的分辨力低于它自身的编码噪声，R 应当废弃，本章退回到只使用二分的“有无未托管类别”这一粗判定。

表 8 R 的最小记录表头

字段	填写要求
计划睡眠区间	明确起止；跨夜按同一时区记录
事件类别	不可用“所有事情”；须拆成可区分触发
期望行动与时限	写明谁做什么、最迟何时做
未回应后果	安全、照护、经济或制度后果
默认责任人	未交接时实际最后回应者
替代持有人／规则	能完成行动、一级判断或可执行延期
覆盖起止	不得以部分覆盖冒充整夜
返还信号	什么条件结束托管
编码依据与复核	保留原始文本、日志与复核意见

二十一·睡眠、休息、麻醉、昏迷与梦：托拍定义的边界测试

一个定义是否锋利，不在它能解释多少，而在它敢排除什么。安静清醒具备外托却没有内转：一个人坐在窗边读书，所有职责已经交班，身体仍维持清醒组织。这是受保护的休息，不是睡眠。疲劳性微睡眠则可能有短暂内转，却来不及形成稳定外托与返拍；它更接近瞬时悬答，尤其在驾驶和监控任务中危险。

麻醉最容易与睡眠混淆。手术团队完全接手外部事件，药物也让患者失去反应，但麻醉是否具有自然睡眠的稳态历史、阶段组织和选择性返还，必须独立检验。若药物状态只能由撤药或医疗干预结束，返还通道并不属于患者自身的睡眠循环。托拍定义因此不把“有人照管的无响应”一概算成睡眠。

昏迷与不可逆损伤缺少迅速可逆性，也缺少对高显著信号的正常选择性返还。它们可以有照护托管，却没有睡眠内转。冬眠和蛰伏具有代谢降低与环境适应，但时间尺度、体温调节和稳态历史不同；是否属于广义托拍需要另建物种标准，不能从人类夜眠直接外推。

梦并不反驳托拍，反而说明睡眠不是空白。快速眼动睡眠中内部世界高度活跃，外部信息既未完全消失，也不会原样进入。梦把身体从即时行动中抽离，却在内部重组记忆、情绪和模拟。选择性返还通道像一扇半开的门：有意义信号可被放大、压制或改写。睡眠完整性的关键不是“意识有没有内容”，而是内容活动是否在外托保护下展开，并能在需要时返回共同世界。

清醒梦与睡眠中双向交流进一步要求定义容纳“睡着而能回应”。少量、受控的回应不必取消睡眠，只要有机体仍处在睡眠内转，回应是被预设通道选择性放行，而不是恢复对所有外界要求的常规承担。这也说明降低反应不是绝对关闭，而是响应权限的重排。

托拍定义由此划出五条边界：没有内转的保护不是睡眠；没有可逆性的无响应不是睡眠；没有外托的内转是悬答眠；没有选择性返还的断联不是完整托拍；有受控回应的睡眠仍可成立。边界越清楚，新概念越不需要靠“真正的睡眠”这类不可裁决词语自保。

二十二·三项最小实验：从概念判断走向联合测量

第一项是“交班先行”自然实验。选择已经存在轮班差异、但不改变安全标准的同一医院科室、照护家庭或技术支持团队，采用被试内交叉设计。每名参与者经历明确替代持有人和仅名义休息两种夜晚；总工作量、床上时间、房间条件、咖啡因、昼夜时相与实际事件次数同步记录。交班内容在入睡前固定并留时间戳。主要结果不是主观满意度，而是交班之后到内转指标变化的时序、责任相关声音的脑电反应、夜间心率变异及次日持续注意。若交班从不领先这些变化，托拍方向受挫。

第二项是“同声异职”实验。选择两段声学参数和熟悉度匹配的声音，在训练阶段随机赋予不同制度角色：一段表示“普通事件已经由系统接手，可以继续睡”，另一段表示“只有出现高等级事件时回应权返回”。随后在受控睡眠中以等响度呈现，并在另一轮交换角色。实验只改变声音在责任序列中的位置，不改变曲目。若脑电加工、微觉醒或次日记忆随角色交换而改变，说明返拍不是固定声学；若只随音高、速度或响度变化，音乐关系命题失败。

第三项是“生态托管”跨物种实验。不能用两种差异巨大的动物作启发式对照，而应在同一物种、同一栖息地与相近能量状态中，利用自然存在的哨兵、庇护等级或风险时窗变异。连续记录至少六个群体或多个同质单元，测量睡眠行为、选择性唤醒、捕食风险与群体警戒分配。若庇护与哨兵只减少实际打扰，却不改变睡眠组织或信号选择性，外托可能只是风险降低；若在没有实际攻击的夜晚仍出现差异，托拍结构获得独立证据。

三项实验都必须先登记事件类别、主要终点和排除规则，再看结果。人类研究不把高风险值守者随机置于无后备状态，只利用既有合规安排或模拟任务；动物研究不人为提高捕食危险。样本量应由先验效应与重复测量结构估算，不能用两个国家、两家机构或两个家庭剂量关系。

联合测量的最低时间轴是：交拍记录时间、入睡与阶段变化时间、外部信号时间、实际响应与持有人、返拍时间、次日结果。只有把这些点放在同一轴上，才能区分“安排得好所以睡得好”“睡得好所以觉得安排可靠”与“交接先改变了选择性响应，随后内转稳定”三种机制。

这三项实验分别对应社会、音乐与生态材料，但共享同一判决：当声学、负荷、时相和实际事件被控制后，下一次回应的去向是否仍改变睡眠的时序与选择性。若答案持续为否，托拍事件应降级为描述性框架，而不是睡眠本体。

二十三·托拍工程：把“早点睡”改写为系统如何接手

多数睡眠建议以个人为动作单位：固定作息、减少咖啡因、降低光照、离开屏幕、练习放松。这些做法在很多情形下有效，但它们默认个人有权关闭接口。对值守者、照护者、平台劳动者和不安全居住者，真正缺失的不是睡前知识，而是没有任何系统接住其缺席。托拍工程把干预单位从“让个人更会睡”改为“让请求在个人睡着时有别的地方”。

第一项设计原则是拦截。关键事件先进入团队、设备或规则，而不是先进入睡者再由睡者决定是否处理。第二项是分级。低价值、可延期和高风险事件使用不同路径，避免每个通知都具有返拍权。第三项是持有。替代者不仅收信，还必须拥有信息、权限与能力完成一级行动。第四项是返还。什么条件把任务交回睡者要预先定义，不能由任何人随时打破边界。第五项是审计。名义交接与真实路由要通过日志比对，失效后改规则而不是责怪睡者。

家庭中的托拍工程可以很小：把“今晚你负责孩子”拆成普通安抚、喂药、体温阈值和急诊决定，并写明何时叫醒另一方。组织中的托拍工程可以是告警路由、时区接力、升级阈值和休息期间的免责规则。技术产品中的托拍工程不是更强的免打扰，而是让系统能够安全处置、可靠延期和透明升级。

这一路线改变睡眠服务的评价。酒店、医院、养老机构或远程工作平台不只提供床、安静和时长，还应说明夜间事件由谁处理。一个房间再安静，若患者必须自己识别每次设备异常，睡眠服务仍未闭合。一个应用再会播放白噪声，若错过消息仍带来惩罚，它只是遮住责任，没有托住责任。

托拍工程不能替代医学。呼吸暂停、疼痛、抑郁、神经疾病和药物问题需要相应诊疗；系统交接也不能使任何人无限睡眠。它解决的是一个更窄而此前无字段的问题：身体准备暂时降低外部响应时，环境是否真正完成接手。窄问题一旦被写清，许多“意志不够”“睡眠卫生差”的归因会被重新分配。

最理想的托拍系统不是让睡者对世界毫无感觉，而是减少无意义返还、保留必要返还，并让每次边界破裂都有记录。它保护的不是沉睡本身，而是睡眠与世界之间一条可逆、可追责的交接通道。

二十四·三层可独立引用的主张：核心倒下时，哪些部分仍然站立

一篇关于“何谓睡眠”的文章最容易把所有判断捆成一束：新定义若被反驳，数据、工具和制度建议一起倒下。本章把主张拆成三层，使每一层拥有自己的证据与失败条件。

第一层是最强、也最容易失败的本体主张：完整睡眠是一种托拍事件，外托、内转与返拍共同构成其最小单位；悬答眠是与完整托拍眠不同的一类存在。这一层要求证明责任去向在控制时长、时相、负荷、焦虑、实际中断和临床机制后仍有独立作用，并且能够解释选择性信号加工与时序。若 F1—F5 持续失败，第一层应撤回。

第二层是分析工具：身体内转与下一拍托管必须分别测量，二维格可以区分受保护清醒、暴露清醒、完整托拍眠与悬答眠。即使“悬答眠”最终没有独立生理后果，这张格仍可能帮助组织发现名义休息与真实值守的错位。第二层的失败条件是，它不能改变任何案例分类，或所有分类都能由睡眠机会和实际中断无损替代。

第三层是最稳的经验与记录主张：现有睡眠账本通常没有同步记录关键回应类别、替代持有人与返还条件；公开睡眠音乐数据也不支持单一镇静声学形态覆盖绝大多数曲目。这一层不依赖托拍本体成立。即使未来证明外托不影响睡眠生理，研究者仍需承认“没有发生中断”与“中断责任已被转交”是两种不同的制度事实；音乐研究也仍需处理睡眠歌单内部的显著异质性。

分层之后，引用者必须说明自己使用哪一层。临床研究若只采用 R 作为流程变量，不能顺手声称发现新睡眠类型；社会学研究若用悬答眠描述值守关系，不能把它当作已经证实的神经状态；音乐研究若引用公开数据复算，只能否定声学充分性，不能证明交拍—返拍机制。

这种分层也保护概念免于自我封顶。本章不因为提出了可证伪条件就自动成为正确理论。最强主张可能倒下，较弱工具仍可保留；较弱工具若也无用，公开数据与记录缺口仍可被独立复核。知识不是一座只能整体接受或整体拒绝的建筑，而是不同承重层分别接受风压。

对后续九篇睡眠研究而言，这一纪律尤其重要。同一问题从 What、How 和 Why 展开时，不能让后续路径与动力文章把本章新名当作不可动的前提。它们应把托拍事件与悬答眠视为待检验对象：

How 文章要给出能改变 R 与内转时序的路径，Why 文章要说明哪一组冲突先推动交接、内转或返还。若后两类研究逼迫本章收缩定义，那不是体系失守，而是概念获得了真实边界。

三层主张还对应三种不同的使用风险。第一层最容易被媒体写成“科学已经重新定义睡眠”，因此每次引用都应附带“待联合测量验证”；第二层最容易被机构变成考核表，因此必须保留伦理与复核；第三层最容易被夸大成因果，因此必须重复数据只能排除什么。把证据等级写进概念的使用说明，能够防止一个有辨别力的判断在传播中重新变成没有边界的口号。

真正的典范并不是一句永不修改的定义，而是一套允许后来者指出它在哪一层失败的结构。托拍事件若最终只在照护与值守场景成立，就应从一般睡眠本体收缩为一种社会—生理耦合类型；若跨物种实验也成立，它才有资格扩展。本章把这条升降通道预先写明，不以概念的新颖性替代证据的等级。

二十五·伦理与反噬：一旦成为考核指标，托拍最容易被伪造

下一拍托管率指向位置与流程，不指向人的品德。R 低不能被解释为个人不善安排、照护者不够坚强或员工抗压不足。它首先问系统为什么仍把最后回应留在这个位置。任何机构不得用 R 筛选“适合值夜的人”，也不得据此惩罚那些必须承担照护的人。

安全关键系统不得为了把 R 做高而随意轮换、制造陌生接手者或降低报警敏感度。航空、急诊、核设施、儿童照护等场景中，错误交接可能比睡眠不足更危险。托管必须服从能力、连续性与安全冗余；无法安全转交时，制度应承认这段时间仍是值守，而不是把它包装为休息。

已经依据 R 做出排班、福利或临床安排的机构，必须公开事件分类与编码规则，并提供复核通道。睡者应能看到哪些义务被列入分母、为什么某项被判为已托管、返还条件由谁设定。没有透明规则的数字只会把责任从可见冲突转移到不可见编码。

理论本身还有一个反噬预言：一旦“托拍”成为绩效要求，最省事的做法不是实际接手，而是在文件上指定一个名义替代者，或把原本关键事件重新分类为“无需即时回应”。这样 R 会提高，真实

安全却下降。本章看不到一个完全消除这种博弈的出口；能做的只有保留原始事件日志、随机复核、让接手者确认其能力，并把实际返还与失效事件公开纳入审计。

音乐干预也有伦理边界。不能把播放列表当作廉价替代，让高负荷个体在未转交责任时“更快睡着”；这可能使其在危险结构中更深地失去响应。任何助眠声音的研究都应同时记录谁承担夜间事件，并允许参与者随时退出。声音可以帮助标记边界，不能替制度完成边界。

二十六·自反与证据边界：这篇文章自己由谁接住下一拍

本章批评睡眠研究没有记录“谁接住下一次回应”，但本章自己的公开数据复算并没有这一字段。睡眠歌单数据只有音频特征与聚类，不知道听者何时播放、谁在值守、是否真正入睡。按照本章判据，这份证据只能位于音乐声学充分性的反证层，不能被拿来证明托拍事件本身。

化学证据也有边界。许多关键研究来自果蝇或小鼠，测量尺度从单细胞转录、胶质钙信号到全脑示踪不等；它们支持“睡眠是主动且多路径的内部重组”，却不能直接证明人类责任转交改变某一分子通路。把跨物种机制写成已经证实的人类因果，会违反本章自己要求的尺度一致。

社会学材料最接近外托，却大多通过访谈、日记、伴侣关系和规范分析呈现。它们清楚显示睡眠被协商、性别化和制度化，但很少同步多导睡眠与精细告警日志。因此，托拍事件目前是一条可检验的命题，而不是已经完成临床认证的变量。它需要新的联合数据：睡眠分期、职责枚举、告警路由、信号意义与实际事件按同一时间轴记录。

本章自身也落在一个“待接手”的位置。作者提出定义、清点规则与反证条件，但不能由提出者同时完成独立确认。下一拍应由不同学科的研究者接住：睡眠实验室检验责任相关声音，组织研究者审计交班，音乐心理学家改变声音的制度角色，动物行为学家检验生态托管。若没有人按可复算方式接手，本章只是一个被写下却没有进入下一轮判断的候选。

证据的最稳一层只有两点。第一，单一内部、社会或声学维度都不足以独占睡眠定义；这一点由各领域自己的反例支持。第二，现有记录普遍缺少“睡眠区间内关键回应由谁承担”的联合字段。更强的主张——这一字段构成新的睡眠类型并具有独立生理后果——仍须接受前述机制检验。

还有一层要认账，而且它正好落在本章自己的判断上。初稿漏掉了值守睡眠这一整支文献——从一九八八年的船舶轮机员脑电研究算起，三十八年，涵盖实验室方案、系统综述与伴侣影响研究。这不是检索疏漏可以打发的：本章通篇主张“身体睡着而人仍在值守”是一类没有账本字段的存在，而它在另一个账本上有整整一栏。

这件事本身是本章命题的一个实例，而且是比初稿任何一个场景都更干净的实例。一件东西在一本账上没有一行，不必是因为它不存在，也可以是因为它被完整地记在另一本账上，而两本账没有共同的事件标识。睡眠医学的账本记分期与时长，职业健康的账本记轮班与呼叫；两边各自记得平，都能出报告，都能通过同行评议，而“同一段睡眠里回应位置在谁手上”这一栏，谁也不必写。本章最初正是站在睡眠医学那一侧读文献，于是把另一本账上的一整栏读成了空白。

按本章自己在第九节末给出的判据，这也说明初稿犯的错与它所描述的失效是同一种：它的失效不产生信号。漏掉一整支文献的稿子与没有漏掉的稿子，在纸面上完全一样——摘要同样完整，分界节同样有十条，证伪条件同样有八项。它被发现，靠的正是本章推荐的那个动作：换一本账，重查一遍。

这一层认账不能用来给本章加分。它只把本章的位置说准：本章不是发现了一个没人看见的现象，而是指出了两本都看得很清楚的账之间，缺一根把它们对齐的时间轴。

结论：睡眠不是世界停止，而是世界暂时改变了向谁索要回答

我们习惯把睡眠看成一个人关上眼睛之后发生在身体里的事。于是，社会条件被放在原因栏，化学变化被放在机制栏，音乐被放在干预栏，而睡眠本身仍被假定为一个独立的内部状态。本章提出另一幅图景：睡眠从交拍开始，在内转中持续，由返拍结束。它是一段由身体与世界共同完成的托拍事件。

这一图景没有否定行为、生理、分子、社会或声学定义，而是重新规定它们各自的位置。行为和化学确认身体是否进入睡眠轨道；社会关系确认关键回应是否找到替代持有人；音乐学提供缺席仍被时间保持、并在特定点返回的结构。任何一维单独都只能看见部分睡眠。

最重要的结果是“悬答眠”的显形：身体已经入睡，下一次有后果的回应却仍无人接手。它可能整夜没有被打扰，可能通过时长、效率和分期检查，也可能在制度上被算作休息。正因为没有事件发生，它的缺口反而最不容易被发现。新的判据不问“昨夜醒了几次”，而问“若事件发生，谁必须起来”。

公开睡眠音乐数据的复算又提醒我们，睡眠没有唯一声学外形。慢与轻是总体倾向，不是本体。下一拍可以被一首快歌、一段讲话、一道门锁规则、一名同事、一个伴侣、一处巢穴或一个自然时窗接住。不同材料共同指向的不是某种固定内容，而是一种关系：睡者暂时不再是所有事情的最后接口，同时仍保留回到世界的路。

因此，对“何谓睡眠”的回答可以压缩成一句话：睡眠不是世界没有下一拍，而是下一拍在睡者不回答时仍被可靠地接住，并在适当时刻交还。谁有能力完成这次交接，谁才真正拥有一段完整的夜晚。

附录 A 公开睡眠音乐数据复算协议与完整读数

数据文件：SPD_unique_withClusters.csv；分析单位：唯一曲目；数据文件 SHA-256：0efd754f2ce0ee46802245c0b09728f795f50481c275d479a9c2c6392e587d7c。预注册文件 SHA-256：02631aa55718d63162ea054dfcfc5b528afd2ea5aa42fef984c2153fe054e3df。

主检验在读取结果前写死四项条件：tempo $\leq$ 90 BPM；energy $\leq$ 0.40；loudness $\leq$ -15 dB；instrumentalness $\geq$ 0.50。四项同时满足才计为“宽松镇静型”。放行线为独立曲目的70%。实际 N=130,150，满足数 39,523，覆盖率 0.3036726854。

敏感性分析遍历 tempo $\in$ 、energy $\in$ 、loudness $\in$ 、instrumentalness $\in$ ，共 81 组。覆盖率区间 0.2082289666—0.3763042643。全样本中位数：tempo=96.5905，energy=0.154，loudness=-19.2465，instrumentalness=0.862。

六簇曲目数分别为 5,437、20,108、17,373、61,771、21,969、3,492。按预注册规则，簇均值的 tempo、energy 和 loudness 同时高于全样本中位数者标为“相对激活型”；第 1、2、3、5、6 簇满足，合计占 0.5253860930。该标签只用于反驳单一低激活形态，不表示曲目在个体听者中必然产生生理激活。

代理变量与限制：Spotify 音频特征由专有算法生成；tempo 存在半拍/双拍识别问题；歌单收录不等于实际助眠效果；没有听者、文化、使用场景、播放时点、责任配置与睡眠结局字段。复算因此不支持因果推断。

附录 B 托拍事件的最小联合研究设计

零情态判据：从这段入睡开始到醒来，哪些事件仍可要求此人回应；这些事件由谁接手；哪一个信号把回应权交还给此人？

最小研究设计建议：在同一机构、同一家庭类型或同一动物群体内部，连续记录至少 14 个睡眠区间；同步采集睡眠阶段或物种适用的行为指标、事件职责清单、交接时间戳、告警路由、实际事件、返还信号与次日功能。优先使用自然发生的轮班或交接变异，避免人为制造安全风险。

预分析顺序：先固定关键事件编码；再计算 R；再排除实际中断或把其作为独立变量；最后检验 R 是否在睡眠机会、时相、负荷、焦虑、噪声与实际中断之外提供增量。若事件分类在看到结果后改变，本轮分析仅作探索，不得充作证伪。

参考文献

- Cage, J. (1961). *Silence: Lectures and Writings*. Wesleyan University Press.
Crary, J. (2013). *24/7: Late Capitalism and the Ends of Sleep*. Verso.

Hasty, C. F. (1997). *Meter as Rhythm*. Oxford University Press.

Lefebvre, H. (2004). *Rhythmanalysis: Space, Time and Everyday Life*. Continuum.

作者与人机协作声明

研究问题、核心概念、理论判断、适用边界与最终学术责任归作者王德生。人工智能用于文献检索辅助、公开数据复算、结构检查与语言编辑。本章不以生成过程替代证据，不把作者参与的内部评估当作独立认证；所有引文、数据与推断均须由后续同行评议和可复算研究继续核验。

版本与字数：v1.0 | 2026 年 8 月 9 日；中文核心文本约 20,174 个汉字，含标点约 23,423 字符（参考文献与英文摘要不计）。

修订说明（v1.0 – v2.0）

本稿相对初稿有九处实质改动，逐条列明，便于对照。

一、把最近的一位排到分界节第一位，并公开承认初稿漏了它。新增第十八节第一条，处理值守睡眠这一支（Torsvall 与 Åkerstedt 1988 脑电研究起，至 Bumpstead 等 2024 指南综述），承认本章核心现象不是新的、本章第十四节第一条预测已被 Wuyts 等 2012 做过，并给出三条同案不同预测的分离线。

二、接受该支文献反过来削本章的一句话。Jay、Aisbett 与 Ferguson 2016 的高可靠闹钟实验显示，在无人接手的条件下，仅把返还通道做可靠也可能足以让身体允许内转。据此把第八节“三环闭环”由定义性主张降为“三者共同构成完整形态”，何种组合足以闭合改判为经验问题。

三、收窄本章主张。新增第九节第五条，把“身体睡着而人仍值守”“选择性唤醒”“陌生环境值守”三块地明确划给既有主人，本章只保留一句：同一段睡眠上，回应去向与身体内转是两笔各自成立而互不登记的账。

四、补最强的竞争解释。新增第十四节第六条，列出陌生度（Tamaki 等 2016）、焦虑中介（本站失眠专文的守夜人机制）、负荷与预期唤醒时间三条更省事的解释，各给分岔点，并指明第二条是本章最容易输掉的一条。

五、音乐维度降格。新增第十七节第五条，声明音乐学在本稿只承担时间语法的命名、不承担经验证据；三环改称“移交—内转—归还”全章预测不变，故不应占材料位。唯一能使其重新承重的是第二十二节第二项实验。

六、更正第十一节复算的身份。该复算是对 Scarratt 等人已有结论的一次预注册复现，不是判决性检验；撤回初稿“比原论文更不利于本章早期设想”一句。

七、补分母类别的粒度规则。新增第二十节第七条：四字段独立可填且互不蕴含才成类；双人独立枚举，报保守 R、宽松 R 与 Jaccard；判定跨线的夜次记为不可判定；并给出一条新的失败方式（编码噪声大于组间差则 R 废弃）。

八、重写制度赌注。第十九节第八条旧版在提出之日很可能已命中（65 份指南已有“呼叫管理”一类），故收紧为“三项同轴登记”，并列五条不算命中的情形。

九、申报同一作者已发论述的自撞五处。新增第十八节第十二条，逐条给出与失眠专文、《交出去的不止那件事》、《挂着》、《叫不来的那一半》及本书前三章的分离线；其中《交出去的不止那件事》的对照使第二十三节“持有”原则补入一句：替代持有人获得行动权限，但复核与规则修订权不随之转移。

另在第二十六节新增一段自反：初稿漏掉整支文献这件事，本身是本章命题的实例——一件东西在一本账上没有一行，可以只是因为被完整记在另一本账上，而两本账没有共同的事件标识。此项不用于给本章加分，只用于把本章的位置说准。

第十三章 没有人违约

没有人违约

——失眠里那一类从未订立、双方不知、可被单方面撤走的支承

摘要 失眠通常被当作一个人自己的事：他的入睡潜伏期、他的信念、他的多基因评分。三种互不相容的看法在这一点上却一致——园林营造之学只看一样东西此刻显露成什么样，遗传之学认为它由一条可继承的规则决定，《道德经》主张属性不为任何一物所持有。三家争的是判据向哪儿取，共同假定的是取的对象是这一个人。本章用遗传学自己的材料推翻这条假定：失眠方差里最大的一块落在非共享环境，由元分析双生子相关算得约 64%。这一块自 1987 年起被判为随机、不可系统研究，而这个判断是在从未向“另一个人”采集过的前提下作出的。本章据此提出一类被三家账本同时漏掉的对象：**默借**——甲的睡眠持续依赖乙的某项例行行为，乙不知道自己在提供，甲不知道自己在依赖；它没有订立时刻，可被单方面终止，终止时两侧都不产生记录。判据是一句可拿去问记录的问话：**这个人开始睡不着的那一周，他身边有谁改变了自己的日常安排？**读数为**他源率 ϕ** 。一次预注册的清点显示：八份权威失眠工具与诊断标准共 78 条计分条目中，问“他人自身行为”的为 0 条，四种编码口径下不动。两处不利结果写入正文并改硬了命题。

关键词 默借；默撤；他源率；失眠；评估工具的信息源；非共享环境

一·三件事从三个方向来

一位造园者站在一片荒地上，判断这里能不能成园。他不问这块地的来历，也不问土里有什么。他只问一件事：站在将来那座亭子的位置往外看，看上去是什么。若远处那座塔恰好落在两株老树之间，这片地就成立了；若塔不在，一样的土、一样的水、一样的匠人，园子就不成。四百年前，计成把这条规矩写进了《园冶》：**园虽别内外，得景则无拘远近**。园子的地界与景的地界不是同一件东西。

一位遗传学者面对一位睡不着的人，问的是完全不同的问题。他不关心此刻这个人在什么位置看什么，他关心的是这个人身上带着的那套编码。2019 年，一项汇集了一百三十三万人的研究在失眠上定位了 202 个位点；2022 年扩到 554 个。双生子研究几十年来给出的遗传度在四成上下。这门学问的立场很清楚：一个人容不容易睡不着，是一件可以在他出生那天就开始算的事。

两千多年前，一个不知道会有失眠这个诊断的人写下了另一句话：**天下皆知美之为美，斯恶已**。定出一个好，同时就造出了一个坏；有和无一起生出来，长和短一起量出来，前和后一起排出来。属性不是一样东西自己带着的行李，是它和它的对头一起被生出来的那个关系。按这句话，向一个人身上去找“他的失眠”，方向从一开始就不对。

这三种说法不能同时成立。造园之学要求一座好园子必须让人看不出它是被谁怎样做出来的——“虽由人作，宛自天开”；而遗传之学的整个事业是让每一处显露都能回溯到它的编码，看不出来路在那里叫做缺口，不叫成就。造园之学的借景要求把不属于自己的东西算作自己的，而《道德经》第二章说，属性根本不为任何东西所持有，谈不上算作谁的。遗传之学把失眠定位为这个人持有的一项性状，而按第二章，没有任何性状是被持有的。

三家吵的是同一个问题：一个人今晚睡不着，这件事的判据该向哪儿去取。

而三家在吵的时候，有一件事谁也没有说出口。

二·三家没有说出口的那一件

把三家的分歧摆开，会看到它们的答案分别是“向他显露出来的样子取”“向组织他的那套规则取”“哪儿也取不到”。三个答案互不相容，但它们指着同一个东西。

取的对象，始终是这一个人。

园林之学说看它现在什么样，那个“它”是这个人；遗传之学说看它的编码，那个“它”是这个人；连《道德经》说“取不到”，说的也是从这个人身上取不到。三家争的是判据放在哪一维，共同假定的是**读数的载体是这个个体**。

这条假定念给三家听，三家都会觉得这还用说吗。它不是任何一家的立场，它是三家脚下的地板。

而失眠这件事上，恰恰有人已经把这块地板的裂缝写下来了——写的人是遗传之学自己，而且写了不少一次。

三·遗传学自己写下的那三句

第一句是1974年Lewontin写的。他在《美国人类遗传学杂志》上指出，遗传度是一个群体参数，不是个体参数：说某个性状的遗传度是四成，不等于说某一个人的这个性状有四成来自遗传。更要紧的是他随手带出的那个推论——**环境越均一，遗传度越高**。也就是说，这个数不是关于基因的，是关于**这批人被放在什么样的环境分布里的**。把同一批基因型放进另一种环境分布，同一个数会变。

第二句是1987年Plomin与Daniels写的。他们问：同一个家庭里的孩子为什么这么不一样？他们列出一种可能，并给它起了名字：“**一种令人沮丧的前景是，起作用的环境可能是一些不成系统的、各人各样的、偶然发生的事件**”——事故、疾病、别的创伤。他们自己接着说，这类事件对研究来说多半是死路一条，所以更值得先去找有系统的来源。

第三句是2000年Turkheimer写的，只有五个字：**那个令人沮丧的前景是真的**。二十年后，最初提出它的人自己也接受了：找了三十年系统性的非共享环境影响，没找到；那些影响就是不成系统的、各人各样的、偶然的，而且效果不持久。

现在把这三句话放到失眠上，做一次纯算术。

一项元分析给出失眠的同卵双生子相关约0.36、异卵约0.16。按标准公式，遗传的那一块是 $2 \times (0.36 - 0.16) = 0.40$ ；家庭共享的那一块是 $2 \times 0.16 - 0.36 = -0.04$ ，负数，按惯例记为零；**剩下的那一块是 $1 - 0.36 = 0.64$** 。

也就是说，在失眠这件事情上：

- 被逐个位点搜寻的那一块，四成；
- 被认为几乎不存在的那一块（同一个家庭共享的），零；
- 被判为随机、各人各样、不可系统研究的那一块，六成四。

再算一次兑现。一百三十三万人的全基因组研究解释了表型方差的2.6%。对着四成的遗传度，这是6.5%的兑现；对着女性的五成九，是4.4%。换个说法：**被判为“研究的死路”的那一块，是全基因组研究至今真正落实的那一块的二十四倍半**。

这三句话在遗传之学自己的领域里回答的是另外的问题。Lewontin那一句处理的是方差分解能不能用来讲因果，Plomin与Turkheimer那两句处理的是非共享环境能不能被测量。没有人把它们放在一起当作关于**读数的载体**的证据。

但它们合起来正是这个证据。因为那个“不成系统、各人各样、偶然”的结论，是在一种特定的采集方式下得到的：**双生子研究、同胞研究、领养研究，采集单元全都是被研究的那个人，或者他的同胞**。非共享环境按定义是“同胞之间不一样的那些东西”，而找它的办法是问同胞双方各自经历了什么。

从来没有一次设计，去问过**同胞之外的第三个人**：“你在这段时间里，改变了自己的哪些日常安排？”

“随机”这个结论，是在这样的前提下写下的。

四·借景：园的边界不是景的边界

回到那位造园者。

《园冶》里最要紧的一篇叫《借景》，开篇一句是：**夫借景，林园之最要者也**。计成把借分成五种：远借、邻借、仰借、俯借、应时而借。借什么？借远处那座山、隔壁那座塔、别人家那几株树、

天上那轮月、这个季节才有的那阵香。他写得很干脆：目力所及之处，俗的挡掉，好的收进来，不管它在谁的地界上，一概化作本园的烟景。

这是造园之学最漂亮的一手，也是它最不设防的一处。

因为那座塔是寺庙的。寺庙什么时候重修、什么时候翻新屋顶、什么时候在前面盖一栋楼，园主没有任何权利过问，通常也不会被告知。反过来，寺庙也不知道自己正在被一座园子借。它的账上没有“本寺正在为某园承担一处对景”这一条，因为寺庙的账是按寺庙的事写的。

于是就出现了一件在两侧的账上都不成立的事：**这座园子的成立，有一部分落在它自己的地界之外、落在一个不知道自己正在承担它的人手里。**

塔一拆，园子的景塌了。而园主的地契一寸未减，园图一笔未改，园记里那些“疏影横斜”“烟云浮动”的句子一个字都不必删——那些句子写的都是这座园子自己的东西。你拿着这座园子的全部文书去核，核不出任何变化。

要留意这里的一层区别，它是本章后面全部内容的接口。**借景是园主有意去借的**：他知道那座塔在，他把亭子的位置挪了三尺好让塔落在两树之间。他知道自己是在依赖。他不知道的只有一件——寺庙不知道自己在被依赖。

也就是说，四百年前这门学问已经处理了一半。它没处理的是另一半：**如果连借的那一方也不知道自己在借呢？**

五·第二章：属性不为任何一物所持有

《道德经》第二章说：有无相生，难易相成，长短相形，高下相倾，音声相和，前后相随。

这几句常被读成辩证的修辞，其实它说的是一件很硬的事：长这个属性不在长的那根棍子里，它在两根棍子之间；高这个属性不在高的那座山里，它在两座山之间。你把短的那根拿走，长的那根不是变成中等，是**不再是长的**——它身上什么也没有变。

第十一章把这条推得更远：车轮上三十根辐条聚到毂上，车之所以能用，用在毂中间那个空处；捏泥做器，器之所以能用，用在器中间那个空处。**有之以以为利，无之以以为用。**有用的那一部分，恰好是这样东西自己没有的那一部分。

第五十七章还有一句，老子用来反对治理，本章要用来说别的：**法令滋彰，盗贼多有。**规矩定得越细，被规矩点名的那一类事就越多。这句话通常被读作对严刑峻法的讥讽，但它同时是一个关于分类的判断：**分类不只是记录，分类还生产被分类的东西。**

这一家在本章里的位置需要交代清楚，否则后面的分量会被读错。《道德经》不提供任何经验证据，它不能被引用来支持或反驳任何一条可查的读数。它在本章里承担的是一个位置：**它是三家里唯一一家从一开始就否认“读数的载体是这个人”的。**遗传之学与造园之学都承认这个载体，只是对该向它的哪一面取读数意见相反。若没有第三家，前两家的争论会一直在同一块地板上进行，永远撞不出地板本身。

本章因此对这一家有一条限制：**它给的是位置，不是证据。**凡是本章提出的可查读数，一条也不以《道德经》为依据。

六·那一格是空的

现在把三家的东西并到一处。

造园之学有一个成熟的概念，讲一样东西的成立依赖于一个不属于它、也不受它控制的外部项——借景。它处理的情形是：借的一方知道自己在借，被借的一方不知道自己在被借。

遗传之学有一块最大的方差，讲**同胞之间那些不成系统、各人各样的事件**——非共享环境。它处理的情形是：这些事件按定义不能被研究者或其同胞报出来，于是被判为随机。

《道德经》有一条本体论，讲**属性不为任何一物所持有**。它处理的情形是：向单个东西取读数取不到。

三样东西合到一起，指着同一个位置。那个位置上有一类事情，形状是这样的：

甲的睡眠持续依赖乙的某项例行行为。乙不知道自己在提供它。甲不知道自己在依赖它。

举几个具体的。

一位男士十年来入睡很快。他不知道的是：同住的伴侣有起夜的习惯，每次路过客厅会顺手把他忘在茶几上的手机调成静音——不是为了他，是因为伴侣自己受不了那点亮光。伴侣从没说过，也不觉得这是一件事。伴侣出差三周，第一周他就开始半夜两点醒来，然后再也睡不回去。他去看医生，主诉是“最近三周入睡困难、维持困难，每周五夜以上”。伴侣回来之后，他的睡眠慢慢好了，而他把这归因为“那阵子工作压力大，后来缓过来了”。

一位老人二十年来睡得很沉。楼下那家便利店整夜亮着，把巷子照得能看清人影。他从没想过这件事，只是从来不需要为楼下的一点响动起身查看。便利店改成晚上十一点关门，两个月后老人开始夜里频繁醒来。便利店老板不知道自己曾经在为楼上的谁提供什么；老人也说不出发生了什么变化。

一位护士在值班表上一直排在第二顺位。第一顺位的那位同事从来没有漏接过一次呼叫，十一年一次都没有。她因此从不需要在夜里“半醒着”。那位同事离职之后，排班表上她的位置一格没动，工时一分钟没变，而她开始在凌晨反复惊醒。她填的问卷里，“工作压力”那一栏和上个月完全一样。

这三件事有一个共同的结构，而且这个结构不是比喻：

- 它没有订立时刻。不是某一天两个人商定的，是在从未被建立的情况下开始生效的。
- 它可以被单方面、无成本、无预告地终止。乙没有违约，因为从来没有约。
- 它终止时不产生任何信号。乙那一侧什么也没发生——伴侣只是出差，店主只是改了营业时间，同事只是换了工作；甲那一侧有事发生，但找不到原因。
- 两侧没有共同的事件标识。即使把两侧的全部记录摆在一张桌上，也没有任何一个字段能把“伴侣出差”和“第三周开始的入睡困难”连起来，因为两侧根本不是用同一种事件在记账。

本章把这类关系叫做**默借**，把它的终止叫做**默撤**。

七·为什么它不是“记录得少”

一个自然的反应是：这不过是记录不够细，把问诊做得更详细一点就能查到。

这个反应值得认真对待，因为它把本章的主张放在了一个错误的档次上。若默借只是“难查”，那么本章只是在推荐一种更勤快的问诊，谈不上是什么新东西。要说清它不是这一档，得看现成的几套本体论各自要求什么。

- 契约与承诺**要求有订立。有订立才有违约，有违约才有救济。默借没有订立时刻，因此没有任何一方违约，也没有任何一方可以被追究。
- 委托与交接**要求有一次交出去的动作，以及一条交还的通道。默借从来没有被交出去，因此也谈不上交还。
- 社会支持**要求提供者能把自己的行为识别为支持。这是它的采集方式决定的，后面会详说。默借的乙识别不出来。
- 依恋**要求依恋对象被指向。婴儿会寻找、会在分离时抗议——抗议正是依赖存在的证据。默借的甲不指向乙，撤走时不抗议，只是睡不着。
- 外部性**要求受影响的一方可被识别、可被定价、产权可被界定。这是它的整套处方（内部化）的前提。默借一旦被界定，它就不再是默借了——这一点在第十六节会变成本章最不堪的一处。

五套现成的本体论，没有一套给“双方都不知道的支承”留了位置。它不是被记录得少的东西，是在这几套账上压根不算一样东西。删掉本章提出的这个读数，它不会变得更难测量，它会不存在——因为把它当作一件事来看，本身就需要一个目前没有的分类。

这是本章与“应该问得更细一点”之间的分界线。问得更细，问的仍然是甲。而默借按定义甲答不出来。## 八·四个格子，其中三个已经有主人

把“知不知情”这一件事拆成两侧，就得到一张四格的表。本章要说的是：这张表上有三格是别人的地，只有一格是空的。

	接受方知道自己在依赖	接受方不知道自己在依赖
提供方知道自己在提供	委托与照护：有约、有交接、有交还	隐形支持：Bolger、Zuckerman 与

	信号。请人替自己守夜、雇夜班护工、把手机交给同伴保管。撤走时双方都知道，且通常有一次告别。	Kessler 在 2000 年命名。伴侣做了事而当事人没察觉，效果反而最好。撤走时提供方知道自己停了。
提供方不知道自己在提供	外部性与搭便车：Pigou 与 Coase 的地。受益方清楚自己在受益，只是没有付钱；处方是把它内部化——定价、征税、界定产权。	默借：两侧都不知道。没有约，没有交接，没有价格，没有可被指认的受益关系。撤走时两侧记录都不动。

左上那一格是本章之前的一篇文章处理过的：一个人能够睡着，往往是因为下一次有后果的回应已经被交给了别人。那种情形有交接、有替代持有人、有把回应权交还回来的信号。它与本章正好落在这张表的对角上——那一篇讲交出去了，本章讲从来没有交出去。

右上那一格值得多说几句，因为它是离本章最近的一位。

Bolger 等人 2000 年的那篇论文，处理的是一个当时的难题：人们普遍相信社会支持有好处，而实际收到支持的记录却查不出好处。他们的解释是，察觉到自己是在被帮助本身有情绪代价，因此**最有效的支持是接受者没有注意到的那一种**。他们用的是伴侣双方各自写的日记，结论建立在一个具体的比对上：**提供者报告了、而接受者没报告的那些事件**。

这个设计漂亮，也正因为漂亮，它把第四格挡在了门外。那个数据能存在，前提是提供者写下了“我今天为他做了什么”。**提供者不知道自己在提供的那一类，在这个范式里不是被漏掉了，是不可能被采集的**——日记上不会有那一行，因为写日记的人不认为发生过什么。

所以本章与隐形支持之间的分界不是“我更强调什么”，而是一条可以去判的线：

向提供方提一个不含“帮助”“支持”“照顾”字样的问题：**你在过去三个月改变了哪些日常安排？**若得到的条目里有一批，其时点与接受方症状起始的时点相匹配，而提供方被追问时说“我不知道这和他有什么关系”——则本章这一格成立，隐形支持的范式不覆盖它。若这些条目全部同时被提供方识别为“我原来是在帮他”，则默借的第二个条件不成立，本章这一格应当并入隐形支持。

第三种情形也要写清：若提供方报不出任何条目，只说明这一次没查到，不说明这一格是空的。

这一格的三个签名，逐条对照，缺一条本章的靶子就选错了：

签名一：它在短期表现为自立。默借是所有支承里最省事的一种——不必开口、不必道谢、不必回报、不占用任何一份人情账。一个人生活里的默借越多，他看上去越像一个不需要别人的人。这与前面那条经验发现是一致的：不被察觉的支持效果最好。**它不表现为脆弱，它表现为独立。**

签名二：它的失效不产生任何信号。乙那一侧没有一件事被记下来，因为乙不觉得发生了什么；甲那一侧记下了症状，却记不下原因。两份记录都完整，都通过审查，而它们之间少的那一条连线，不在任何一份表格的字段里。

签名三：它越正规化越严重。把失眠讲成个人的事——睡眠卫生、认知行为治疗、可穿戴设备的综合评分、多基因风险评分——每一样都在教这个人更仔细地观察自己。观察得越仔细，他越不会去问“我身边谁改了什么”；而问诊表也不问。**每一次评估流程的改进，都在把这一格推得更远。**

这三条要放在一起读。一个一眼就能看出来的坏格子（该做的没做、该问的没问）不需要任何新框架，谁都看得见。第四格通过全部形式审查：量表信度效度俱佳，诊断标准逻辑自洽，医生尽职，患者诚实，记录不假。

九·一句可以拿去问记录的问话

一条判断若只能拿去问人的看法，它就还不是判据。本章的判据是一句可以拿去问一份文件的问话：

这个人开始睡不着的那一周，他身边有谁改变了自己的日常安排？

这句话里没有“应当”，没有“有意义”，没有“充分”，没有“真正”。它只问一事实：谁、什么时候、改了什么。

把它拿到五个具体的地方去问，答案互不相同：

1. 门诊初诊问卷：没有这一栏。（这一条是查出来的，见下一节。）
2. 共识睡眠日记：有一栏可以写“昨夜发生了什么”，但填写者是患者本人；他能写下的是他察觉到的事。
3. 行为遗传学的非共享环境测量：问同胞双方各自经历了什么。第三方不在采集范围内。（这一条也是查出来的。）
4. 值守睡眠研究：有——这是唯一一支认真去查外部安排的文献，它研究的是“今晚要待命”这件事本身如何损害睡眠。但待命是**已经告知**的安排，落在这张表的左上格。
5. 园林修缮档案：寺塔的重修记录在寺庙，园记在园主。两份档案都保存完好，而它们没有共同的事件编号。

第1条和第3条不是想出来的，是查出来的，而且它们反过来改了本章——怎么改，下一节和第十一节交代。

十·一次预注册的清点

本章的第三层主张需要一个数：现行失眠评估里，究竟有多少条目问的是“患者之外的人做了什么”。

这个数可以当场清点。清点之前，本章把假设、纳入清单、编码规则、判定阈值和撤销规则一次写死并存档，之后不再改动。

纳入的八份（编码前一次定死，中途不增不删）：匹兹堡睡眠质量指数（PSQI）、失眠严重指数（ISI）、雅典失眠量表（AIS）、睡眠状况指标（SCI）、共识睡眠日记核心版（CSD）、DSM-5失眠障碍标准 A-F、ICSD-3 慢性失眠障碍标准 A-F、睡眠信念与态度量表（DBAS-16）。

读数（他源率 φ ）：

$\varphi = (\text{信息源是患者以外的人，且所问内容是那个人自己的行为或安排，且进入正式计分或正式判定的条目数}) \div (\text{全部计分条目数})$

清点手册：

项	写死的规定
分母粒度	诊断标准以一个字母条款为一条；量表以官方编号的一个可评分应答格为一条
边界例一	“你的睡眠问题在多大程度上被别人注意到”——信息源是患者，所问对象是患者自身。 不计入分子。
边界例二	“同床者是否观察到你打鼾”——信息源是他人，所问对象仍是患者自身。 不计入分子。
边界例三	PSQI 官方说明称“19 个自评条目”，逐格枚举得 18 条。差的一条来自自填项算一格还是两格。本章一律按“一个可评分应答格=一条”计，记 18，并在稳健性里把 19 也算一遍
不适用	不评价任何量表的心理测量学品质；不复制任何条目原文，只记录结构属性

预注册的三条假设与撤销规则：

- P1：八份工具中 $\varphi=0$ 的份数为 8。**撤销线：若 ≤ 6 ，本章第三层主张作废，不许换工具重试。**
- P2：至少有一份工具包含由他人填答的条目。（这一条若命中，是**对本章不利**的结果。）
- P3：在两套诊断标准中，涉及患者之外条件的条款，位置全部是排除条款。

清点结果：

工具	条目	计分条目	信息源为他人	所问为他人行为	φ
PSQI	23	18	5	0	0.0000
ISI	7	7	0	0	0.0000
AIS	8	8	0	0	0.0000
SCI	8	8	0	0	0.0000
CSD	9	9	0	0	0.0000

DSM-5	6	6	0	0	0.0000
ICSD-3	6	6	2	0	0.0000
DBAS-16	16	16	0	0	0.0000
合计	83	78	7	0	0.0000

P1 命中：八份全部为零，撤销线未触发。

口径稳健性。这个数会不会是编码规则做出来的？换三种都说得通的口径重算：

- 把 PSQI 按官方的”19 条”作分母：分母 79，分子仍是 0。
- 把”物理环境”也算作患者之外的内容：这类计分条目共 2 条（DSM-5 的 E 条与 ICSD-3 的 C 条），其中位于病因条款的 0 条。
- 把 PSQI 那 5 条同床者条目强行计入分母：分母 83，分子仍是 0，因为那 5 条问的是患者自己的身体表现。

四种口径下 φ 都不动。

这个数与既有指标的关系。 φ 与量表的信度、效度、被引次数彼此不相干，而且方向可以完全相反：PSQI 是这个领域被验证得最充分的工具之一， $\varphi=0$ ；一份心理测量学上很粗糙的自制表，只要加上两条”请与你同住的人回答：你在过去一个月改变了哪些日常安排”， φ 就大于 0。一个工具可以在所有既有指标上都很好看，而这个数是零。

十一·两处不利结果，以及它们把命题改硬的地方

预注册的意义在于，写下的预测有可能不成立。这一次有两处不成立，都写在这里。

不利结果一：P2 命中——本章不能说”没有人想过向第二个人取信息”。

PSQI 有 5 条由同床者或室友填答的条目。也就是说，1989 年这份量表被造出来的时候，作者们明确想到了要向第二个人取信息，并且把它做进了工具里。任何一篇声称”这个领域从来没想过问别人”的文章都会当场被这五条驳倒。

但接着看这五条被怎样安置。第一，按官方评分说明，**它们不计入总分**，只作临床参考。第二，也是更要紧的一点：**它们问的是患者自己的身体表现——打鼾、呼吸停顿、腿部抽动、迷糊、其他不安**。同床者在这里被当作一台观察患者的仪器，不是一个自己也在做事的人。

于是本章的命题从一个较弱、较响的版本，改成了一个更窄、更硬的版本：

不是”没有人想过向第二个人取信息”，而是——**问过第二个人的那一次，问的仍然是这个人。**

不利结果二：P3 只对了一半。

ICSD-3 的 A、B 两条明文写着，相关困难可以由患者报告，**或者由患者的父母或照护者观察到**。也就是说，第三方在诊断标准里是被正式承认的信息源。本章原来打算写的”信息源只能是患者”这句话不成立，撤回。

P3 的主体仍然成立，而且方式很清楚。DSM-5 的 E 条要求困难发生在**有充足睡眠机会**的前提下；ICSD-3 的 C 条要求主诉**不能纯粹由睡眠机会不足或环境不适解释**，并把”环境不适”具体化为四个词：安全、黑暗、安静、舒适。

请注意这四个词。它们全部是物理环境的静态属性。伴侣出差、便利店关门、值班表上第一顺位换了人——**这三件事没有一件会让房间变得不安全、不黑暗、不安静或不舒适**。它们不在 C 条的射程内，因此不会被这一条排除掉；同时它们也不在 A、B 两条的射程内，因为那两条问的是患者的困难，不是别人的行为。它们从两条之间掉了下去。

而更要紧的是这两条的**功能**：外部条件在这两套标准里唯一出现的位置是**排除条款**。它的作用是——如果外面的事解释得了，那就不算失眠。于是”失眠是什么”这个问题在被提出之前，答案的范围已经被这一条限定在这个人身上了。

综合这两处，本章的第三层主张收窄为一句可以直接去核的话：

八份权威工具与诊断标准共 78 条计分条目中，问”患者身边的人自己做了什么”的条目为 0 条；被允许发声的第三方，被允许说的仍然是关于患者的事。

这句话不依赖默借这个概念成不成立。即使读者完全不接受第六节的那类对象，这一句仍然是真的，而且今天就可以复核。

还有一件事必须记在这里。这两处结果都是清点做出来的，不是想出来的；它们各自削掉了本章一句更响亮的话，而留下来的那句更难被推翻。第一处把本章从”这个领域忽视了他入”改成”这个领域向他入问的是患者”；第二处把本章从”外部条件缺席”改成”外部条件在场，位置是排除项”。**两句改后的话都比改前更窄，也都比改前更硬。**

十二·与最近的几种说法的分界

一个新的说法最容易出的毛病，是它其实早有名字。下面逐条与最接近的几种说法划清，每一条都给一条能分出对错的线。

隐形支持（Bolger、Zuckerman、Kessler，2000）。上一节已详述：它的采集方式要求提供者能识别自己在提供。分界线在第八节，此处不重复。它是本章最近的一位，而它不能吸收本章，理由不是本章更深入，是它的数据来源在结构上取不到第四格。

外部性（Pigou，1920；Coase，1960）。经济学一百年前就处理了”一方的行为对另一方产生了未经交易的影响”。分界有两条。第一，外部性理论只要求受影响的一方能被识别，不要求施加影响的一方不知情；本章要求两侧同时不知情。第二，也是更关键的一条：**外部性理论的全部处方是内部化**——定价、征税、界定产权、让双方谈判。而默借一旦被内部化，它就不再是默借。这不是本章的巧辩，这是一条可判的预测：若把默借关系公开并加以安排，接受方的睡眠改善程度**不高于**未公开的对照组，则本章对；若公开并安排之后改善更大，则外部性框架对，本章的整套主张应当被它吸收。

共识主动性（Grassé，1959）。白蚁不互相通信，各自对环境中原有的痕迹作反应，于是巢就建起来了。这是”无需知情的协调”最经典的模型。分界线：共识主动性里的痕迹是**在同一件工作上**留下的，后来者把它读作任务信号；默借里的乙根本不在做任何与甲的睡眠有关的事，他的行为不承载信号，也不是留给任何人的。判定形式：若在一批案例中，被指认的提供方行为都能被解释为对同一项共同任务的贡献，则本类应当并入共识主动性；若这些行为与接受方的事务毫无共同任务，则不能。

安全基地（Bowlby，1969）。依恋对象在场，哪怕不互动，也使探索成为可能——这与”乙在场，甲才睡得着”形状极近。分界线在**指向**：依恋理论的核心观察指标是分离时的寻找与抗议，那正是依赖存在的证据。默借的甲不指向乙，撤走时不抗议，只是睡不着并且归因于别处。判定形式：若在默借被撤走的案例中，接受方在被提示后普遍报告”我一直知道我需要在”，则本类属于依恋；若他们报告的是”我从来没往那儿想过”，则不属于。

目的性社会行动的未预期后果（Merton，1936）。分界线很短：Merton 讲的是**行动者自己的行动**带来了他没预料到的后果；本章讲的是**别人的行动**构成了你的条件，而双方都不知道。

影子工作（Illich，1981）。未被计价、却是有酬劳动之前提的那部分劳动——自己去超市推车、自己填表、自己排队。分界线：影子工作有一个**施加者**，是被受益的体系推给个人的；默借没有施加者，乙没有被任何人要求提供任何东西。

撤走看还剩什么，这是方法上最近的一位。基因敲除、条件性敲除、机器学习的消融实验，共用同一个动作：把一个部件拿掉，看整体还成不成立。本章的默撤在形式上就是一次消融。分界线是**谁执行了这次撤走**：消融实验里执行者是研究者，撤走的时点、对象和范围都由他写在方案里；默撤的执行者是乙，他不知道自己在做实验，也没有人记下他撤了什么。判定形式：若这类事件都能在某一方的记录里找到一条对应的”变更登记”，则它是一次可追溯的移除，方法学上已有成套处理；若找不到，则现有的消融范式对它无能为力——因为消融的前提是知道自己拿掉了什么。

目标置换与手段吞噬目的这一族，以及”越努力越糟”这一类反转模板，本章都不落在里面。本章的形状不是反转，是**归属错配**：一件事的成立条件跨在两个主体上，而账按主体划。

生活事件研究（Holmes 与 Rahe 一路）。这是本章最强的竞争解释——其实不就是”最近发生了些事”吗？分界线是可判的：生活事件量表记录的是**患者本人经历并能报出的事件**。默借要求的是**他人的例行变更，且患者未察觉**。控制掉患者已报告的生活事件之后，若”仅由第三方报出的他人例行变

更”与失眠起始时点之间不存在关联，本章的机制为伪，届时应当归因于患者可报告的应激负荷，而不是归因于支承结构。

共睡与伴侣睡眠同步这一支（Troxel 一路）。它研究同床两人的睡眠如何互相影响。分界线：那一支的两个人互相知道对方在场，睡眠影响通过身体、动作、体温、呼吸直接发生；默借的提供方可以完全不在场，甚至素不相识——便利店的店主与楼上的老人从没见过面。

值守睡眠这一支。这是这道题上必须点名的实务正主，它研究“今晚可能被叫”如何独立地损害睡眠，哪怕整夜零呼叫。它与本章的关系是**正面对照而不是对手**：它证明了外部安排确实能独立影响睡眠，而它研究的外部安排是**写在排班表上、当事人明确知道的**。本章关心的那一类不在任何表上。

几处必须点名的正面对照。三个别的领域已经把“外部支承”做成了字段：园林之学有借景，并且把借来的项分成远、邻、仰、俯、应时五类；生态学有交错带；天文学有分级的曙暮光。它们各自都把“边界之外的那一块”当作一个可以命名、可以清点、可以立项研究的对象。失眠评估里没有对应的东西——不是因为没人想到边界之外，是因为可数的那些外部项全被写进了排除条款。

最后是与本章作者自己已发表的几篇的分界，这一栏最容易被略过，而它们其实是离本章最近的邻居。

一篇讲失眠为什么越想睡越睡不着的旧文，整篇处理的是**这个人自己**做了什么：他派出意志去监督入睡，而入睡要求意识退场。本章讲的是**别人停了什么**。判定形式：一个从不监督自己入睡、也不为睡眠焦虑的人，忽然开始失眠——那篇的机制预测这不会发生，本章预测可以发生。这里还须承认一处紧张：本章第十六节的处方一旦回传给患者，就会制造那篇所说的守夜人。这一点本章接受，并据此把处方限制成程序性的。

另一篇讲完整的睡眠需要有人接住下一拍，主张回应位置被转交出去之后身体才允许自己重组。那一篇与本章正好在第八节那张表的对角上：**它讲交出去了，本章讲从来没有交出去**。两篇合起来才是完整的一对。

还有一篇讲每一次委托都同时委托掉了验收能力，方向与本章相反：那篇里的关系是被订立的，代价是能力萎缩；本章里没有订立，代价是撤走时无从追溯。

再有一篇讲撤销一个名字之后，空着的那个位置本身是第三个位置——它没有内容，只有形状，因此针对内容的补救对它无效。本章的分界很短：**默借撤走之后，那个位置连形状也没有**，因为它从来没有被画来过。

最后，本章之前几篇讲睡眠记录的文章，缺口都在一份记录内部，或在同一个人的两本账之间。本章的缺口在**两个不同的人的记录之间**，而这两份记录没有共同的事件标识。

那么，是不是所有这些邻居底下有同一块地板？一部分是。以人为报告单位的那些学科——临床评估、行为遗传学、社会支持研究、精神诊断——共同把一件事当作给定：**一个原因要算得上原因，得能出现在某一方的自陈范围内**。承认有一类原因两侧都自陈不出来，这些学科的采集工具就同时失去分母。

但这块地板不是普遍的，本章不打算把它说得比实情更大。环境流行病学从来不靠自陈：空气监测站、行政数据、气象记录都不问任何人。社会流行病学处理收入、住房、政策，也不靠当事人报告。**这块地板只在以人为报告单位的学科里成立**——而失眠评估的全套工具，恰好一件不落地属于这一类。## 十二之二·在别的地方，同一格也认得出来

一个说法若只在提出它的那个题目上说得通，它多半是那个题目的一种修辞。这一节把第四格拿到十二个别的地方去，每一处都给一件**可以去查的具体事**，不是一句“这也很像”。

一、遗传学的动物实验。同笼伙伴是一个从不被记录的自变量。预测：同一批基因型的小鼠，在同笼伙伴构成不同的笼子里，睡眠表型的组间方差应当大于单独饲养时的方差；而现行报告只记品系、日龄、光照周期与温度，不记同笼伙伴在实验期间有没有被换过。这一条今天就能做，成本是记录一份换笼日志。

二、园林的保护档案。预测：取一批列入保护名录的历史园林，把它们的“景观退化”记录与园界之外三百米内的建设施工档案摆在一起，两份档案将无法按事件对齐。原因很直白：前者在文物或

园林管理单位，后者在建设与管理单位，二者没有共同的事件编号。若能对齐，说明借景依赖项已被某个制度记过一次，本章关于这一格的说法在园林上不成立。

三、**临床试验的安慰剂组**。安慰剂组里那些“自发缓解”的病例，一部分可能对应着某个同住者行为的恢复。预测：在失眠药物试验中加一条向同住者提的问题（只问他自己的日常安排有无变更），安慰剂组结局的组内方差会下降；若不下降，本章在这个场景里没有解释力。

四、**关键岗位的排班**。预测：某个关键岗位第一顺位人员离职后的三个月内，第二顺位人员的睡眠主诉上升，而其工时、排班表、薪酬记录一格未变。这一条可以用人事数据与员工健康问卷直接查，而且这两份数据通常已经存在于同一个机构里，缺的只是把它们按时间对齐。

五、**独居老人**。预测：独居老人的睡眠恶化事件，与其所在街区商铺营业时间、社区服务时段的变更之间存在时间关联。这两类数据一份在卫生系统，一份在市政与工商系统。

六、**政策性的外生冲击**。路灯改造、便利店营业时间调整、公交末班车时刻变更，都是外部安排的自然实验。预测：受影响街区的助眠药物处方量在变更后的一到三个月内上升，而未受影响的相邻街区不变。这是纯行政数据研究，不需要招募任何人。

七、**夜间照明的评估指标**。目前评估夜间照明改造，看的是能耗与治安。预测：把睡眠主诉加进去，会出现一批两个指标方向相反的分区——能耗降了、治安没变、而睡眠主诉上升。这类分区在现行评估里不会被发现，因为没有人同一张表上看这两件事。

八、**软件的依赖关系**。一个开源组件的维护者停止维护时，下游的故障记录会出现，而维护者那一侧没有任何事件。预测：把公开代码库的“废弃公告”时点与下游项目的问题记录时点对齐，会发现相当比例的下游故障发生在没有任何公告的停更之后。依赖图是画得出来的，而“谁在依赖我”这一格在维护者一侧没有字段。

九、**生态学的关键种移除**。这一处是正面对照而不是兑现：生态学已经把“边界之外的那一块”做成了字段——交错带有宽度、边缘物种有名字、关键种移除有成套的实验设计。预测：把同一套设计的字段表拿来对照失眠评估工具，生态学那边六项俱全，失眠这边五项为空。同一张清单，两个相反的结果，差别不在对象，在账本。

十、**教育**。预测：一位从不缺勤的班主任调离之后，班上某些学生的表现在一个学期内下降，而学生档案、课程表、师资配置记录均无变化——因为调离记录在人事，成绩记录在教务。

十一、**医疗质量**。预测：某个病区一位资深护士离职之后，该病区用药差错率的变化能被查到，而这次离职与这些差错在记录上不构成因果链：离职在人事系统，差错在质量控制系统，两者没有共同的事件标识。

十二、**家庭研究**。出差与两地分居的研究通常测“分离带来的压力”。预测：控制掉留守方报告的主观分离压力之后，出差方日常安排的具体条目（例如夜里起不起身、几点关灯）仍能预测留守方的睡眠变化。若控制之后这种预测消失，那么本章这一格在家庭场景里应当并入主观分离压力，不必单设。

十二处里，有六处所需的数据已经存在于某个机构的记录中，缺的只是把两份记录按时间对齐；有四处需要新加一份日志；只有两处需要新的实验。这个分布本身是一条信息：这一格之所以空着，多半不是因为难，是因为要对齐的那两份记录分属两个单位。## 十三·三家反过来限制本章

一个说法若只从三家取好处，不接受三家的约束，它就只是借了三张皮。这一节写三家各自削掉了本章原本打算写下的哪一句。

遗传之学削掉的那一句：本章原本会写“失眠的主因在他人”。

它写不成。同卵与异卵双生子的相关差是真的，遗传度四成是真的，两百多个位点是真的。这四成不会因为本章提出了一类新对象而缩小一分。本章因此把适用范围压到一个很窄的地方：它只主张，在那块被判为随机的六成四里，存在一个可以被系统辨认的亚类。它不主张这个亚类占那六成四的多大比例——这个比例至今没有任何人测过，也包括本章。它更不主张失眠主要是外部的。

这一削动的是适用范围，而且动得很实：删掉它之后，本章从一个关于失眠病因的主张，退成一个关于采集方式的主张。

造园之学削掉的那一句：本章原本会写“应当在问诊表里加一栏‘默借项清单’，让患者填写他依赖的外部条件”。

《园冶》的题词里有一句话挡住了它：**园有异宜，无成法，不可得而传也**。造园之学自己承认，它给不出可搬运的规格——每一处的合宜由那一处的条件定，写不成一份通用清单。借景更是如此：借什么、借得成借不成，只有站在那个位置上才知道。

把这条搬到本章上，结论很直接：**默借按定义不能由甲枚举**。他若能列出清单，那些项就不是默借。一份让患者填写“我依赖谁”的表格，只会生产两样东西——猜测和焦虑。而猜完之后他多半会去向被猜到的那个人求证，那一刻这段关系就变了性质。

于是本章的处方被这一削压成了纯程序性的一条，见第十六节。这一削动的是价值排序：**本章不主张“知道自己被谁支撑”是好的**。

《道德经》削掉的那一句：本章原本会把默借的数量当作一个可以比较人与人的指标。

第二章那句“皆知美之为美，斯恶已”直接堵住了这条路。把默借命名出来、数出来，就同时造出了“默借少的人”这一类，而这一类会立刻被读成“更自立”或者“更脆弱”——两种读法都会有人用，方向相反，而且都不是本章的意思。第五十七章那句“法令滋彰，盗贼多有”说的正是这件事：**分类不只是记录，分类还生产被分类的东西**。

这一削的后果写在第十七节的三条限制里，它们是硬的：**本章的读数指向的是一份工具，不是一个人**。

十四·不适用的边界

以下几种情形，本章管不着，写在这里以免被误用。

- **有明确诱因的失眠**。倒班、时差、疼痛、新用药物、丧亲、明确的应激事件——这些有现成的解释，也有现成的处置，本章不与它们竞争。
- **有交接的情形**。值班表、请人守夜、雇护工、把孩子交给伴侣带一夜。这些落在第八节那张表的左上格，已有成熟研究。
- **物理环境改变**。搬家、装修、噪声源出现或消失。这些是 ICSD-3 C 条要处理的东西，处理得也对。
- **提供方知情的支持**。落在右上格，属于隐形支持那一支。
- **儿童**。儿童的睡眠高度依赖照护者的安排，而那些安排通常是有意的、被照护者清楚知道的，因此不满足默借的第二个条件。把本章套到儿童身上会立刻退化成常识。
- **本章不主张默借是失眠最常见的成因**。它的出现频率至今没有任何数据，第十五节的 F4 就是为了得到第一个这样的数。

十五·可以使本章为假的条件

F1（已经跑过，且判了本章一部分错） 第十节那次清点。它同时验证了第三层主张，也推翻了本章两句更响的话。任何人可以用第十节写下的编码规则重跑，若得到 $\varphi > 0$ 的工具在三份以上，本章的第三层主张作废。

F2（正向的顺序读数） 在新发失眠的病例中，若“他人日常安排的变更”与“患者自身可查变更”都被采集，前者的时点应当早于后者。若外侧变更的中位时点不早于内侧，本判断为伪。顺序比相关难被巧合满足，这一条是本章最愿意被检验的。

F3（控制掉最强竞争解释之后的机制检验） 最强的竞争解释是“其实不就是最近压力大吗”。控制掉患者本人已报告的全部生活事件与应激负荷之后，若“仅由第三方报出的、患者未察觉的他人例行变更”与失眠起始之间不存在剂量—反应关系，则本机制为伪。样本要求：至少六例以上、在同一制度环境内（同一家门诊、同一地区、同一职业群体），不接受两个国家或两个地区之间的对比充数。

F4（一个当场就能做的最小实验） 在一家睡眠门诊连续三十例新发失眠中，经患者知情同意，向一位同住者或密切共事者提一个问题，不含“帮助”“支持”“照顾”字样：**你在过去三个月改变了**

哪些日常安排？统计有多少例得到至少一条时点落在症状起始前三十天内、且患者本人未曾提及的条目。若低于 10%，本章的经验落点不成立。这一条所需的全部资源是三十次电话。

F5 若这类条目全部同时被提供方识别为“我原来是在帮他”，则默借的第二个条件不成立，本章应当并入隐形支持。

F6 若把默借关系公开并加以安排之后，接受方的睡眠改善不低于未公开的对照组，则内部化有效，本章的整套主张应当被外部性框架吸收。

F7 若 $\varphi > 0$ 的工具（一旦有人加了这一栏）在预测失眠起病上并不优于 $\varphi = 0$ 的工具，则本章的第三层主张虽然事实上成立，但没有用处。这一条比前面几条更难通过，本章承认。

若本章被证伪，会学到什么。若 F3 或 F4 失败，学到的不是“外部条件不重要”，而是“**两侧都不知情**”这个筛子筛出来的东西太少，不值得单设一格——那么正确的做法是回到隐形支持那一格，把力气花在向提供方采集上，而这仍然比现在只问患者要多一个信息源。

十六·处方，以及它的反噬

被三家削过之后，本章能给的处方只剩一条，而且是程序性的：

在患者知情同意的前提下，把评估的信息源从一个人扩到两个人；向第二个人提的问题不问关于患者的任何事，只问他自己的日常安排在这段时间里有什么变更。

这条处方刻意不含任何内容上的指导。它不列清单，不给分类，不问“你有没有在帮他”。它只改采集单元。

现在写它的反噬，这是本章最难堪的一段。

一旦这条被写进流程，最省事的实现方式是让患者自己去填。因为向第二个人取信息麻烦、要额外的同意、要额外的时间，而在患者面前加一栏“请列出你依赖的外部条件”几乎不花成本。于是流程照样“实施”了，而填出来的是一份猜测清单。

接着，患者会拿着这份清单去求证。他会去问伴侣“你是不是每天晚上都会帮我把手手机静音”。那一刻，这段关系的第三个条件被摧毁了——他知道了。而按隐形支持那一支的实证，支持一旦变得可见，效果会下降。也就是说，本章最省事的实现方式，会摧毁本章要保护的那样东西。

更糟的一层：被问到的那个人，从此知道自己“在承担一件事”。他原本只是顺手关个灯，现在这变成了一项责任。他可能因此做得更周到，也可能因此觉得被绑住而停止——而无论哪一种，那件事都不再是原来那件事了。

我看不到出口。唯一在逻辑上说得通的做法是：只向提供方采集，不把采集内容回传给患者。但这在知情同意上几乎不可能成立——你要向一个人打听另一个人的日常，还不能告诉后者为什么问，也不能把结果告诉他。本章不认为这个做法可以推行，把它写在这里是为了说明：这一格之所以至今空着，可能不完全是疏忽，有一部分是因为凡是能把它填上的办法，都会同时把它毁掉。

这也是外部性那一节里那条分界的真正分量所在。经济学处理外部性的全部智慧是内部化，而这里有一类东西，内部化会消灭它。

十七·关于这个读数的三条限制

φ 是一个可以拿去比较的数，因此它一定会被拿去考核。以下三条是硬的：

一、这个读数指向的是一份工具，不是一个人。说某份量表的 $\varphi = 0$ ，是关于那份量表的条目结构的陈述。它不评价编制者，不评价使用它的医生，更不能被拿来算某个患者的“默借数量”——本章没有提出过任何个体层面的读数，也不主张有这样一个数。

二、不得为了把这个数做好看，而在评估流程里增设向他人提问的条目。一条为了抬高 φ 而加的条目，会让第二个人被例行询问而不知道为什么，这既不改善评估，也侵入了一个没有就诊的人的生活。 φ 是用来格式化的，不是用来做业绩的。

三、若已有机构据此改动流程并作出了对人不利的安排（例如把某类患者归为“外部依赖型”），必须公开编码规则并给出复核通道。本章提供的编码规则见第十节，可以完整复算；任何在此之外的分类都不是本章的。

十八·三层可以分开引用的主张

本章的内容不是一整块。它有三层，强度依次递减，而稳固程度依次递增。读者可以只接受其中任何一层。

第一层（最强，也最容易倒）：默借是一类存在物，它是失眠的一种独立发生方式。这一层依赖第六节那三个条件同时成立，也依赖第十五节 F4 那次尚未进行的清点。**在 F4 做出来之前，这一层是一个假设，不是一个发现。**

第二层（分析工具）：第八节那张四格表。它的价值不在于第四格，而在于四格里有三格已经有主人——把委托、隐形支持、外部性摆在同一张表上，本身就说明了它们各自处理的是“知情”的哪一种组合。即使第四格最终被证明是空的、或应当并入右上格，这张表仍然可用。

第三层（最稳，也最便宜）：八份工具、78 条计分条目、问他人自身行为的条目 0 条；被允许发声的第三方，被允许说的仍是关于患者的事；外部条件在两套诊断标准里唯一出现的位置是排除条款。这一层不依赖前两层。它今天就能复核，编码规则和边界例都写在第十节。

引用纪律一条：若有研究只采用了 φ 作为流程指标，不得顺带声称发现了一种新的失眠类型。第三层成立不蕴含第一层成立。

十九·本章自己的信息源

按本章自己的判据检本章一遍。

本章的全部材料来自公开文献、量表的公开评分说明和公开的诊断标准结构。本章没有向任何一个“被借的人”提过一个问题。第六节那三个例子是构造出来的，不是采集来的；第十五节 F4 那三十通电话，本章一通也没有打。

这意味着：**本章的信息源结构，与它所批评的那些工具完全相同。**它们只问患者，本章只问文献。本章对第四格的全部描述，都停留在第二层和第三层；第一层是一个至今没有被任何一次采集触及过的对象。

而这件事本身正是第八节签名二的发作：一篇没有采集过供给方的论文，读起来和采集过的完全一样。参考文献一样长，论证一样连贯，表格一样齐整。若不是本节写出来，读者没有任何办法从文本上分辨这两种论文——这恰恰是本章声称的那种失效方式：它不产生信号。

还有一层要认。本章的写成本身依赖着一批默借项。用来复算的公开数据库、期刊的开放摘要页、量表的公开评分说明、把这些东西挂在网上二十年而无人道谢的那些维护者——他们不知道本章在依赖他们，本章在动笔时也没有把他们记为条件。取消其中任何一项，本章都写不出来，而参考文献表一行不会变：文献标记的是被引用的作品，不是使这次引用成为可能的那些安排。

二十·一个写死日期的赌注

到 2031 年 12 月 31 日，至少有一份由国际睡眠医学学会或同级机构发布的成人失眠评估指南，包含一条向患者以外的人询问“该他人自身日常安排的变更”的条目，并且该条目进入正式记录。

写死什么不算命中：

1. 询问他人对患者睡眠的观察不算。PSQI 那五条与 ICSD-3 的照护者条款已经属于这一类，它们在 1989 年和 2014 年就存在了。
2. 询问患者“你身边有谁改变了什么”不算。信息源仍是患者。
3. 只出现在研究方案、共识声明或综述建议里不算，必须进入评估工具或诊断流程的正式条目。
4. 由本章作者或其所在机构推动写入的不算。

若到期未命中，本章第三层所描述的那个缺口就不只是一个疏忽，而是一件有人试过、没做成的事——那时应当去查为什么做不成，而第十六节那段反噬多半就是答案的一部分。

注释与材料

本章的清点。第十节的编码在任何工具被查看之前完成预注册，内容包括纳入清单、编码规则、读数定义、三条假设、判定阈值和撤销规则，之后未作改动。编码脚本与结果文件随文提供，任何人都可按第十节的手册重跑。清点对象是各工具的**结构属性**（条目数、信息源、所问对象、是否计分、条款位置），不涉及条目原文。

一处必须写明的局限。本章的清点覆盖八份工具，不是全部。选取标准是“在临床与研究中被广泛使用”，这个标准由本章自定，可以争议。若有人举出一份被广泛使用而 $\phi > 0$ 的工具，本章第三层的措辞应当相应收窄；举出三份，则第三层作废。

关于遗传学那几个数。双生子相关 0.36 与 0.16 取自公开发表的元分析，方差分解按标准公式计算，过程写在第三节，可以逐步复算。需要说明的是，1-r(同卵) 这一项同时包含非共享环境与测量误差，本章没有把两者分开，也没有能力分开；因此“六成四”是一个上界，不是一个精确的份额。这一点削弱了第三节的力度，但不改变它的方向：**这一块无论如何都比遗传的那一块大，也比全基因组研究已落实的那一块大一个数量级以上。**

关于《道德经》。引文依通行本章次（第二章、第十一章、第五十七章）。本章不参与关于文本年代与版本的争论，也不主张这些句子的原意是本章所用的意思。这一家在本章里承担的是一个位置，不是一条证据；本章提出的任何可查读数都不以它为依据，这一点在第五节已经写明。

关于《园冶》。引文出自明代计成《园冶》，成书于崇祯四年，另引郑元勋题词。所用的是“借景”“园无成法”这两条，不涉及该书的营造技术部分。

参考文献

- 12. Illich, I. (1981). *Shadow Work*. London: Marion Boyars.
- 17. Pigou, A. C. (1920). *The Economics of Welfare*. London: Macmillan.
- 27. 计成（明）：《园冶》，崇祯四年成书。郑元勋题词。
- 28. 《道德经》通行本，第二章、第十一章、第五十七章。

人机分工声明

本章的问题、三个来源的选定、承重判断与全部结论由作者确定。人工智能用于：文献检索与出处核对、第十节清点脚本的编写与执行、数值复算、结构检查与文字编辑。第十节的预注册内容由作者在清点之前确定；两处不利结果由清点直接产生，未经改写。文中所有可查读数均可按第十节的编码规则独立复算。

字数：正文约 1.6 万汉字。**版本：**第一稿。

第十四章 办法都在夜里

办法都在夜里

论睡眠干预点的分布由症状可见的位置决定，而不由因果强度决定，以及现有记录为何分不出哪个时点更要紧

Where the Remedies Sit: On Symptom-Site Bias in Sleep Intervention Timing

摘要

把所有关于睡眠的办法摆到一条二十四小时的时间轴上，会看见一个没有人解释过的分布：它们几乎全部堆在熄灯前后那两三个小时里。而睡眠压力在清醒的十六小时里累积，昼夜相位由白天见到的光设定。因在白天累积，果在夜里发生，办法压在果那一端。

三门学问给出三种解释：时间生物学说这是长滞后造成的常态误解；控制工程说这是一个被明令禁止的动作——纯滞后大的系统在输出端做反馈必然失稳；照明工程说条件场本身不是为授时造的。三家共同踩着一句没说出口的话：干预应当施加在问题出现的地方。推翻它的材料来自控制工程自己：输出端在一整类系统上根本不是合法的干预点。

本章用七十八名受试者的判读记录做了一次预注册检验，造了预测力—滞后剖面。三条先验预测全部命中，而本章一条也不据此立论：整条剖面是平的，年龄分层后最大只剩 0.18；而白天代理薄得可怜——八个窗里有七个的窗内睡眠中位数是零分钟。

本稿另补三件。其一，八个窗的年龄相关全部报出（前一稿只报两个近端窗）：熄灯前八至十小时那窗是 +0.417，几乎等于睡前两小时的 +0.433——不是“睡前打盹”特别是年龄的标志物，是任何时段的打盹都是。其二，算出三条预测在零假设下的命中概率：P1 99.98%、P2 99.06%，而撤销规则触发率只有 0.50%——前一稿的预注册在设计上就不可能失败。本章据此提出一个可带走的装置：风险度，并与 Meehl 的弱检验批评、Mayo 的严格性划界。其三，补上初稿整支漏掉的两位祖宗：公共卫生的上游／下游、观察偏误里的路灯效应。

落点：在现有记录的分辨率上因果强度分不出时点，而所有手段仍押在同一个时点上；那么这条分布不是被因果强度决定的，是被症状可见的位置决定的。本章把这类配置命名为症状位偏置，给出一张可核对任何建议清单的干预点清单（在近视与龋齿上各跑一遍），做十七条分界与四处反向削。

关键词：症状位偏置 | 滞后距 | 风险度 | 干预点清单 | 输出端反馈 | 上游与下游

一·先做一次清点

本章的起点不是一个理论，是一次数数。

把一个人能为睡眠做的事全部列出来，然后给每一条标上一个东西：它必须在几点钟被执行。

做法	必须执行的时点
服褪黑素	睡前 0.5-2 小时
服助眠药	睡前
睡前热水澡	睡前 1-2 小时
白噪音、耳塞	上床时
遮光窗帘	上床时
调低卧室温度	上床前
换床垫、换枕头	上床时
睡前拉伸、冥想、呼吸法	睡前
把手机放到客厅	睡前
睡前不看屏幕	睡前 1-2 小时
刺激控制：躺二十分钟睡不着就离床	夜里
数羊、身体扫描	床上
睡前不喝酒	傍晚起
下午之后不喝咖啡	午后
白天规律运动	白天
早晨见光	清晨
固定起床时间	清晨
睡眠限制：压缩卧床时长	全天日程

一共十八条。前十二条的横坐标全部落在熄灯前后两小时内，第十三条落在傍晚，只有最后五条落在白天。而这个比例还是偏乐观的——真正被人执行、被人购买、被人反复推荐的，绝大多数是前十二条。市面上有卖褪黑素的，有卖遮光帘的，有卖床垫的，有卖白噪音机的；没有人卖“早晨见光”。

现在把因那一边也摆上来。睡眠压力在清醒的十六个小时里单调累积；昼夜相位由视网膜见到的光设定，而光的作用取决于它落在相位的哪一段；核心体温曲线的下降沿由白天的活动与进食塑造。因在十六小时的那一侧，办法在两小时的这一侧。

这个分布本身就是本章要问的东西。它不是一个错误——本章第八节会说明它可能完全合理。它是一个从来没有人问过它由什么决定的事实。

二·三家对这条分布的三种解释

2.1 时间生物学：这是长滞后造成的常态误解

第一家的回答最直接：因与果隔得太远，人只看得见果。

这一家的硬材料是相位反应曲线。Khalsa 等人在《生理学杂志》上给出了人对单次强光脉冲的相位反应曲线（Khalsa, Jewett, Cajochen & Czeisler, *The Journal of Physiology* 549(3): 945–952, 2003）。这条曲线的形状本身就是本章要的证据：昼夜起搏器对视网膜光暴露的重置效应的敏感性，取决于光暴露发生在哪一个昼夜相位。后续文献把它的临床含义写得更直白：在生物学夜晚早段（傍晚）的光暴露造成相位延迟，把节律推后；而在深夜或清晨的光暴露造成相位提前，把节律拉早。这些相位移动在临床上是关键的：晨光被用来把延迟睡眠相位障碍的昼夜时点提前，而晚光被用来延迟提前睡眠相位障碍的节律。

同样东西，同样的强度，给在不同时刻，效果符号相反。这是一个纯粹的时点效应，而它的作用窗口全部落在夜里之外。

这一家的单因主张是：睡眠的好坏由昼夜相位与稳态压力这两个在白天被设定的量决定；夜里那两小时只是它们的显示屏。

它自曝的地方也写得很清楚：相位反应曲线是在受控实验室、暗光基线、精确相位测定的条件下测出来的。它不告诉你在自由生活中，一个人的白天光照史对当夜睡眠有多大的预测力。

2.2 控制工程：这不是误解，是一个被禁止的动作

第二家的回答比第一家重。它说：把调节动作施加在一个大滞后系统的输出端，不是效率低，是不稳定。

一个带纯滞后的被控对象，反馈回路里的相位滞后随频率线性增长；滞后一大，开环相位裕度被吃光，闭环必然振荡甚至发散。控制工程对此的标准处置不是“加大增益”或“更用心地调”，而是换结构：要么前馈，在扰动进入系统时就补偿掉它；要么滞后补偿，用被控对象的模型把延迟从反馈回路里“移出去”，再对无延迟的预测量做闭环——这就是史密斯预估器（Smith, 1957）。

这一家的单因主张是：能不能在某一点上有效地调节一个系统，由这一点相对于滞后环节的位置决定，与那一点上症状有多明显、有多痛苦、有多容易观察，全无关系。

它自曝的地方是本章第八节要用到的：控制工程的这条结论有一个前提——你得知道滞后有多大、模型是什么。模型不准的时候，史密斯预估器反而比普通反馈更糟。而人体睡眠的滞后结构，没有人有一个可用的模型。

2.3 照明工程：条件场根本不是为这件事造的

第三家把前两家的对象放回它们所在的建筑里。

现代室内光环境标准是按视觉作业制定的：多少照度能让人看清文件、辨认颜色、不产生眩光。办公室三百到五百勒克斯，对阅读绰绰有余。而昼夜系统的输入量级完全不同——户外阴天上万勒克斯，晴天数万；室内五百勒克斯对授时的作用接近于零。

这一家的单因主张是：一个人白天见到多少有效光，由他所处的建筑决定，而建筑是按看得清造的，不是按授时造的。因此白天那一端的自变量，对绝大多数人根本不是一个可以自由调节的变量——它被墙、窗、工位与通勤路线锁死了。

它自曝的地方很关键：这一家其实在替”办法都在夜里”辩护。如果白天那一端在结构上不可调，那么把办法压在夜里就不是错误，是理性。这一处第十节还要用。

2.4 三家两两之间是真打架

时间生物学与控制工程打的是这条分布的性质。前者说这是认识错误——人不知道因在白天，教育可以纠正；后者说这是结构错误——就算你完全知道因在白天，只要你的调节动作施加在输出端，它照样不稳定。一个是知识问题，一个是拓扑问题，处方完全不同。

控制工程与照明工程打的是可调性。控制工程说：把动作移到扰动进入的那一端去。照明工程说：那一端不在你手里，它在物业和建筑规范手里。前者假定输入端可达，后者说输入端被锁死——而”输入端是否可达”正是前馈控制能否成立的前提。

照明工程与时间生物学打的是谁在解释谁。照明工程说人的光照史是建筑的函数；时间生物学说建筑之所以那样造，是因为人只按视觉需要提要求，而人之所以只提视觉需要，是因为他不知道自己还有一个授时系统。因果箭头相反。

顺带挡掉第四家。临床指南、睡眠卫生手册、可穿戴设备的建议列表不构成一种立场——它们是上面三种主张各自落地的场所，也正是第一节那张表的来源。本章后面反复使用公开判读记录，那是兑现材料。

三·三家共同踩着的那块地

把三家的争论压成一句：三家争的是这条分布该由什么解释——由认识、由结构、还是由建筑。而”干预点”这样东西，三家都默认它是这样一种东西：

干预应当施加在问题出现的地方。

时间生物学默认它，所以它把”办法都在夜里”读成一个需要被纠正的错误——纠正的方向是把干预点移到因那一边，而这仍然假定干预点应当由因果强度决定。照明工程默认它，所以它替夜里那一端辩护的理由是”另一端不可调”——这也仍然是在按因果强度和可达性排序。

念给谁听都会说这还用说吗。而这正是三家共同踩着的那句话的验收标志。

四·控制工程自己拆掉了它

从外面搬一条理由来推翻它，只会得到第四种立场。要取消它，材料必须是三家之一自己早已印出来、只是没往这个方向使的东西。

它在控制工程那里，而且不是边角料，是任何一本过程控制教材都会写的一条：

对一个纯滞后占主导的对象，在输出端做反馈，不是效果差一点，是会失稳。

请注意这句话的形状。它说的不是”在输出端调节效率较低”，也不是”最好把干预点前移”。它说的是：输出端这个位置，在一整类系统上根本不是一个合法的干预点——不管那里的症状多明显，不管那里多容易观察，不管在那里动手多方便。

于是三家共同踩着的那句话塌了一半：干预点的合法性不由”问题在哪里出现”决定。一个问题出现的地方，可以恰好是最不该动手的地方。

另一半由照明工程拆掉。它说白天那一端在结构上不可调。如果这是真的，那么干预点也不由”因在哪里”决定——因在哪里和你能不能在那里动手，是两件事。

两半都塌之后，问题就换了：

如果干预点既不由症状的位置决定，也不由因的位置决定，那么现有的这条分布，究竟是由什么决定的？

五·症状位偏置，与一张可以拿去核对的清单

5.1 承重命题（先给出，第七、八节会按数据改写）

睡眠干预点的分布，不是按因果强度排出来的，也不是按可调性排出来的，而是按**症状可见的位置**排出来的：夜里是唯一有人醒着抱怨的时段，于是全部手段都落在那里。本章把这种配置命名为**症状位偏置**。

5.2 一张干预点清单

要判断一份建议清单有没有症状位偏置，本章不给分类格，给一张逐条核对的清单。对清单上每一条建议问五个问题：

① **它必须在几点被执行？**（写下钟点，不接受“每天”）② **它作用的那个量，在几点被设定？**（写下钟点）③ **①与②相差几小时？**（这是这条建议的**滞后距**）④ **它有没有一个可以在执行当时观察到的效果？**（有 → 它落在症状位）⑤ **有没有人卖它？**（有 → 它落在症状位的可能性更高）

对第一节那十八条逐条核对，得到的结果是：滞后距为零的十二条，全部有可当场观察的效果，其中九条有商品；滞后距在八小时以上的五条，没有一条有当场可观察的效果，其中一条有商品（遮光/助眠灯具，而它恰恰又被搬回了夜里）。

这张清单本身就是本章的**辨别装置**，它可以被拿去核对任何一份健康建议：不只是睡眠。滞后距为零而有当场效果的条目占压倒多数，就是**症状位偏置**。

5.2.1 这个类不是为睡眠定制的：另外两个成员

一个只有一个成员的类，是一个改了名的个例。本章举两个别处的成员，读者可以自己把五问跑一遍。

近视。症状位那一侧：眼镜、护眼台灯、坐姿矫正器、眼保健操——滞后距为零，当场有可观察的效果（看得清了、坐直了），且**全都有人卖**。而这些年证据最硬的那一条是**白天的户外时间**：它作用于一个在白天被设定的量，执行时看不见任何效果，效果要以年计才显形，而且**没有人卖**“每天两小时户外”。这个例子与本章的“早晨见光”结构完全相同——**同一个物理量（光）、同一个时段（白天）、同样没有商品**。

龋齿。症状位那一侧：牙膏、漱口水、电动牙刷、牙线——滞后距接近零，当场有可观察效果（口气、光滑感），且是一个巨大的市场。而公共卫生上效应最大的两条是**含糖食物的进食频次与饮水加氟**：前者作用的量分散在一整天，后者干脆不由个人执行，两者执行时都毫无可观察效果，也都没有对应的消费品。

两个例子都不在睡眠里，也都不依赖本章的任何一个数。这说明**症状位偏置**是一个可以被别人拿去用的类，而第五节那张五问清单是它的检验程序。△ 两例的证据强度本章未做系统核验，它们在这里只承担“这个类有别的成员”这一件事，不承担任何临床主张。

5.3 一句不含程度词的判据

在这份清单里，滞后距最大的那几条，和滞后距为零的那几条，数量之比是多少？如果比值远小于一，是什么支持了这个比例——是因果强度的证据，还是别的什么？

这句话可以拿去问指南、问产品目录、问医嘱模板，不用问任何人的判断。而它的后半句，正是第六节要去检验的东西。

六·可判定：预测力—滞后剖面

“办法都在夜里”这条分布，如果真是按因果强度排的，那么应当有一条证据支持它：**越靠近熄灯的时段，对当夜睡眠的影响越大**。

这是一条可以被数据判错的话。本章造一个读数来判它。

预测力—滞后剖面：把熄灯之前的十六小时切成八个等长的两小时窗，逐窗计算这个窗里发生的事对当夜睡眠的预测力。等长是关键——它挡掉了“白天比夜里长所以信息多”这条最省事的反驳。

— 若分布是按因果强度排的，剖面应当**单调下降**：越靠近熄灯，预测力越强。

— 若不是，剖面就不该是这个形状。

预注册（跑统计之前封存，其最后一字未动）

数据：Sleep-EDF Expanded 1.0.0 的 sleep-cassette 部分，七十八名二十五至一百零一岁健康受试者的二十四小时判读标注，每人两份连续记录，与本书前六章同一批；年龄与熄灯时刻取自随库发布的受试者表。

构造：把同一受试者的两份记录按 EDF 头里的起始日期时间拼到一条绝对时间轴上。令 T_0 为第二夜的熄灯时刻。结局为窗 T_0, T_0+8h 内的深睡百分比与睡眠效率。预测窗为等长两小时窗，滞后 $k = 0, 2, \dots, 14$ 。预测量为该窗内的睡眠分钟数——判读标注只能给出这一个白天变量。纳入条件：两份记录起点间隔在二十至三十小时之间；该滞后窗的判读覆盖率不低于九成。

先验预测：

- **P1** 剖面不是单调下降的。
- **P2** 至少存在一个 $k \geq 6$ 的窗，其预测力不低于 $k = 0$ 的窗。
- **P3** 年龄是主要混杂：按年龄三分位内标准化后，全部预测力下降。（这是一条自我削弱的预测，写在跑之前。）

撤销规则：若 $k = 0$ 那个窗的预测力对两项结局都严格大于所有其他窗，则症状位偏置这个那一格不成立，本章撤回主张，不得改口为”至少某几个窗”。若年龄分层后全部预测力低于 0.10，则本剖面不承载任何超出年龄的信息，须如实报告。

写在跑之前的四条局限：判读标注里没有光照、活动、进食、咖啡因，白天在本章里只由”白天睡眠分钟数”这一个薄代理表示，它绝不等于白天；观察性数据，一切结论只是相关；每人两夜，个体内重复只有一次；白天睡眠与年龄强相关，年龄混杂必然存在。

七·剖面跑出来是平的

纳入六十六人，年龄二十五至一百零一岁，中位五十七岁。当夜深睡百分比中位 0.113，睡眠效率中位 0.819。

滞后窗	n	窗内睡眠中位	vs 当夜深睡%	vs 当夜效率
熄灯前 0-2 h	66	0.0 分	-0.144	-0.124
熄灯前 2-4 h	66	0.0 分	-0.154	-0.029
熄灯前 4-6 h	65	0.0 分	-0.030	+0.181
熄灯前 6-8 h	43	0.0 分	-0.087	+0.062
熄灯前 8-10 h	24	0.0 分	-0.258	-0.105
熄灯前 10-12 h	60	0.0 分	-0.102	-0.218
熄灯前 12-14 h	66	0.0 分	-0.121	-0.099
熄灯前 14-16 h	66	3.2 分	-0.030	+0.157

三条先验预测全部命中。剖面对两项结局都不是单调下降的 (P1)；存在 $k \geq 6$ 的窗其预测力不低于 $k = 0$ ——深睡那一列是 $k = 8$ ，效率那一列是 $k = 10$ 与 $k = 14$ (P2)； $k = 0$ 不是严格最大，撤销规则未触发。

⚠ **P2 在深睡那一列的命中，靠的是 $k = 8$ 那个窗，而它只有二十四人——本章在 7.1 说它“不足以支撑任何结论”，却让它进了命中判定。这一处前后不一致，记在这里。去掉那个窗之后，深睡列 $k \geq 6$ 的最大值只剩 -0.121，仍不低于 $k = 0$ 的 -0.144（按绝对值），P2 在深睡列仍然命中，但幅度更接近于零。而七之三节会说明，这条命中本来就不携带信息。**

按年龄三分位内标准化后重算：

滞后窗	n	深睡% (分层后)	效率 (分层后)
0-2 h	66	-0.088	+0.077
2-4 h	66	-0.094	+0.062
4-6 h	65	-0.053	+0.093
6-8 h	43	-0.018	-0.037
8-10 h	24	+0.011	+0.177
10-12 h	60	+0.023	-0.035
12-14 h	66	-0.054	+0.033
14-16 h	66	+0.123	+0.170

全部下降，最大的一个只剩 0.177——**P3 也命中**，而它是本章自己写下的自我削弱预测。

7.1 三条都命中，而本章一条也不打算据此立论

因为效应量太小。未分层时最大的一个是 -0.258，而它出现在 $n = 24$ 的那个窗上；分层后最大 0.177。八个窗的预测力互相之间没有可辨的差别——这条剖面不是”远端更强”，它是平的。

更要命的是预测量本身。八个窗里有七个的窗内睡眠中位数是零分钟。也就是说，本章用来代表”白天”的那个变量，在大多数人的大多数窗里根本没有方差。用一个没有方差的量去测预测力，测出来的平，不能证明白天不重要，只能证明”白天睡眠分钟数”这个薄代理没有预测力。

本章把这一条写在正文里，不写在局限一节：本章没有测量白天。本章测量的是白天打盹的分钟数，而那不是同样东西。

7.2 顺带一条，它比剖面本身更有用

前一稿在这里只报了两个窗的数（0-2 小时 +0.433、2-4 小时 +0.303），并据此说“睡前别打盹”瞄准的是年龄的标志物。这是一次选择性呈现，本稿补全八个窗：

滞后窗	年龄 vs 窗内睡眠分钟数
0-2 h	+0.433
2-4 h	+0.303
4-6 h	+0.083
6-8 h	-0.100
8-10 h	+0.417
10-12 h	+0.327
12-14 h	+0.086
14-16 h	+0.262

熄灯前八到十小时那个窗是 **+0.417**，几乎和睡前两小时一样高。于是结论必须改写，而且改写后更准：

不是“睡前打盹”特别是年龄的标志物，是白天任何时段的打盹都是。本章那个薄代理测到的东西，主要是年龄本身。

这对本章是一次自我削弱，也是对第 7.1 节那句话的坐实：用一个几乎没有方差、而与年龄强相关的量去测预测力，测出来的一切主要是年龄在两处的投影。年长者更容易在任何时段打盹（多数窗 ρ 为正），也更少深睡（年龄 vs 深睡百分比 $\rho = -0.324$ ）；把任意一个窗与深睡连起来，都很容易得到一个看起来像因果的关联。分层结果支持这个警告：分层之后，熄灯前 0-2 小时那个窗的预测力从 -0.144 掉到 -0.088。

“睡前别打盹”这条建议仍然可能是对的，本章没有反对它——本章只是指出，本章这份数据不能用来支持它，也不能用来反对它。

七之三·三条预测全都命中，而它们本来就不可能不命中

7.1 说“三条都命中，而本章一条也不打算据此立论”，理由是效应量太小。那个理由是对的，但它不够——还有一个更难看的理由，前一稿没有查：

这三条预测，在完全没有任何关联的情况下，也几乎一定会命中。

本节把它算出来。做法很简单：把结局变量随机打乱五千次（打断一切真实关联），每次重算八个窗的相关系数，看这三条预测还中不中。随机种子写死。

表 4 三条先验预测在零假设下的命中概率

预注册条款	零假设下命中率	风险度
P1 剖面不是单调下降	99.98%	0.0002
P2 存在 $k \geq 6$ 的窗，其预测力不低于 $k = 0$	99.06%	0.0094
撤销规则： $k = 0$ 对两项结局都严格最大	触发率 0.50%	—

三行都要认账：

P1 是一条几乎不可能失败的预测。八个含噪的相关系数恰好排成严格单调下降，概率约万分之零点二五。写下“剖面不是单调下降”，等于写下“太阳明天还会升起”。它命中了，什么也没证明。

P2 同样。在一堆围绕零的噪声里，总会有某个远端窗的绝对值不低于第一个窗。

而撤销规则也不是一道真的闸。它在零假设下只有 0.50% 的触发率——也就是说，即便本章完全错，这道闸也几乎一定不会响。一道 99.5% 的情况下都不会响的闸，不是安全装置，是装饰。

本章因此在这里写下一句对自己不利的总账：

前一稿的预注册，在设计上就不可能失败。三条预测全部命中这件事，不构成任何证据；而“撤销规则未触发”这句话，也不构成任何证据。第七节真正携带信息的只有一样东西——效应量本身很小，而预测量几乎没有方差。

七之三之一·由此得到一个可以带走的东西：预测风险度

上面那个算法不依赖睡眠，也不依赖本章。任何一份预注册，都可以在跑数据之前把这个数算出来：给定零假设与该条预测的判定规则，用置换或模拟求出它的命中概率；一减它，就是这条预测的风险度。

三条纪律，防止它变成又一个可以填的指标：

其一，风险度必须在跑数据之前算，且与预注册文本一并封存。事后算出来的风险度可以被挑着报。

其二，风险度低不等于预测无效，等于“它命中时不提供证据”。一条风险度 0.0002 的预测，仍然可以在它罕见地失败时提供极强的证据——它是一道单向的闸。

其三，撤销规则的触发率必须一并报告。一份预注册若其撤销规则的触发率低于 5%，应当在正文里写明这一点，而不是让读者以为作者给自己上了一道锁。

本章全章最可迁移的东西，可能是这一节而不是症状位偏置。症状位偏置需要接受本章对分布的解释；风险度不需要——它只是一个把已有的预注册规范补齐一格的计算。

七之三之二·这一节的祖宗必须点名

这条思路有两位很硬的前人，本章初稿一位也没提。

Meehl (1967, 1978) 早就指出：在社会科学里，一个只作方向性预测的检验，其先验命中率接近一半，因此通过它几乎不提供支持；而随着测量精度与样本量提高，这类检验反而更容易通过。本章表 4 那三行，正是 **Meehl** 那条批评的一次具体计算——只不过本章的三条预测比方向性预测还弱，先验命中率不是 50% 而是 99%。

Mayo 的严格性 (severity) 给出了这件事的规范形式：一个假设通过检验，只有在“若该假设为假则此检验极可能不通过”时，才算得到支持。表 4 三行的严格性都接近于零。

分离线（两条，而且它们反过来给本章加力）：

1. 严格性通常是对一次已完成的推断、针对一个具体的替代假设来评估的；风险度是只需要零假设与判定规则、在登记时就能算出的一个数。它比严格性弱，但它有一个严格性没有的性质：它可以被写进预注册模板，作为一个必填项。

2. **Meehl** 的批评针对的是统计显著性检验的用法；本章这三条预测里没有一个用了 p 值——它们用的是效应量与序关系，而问题一样发生。这说明弱检验不是显著性检验特有的病，是判定规则宽松这件事本身的病。

判定形式：若在一份预注册模板里加入“逐条报告风险度”之后，作者所写预测的平均风险度没有上升，则这个装置不改变行为，只是多一列数字，本章这一节多余。

八·平的剖面意味着什么

一条平的剖面有三种读法，本章只能选其中一种。

读法一：白天更重要，只是本章测不出来。这是本章原本想要的结论。**本章不选它**，因为数据不支持——支持它需要远端窗的预测力显著高于近端，而实测是全都很低。

读法二：夜里更重要，剖面平只是噪声。本章也不选它，理由同样是数据： $k=0$ 那个窗不是最大的，撤销规则没有触发；而且它的预测力在分层后掉了四成。

读法三：在现有记录能给出的分辨率上，各时点根本分不出高下。这是数据唯一支持的一条，也是本章选的那一条。

而选定它之后，本章的承重命题必须改写，且改写后的版本比原来的更硬：

改写后的承重命题：睡眠干预点的分布集中在夜里，这件事**没有因果强度的证据支持**——不是证据反对它，是在现有可得的记录上，因果强度分不出时点。既然分不出，那么这条分布就必须由别的东西来解释；而唯一在夜里独有、在白天没有的东西，是**症状在那里可见**。本章因此主张：**干预点是被症状可见的位置定下来的，而这条定法从未被要求给出理由。**

这个版本不需要证明白天更重要。它只需要两件事：其一，夜里更重要这条没有证据；其二，分布仍然压倒性地在夜里。后者是第一节那次清点，站得住。

而前者要说得更小心，因为七之第三节刚刚拆过本章自己。本章并没有“证明”夜里更重要这条没有证据——本章做的检验风险度接近于零，它无论如何都会得到现在这个结果。**准确的说法是：本章去找过支持那条分布的因果证据，在现有可得的通道上没有找到，而本章这次寻找的灵敏度本身也很低。**两件事叠起来仍然足以支持本章的落点，因为本章的落点不是“夜里不重要”，而是“**这条定法从未被要求给出理由**”——这一条不需要本章的检验成功，只需要那次清点成立。

这正是本章分层引用的用处：第一节那次清点不依赖第七节任何一个数，而第七节全部塌掉，第一节仍在。

而它也因此背上了一条更重的举证责任：如果将来有人用带光照与活动的记录测出剖面确实单调下降，本章这一条整个作废。第九节把它写成了本章的第一顺位待办。

九·与最近的十七种说法的分界

逐条列三样：它走到哪一步、线划在哪里、以及一个能把“本章对”与“它对”分开的具体判法。只写得程度比较的，算被它覆盖，不留。

排序按危险程度。**头两条排在最前，因为初稿把它们整支漏了，而它们是本章命题在别处的成名版本。**

9.1 上游与下游——本章命题在公共卫生里的成名版本

它走到哪一步：麦金利在一九七九年讲过一个此后被反复引用的寓言：一个医生忙着从河里把落水者一个个捞上来救活，忙到没有工夫上游去看看，究竟是谁在把人往河里推。这个寓言把干预分成下游（对已经出现的后果动手）与上游（对产生后果的条件动手），并成为公共卫生四十多年来的一条基本区分；一级、二级、三级预防的分级，正是按“距病因多远”排的。

本章初稿一句也没提，而第一节那张表——十二条压在夜里、五条在白天——就是一次标准的“全在下游”清点。这是本章最大的一处缺席。

分离线（两条，且都反过来给本章加力）：

1. 上游那一支预设了上游更要紧，处方是“到上游去”；本章去测了因果强度，测出来分不出高下。本章因此不能说“到白天去”，只能说“**连夜里更要紧都没有证据**”。这个主张比上游那一支更弱，也更难反驳——它不要求你相信任何一端更重要。

2. 上游/下游是因果链上的远近，本章的滞后距是时间轴上的钟点差。二者不重合：一条干预可以在因果链上很上游、而滞后距为零（例如“上床前把手机放到客厅”作用于当下的光与唤醒，链短而钟点近）；也可以在链上很下游、而滞后距很大（例如“白天补觉”作用于当夜的睡眠压力）。滞后距是可以逐条填表算出来的，因果远近不能。

判定形式：若把第一节那十八条按“上游/下游”而不是按滞后距分类，得到的分组与本章的分组一致，则滞后距是上游/下游的一个换算，本章这一格并入那一支；若两种分法给出不同的分组（本章认为会），则它们测的是两件事。**这一条可以当场做完，本章未做。**

9.2 路灯效应——同形物，而且有名字

它走到哪一步：一个人在路灯下找钥匙，别人问他钥匙是不是掉在这儿，他说不是，掉在那边的草丛里，但**这儿有光**。这个寓言被用来指称一类系统性偏误：探究的位置由可观察性决定，而不由目标所在的位置决定。

分离线：路灯效应讲的是**找**——一次认识活动，代价是找不到；本章讲的是**动手**——一次干预配置，代价重得多：按第四节引的那条控制工程结论，**在输出端动手不只是效率低，是被明令禁止的**。一个在路灯下找钥匙的人只是白费工夫；一个在纯滞后系统的输出端做反馈的人会把系统弄不稳。

第二条分离线：路灯效应没有可算的量。本章给了两个——**滞后距**（第五节）与**风险度**（七之第三节）。

判定形式：若把“症状位偏置”换成“路灯效应”，第五节那张五问清单里有任何一问失去意义，则两者不同；若五问原封不动都成立，则本章只是给一个旧寓言换了名字。**本章认为第四问与第五问（当场可观察效果、有没有人卖）在路灯效应里没有对应，因为找钥匙这件事没有商品。**

9.3 婴幼儿行为睡眠干预的内容学综述（2026）——问题域内最危险的一位

它走到哪一步：一份关于零至五岁儿童行为睡眠干预的系统综述，统计了这些干预的内容与所测结局。它给出一条与本章几乎同形的批评：主要针对夜间睡眠的干预，可能只捕捉到儿童睡眠生态的一部分，从而低估了小睡时点、白天刺激与傍晚常规如何在全天尺度上共同塑造睡眠。该综述同时统计了结局的分布：所报告的睡眠结局主要关乎夜间睡眠模式，夜间睡眠时长与夜醒是最常被评估的两项，而白天睡眠结局被评估的频率低得多。

这不是一处遗漏，是一处必须承认的先后到。“夜间偏置”这个批评，在儿童睡眠领域已经被明确提出，而且配了内容学统计。

分离线在被诊断的对象与给出的解释。该综述诊断的是**研究与干预的覆盖不全**，处方是“把白天也纳进来”；本章诊断的是**分布由什么决定**，并且给出一个不同的答案——不是覆盖不全，是定位规则本身从未被要求给出理由。更要紧的是本章第七节：本章去测了因果强度，测出来是分不出的；**“把白天也纳进来”这条处方，在本章的数据上得不到支持。**

判定形式：若在成人身上，把白天变量（光照、活动）纳入之后，剖面变成单调远端更强，则该综述的处方对而本章这一格多余；若纳入之后剖面仍平，则“覆盖不全”不是问题的形态，本章的诊断更接近。这一条本章没有做——本章的数据里没有光照与活动。

9.4 日常经历与睡眠的当夜关联（2026）——直接对着本章的反证

它走到哪一步：一项大样本、多日、被试内设计的研究用移动脑电与日间日记做了本章想做的事：主要发现是，日常经历主要影响的是人什么时候入睡，而不是睡得好不好。并且给出了一条更重要的结论：除了长时间清醒带来的稳态效应之外，睡眠对日常经历相当有韧性，只被少数几种独特的白天经历所影响；意在改变睡眠的非药物干预，可能需要用超出日常波动范围的行为改变才行。

这一条直接反对本章原来那句“原因在白天”。它有脑电、有一千九百多夜、有被试内设计，比本章的薄代理强得多。本章接受它。

但它同时把本章逼准了。请注意它划出的那个例外：**长时间清醒带来的稳态效应被排除在外之后**，剩下的白天经历影响才很小。而清醒时长本身就是一个白天的量，而且是最大的那个。于是白天真正起作用的东西，不是行为清单上的那些条目，是两个结构量——**你醒了多久，你什么时候见了光**。而第一节那十八条建议里，直接作用于这两个量的只有三条（固定起床时间、睡眠限制、早晨见光）。

分离线：那项研究测的是**白天经历对当夜睡眠的效应量**；本章测的是**干预点分布由什么决定**。它可以完全对，而本章的问题依然存在：即使白天经历确实影响很小，也没有任何证据说明夜里那两小时的影响更大——而手段仍然全在那里。

判定形式：若同一批数据显示睡前两小时的行为对当夜睡眠结构的效应显著大于白天任一等长时段，则夜里那条分布有了证据，本章的那一格不成立。该研究报告的是时点而非结构，未直接回答这一问，如实记为未判。

9.5 睡眠卫生行为的日水平关联

它走到哪一步：一项两周重复测量、带客观睡眠测量的研究，评估了一批睡眠卫生行为与睡眠的日水平关联。其人际间效应给出一条对本章有用的结果：就寝时刻与起床时刻是睡眠效率最强的预测因子；白天小睡频率、咖啡因、酒精等也与睡眠效率相关。

分离线：最强的那两个预测因子是**时刻**，不是**睡前做了什么**。就寝与起床时刻是日程结构，它们被设定在白天（起床时刻尤其）。这一条实际上站在本章这一边，但本章不把它算作支持——它是人际间效应，混杂极多。

判定形式：若在被试内设计中，就寝/起床时刻的效应仍强于睡前行为的效应，则日程结构是主要抓手，第一节那张表的偏置更严重；若被试内效应消失，则那是人群差异不是干预抓手。

9.6 相位反应曲线

它走到哪一步：见 2.1。它证明了同一刺激在不同时点效果符号相反。

分离线：相位反应曲线讲的是**同一种干预在不同时点的效果不同**；本章讲的是**不同干预在时间轴上的分布由什么决定**。前者是一条剂量—时点曲线，后者是一个配置事实。相位反应曲线完全可以是对的，而分布依然可以由症状位决定——事实上正是如此：知道晨光有效的人很多，把晨光当作产品买的人很少。

判定形式：若把晨光干预的采纳率与褪黑素的采纳率相比，两者之比接近其证据强度之比，则分布是按证据排的，本章错。就现有的销售与处方数据而言，差距悬殊。本章未做正式统计，如实记为未做。

9.7 史密斯预估器与长滞后控制

它走到哪一步：见 2.2。它给出本章的推翻材料。

分离线：这一支处理的是**已知模型的工程系统**；本章处理的是**没有可用模型的人体**。控制工程自己规定，模型不准时预估器反而更糟——因此本章不能把“改用前馈”当作处方，只能把“输出端不是合法干预点”当作一条否定性结论。

判定形式：若有人给出一个可用的睡眠滞后模型并据此做出前馈干预且优于夜间干预，则本章的类比可以进一步使用；否则它只提供那条否定。

9.8 光环境标准

它走到哪一步：见 2.3。室内照度标准按视觉作业制定。

分离线：这一支解释的是**为什么白天那一端难以调节**，它替夜间干预辩护；本章问的是**为什么没有人把这件事说出来**。第一节那张表里，没有一条建议写着“你白天的光照被你的办公室锁死了，这一条你改不了”。**难以调节与不被提及，是两件事。**

判定形式：若在建筑光环境可调的人群（户外工作者、居家办公者）中，干预点分布向白天移动，则分布由可调性决定，本章错；若分布不变，则它由症状位决定。这一条可做，本章未做。

9.9 轮班工作的干预分布

它走到哪一步：一份轮班工作干预方案里写着一句与本章同形的观察：大多数干预集中在移动生物钟的时点或急性提高工作时的警觉度上；而被这些策略忽略的是，夜班工人通常在下班后不久的早晨睡觉。

分离线：那是在**一种特殊作息里**指认干预点选错；本章是在**常规作息里**指认分布本身。两者可以互为例证，但对象不同。

判定形式：若轮班人群的干预分布同样堆在他们各自的“上床前两小时”，则症状位偏置跨作息成立；若它跟着生理相位走，则本章只适用于常规作息。

9.10 自家六处

(一) 同作者的失眠专文(站上 /column/insomnia/)——本章在自家这一侧最该先划的一条线。那一篇把靶子从“用力睡”挪到“停止阻止入睡”，是一次实打实的推进。而它移过去之后，动作仍然发生在床上。本章对它是再推一步：连“停止阻止”也是在结果端执行的动作，它的滞后距同样是零。分离线：那一篇处理的是动作的性质(使劲还是撤退)，本章处理的是动作的时点。一个完全学会了撤退的人，他的干预点照样在夜里。

(二) 本系列 Why 组第一篇(年老为何难深睡)。那一篇的核心是“一天睡多少”这个量随测量窗口长度改变；本章的核心是干预点在时间轴上的位置。词面很近，必须分开：一个是量的窗口，一个是动作的坐标；前者改的是被测之物，后者改的是施力之处。一份窗口定义完全正确的记录，照样可以配一份全部压在夜里的干预清单。

(三) 本系列 How 组第一篇(下去的那一段没有人记)。那一篇讲一个量未被记录也未被裁决；本章讲一批动作放错了位置。分离线：那一篇的缺口在记录，本章的偏置在配置。而且本章第七节的结局与那一篇相似——都是“测不出来”，但性质不同：那一篇是测出来没有信息，本章是测出来各时点没有差别，而差别恰恰是本章要否定的东西。

(四) 第二篇《住在线上》(第二章)的邻记(一片区域的产出记在两侧名下)：本章不涉及记账归属，只涉及施力位置。(五) 第一、三、四篇(段落账、判余、托拍)同属“记录制度”全书赌的那句话，本章的那一格是干预配置，是本系列第一次离开记录制度。(六) /column/rhythm-deprivation/《节律剥夺》讲同质化时钟碾平各种节律；本章讲手段在时间轴上的堆积，两者可独立发生。

9.11 这些邻居共同站着的那块地

逐个问“它结构性地看不见什么”：

— 相位反应曲线把干预已经被施加当作给定，因此看不见谁会去施加它；一旦承认采纳率是问题的一部分，它的临床价值就要重新计算。

— 睡眠卫生清单把每一条独立有效当作给定，因此看不见清单整体的时间分布；一旦承认，清单就要多一列坐标。

— 儿童干预综述把覆盖不全当作诊断，因此看不见分布可能有一个独立于覆盖的成因；一旦承认，它的处方就要从“补上白天”改成“先问分布由什么决定”。

— 控制工程把模型已知当作给定，因此看不见没有模型时该怎么办；一旦承认，它那条否定性结论就是它在这里唯一能给的东西。

背后是同样被当作给定的东西：一件事该在哪里做，由它在哪里出问题决定。

本章不主张自己是第一个看见它的——第一节那张表任何人都能列。本章主张的只是：这条定法从未被要求给出理由，而当本章去为它找理由时，在现有记录上找不到。

十·四家反过来给本章加的约束

四处，每一处写明本章原想写哪句更强的话、哪一家不许、削掉之后剩什么。

10.1 时间生物学削掉的：本章不能说“因在白天”

原想写：睡眠的原因在白天，所以办法也该在白天。

时间生物学不许，而且是用 9.2 那项研究不许的：白天经历主要改变入睡时刻，不改变睡眠结构；除稳态效应外，睡眠对日常经历相当有韧性。本章自己的剖面也不支持——远端窗的预测力并不高。

削掉之后剩：本章不主张因在白天。本章只主张没有证据说明因在夜里。这弱得多，但它是数据支持的那一句，而且它已经足以让第一节那张表失去理由。

10.2 控制工程削掉的：本章不能开出“改用前馈”这条处方

原想写：既然输出端不是合法干预点，就应当改成前馈控制——在白天补偿。

控制工程不许。前馈与滞后补偿都要求一个足够准的对象模型；模型不准时，它们比普通反馈更糟。而人体睡眠的滞后结构没有可用模型——本章第七节就是一次失败的建模尝试。

削掉之后剩：控制工程在本章里只提供一条否定(输出端不是合法干预点)，不提供任何替代方案。本章因此没有处方，只有一个诊断。这一处动的是本章的实用价值，是最大的一次损失。

10.3 照明工程削掉的：本章不能把这条分布判为不理性

原想写：办法都堆在夜里，这是一个错误。

照明工程不许。如果白天那一端在结构上被建筑锁死，那么把手段压在唯一可动的那一端，是理性的。一个在无窗办公室工作、通勤在地下的人，“早晨见光”对他不是一条建议，是一句风凉话。

削掉之后剩：本章不判这条分布为错。本章只主张它没有被要求给出理由，而当理由被要求时，能找到的是可调性而不是因果强度——这两者是不同的辩护，而现在使用的是前者的效果、说的是后者的话。这一处动的是价值排序。

10.4 本章自己的数据削掉的：本章没有测量白天

原想写：本章用四十夜数据测出白天与夜里的因果强度没有差别。

本章自己的数据不许。本章的白天代理是“白天睡眠分钟数”，而八个窗里有七个的中位数是零。用一个几乎没有方差的量测出来的“没有差别”，其信息量接近于零。

削掉之后剩：本章的检验只排除了一件事——在判读标注这一个通道上，各时点分不出高下。它不能排除在带光照与活动的记录上能分出高下。第一顺位待办因此不是理论工作，是一次数据工作：拿一份带光照与活动的多日记录，重跑同一张剖面。若剖面变成单调远端更强，本章这一格作废。

十一·那么现在能说什么

本章有责任回答第一节那张表该怎么办，而不是只说它没有理由。能说的分三层，一层比一层弱，也一层比一层稳。

第一层（最稳，且不依赖第七节任何一个数）：给每一条建议标上它的滞后距。这不需要任何新证据，只需要在现有清单上加一列。加完之后，一份清单的分布是一眼可见的，而分布本身就成了一个可以被讨论的对象。这是本章唯一一条不会塌的建议。

第二层（有证据，但不是本章的）：那三条滞后距最大的建议，恰好是证据最硬的。固定起床时间与睡眠限制是失眠行为治疗里效应最大的成分，早晨见光有相位反应曲线支撑。它们的共同点是：执行时看不见任何效果，效果要到十几小时以后才出现，而且不是每次都出现。这解释了为什么它们最难被坚持——也解释了为什么没有人卖它们。

第三层（本章自己的，且是一条警告）：不要把“这条建议我做了当场就有感觉”当作它有效的证据。滞后距为零的建议，其当场感觉恰恰是它落在症状位的标志，而不是它作用于原因的标志。第七节 7.2 给了一个具体的例子：睡前打盹与深睡减少之间那条看起来很像因果的关联，在年龄分层后掉了四成。

十二·用本章的判据检本章自己

本章的判据是：这条建议必须在几点执行，它作用的量在几点被设定，两者相差几小时。把它转过来对准本章。

先对准本章的检验。本章声称要测因果强度，实际测的是“白天打盹分钟数”——一个在七个窗里中位数为零的量。本章自己的干预点也没有落在因上，落在了手边有的那份数据上。这不是疏忽，是本系列一直在做的同一个动作：用现成的公开记录去问一个现成记录回答不了的问题，然后如实报告它停在哪里。但必须承认：本章在批评别人把手段放在方便的地方时，自己的方法也放在了方便的地方。

再对准本章的预注册。七之三节算出来的三个数，是本章这一稿最难看的一处：本章写下的撤销规则，在零假设下只有 0.50% 的触发率。也就是说，本章给自己上的那把锁，在本章完全错的情况下也几乎一定不会响。一个声称“干预点应当被要求给出理由”的人，自己给出的检验没有被要求给出理由。这一条不是补充说明，它与本章的命题同形：症状位偏置的第五问是“有没有人卖”，而预注册的对应问题是“这条预测有没有人可能不通过”——两处都是在问，这件事有没有可能失败。

再对准本章这一篇的位置。本章属于一个九篇的系列，前六篇的那一格全部是记录制度——账上少一栏、边界没有地、速率没被裁。本章是第一次离开那一族，去处理配置而不是记录。而本章的结论恰恰是：本章所属的那个全书赌的那句话，在这里帮不上忙。第七节测不出高下，不是因为记录里

少了一栏——白天的那些窗都在记录里，判读得清清楚楚；是因为记录的通道里没有光和活动。少一栏和少一个通道，是两件不同的缺。这一条本章写在这里，供后续几篇划界。

最后一句留给那件本章没能做成的事。要判定这条分布有没有理由，只需要一份带光照与活动的多日记录，把第六节那张剖面重跑一遍。这样的数据集是公开存在的。**本章可以被引用的最稳的一层，不是那个新名字，也不是第七节那两张表——那两张表里，一张的预测量几乎没有方差。最稳的是那张清单：给每一条建议标上它必须被执行的钟点，和它作用的那个量被设定的钟点。这件事不需要任何研究经费，也不需要等任何人同意。**

注释与材料

本章取的是与前六篇同一个公开库，但取法不同：这次用的是 sleep-cassette 的全部一百五十三份判读标注，覆盖七十八名二十五至一百零一岁的受试者，而不是此前几篇用的前四十份。年龄与熄灯时刻来自随库发布的受试者表。第六节的预注册文本在取数之后、跑任何统计之前写定，其后一字未动；三条先验预测的结果与效应量按封存时写下的规则原样报告，包括那条自我削弱的第三条。解析与统计脚本随文提供，可从原始 EDF+ 标注起逐步复算。

一处必须声明的实现更正：首次实现时把“逐窗覆盖率”误写成了“全部窗联合覆盖率”，导致纳入人数为零。预注册原文写的是“该滞后窗的记录覆盖率”，即逐窗纳入；本章按预注册原文修正了实现，此更正发生在任何统计量被计算之前。因此各滞后窗的样本量不同（二十四至六十六人），熄灯前八至十小时那一窗只有二十四人，其读数不足以支撑任何结论，本章未据此立论。

关于文献核验：本稿写作过程中核对了相位反应曲线（Khalsa 等，2003）、婴幼儿行为睡眠干预综述（2026）、日常经历与睡眠的当夜关联研究（2026）、睡眠卫生行为日水平关联研究四条的来源页面。其余各条依通行出处引用，未在本轮逐页核验。

人机分工声明：本章的三家选源由人指定；第一节那张干预点清单由人机协作列出并逐条标注时点；文献检索与核对、检验设计、脚本、真跑与初稿由人机协作产出。三条先验预测与撤销规则在拿到任何结果之前即已封存；效应量偏小这一不利结果按封存时写下的规则原样报告，未作补救性重述。

参考文献

9. Kemp, B., Zwinderman, A. H., Tuk, B., Kamphuisen, H. A. C., & Oberyé, J. J. L. (2000). Analysis of a sleep-dependent neuronal feedback loop: the slow-wave microcontinuity of the EEG. *IEEE Transactions on Biomedical Engineering*, 47(9), 1185–1194. (Sleep-EDF 数据库来源)
11. Åström, K. J., & Hägglund, T. (2006). *Advanced PID Control*. ISA. (纯滞后对象的稳定性与滞后补偿)
12. Boyce, P. R. (2014). *Human Factors in Lighting* (3rd ed.). CRC Press. (室内光环境标准的视觉作业依据)

修订说明（本稿相对初稿的六处改动）

初稿完成后经一次外部评审。评审方独立复算了第七节的全部结果——两张剖面表的各窗样本量（66/66/65/43/24/60/66/66）、相关系数、以及年龄相关，逐项吻合。评审并指出两处整支缺席与一处选择性呈现。逐条列明：

1. **新增 9.1“上游与下游”与 9.2“路灯效应”，排在分界节最前，并承认初稿把这两支整个漏了。**前者是本章命题在公共卫生里四十多年的成名版本（麦金利的水落者寓言、一二三级预防的分级），后者是同形的观察偏误。各给两条反向加力的分离线：上游那一支预设了上游更要紧而本章测出分不出高下，故本章的主张更弱也更难反驳；滞后距是钟点差、可逐条填表算出，因果远近不能。路灯效应讲的是找（代价是找不到），本章讲的是动手（代价是在被明令禁止的位置上动手）。

2. **★新增七之三节“三条预测全都命中，而它们本来就不可能不命中”。**把结局随机打乱五千次重算：**P1 在零假设下命中率 99.98%，P2 为 99.06%，而撤销规则触发率只有 0.50%。**前一稿的预注册在设计上就不可能失败；“三条全部命中”与“撤销规则未触发”都不构成证据。

3. 由此提出可迁移的装置“风险度”（七之三之一），配三条使用纪律；并在七之三之二点名两位祖宗——Meehl 的弱检验批评（方向性预测先验命中率近半，故通过它几乎不提供支持）与 Mayo 的严格性，给出两条分离线（风险度只需零假设与判定规则、可在登记时算出，可作预注册模板的必填项；且本章三条预测都没用 p 值，说明弱检验不是显著性检验特有的病）。

4. 7.2 补全八个窗的年龄相关（初稿只报了 0-2 h 与 2-4 h 两个近端窗）。熄灯前 8-10 h 是 +0.417，几乎等于睡前两小时的 +0.433 $\Rightarrow$ 结论由“睡前打盹是年龄标志物”改写为“任何时段的打盹都是”，并明写本章这份数据既不能支持也不能反对“睡前别打盹”这条建议。

5. 第七节补一处前后不一致的认账：P2 在深睡列的命中靠的是 k=8 那个只有二十四人的窗，而本章在 7.1 说它“不足以支撑任何结论”。去掉它之后 P2 仍命中但幅度更接近零。

6. 新增 5.2.1“这个类不是为睡眠定制的”：近视（户外时间对眼镜与护眼灯）与龋齿（进食频次、饮水加氟对牙膏与漱口水）两个成员，读者可自行跑那五问。 Δ 两例的证据强度本章未做系统核实验，只承担“这个类有别的成员”。第八节与自反节同步收窄：本章并没有“证明”夜里更重要这条没有证据——本章这次寻找的灵敏度本身极低；本章的落点不需要检验成功，只需要第一节那次清点成立。

仍未做的三件：拿一份带光照与活动的多日记录重跑剖面（第一顺位）；把十八条按上游/下游分一次类，看是否与按滞后距的分组一致（当场可做）；在一份真实的预注册模板里加入逐条风险度，看作者所写预测的平均风险度会不会上升。

编五·合论

全书层

前四编各自成立，但它们合起来欠着一件事：这十四个读数，是不是同一个读数换了十四个名字？

这一编用五章回答它，而五章交回来的对本书不利的部分比有利的多。

第十五章把判据写死在前面——数学对象是否相同、在同一份记录上是否高度相关、四个层级是否互不蕴含——然后按判据走，并做了一次全新的粗化实验去检验“层级不能互相还原”。第十六章把十四个位置摆在一起，为每一个配一个在记录上与它完全同形而无害的对照格，并给出三条共同的签名：短期表现为改善、失效不产生任何信号、越正规化越自我加固。第十七章做了一件别的书做不成的事——本书有十一章用的是同一份公开判读记录，那一章把能算的读数在同一份记录上一次算齐。第十八章处理家族级的对手，十族里有八族是十四章全部写完之后才检索出来的。第十九章是与判读手册、与概率化睡眠图、与三个“别处已经做了”的领域的对话。

这一编最要紧的一句写在第十七章：本书最想成立的那条预测在检验里命中了，而它的风险度是零——它在设计上就不可能失败。那条命中不是证据。

第十五章 十四个读数，是不是同一个读数

十五之一·为什么这一问必须问在前面

前面十四章各提出了一个读数，每一章都为自己那一个辩护过：它测的是什么、它在什么条件下失效、它与本领域最近的几种说法怎么分开。十四章因此各自站得住。

但没有一章问过一个更靠前的问题：**它和另外十三个，是不是同样东西换了十四个名字？**

这个问题不问，本书就不是一本书，是一本词典——十四个词条，各有定义，彼此不发生关系。而如果十四个读数里有几个其实是同一个量的不同口径，那么全书的分量就要按去重之后的数目重算，不是按名字的数目。

更要紧的是：这件事**不能事后判**。一旦看过了数据，任何两个读数之间的相关，无论多高，都能被解释成“侧重不同”。所以本章的顺序是这样的：

先在这里把十四个读数摆成一张表，先在这里写死一条合并规则，然后由第十七章去执行它——而第十七章的结果，不管对本书有利还是不利，都写回本章第七节。

本章因此有一件不太常见的事要做：**它同时是提出规则的一章，和被规则判过的一章**。判决的结果对本书不利的部分比有利的部分多，这些都在第七节。

十五之二·十四个读数的总表

表 15-1·全书十四个读数

章	读数	它数的是什么	分母是什么	强制作答发生在
一	未追认段比例 U	没有被后续事件延续下来的那些段，占全部段的多少	这一夜的段数	每一格必须给一个标签
一	追认延迟 L	一段时间从发生到被认下来，隔了多久	无（是时长）	同上
二	成簇承载比 C	转换有多大比例挤在少数几处	这一夜的转换数	同上
二	界带宽度	睡与醒之间那片区域有多厚	无（是时长）	同上
三	判余	被选中的那个标签，比第二名高出多少	无（是概率差）	同上
三	体例折叠损失 F	换一套同样正确的手册，少掉多少次转换	旧手册下的转换数	同上
四	残余率 Res	一条被规则压低的转换，还剩多少	同一动作的另一条可比路径	同上
五	索验比 $\Delta/6$	索求的间距，比验收的分辨率大几倍	使用者的验收分辨率	每次改善必须给一个分数
六	自噪底	同一个人两个可比夜之间，同一读数自然差多少	无（是一个差）	同上
七	端权对	一个差值读数里，两端各解释了多少变异	无（是一对份额）	同上
八	下降时长 T	从入睡到第一次进入深睡，用了多久	无（是时长）	同上
九	双向漏出率	两套划分互相漏掉了多少人	各自那一套的人数	每个人必须归进一类
十	长断占比 Ib	一夜最长的一次中断，占全部中断的多少	这一夜的入睡后清醒总量	同上
十一	位重复度	同一个人两夜之间，这个量的位置有多可重复	无（是一个跨次相关）	同上
十二	下一拍托管率 R	已被转交的关键回应类别，占全部类别的多少	预先枚举的关键类别数	每笔账必须记在一个人名下
十三	他源率 ϕ	问对面那个人自己做	全部计分条目数	同上

		了什么的条目，占多少		
十四	预测力—滞后剖面	白天各时段对当夜的预测力，随滞后怎么变	无（是一条曲线）	同上

十四章共提出十七个读数（第一、二、三章各有两个）。表里最要紧的一列是最后一列：**它把十四章分成了四组，而分组的依据不是题目，是“必须给出一个答案”这条规矩落在哪一层。**第一层落在三十秒的格子上，第二层落在被公布的那个数上，第三层落在一个人被归进哪一类上，第四层落在这笔账记在谁名下。

第六节要论证的就是：**这四层不能互相还原。**

十五之三·它们甚至不是同一种数学对象

在问“是不是同一个”之前，得先问“能不能放在一起问”。

第十七章按可算性把十七个读数逐条判过一遍，结果不是“能”与“不能”两类，是三类：

- 每夜一个数（八个）**：U、C、F、Res、T、H（命中界线数，双向漏出率的逐夜形式）、Lb、π（窄口位置，位重复度的逐夜形式）。这八个能进同一张相关矩阵。
- 是别的数学对象（三个）**：自噪底是一个跨夜配对差，样本单位是人不是夜。端权对是一对解释份额，它是可以施加到任何差值读数上的装置，本身不是读数。预测力—滞后剖面是一条八点曲线，**把一条曲线压成一个数正是第五章说的那件事，故不做。**
- 在这份记录上算不出来（三个）**：索验比的分母、下一拍托管率、他源率。另有三个从属读数同样算不出来：追认延迟、界带宽度、判余。

八／三／三。这个分布本身是本章的第一个结果，而它有一个不舒服的读法。

算不出来的那三个是索验比的分母、下一拍托管率、他源率。它们算不出来的理由，逐条就是那三章各自的论点：

使用者的验收分辨率从来没有被测过，所以第五章说那个要求无法被验收；

记录里没有谁接住下一拍这一栏，所以第十二章说两笔账互不登记；

计分条目里没有一条问对面那个人自己做了什么，所以第十三章说账按主体划。

这三章的读数在同一张表上算不出来，不是这次核对的失败，是那三章的命题在这份记录上原样发作了一次。

推论也得写下来：**能进那张表的八个读数，恰恰是最不激进的八个——它们全部可以从现成的判读标注里算出来，也就是全部落在现行记录制度已经愿意登记的范围之内。**本书最想指出的那几样，正因为记录里没有它们的位置，也就进不了这次核对。**这次核对的效力，在本书最要紧的地方最弱。**

十五之四·合并规则

规则写在这里，第十七章执行它。

任何两个读数，若在同一份记录上的 Spearman 相关的绝对值超过 0.80，对应的两章必须合并，或明写为同一个量的两种口径。不许用“侧重不同”保留两个名字。

若出现三对以上超线，“十四个读数是十四样东西”这条主张整个撤回，第五编改写为《**十四个名字，K 个量**》，且不许换定义重试。

配三条纪律：

1. **不许事后换定义。** 一个读数被判为与另一个重合之后，改它的分母、改它的分段阈值、改它的窗口，然后重新算一遍相关——这是本书在第三章批评过的动作，本书自己不做。
2. **不许用“侧重不同”保留两个名字。** 两个高度相关的量当然可以有不同的侧重，问题在于侧重不能被观察到，而相关能。判据只认能被观察到的那一样。

3. 风险度必须与结果同表公布。这条来自第十四章：一条预测在零假设下的命中概率如果接近一，那么它命中不提供任何证据。合并规则也要交这份账。

同时把这条规则自己的两处毛病写在这里，而不是等它出结果之后再写：

· **0.80 这个数是任意的。** 它没有理论根据，只是一个惯例上的强相关线。如果实测最强的一对落在 0.75 附近，这条规则给出的答案就完全依赖于这个任意选择。

· **相关不是关系的唯一形状。** 两个量可以在相关系数上接近零，而其中一个另一个的确定性函数（例如二次关系）。这条规则查不出那种情形。

十五之五·判决性反例：八对最像的

对每一对最容易被怀疑是同样东西的读数，这里给一句“在什么观察下它们分得开”。只能写出“本书更强调”或“视角不同”的，视同承认被那一条覆盖。

其一·未追认段 **U** 与成簇承载 **C**（第一章 × 第二章）。两者都在数短促而密集的结构。判定形式：把由一窗段造成的转换从成簇的分子分母里全部剔除，若两者仍相关，则它们是两件事。**第十七章跑了：原始相关 +0.735，剔除后 +0.033。**⇒ 相关的九成半由一条算术解释——一个一窗段进一次出一次，本身就构成一簇。这一对没有通过自己的判定形式，处置见第八节。

其二·未追认段 **U** 与残余率 **Res**（第一章 × 第四章）。两者都在数罕见的东西。判定形式：残余率数的是哪一条转换被一条明文规则压低，未追认段数的是段有多短；若把判读手册里“深睡之后不判第一期”那条条款去掉重判，残余率应当大变而未追认段几乎不动。这条实验本书没有做，但两者在这份记录上的相关是 **+0.051**——近乎正交，与判定形式的方向一致。

其三·体例折叠损失 **F** 与判余（第三章内部的两个读数）。两者都在讲“格式扔掉了什么”。判定形式：体例折叠是两套手册之间的差，判余是一套手册内部取最大值那一步扔掉的量；若软输出能吸收体例差异，则两者不是同一个量。**第三章自己跑了这一条，并因此撤回了自己的一条分离线：**同一份特征只换标签体例，折叠后硬标签不一致率仅 0.59%，概率向量总变差距离中位 0.0027。⇒ 两者确实不是同一个量，而且分开它们的那次实验判了第三章原来的说法错。**这是全书唯一一对被真跑分开、且分开的方式对提出者不利的读数。**

其四·自噪底与端权对（第六章 × 第七章）。两者都是“一个数背后其实有两端”。判定形式：自噪底问的是同一个人的两个可比夜之间这个数会自然差多少，端权对问的是同一个数的两个构成端各自解释了多少变异。判定：给一个只有单夜记录的数据集，端权对算得出来，自噪底原理上算不出来。二者对“多记一夜”这个动作的预测相反——自噪底因此成立，端权对不受影响。

其五·索验比与自噪底（第五章 × 第六章）。两者都在讲分辨率。判定形式：索验比的分母是使用者的验收分辨率，自噪底是记录的夜间自然变异。把仪器做到无限精确，自噪底不会降到零，因为它是人的夜与夜之差；索验比的分母也不会变小，因为它是人手里那把尺。两者对“把仪器做细”的预测相同，都说不管用；但对“多记一夜”的预测相反——自噪底能被测出来，索验比的分母不能。**这一对最容易被合并，而分开它们的正是那个不对称。**

其六·下降时长 **T** 与窄口位置 **n**（第八章 × 第十一章）。两者都在讲夜的前半段。一个自然的怀疑是：下降段本身就是一段没有深睡也没有快速眼动的时段，所以下降越长，最长的那段区间越可能就是它，位置自然越靠前——若真如此，这一对是定义蕴含。判定形式：控制住那段区间的长度，相关应当塌掉；并直接数一数两段是不是同一段。**第十七章跑了：原始 -0.541，控制长度后仍 -0.461；两段几乎同一段的夜只有 $24/135 = 17.8\%$ 。**⇒ 不是定义蕴含，是实质相关，而且是本书十四章无一预言过的一件事。见第九节。

其七·命中界线数 **H** 与长断占比 **Lb**（第九章 × 第十章）。两者都在数同一批人的失眠。判定形式：命中界线数问“这一夜有没有越过那三条线”，长断占比问“这一夜的清醒是集中在一次还是摊开在很多次”。一夜可以三条界线全中而长断占比很低（清醒摊得很开），也可以一条不中而某一次中断特别长。**实测相关 -0.033，近乎正交。**⇒ 两条独立的失眠读数，而现行报告只出前一类。

其八·下一拍托管率 R 与他源率 φ （第十二章 × 第十三章）。两者都在讲第二个人。判定形式：托管率问的是这一夜有哪些必须作答的事被转交了，对象是事件；他源率问的是评估工具有没有一栏问对面那个人自己做了什么，对象是条目。判定：一个人完全可以托管率为一而他源率为零——事实上这正是现状，也正是第十三章那次清点的结果（七十八条计分条目，问对面的零条）。⇒ 两者不可能相关，因为它们根本不在同一份记录上。

十五之六·四个层级能不能互相还原

前一节是逐对分开，这一节要问一件更大的事：把十四章分成四编的那条轴，是不是也只是一种叙述顺序？

本书的说法是：不是。四个层级不能互相还原——把判读格子做细，不会解决读数层的问题；把读数做准，不会解决归属层的问题；把归属做对，不会解决跨主体的问题。

这句话至今只有论证，没有证据。而它有一个可以当场执行的检验：

改变判读格子的粗细，只应当动到第一层的读数，不应当动到第三层的读数。

若两层的读数在同一次粗化下动得一样多，那么“层”这个区分在这份记录上就得不到支持。

预注册文件写在跑之前并封存，SHA-256 为 `df9f2034...0211066`。三条预测：第一层四个读数的稳定度中位低于 0.80；第三层三个读数的稳定度中位高于 0.80；两者相差至少 0.15。撤销规则写死：若差为负，第六节整节改写为“本书的分层是论证上的，没有经验支持”，且不许换刻度或换读数重试。

稳定度的定义是：某读数在粗化刻度上的一百五十三个夜值，与它在三十秒刻度上的一百五十三个夜值之间的 Spearman 相关。粗化按占优规则合并，并列取最早出现的那一个——与判读手册处理一个窗内多分期的办法一致。

表 15-2·粗化下的稳定度（与三十秒刻度的相关）

读数	编	六十秒	九十秒	一百二十秒
未追认段比例 U	一	+0.533	+0.326	+0.295
成簇承载比 C	一	+0.774	+0.406	+0.443
体例折叠损失 F	一	+0.975	+0.946	+0.909
残余率 Res	一	+0.550	+0.487	+0.401
下降时长 T	二	+0.985	+0.882	+0.871
命中界线数 H	三	+0.954	+0.937	+0.902
长断占比 Lb	三	+0.946	+0.864	+0.840
窄口位置 π	三	+0.910	+0.663	+0.619

第一层中位 **+0.510**，第三层中位 **+0.902**，差 **+0.392**。三条预测全部命中，撤销规则未触发。

表 15-3·三条预测与它们的风险度（五千次独立打乱）

预测	放行线	实测	判定	零假设命中率	风险度
第一层中位 < 0.80	—	0.510	命中	1.0000	0.0000
第三层中位 > 0.80	—	0.902	命中	0.0000	1.0000
两者相差 ≥ +0.15	—	+0.392	命中	0.0004	0.9996
（撤销规则触发）	差为负	未触发	—	0.4946	—

第一条的风险度是零——打乱之后一切稳定度都趋近于零，它当然低于 0.80。这一点在预注册第四节里就写明了，因此它的命中不作证据。本节的承重全部落在后两条上，而后两条的风险度是 1.0000 与 0.9996：在零假设下，第三层的稳定度几乎不可能高于 0.80，两层之差也几乎不可能达到 0.15。这两条是真检验，而本书通过了。撤销规则在零假设下的触发率是 0.4946，接近掷硬币——它是一次真正的风险，不是形式。

一处必须认账的反常

表 15-2 里有一行不合群：**体例折叠损失 F 在第一层，却像第三层一样稳**（ $0.975/0.946/0.909$ ）。它的机制不难看出：这个读数是两个计数的比值，而粗化同时缩小分子与分母，比值因此近乎守恒。也就是说，**F 之所以稳，不是因为它不依赖格子，是因为它依赖格子的方式在分子分母上抵消掉了。**

这与第十七章那次端权对的结果对得上：F 的分子端解释了 0.908 的变异，分母端只有 0.022。分母几乎不说话，那么分母随格子怎么变，也就影响不到这个比值。

把 F 从第一层剔除重算：**第一层中位由 0.510 降到 0.443，与第三层的差由 +0.392 升到 +0.459。**结论方向不变而更强。但本节不据此重报主结果——**剔除一个对自己不利的数据点然后重算，正是本节的规则第一条禁止的动作。**主结果仍是 +0.392，F 那一行照原样留在表上。

预注册第五节那条替代解释的下场

预注册事先写明了一处：第三层的三个读数里有两个（命中界线数与长断占比）是以分钟为单位的，粗化对它们的影响天然小于以窗为单位的读数。**第三层更稳，可能只是单位的后果，不是层级的后果。**

这条替代解释可以被检验：第三层里的窄口位置是一个无量纲的相对位置，不含任何时间单位。只留它一个来算：

只留窄口位置：中位 +0.663，与第一层的 +0.510 相差 +0.153。

放行线是 +0.15。差 **+0.153，勉强过线。**

这句话要说得准确：本节的结论在剔除了单位这条替代解释之后仍然成立，但只剩下三厘的余量。若放行线当初写的是 +0.20，本节此刻就该判未命中。**因此本节的正确读法是：层级区分在这份记录上得到了支持，而支持的强度比表 15-2 那个 +0.392 看起来的弱得多。**

（两处收紧同时做——剔除 F、只留窄口位置——差是 +0.220。这个数比 +0.153 好看，但它包含了那个被禁止的动作，只作附注，不作结论。）

这一节能说与不能说的

能说：改变判读格子的粗细，明显地动到了第一层的三个读数（未追认段从 +0.533 掉到 +0.295，成簇从 +0.774 掉到 +0.443），而第三层的读数几乎不动。**把格子做细，解决不了第三层的问题；把格子做粗，也毁不掉第三层的读数。**这两句是同一件事的两面。

不能说：本次检验只分开了第一层与第三层。**第二层在这份记录上只剩一个可算的读数（下降时长），第四层一个也没有，**因此四个层级里有两个没有被检验过。而第二层的读数为什么系统性地进不了这张表，本身是一件要紧的事——**第二层讲的正是“一个被公布的数”，而它的四个读数里有一个根本不是一个数（一条曲线、一个跨夜差、一对份额）。**这个巧合值得记下来，但本书不把它当证据。

还有一件更根本的：粗化只是改变格子的一种方向。**把格子做细——例如五秒判读——才是第一层真正的处方，而这份记录做不到。**本次检验是那个动作的一个反向代理，不是那个动作本身。

十五之七·合并规则的下场，以及它对本章的三处判决

第十七章执行了第四节那条规则。结果如下，三处都对本章不利。

其一，合并规则形同虚设。八个读数两两相关，越过 0.80 的对数是零，撤销规则未触发。按字面，本书的主张没有被推翻。但把八个读数在夜与夜之间彻底打乱、破坏一切配对关系，五千次里五千次仍然满足“两两不越 0.80”。**这条规则的风险度是零：在这个样本量与这些边缘分布下，它根本不可能失败。**

按第十四章立下的第三条纪律——撤销规则触发率低于百分之五须在正文写明——**在此写明：本章第四节那条合并规则，是一条不可能触发的规则。**它命中，不构成“八个读数是八样东西”的任何证据。

这件事本身值得停一下。本章在第四节写这条规则的时候，写全了形式：写在数据之前、有放行线、有撤销规则、有三条纪律、连它自己的两处毛病都事先写明了。它通过了一切形式审查，而它一条证据也不提供。通过形式审查与提供证据是两件事，而现行的预注册制度只检查前一件。

其二，编一头两章没有通过自己的判定形式。见第五节其一：未追认段与成簇承载相关 +0.735，其中九成半是算术。0.735 没有越过 0.80，因此按规则不必合并——而 0.80 是本章第四节自己承认的任意数。若线画在 0.70，这一对此刻就该合并。这是本章那两处毛病里的第一处，在这一对上真发作了。

其三，四编的分层在数据上是方向正确、幅度不足。同编内读数对的平均相关 0.202，跨编 0.135，差 +0.067；置换检验单侧 $p = 0.0008$ ，方向为正且极显著。而第十七章预注册要求的放行线是 +0.10，按纪律判未命中。

两句都要说，且不许用后一句掩盖前一句：同编内的读数确实比跨编更相关，但这个差距比本书预设的小。导论因此不得写“四编在数据上得到支持”；能写的是——四编的分层主要是论证上的，不是统计上的；它在第六节那次粗化检验里得到的支持，比在相关矩阵里得到的强。

十五之八·由此改判的四件事

其一，第五编第一句要改。原来的问法是“十四个读数是不是同一个读数”。现在应改为：十四个读数里有八个能被放到一起问这个问题，剩下六个连问都问不了。后半句比前半句要紧——因为问不了的那三个，理由就是提出它们的那三章的论点。

其二，第二章的成簇承载比降级。不与第一章合并，但从一条读数降为“第一章那条读数在转换侧的一个投影”。两章正文各加一句公开声明：同一份记录上相关 +0.735，其中九成半由“一个一窗段本身就是一簇转换”这条算术解释。

这个处置对第二章的损失比看起来小。第二章的打磨稿早已用正确的零模型把成簇砍到只剩三个百分点，并写明“原先报告的成簇，约十九分之十八是段长分布的后果”。本次核对从完全不同的路径——不用零模型，只做一次剔除——得到同一个方向的结论。两条独立路径撞到同一处，比任何一条单独说都硬。第二章真正无主的那一格是邻记，以及十五秒对三十秒判读窗那条纯算术，两者都不受影响。

其三，第四章的残余率必须报端权对，并考虑改分母。第十七章把端权对推广到比例型读数之后查出：残余率的变异主要来自分母（深睡转入第二期的次数，解释份额 0.668），而不是来自分子（残余本身，0.074）。也就是说，第四章报告残余率高低时，实际上主要在报告这一夜从深睡浅化了多少次。而第四章那条承重指纹——规则禁转的特征不是零，是一个由例外条款决定的稳定残余——需要的是分子端说话。建议改用绝对计数或另设分母，并在正文报告端权对。△ 该端权对只有四十八夜可算，其余夜分子为零，样本本身有偏，结论强度以此为限。

其四，导论的措辞。见第七节其三。

十五之九·两条本书没有预言的命题

这次核对逼出两件十四章无一提过的事。它们不是本书的结论，是本书的待办。

命题甲·下降得越慢的夜，夜里最长的那段浅睡区间越靠前。下降时长与窄口位置相关 -0.541 ($n=135$)，控制那段区间的长度之后仍为 -0.461 ，而两段几乎重合的夜只有 17.8%。这条相关跨编，第八章与第十一章各自都没有预言过它，四编的分层也解释不了它。可检验的形式：在另一批记录上重算这一对，若相关消失或方向翻转，本条作废。

命题乙·年纪越大，换一套手册损失的转换越少。体例折叠损失与年龄相关 -0.425 ($n=153$)。机制上说得通——年长者深睡少，旧手册第三期与第四期之间的分界本来就少——但第三章正文里根本没有年龄这个变量。可检验的形式：在按年龄分层的样本上重算，若各层内部相关都接近零，则这条相关由别的量中介。

两条命题一并写进第五编的待办，理由与第十四章那条纪律相同：一个从这次核对里掉出来的相关，如果不被写下来，它下一次出现时会被当成新发现。

十五之十·本章自己的四处缺陷

其一，合并线 **0.80** 没有理论根据，而实测最强的一对恰好是 0.735——落在线下方，且拆开来看九成半是算术。若线画在 0.70，本章第八节的处置就该是合并而不是降级。这是本章最脆弱的地方，写在这里而不是附注里。可行的改进：下一版把判据改为“扣掉可由定义解释的部分之后的残余相关”，而不是原始相关——第五节其一次拆解正是这个判据的一次演示。

其二，**相关不是关系的唯一形状**。本章的规则查不出确定性的非单调关系。八个读数里有没有这种关系，本章没有查。

其三，**本次核对只在一个队列上做**。一百五十三夜全部来自同一项研究、同一套设备、同一年代。八个读数之间的相关、四个层级的稳定度差，是否在别的队列上复现，本章答不了。这是第五编的第一顺位待办。

其四，**第二层与第四层没有被检验过**。第六节那次粗化只分开了第一层与第三层；第二层在这份记录上只剩一个可算的读数，第四层一个也没有。本书主张四层不可互相还原，而本次实证只覆盖了其中两层之间的一对关系。

十五之十一·自反

按本书的判据检本章一遍。

本章做的事，是把两个读数之间的关系压成一个相关系数，再拿它与一条线比。**两个量之间是什么关系，可以有很多种形状——线性的、条件的、只在某个区间里成立的、由第三个量中介的——而本章把这一切压成一个数，然后问它超不超出 0.80**。这正是第五章描述的那个动作：一个多维的东西被压成一维，再被一个阈值裁掉。第五章会说：这个判定的分辨率，比它要判的那件事粗。

第六章会接着问：本章没有为这个判定设一个噪声底。**我不知道同一批读数在另一批可比的夜上，两两相关会自然波动多少**，因此我也不知道 0.735 与 0.80 之间那 0.065 是不是噪声。而本章的全部处置——降级而不合并——恰恰挂在这 0.065 上。

还有一层。第四节那条合并规则的风险度是零。而我在写它的时候，是先想清楚了“我希望结论是什么”，才去写那条规则与它的放行线的。第十四章已经指认过这种毛病，我在第六节的预注册里甚至专门把风险度这个装置写了进去，**而我仍然在第四节写出了一条不可能失败的规则**。

知道一个陷阱存在，与不掉进去，是两件事。

最后一层最不舒服。本章第六节那次粗化，第一条预测的风险度也是零，我在预注册里事先写明了它不作证据；后两条是真检验，本书通过了。**但如果我当初只写了第一条，本章此刻就会有一条“命中”的记录，看起来和一条真证据完全一样**。分开它们的不是数据，是我在跑之前多写的那两行。

这两行如果没写，它的缺席不会产生任何信号。

本章脚本与存档：PREREG15.md (SHA-256 df9f2034...0211066，跑前封存)·run15.py (粗化实验与风险度)·coarse.json (四个刻度上一百五十三夜的逐夜读数，供复核)。第七、八节引用的结果全部来自第十七章，脚本与存档见该章。数据为公开集，任何人可从原始文件完整复现。

第十六章 十四种看起来完整的记录

十六之一·一张表，与它的读法

本书十四章各指认一个位置。本章把十四个位置摆在一起，为每一个配一个健康对照格，并给出三条共同的签名。

先说这张表怎么读，因为读错的方式只有一种，而它很自然：把它读成一张“睡眠研究做错了十四件事”的清单。那样读会让整本书变成一次抱怨，而且是站不住的抱怨——十四处里没有一处是违规，没有一处是疏漏，没有一处是可以靠更用心的工作避免的。

正确的读法是把每一行读成一对：左边是那个位置上正在发生的事，右边是与它在记录上完全同形、而实际上无害的那种情形。一个位置之所以要紧，不是因为它坏，是因为它与无害的那一格在纸面上分不开。

表 16-1·十四个位置与它们的健康对照格

章	那里正在发生的事	与它同形而无害的那一格	分开二者需要什么
一	一段时间没有被后续事件追认，因而不构成一个可数的段	那段时间确实什么也没发生	判读中间态；或一份写明追认规则的手册
二	一段有厚度的区域被记成一条分界，产出记在两侧名下	那里确实只是一条无厚度的分界	亚窗分辨率；或一个专给边界的字段
三	几个同时成立的候选里，格式选了第一名而扔掉分差	第一名压倒性领先，分差本来就无意义	概率向量；或允许弃答的格式
四	一条转换被规则禁止这样记，于是在记录上近乎为零	那条转换在生理上真的不可能	换一套规则重判同一份原始信号
五	三个不同的要求在使用者的验收工具上映射为同一个读数	那三个要求本来就是同一件事	测一次使用者的验收分辨率
六	一种办法开一条通路的同时封另一条，净效果为正	那种办法只开不封	同一人无干预两夜的噪声底
七	一个差值读数的变异几乎全部来自其中一端	两端各贡献一半	把两端的方差与协方差一并公布
八	一个量可算、无栏、从未被裁决过	那个量已经被裁过，结论是无效应	一把细于三十秒的刻度+一次标准化扰动
九	主诉与客观给出两套互不细化的划分，而用同一个词	一套确实细化另一套	同一批人身上同时有细致主诉量表与多导记录
十	一次事件被读成一个人身上的一种能力	那确实是一种能力	同一个人的多夜记录
十一	稳定的那个量没有位置，而清单总能找到一个位置	位置确实稳定，只是还没找准	同一人跨夜的位置重复度
十二	关键回应仍由睡者承担，而分期与时长完全正常	回应确实已被转交	在睡眠记录上加一栏“由谁接手”
十三	一个人的睡眠依赖另一个人的例行安排，两侧都不知道	那个依赖不存在	向第二个人问他自己的日常安排变更
十四	手段全部押在症状可见的那一端，而因果强度分不出时点	那一端确实是因果强度最大的时点	一份带光照与活动的多日记录

这张表的第三列是本书全部的实操内容。十四行里，有九行只需要在现行记录上加一个字段或多记一次；只有三行需要新的测量能力（判读中间态、亚窗分辨率、使用者验收分辨率）。这是本书主张这些缺口“不是技术难题”的全部根据。

十六之二·签名一：短期表现为改善

十四个位置有一个共同的、令人不安的性质：**每一次损失都以一次简化的形式发生，而简化在当下总是被算作进步。**

- 把旧体例的深睡两期合并成一期，逐夜转换数中位下降百分之一一点七、四十夜合计下降百分之十六点三，记录变干净了，判读者之间的一致性也提高了（第三章）。
- 用占优规则给每个窗一个标签，判读速度快得多，报告可比性好得多（第一、二章）。
- 把三个要求压成一个综合分，评价变得可比、可排序、可追踪（第五章）。
- 把“深睡多少分钟”作为主结局，试验的样本量算得出来、终点写得清楚（第七、八章）。
- 把办法集中在睡前那两三个小时，执行成本最低、当场就有感觉（第十四章）。

没有任何一条是错的。每一条都解决了一个真实的问题，而且解决得很好。这就是为什么它们能够稳定下来，成为标准，写进手册。

推论有两条，都要写清楚：

第一条，本书不主张回退。把 N3 拆回 S3 与 S4、把综合分拆成三个分数、把办法搬到白天——本书一条也不建议。第十四章那一处最明白：控制工程只允许它诊断，不允许它开处方，因为前馈控制在模型不准时更糟。

第二条，也是要紧的那条：一个装置若在短期表现为改善，那么“它带来了改善”这件事，**不能被用作“它没有代价”的证据。**这两句在实践中总被当成一句。

十六之三·签名二：失效不产生任何信号

这是三条里最要紧的一条，因为它解释了为什么这些位置能存在几十年而无人指认。

一个普通的失败会留下痕迹：数据缺失、异常值、检验不显著、判读者意见不一。而本书这十四处的共同点是——**失败之后，记录与没有失败时完全一样。**

- 一段时间没有被追认，记录上不会出现空格，只会出现相邻分期的一个延续（第一章）。
- 判余被扔掉之后，记录上留下的是一个和别的标签一模一样的标签（第三章）。
- 一条转换被规则涂抹之后，记录上是一个接近零的数字，而“真的不可能”也是一个接近零的数字（第四章）。
- 一次改善落在使用者分不开的地方，他会照常给出一个评分，而那个评分和真的改善了完全一样（第五章）。
- 一条通路被封掉之后，读数上什么也不会发生，因为另一条通路补上了（第六章）。
- 一段支承被撤走之后，两侧都不产生记录——**没有人违约**（第十三章）。

这一条在第十三章有一次最狠的兑现，值得原样抄在这里：

一篇没有采集过供给方的论文，读起来和采集过的完全一样。参考文献一样长，论证一样连贯，表格一样齐整。若不是那一节写出来，读者没有任何办法从文本上分辨这两种论文。

而它在本书自己身上也发作过一次，写在第十七章：本书最想成立的那条预测在检验里命中，事后算出的风险度是零——**它不可能失败，因此它的命中不是证据。**若不是那一章事先写死了要算风险度，那条预测会以“命中”的形式进入结论，看上去和一条真证据完全一样。

它不产生信号。

十六之四·签名三：越正规化越自我加固

第三条签名是三条里最反直觉的：**这些位置不会因为工作做得更规范而减少，反而会因此变得更稳定、更像一条规律。**

机制在第四章讲得最清楚。一条被判读手册压低的转换，其残余高度由例外条款的宽窄决定。**如果提高判读者之间的一致性，随机的分歧会减少，而规则造成的那个差距不变。**于是一致性越高，规则造成的结构反而越干净、越不像噪声、越像一条生理规律。

这条签名给出了三处可判定的形式：

- 第四章：**提高编码者一致性，误差型的偏差减小，规则型的残余率不变甚至更稳定。

·第十四章：一条在零假设下也有百分之九十九命中率的预注册预测，通过得越“漂亮”，越说明判定规则宽松——通过形式审查与提供证据是两件事，而现行制度只检查前一件。

·第十一章：清单越长，能找到的“原因位置”越多，每一条都“有时管用”；于是“还有一个原因没找到”这个信念被越加固。

三处的形状相同：正规化提高的是可重复性，而这十四个位置上的问题不是可重复性问题。

十六之五·三条签名合起来意味着什么

单独看，三条签名各自都不足以立论。合起来它们给出一条可以拿去核对任何领域的判据：

若一个装置同时满足三条——它带来的改善在短期可见、它的失效不产生任何读数、它随着正规化而更加稳固——那么它所在的位置上，很可能有一样东西在被持续地扔掉，而没有任何现行程序会发现这件事。

这条判据可以被证伪，方式有两种：

其一，找到一个同时满足三条签名、而实际上什么也没扔掉的装置。若这类例子很多，三条签名就只是“有用的东西的一般性质”，不指向任何缺口。

其二，找到一个扔掉了东西、却不满足三条签名的位置。若这类例子很多，三条签名就不是必要条件，本书的判据会漏掉大量真实缺口。

本书没有系统地找过这两类反例。这是全书最大的一处方法论空缺，写在这里而不是附注里。

十六之六·一句必须说的话：这些缺口补上之后，可能更糟

本书十四章给出了十四条处方，全部是程序性的：加一栏、标一个版本号、把两端分开公布、给每条建议标上时点、基线期多记一夜、向第二个人问一句。

这些处方看上去成本很低。但它们合起来是一次记录负担的增加，而记录负担是有代价的——判读时间变长、量表变厚、患者的填答质量下降、临床医生更不愿意用。若补上这些缺口之后，睡眠研究并没有变得更好，甚至因为负担加重而变差，那么本书的处方是净损失。

本书没有能力事先排除这一条。能做的只有三件：

第一，把处方全部限制在程序性的层面，不动任何分类、不改任何界线、不要求任何新设备。第二章那条“另账登记”是本书处方的标准形态：不动现行分类与份额恒等式，另设一份不参与加总的登记。

第二，把每一条处方的成本写出来，供决定的人权衡。附录一逐条列出。

第三，也是最要紧的一条：本书的处方全部是可撤的。加了一栏发现没用，删掉即可，不留后果。这与本书批评的那些装置正好相反——那些装置一旦立起来，扔掉的东西就不再有痕迹，撤不回来。

这一条不对称，是本书愿意提出这些处方的唯一理由。

十六之七·十四个位置合起来是不是同一个位置

最后处理一个必然会被问到的问题：既然十四处的形状如此一致，它们是不是同一个位置的十四次投影？

本书的回答分两半。

一半是：在某个层面上，是的。十四处共用同一个机制——某个装置被要求给出恰好一个答案，而给出答案这一步扔掉了别的候选，且扔掉不留痕迹。这个机制就是全书的全书赌的那句话，它不是十四条，是一条。

另一半是：在可操作的层面上，不是。第十五章给出了四条互不蕴含：补上判读窗层不会补上读数层，补上读数层不会补上个体层，补上个体层不会补上主体层，补上主体层不会动到判读窗层。一个共同的机制，与四个必须分别处理的位置，这两件事不矛盾。

打个不太精确的比方：所有的漏水都是同一件事——水往低处走；而堵屋顶的漏、堵管道的漏、堵地基的漏，是四件不同的工程。

⚠️ 但这里必须补上第十七章那条对本书不利的结果：**四层的分层在数据上得到的支持是方向正确、幅度不足**（同编内相关比跨编高 0.067，置换检验极显著，而预注册要求 0.10，判未命中）。因此上面那个“四件不同的工程”是论证上的判断，不是统计上的结论。**本书在此不主张更多。**

第十七章 一次同批复算

十七之一·为什么这件事只有这本书做得成

前面十四章各自提出了一个读数。每一章都用公开数据跑过自己那一个，每一章都报告了它的分布、它的稳健性、以及它在哪些条件下会失败。十四章因此有十四条各自成立的证据链。

但十四条各自成立的证据链，不构成一本书。它们合起来还欠一件事：**这十四个数，是不是同一个读数换了十四个名字？**

这个问题在别的书里问不出来，因为别的书的十四篇文章各用各的数据。本书的情况不同：十四章里有十一章用的是同一份公开判读记录——Sleep-EDF Expanded 的老化队列，七十八名受试者、一百五十三夜、年龄二十五到一百零一岁。同一批夜被十一条不同的产线各切过一遍，而没有任何一次是**把它们放在同一张表上算的。**

这一章做的就是那件从没做过的事。在同一份记录上，把能算的读数一次算齐，给出它们两两之间的相关。而在看到相关之前，先写死一条合并规则：**任何两个读数若相关过高，对应的两章必须合并，不许用“侧重不同”保留两个名字。**

这条规则是本章存在的全部理由。没有它，这一章就只是一次汇总；有了它，这一章有可能拆掉本书自己的编次。

本章的结论提前说在这里，因为它对本书不利的部分比有利的部分多：

- 十四个数里，**只有八个是同一种数学对象**，能被放进同一张表。
- 八个读数里没有任何一对越过合并线。但这条**“没有越线”**的预测，经事后检验，在设计上就不可能失败——它的风险度是零。
- 本书按“强制作答发生在哪一层”分的四编，在数据上得到了方向正确但幅度不足的支持：同编内的相关确实高于跨编，差值 +0.067，置换检验单侧 $p = 0.0008$ ，而预注册要求的放行线是 +0.10。**按纪律，这一条判未命中。**
- 编一头两章的两个读数在同一份记录上相关 **+0.735**，而其中 **95% 由一条算术解释**：一窗段本身就是一簇转换。两章各自独立提出、各自命名、各自与全球近邻划界，**从未发现这一点。**
- 八个读数与总睡眠时长的相关中位数是 **0.038**，最大 0.142。这是本章唯一一条对本书有利而且站得住的描述性结果。

十七之二·预注册

本章的预注册文件 PREREG17.md 在脚本第一次运行之前写定并封存，SHA-256 为：

、

5455ec98871dd2aefb722c4f4eb02a480559622b5b6f5e3709dde4bed7c0b926

、

 预注册第二节那张表里的编号是**成书前的草案章序**，与定版章序在第九至十四章之间有偏移；该表逐行写明了章名，**以章名为准**，本章正文一律用定版章号。

文件第零节按本书通例先列“写定时我已经知道什么”：各章已发表的单读数数值全部已知（未追认段占段数百分之三十二点四、体例折叠使转换减少百分之十六点三、成簇承载百分之七十五点三、窄口位置跨夜相关负零点零七九而深睡百分比正零点八二三，等等），解析器也已与第一章前四十夜的数字对过表。**未知的是任何一对读数之间的相关——本章的四条预测全部关于这些相关。**

四条预测与它们的判定线：

编号	预测	放行线	撤销规则				
M1	八个读数两两 \	$\rho \backslash$	全部不超过 0.80	0 对越线	出现 ≥ 3 对越线，则“十四个数是十四样东西”整		

					条撤回，第五编改写为《十四个名字，K个量》，且不许换定义重试		
M2	同编内读数对的平均 \	$\rho \backslash$	高于跨编平均 \	$\rho \backslash$		差 $\geq +0.10$	—
M3	八个读数与总睡眠时长的 \	$\rho \backslash$	中位数不超过 0.40	≤ 0.40	若超过，本书大部分读数只是“这一夜睡了多久”的改写，第五编须整章重写		
M4	比例型读数中至少三个的两端解释份额明显不对称	≥ 3 个	—				

预注册第五节另写死一条：**每一条预测在看到结果之前必须先算它的风险度**——把各读数在夜与夜之间独立打乱、保留各自的边缘分布、破坏跨读数的配对关系，重复五千次，看该预测在零假设下仍然命中的比例；风险度等于一减去这个比例。这个装置来自第十四章，那一章用它查出自己的预注册“在设计上就不可能失败”。本章把它调转过来对准自己。

预注册第六节还事先写明了本章可能出的三处错，其中第二条是：五个比例型读数天然共享“分母是这一夜有多长”这个成分，因此 M3 对本书不利的方向是先验合理的；**若 M3 不成立，那不是本书赢了，是我这条预测设得太松**。这句话在结果出来之后变得很重要，见十七之六。

数据口径全部写死在预注册第一节，此处只重复三条最要紧的。其一，分期折叠按现行体例，第三期与第四期合并。其二，剪裁去首尾连续的清醒与非五分期窗，首尾各回留三十分钟。其三，**入睡后清醒只计入睡点到最后一次入睡之间，最后一次醒来之后的时间不计**——这是第十章那次自查的结论，本章沿用。

解析器与第一章前四十夜的对表附在预注册末尾：窗数 43330 对 43329，时长逐项吻合（差不超过一分钟），转换数 4405 与第一章原稿逐位相同。**段数差三十二（百分之零点七），来自跨非五分期窗处是否断段的口径差，本章不追平，如实并报。**

十七之三·十四个读数里，只有八个是同一种数学对象

这一节是本章最要紧的一节，而它的内容在跑任何统计之前就已经确定。

预注册第二节要求逐个判定十四章的读数能不能进同一张表。判定结果不是“能”与“不能”两类，是三类：

第一类·每夜一个数（八个，进主表）

章	读数	定义
一	未追认段比例 U	长度为一窗的段数 ÷ 总段数
二	成簇承载比 C	落在含两次以上转换的八窗粗窗里的转换数 ÷ 总转换数
三	体例折叠损失 F	(旧体例转换数 - 现行体例转换数) ÷ 旧体例转换数
四	残余率 Res	深睡直接转入第一期的次数 ÷ (该次数 + 深睡转入第二期的次数)
八	下降时长 T	入睡点到首个深睡窗的窗数
九	命中界线数 H	三条常用界线命中条数，取值零到三
十	长断占比 Lb	一夜最长的一次中断 ÷ 入睡后清醒总

		量
十一	窄口位置 π	睡眠期内最长的连续“无深睡无快速眼动”区间中点的相对位置

第二类·是别的数学对象（三个，进不了主表）

- 第六章的**自噪底**是一个跨夜配对差，不是每夜一个数。它需要同一个人的两个可比夜，样本单位是人不是夜。
- 第七章的**端权对**是一对解释份额，它是一个可以对任何差值读数施加的装置，本身不是读数。本章在十七之九把它推广到本书的比例型读数上。
- 第十四章的**预测力—滞后剖面**是一条八点曲线。把一条曲线压成一个数，正是第五章说的那件事；本章不做。

第三类·在这份记录上算不出来（三个）

- 第五章的**索验比**分母 δ 是使用者的验收分辨率。它从来没有被测量过一次。
- 第十二章的**下一拍托管率 R** 需要“这一夜有哪些必须作答的事、由谁接手”的字段。睡眠记录里没有这个字段。
- 第十三章的**他源率 ϕ** 根本不是睡眠记录的量，它是对评估工具计分条目的清点。

另有三个从属读数同样算不出来。第一章的**追认延迟 L** 需要判读过程的中间态，至今未被测量过一次；第二章的**界带宽度**需要亚窗分辨率；第三章的**判余**需要模型的概率向量，第三章用原始信号跑过一次，但那不是判读标注里的东西。

于是分布是 八 / 三 / 三（另加三个从属读数）。

这个分布本身是本章第一个结果，而且它对本书是有利的——但有利的方式不是我原先设想的那种。请看第三类那三个：**索验比的分母、下一拍托管率、他源率**。它们算不出来的理由，逐条就是那三章各自的论点：

使用者的验收分辨率没有被测过，所以第五章说“这个要求无法被验收”；

记录里没有“谁接住下一拍”这一栏，所以第十二章说“两笔账互不登记”；

计分条目里没有一条问对面那个人自己做了什么，所以第十三章说“账按主体划”。

这三章的读数在本章算不出来，不是本章的失败，是那三章的命题在本章的记录上原样发作了一次。我在写预注册第二节时并没有预见到这一点；它是把可算性判定逐条写下来之后才显出来的。

而这件事有一个不舒服的推论，写在这里：**本章能够放进同一张表的那八个读数，恰恰是最不激进的八个**——它们全部可以从现成的判读标注里算出来，也就是说，它们全部落在现行记录制度已经愿意登记的那个范围之内。本书最想指出的那几样东西，正因为记录里没有它们的位置，所以也进不了本章这张验证表。**本章的验证力，在本书最要紧的地方最弱。**

十七之四·同一份记录上的八个读数

一百五十三夜全部纳入，零剔除；受试者七十八人。

表 17-1·八个读数与两个对照量的分布（153 夜）

读数	有效 n	中位	四分位（下 / 上）
未追认段比例 U	153	0.3146	0.2697 / 0.3676
成簇承载比 C	153	0.8175	0.7458 / 0.8658
体例折叠损失 F	153	0.0169	0.0000 / 0.1481
残余率 Res	135	0.0000	0.0000 / 0.0488
下降时长 T （窗）	135	76.0	46.5 / 187.0
命中界线数 H	153	2.0	1.0 / 2.0
长断占比 Lb	153	0.3876	0.2517 / 0.6027
窄口位置 π	153	0.5419	0.2422 / 0.7887
（对照）总睡眠时长（分）	153	428.0	387.0 / 464.0

(对照) 深睡百分比	153	7.98	1.36 / 16.74
------------	-----	------	--------------

三处须先说明：

其一，残余率的中位数是零。一百三十五个可算的夜里，超过一半的夜从来没有出现过“深睡直接转入第一期”这件事。第四章报告的百分之二点七五是把全部夜的转换汇总之后的比例；受试者层面的中位数是零。这与第四章自己在补跑里得到的教训一致——汇总计数不能当配对证据用——但第四章正文里那个二点七五，读者仍然会当作“一个典型的夜大约是这个数”。它不是。建议第四章正文补一句中位数。

其二，命中界线数的中位是二。一百五十三夜里的中位数情形是：三条常用临床量化界线，命中两条。而这批人在库说明书里载明为健康、不服任何与睡眠有关的药物。第九章与第十章各自从不同角度报告过这件事，本章只是在同一张表上把它再确认一次。

其三，体例折叠损失的下四分位是零。四分之一以上的夜，换一套体例一次转换也不会少——因为那些夜里根本没有旧体例第三期与第四期之间的分界。第三章报告的百分之十六点三是四十夜合计的比例，逐夜中位数只有百分之一点七。这个差距比残余率那一处还大。第三章的论证不依赖这个数的大小（它依赖的是“这些转换在两份都正确的记录里一份有一份没有”这件事本身），但正文那个十六点三同样容易被读成典型值。

十七之五·相关矩阵与合并判定

表 17-2·八个读数两两 Spearman 相关（成对删除，n = 135-153）

	U	C	F	Res	T	H	Lb	π
U 未追认段	—	+0.735	+0.210	+0.051	-0.145	+0.014	+0.012	+0.092
C 成簇承载	+0.735	—	+0.152	+0.207	-0.041	+0.092	-0.071	+0.099
F 体例折叠	+0.210	+0.152	—	+0.122	-0.275	-0.247	-0.014	+0.119
Res 残余率	+0.051	+0.207	+0.122	—	-0.081	+0.131	-0.046	+0.094
T 下降时长	-0.145	-0.041	-0.275	-0.081	—	+0.253	+0.203	-0.541
H 命中界线数	+0.014	+0.092	-0.247	+0.131	+0.253	—	-0.033	-0.132
Lb 长断占比	+0.012	-0.071	-0.014	-0.046	+0.203	-0.033	—	-0.176
π 窄口位置	+0.092	+0.099	+0.119	+0.094	-0.541	-0.132	-0.176	—

M1 判定：越过 0.80 的对数为零，撤销规则未触发。

按预注册，这意味着“十四个读数是十四样东西”这条主张在本次检验里没有被推翻。

但这条命中不提供任何证据。见下一节。

两对值得单独处理的相关：

- 未追认段 × 成簇承载 = +0.735（编一内，n=153）。这是全表最强的一对，离合并线只差 0.065。我把线画在 0.80 是任意的，而 0.735 已经高到不能当作“两个不相干的量”来写。十七之七拆它。
- 下降时长 × 窄口位置 = -0.541（编二 × 编三，n=135）。这是跨编最强的一对，也是本表唯一一个跨编的强相关。十七之七也拆它。

十七之六·四条预测的下场，以及一条对本章自己的判决

表 17-3·四条预测的结果与风险度（5000 次独立打乱）

预测	放行线	实测	判定	零假设命中率	风险度
M1 两两不越 0.80	0 对	0 对	命中	1.0000	0.0000
M2 同编内高于跨编	差 $\geq +0.10$	+0.067	未命中	0.0000	1.0000
M3 与总睡眠时长中位 ≤ 0.40	≤ 0.40	0.038	命中	1.0000	0.0000
M4 至少三个两端不对称	≥ 3 个	2 个	未命中	—	—
(撤销规则触发)	≥ 3 对越线	0 对	未触发	0.0000	—

这张表要读的不是那一列“判定”，是那一列“风险度”。

M1 的风险度是零。把八个读数在夜与夜之间彻底打乱、破坏一切配对关系，五千次里有五千次仍然满足“两两不越 0.80”。也就是说：**本书最想成立的那条预测，在这个样本量与这些边缘分布下，根本不可能失败。**它命中，不构成“八个读数是八样东西”的任何证据。

这与第十四章查出的病是同一个病。第十四章那次是作者自己的两条预测在零假设下命中率百分之九十九点九八与百分之九十九点零六，据此提出了风险度这个装置。**本章拿这个装置对准本书自己最想要的那条结论，结果比第十四章更糟：那两条还有万分之二与百分之一的失败余地，本章这一条是零。**

而撤销规则的触发率同样是零。按第十四章立下的第三条纪律——“撤销规则触发率必须一并报告，低于百分之五须在正文写明”——**本章的撤销规则形同虚设，在此写明。**

M3 的风险度也是零，理由相同，而且预注册第六节第二条已经事先写到过类似的担心。打乱之后一切相关趋近于零，中位数当然不超过 0.40。**M3 这条预测同样不可能失败。**

于是本章四条预测里，**两条命中的都零信息量，两条未命中的都有信息量。**这句话是本章最诚实的一句话，也是本章对第五编的第一条贡献：

一个能通过全部形式审查的预注册——写在前、封了哈希、有放行线、有撤销规则——仍然可以一条证据也不提供。**通过形式审查与提供证据是两件事，而现行的预注册制度只检查前一件。**

这句话不是本章的发现，是第十四章的发现；本章的贡献是在自己身上又发作了一次，而且是在本书最想要的那条结论上。

M2 未命中，但方向是真的

M2 的实测差是 +0.0667，未达 +0.10 的放行线，**按纪律判未命中。**

但 M2 的风险度是 1.0000——打乱之后它一次也不命中。于是有必要问：+0.0667 在零分布里有多极端？补跑（探索性，非预注册）给出：

零分布中位 -0.0017，百分之九十区间 [-0.0338, +0.0331]，**实测 +0.0667 的单侧 $p = 0.0008$ 。**

方向为正、极显著，幅度不足。两句都要说，而且不许用后一句掩盖前一句：预注册写的是 +0.10，实测是 +0.067，**这条预测判未命中。**本书按“强制作答发生在哪一层”分的四编，在这份数据上得到的支持是**真实但弱**的：同编内的读数确实比跨编更相关，只是没有我预设的那么多。

对编次的后果写在十七之十。

M4 未命中

五个比例型读数里，只有两个呈现预设的不对称（一端解释份额 ≥ 0.60 、另一端 ≤ 0.20 ），预注册要求至少三个。**判未命中。**但那两个的内容比命中与否要紧，见十七之九。

十七之七·两对最强相关的拆解

未追认段 × 成簇承载 = +0.735，其中 95% 是算术

第一章的未追认段比例数的是**长度为一窗的段占全部段的比例**。第二章的成簇承载比数的是**转换有多大比例挤在含两次以上转换的粗窗里**。

两章各自独立提出、各自命名新对象、各自与全球近邻逐条划界。第一章与周期性交替模式、可变时间单元问题、通约化划界；第二章与生态交错带、睡眠转换动力学、可变时间单元问题划界。**两章都没有与对方划界，因为两章都不知道这件事：一个一窗段，本身就是两次转换。**

一个长度为一窗的段，进去一次、出来一次，这两次转换必然落在相邻的位置上，因此必然构成一个簇。**未追认段多的夜，成簇承载比必然高——这不是发现，是算术。**

补跑（探索性，非预注册）把它量化了：把“起点或终点属于某个一窗段”的转换从成簇承载比的分子与分母里全部剔除，重算相关：

原始 $U \times C$: $\rho = +0.735$ ($n=153$)

剔除一窗段造成的转换后: $\rho = +0.033$ ($n=153$)

相关的百分之九十五由这条算术解释。

这一刀对两章的后果不同：

对第一章无损。第一章的承重不是这个相关，是“段落账与时长账同时成立而互相看不见”那对数（未追认段占段数百分之三十二点四、占时长百分之三点三五）。这一条不受影响。

对第二章是又一次打击，而且方向与第二章自己的打磨稿一致。第二章的打磨稿已经用正确的零模型把成簇那条读数砍到只剩三点五个百分点，并明写“本章原先报告的成簇，约十九分之十八是段长分布的后果”。本章从完全不同的路径——不用零模型，只做一次剔除——得到同一个方向的结论：**成簇这条读数扣掉一窗段之后几乎不剩什么**。两条独立路径撞到同一处，这比任何一条单独说都硬。

处置建议（第十五章执行）：不合并两章，但**第二章的成簇承载比必须降级**，从一条读数降为“第一章那条读数在转换侧的一个投影”。两章正文各加一句公开声明：同一份记录上 $\rho = +0.735$ ，其中百分之九十五由“一窗段本身即一簇转换”这条算术解释。第二章真正无主的那一格——邻记，以及十五秒对三十秒判读窗那条纯算术——不受影响，第二章的打磨稿已经把承重压到那两样上。

下降时长 × 窄口位置 = -0.541，不是定义蕴含

第八章的下降时长数的是**入睡点到首个深睡窗有多久**。第十一章的窄口位置数的是**夜里最长那段“无深睡无快速眼动”区间的中点落在哪里**。

一个自然的怀疑是：下降段本身就是一段“无深睡无快速眼动”，所以下降越长，最长的那段区间越可能就是下降段， n 自然越小。若真如此，这一对是定义蕴含，不是发现。

补跑否掉了这条：

控制窄口长度之后的偏相关: $\rho = -0.461$ ($n=135$)，只掉了 0.08。

“最长无深睡无快速眼动区间”与“下降段”几乎是同一段的夜: $24 / 135 = 17.8\%$ 。

八成以上的夜里，那两段不是同一段。这一对相关是实质的：**下降得越慢的夜，夜里最长的那段浅睡区间越靠前**。

这条相关跨编（编二 × 编三），也是本表唯一一个跨编的强相关。它对本书的编次是个麻烦——见十七之十。它同时也是本书十四章里**没有任何一章预言过的一件事**，值得单列为第十五章的一条待检验命题。

十七之八·年龄不是共同驱动

老化队列的年龄从二十五岁到一百零一岁。最明显的竞争解释是：八个读数的一切相关都由年龄驱动。

表 17-4·八个读数与年龄的相关

读数	ρ (对年龄)	n
命中界线数 H	+0.511	153

体例折叠损失 F	-0.425	153
下降时长 T	+0.203	135
成簇承载比 C	+0.085	153
未追认段比例 U	+0.063	153
残余率 Res	+0.039	135
窄口位置 π	-0.015	153
长断占比 Lb	-0.004	153

八个里只有两个与年龄有实质相关。命中界线数那一条是意料之中的——第九章与第十章各自都报过它。**体例折叠损失那一条是新的**：年纪越大，换一套体例损失的转换越少。机制上说得通（年长者深睡少，旧体例第三期与第四期之间的分界本来就少），但**这一条此前没有任何一章提出过**，第三章正文里没有年龄这个变量。它是本章第二条值得单列的待检验命题。

控制年龄之后的最强六对：

对	原始 p	控年龄后	Δ
未追认段 × 成簇承载	+0.735	+0.734	-0.001
下降时长 × 窄口位置	-0.541	-0.547	-0.006
体例折叠 × 下降时长	-0.275	-0.218	+0.056
下降时长 × 命中界线数	+0.253	+0.180	-0.074
体例折叠 × 命中界线数	-0.247	-0.038	+0.209
未追认段 × 体例折叠	+0.210	+0.262	+0.052

两条结论：

其一，年龄不是共同驱动。 最强的两对控年龄后几乎不动（ Δ 分别为 -0.001 与 -0.006）。这条最明显的竞争解释被排除。

其二，有一对被年龄完全解释掉了。 体例折叠 × 命中界线数从 -0.247 掉到 -0.038。这一对的相关全部来自“两者各自与年龄相关”。如实报告。

十七之九·端权对推广

第七章提出端权对：一个差值读数里说话的永远是绝对离散度大的那一端，而哪一端更散，由构造者挑了什么量决定。**因此报告一个差值读数而不报端权对，等于把一个数交给读者，而不告诉他这个数在说哪一端。**

第七章的对象是差值。本书还有五个比例型读数。比例取对数之后就是差值，因此端权对可以直接推广过去：对读数 a/b ，取 $x = \log a$ 、 $y = \log b$ 、 $D = x - y$ ，报告 x 与 y 各自对 D 的解释份额。

表 17-5·五个比例型读数的端权对

读数	n	分子端 R^2	分母端 R^2	两份额和	
未追认段比例 U	153	0.745	0.383	1.128	
成簇承载比 C	153	0.705	0.539	1.244	
体例折叠损失 F	84	0.908	0.022	0.930	★不对称
残余率 Res	48	0.074	0.668	0.742	★不对称
长断占比 Lb	153	0.450	0.154	0.605	

M4 判未命中（只有两个不对称，预注册要求三个）。但这两个的内容比命中与否重要，而且方向相反：

体例折叠损失说话的完全是分子。 分子端解释份额 0.908，分母端 0.022。也就是说，这个读数的全部变异来自“被折叠掉了多少次转换”，而“旧体例本来有多少次转换”几乎不说话。这对第三章有利：**这个读数确实在测它想测的那件事。**

残余率说话的完全是分母，这对第四章不利。 分子端 0.074，分母端 0.668。第四章把这个读数命名为“残余率”，意思是“一条被规则压低的转换在记录中保留的出现比例”，承重在**残余**。但在同

一份记录上，这个比例的变异主要来自残余的多少（深睡转入第一期的次数），而来自参照量的多少（深睡转入第二期的次数）。

换句话说：第四章报告“残余率”的高低时，实际上主要在报告“这一夜从深睡浅化了多少次”，而不是“规则的例外条款留下了多少残余”。

这一条对第四章是实质性的。第四章那条承重指纹——“规则禁转的特征不是零，是一个由例外条款决定的、稳定的非零残余”——需要的是分子端说话；而数据说分母端说话。建议第四章改用绝对计数（深睡转入第一期的次数）或另设一个分母，并在正文报告端权对。第四章本来就是本批评分最低的一篇（一四四六），这一处是它的又一个必修项。

另须注明：残余率的端权对只有四十八夜可算（其余夜分子为零，取不了对数）。这个样本本身是有偏的——只有那些“规则的例外条款真的发作过”的夜才进得来。这一处限制写在这里，结论的强度以此为限。

十七之十·这次复算对编次的六条后果

其一，四编的分层得到方向正确、幅度不足的支持。同编内平均 $|p|$ 0.202，跨编 0.135，差 +0.067，置换单侧 $p = 0.0008$ ，而预注册放行线是 +0.10。判未命中。编次可以保留，但导论里不得写“四编在数据上得到支持”这类话；能写的是：“同编内的读数确实比跨编更相关，但这个差距比本书预设的小；四编的分层主要是论证上的，不是统计上的。”

其二，编一的头两章必须正面处理那条算术。处置见十七之七：不合并，第二章的成簇承载比降级为投影，两章各加一句公开声明。

其三，编二与编三之间有一条本书没预言的实质相关。下降时长 $\times$ 窄口位置 -0.541，控制窄口长度后 -0.461，八成以上的夜两段不重合。这条相关跨编，编次解释不了它。建议不掩盖，写成第十五章的一条待检验命题：下降得越慢的夜，夜里最长的浅睡区间越靠前。

其四，编二在本表上只剩一个读数，M2 的检验力因此很弱。编二四章里，第五章的分母算不出、第六章是跨夜差、第七章是装置，只有第八章的下降时长进了主表。这一条在预注册第六节第三条已经事先写明，命中与否都不构成层级结构的证据。编二的读数为什么系统性地进不了这张表，本身值得第十五章处理——编二讲的正是“一个被公布的数”，而它的四个读数里有三个根本不是一个数。

其五，第五编的第一句话要改。原设想是“十四个数是不是同一个读数”，现在应改为“十四个数里有八个能被放到一起问这个问题，而剩下六个连问都问不了”。后半句比前半句要紧。

其六，本章不支持任何一次合并，但也不支持“十四样东西”这个说法。M1 命中而零信息量，M4 未命中，M2 未命中。本章能诚实说出的最强一句是：在这份记录上，没有找到必须合并的两章；而这句话的证据力，比它听起来弱得多。

十七之十一·本章自己的三处缺陷

其一，八个读数里有五个是比例，而比例的两端在本章之外没有被分开公布过。表 17-5 是本书第一次为这些读数算端权对；在此之前，前面十四章报的全部是比值本身。按第七章自己的纪律，那些报告都欠着两端。这一处是本书对自己的一笔欠账，不是本章补出来的功劳。

其二，本章用的合并线 0.80 是任意的。预注册里写死了它，因此它是一条正当的预注册条款；但它没有理论根据。而实测最强的一对是 0.735——恰好在线下方，且拆开来百分之九十五是算术。如果线画在 0.70，本章的结论会完全不同。这一点是本章最脆弱的地方，把它写在这里，而不是写在附注里。可行的改进：下一版改用“扣掉可由定义解释的部分之后的残余相关”作判据，而不是原始相关——本章十七之七那次拆解正是这个判据的一次演示。

其三，本章只在一个队列上做。老化队列一百五十三夜，全部来自同一项研究、同一套设备、一九八七至一九九一年采集。药物队列的四十四夜本章没有并入主表（受试者不同、记录方案不同、年

代不同，第六章已经指出过这种对照的不完美）。八个读数之间的相关是否在别的队列上复现，本章答不了。这是第十五章的第一顺位待办。

十七之十二·自反

按本书的判据检本章一遍。

本章做的事，是把十四个读数放进一张表，用一条事先写死的规则判它们该不该合并。而这条规则要求每一对读数给出一个数——一个相关系数——然后拿这个数与一条线比。这正是本书第六章描述的那个动作：把一个多维的关系压成一维，再拿一个阈值去裁。

两个读数之间是什么关系，可以有很多种形状：线性的、条件的、只在某个区间里成立的、由第三个量中介的。本章把这一切压成一个 Spearman 相关系数，然后问它是不是超过 0.80。第五章会说：这个判定的分辨率，比它要判的那件事粗。而第六章会说：本章没有为这个判定设一个噪声底——我不知道同一批读数在另一批可比的夜上，两两相关会自然波动多少，因此我也不知道 0.735 与 0.80 之间那 0.065 是不是噪声。

还有一层。本章的八个读数之所以能进主表，是因为它们都能从现成的判读标注里算出来。而“能从现成记录里算出来”，正是本书用来定义算熟未裁（第八章）的三个条件里的第一个。本章验证的这八个读数，全部是可算的那一类；而本书真正想指出的那几样，恰恰因为记录里没有它们的位置而无法进入本章。本章因此在本书最要紧的地方最没有话语权。

最后一层，也是最不舒服的一层。本章最想成立的那条预测，风险度是零——它在设计上就不可能失败。而我在写预注册的时候，是先想清楚了“我希望结论是什么”，才去写那条预测与它的放行线的。第十四章已经指认过这种毛病，我在写本章预注册第五节时甚至专门把风险度这个装置写了进去；而我仍然写出了一条不可能失败的预测。

这说明：知道一个陷阱存在，与不掉进去，是两件事。本书用了十四章讲“失效不产生信号”，而本章这条零风险度的预测，如果不是我事先写死了要算风险度，它就会以“命中”的形式进入第五编，看起来和一条真证据完全一样。

它不产生信号。

本章脚本与存档：PREREG17.md (SHA-256 5455ec98...c0b926, 跑前封存)·edfio.py (判读标注解析)·run17.py (主表、相关矩阵、风险度、端权对)·run17b.py (四问补跑，探索性)·readings.json (一百五十三夜逐夜八读数，供复核)。数据为 Sleep-EDF Expanded 公开集，任何人可从原始文件完整复现本章全部数字。

第十八章 这套说法共同的对手

十八之一·为什么要在书的层面再划一次界

十四章每一章都做过自己的分界节，逐条与最接近的既有说法划清。但章级的分界有一个系统性的盲区：一位反对者可以不反对任何一章，而反对整本书。他的话大致是这样的：

你这十四处，各有各的名字，可它们合起来无非是几件早就有人说过的事：分类会引入伪结构；差值读数不可靠；测量有分辨率下限；从来没观察到不等于不可能。这些在统计学、语言学、心理测量学里各有几十年到上百年的文献。你换了一个领域，换了一批名字。

这一节的对手不是任何一章的对手，是本书这个层级上的对手。本章列十族，每一族给三样东西：它已经说到哪一步；本书与它的分离线；以及这一族反过来削弱本书的一条让步。

第三样是关键。一条分界如果只削对方不削自己，那它多半不是分界，是辩护。

⚠ 全书须先认一件事：本章十族里有八族，是在十四章全部写完之后才被检索出来的。也就是说，十四章各自写作时并不知道它们。这件事本身是本书命题的一次发作——一个缺口没有在任何一章的记录上留下条目，不是因为它不在，是因为那十四次检索的半径一律只到“方法学加睡眠学”。凡是“甲与乙在记录上同形”这一族的构念，正确的做法是先去统计学与语言学查这个二分有没有被起过名。这条教训写在这里。

十八之二·第一族·结构零与抽样零（统计学，一九七五年起）

它说到哪一步。列联表里的零分两种：结构零是逻辑上不可能出现的格（一个三岁的孩子不可能在读中学），抽样零是本来可能出现、只是这次没抽到的格。这是分类数据分析的标准区分，见于该领域的基础著作。更要紧的是，二〇〇五年已有工作用似然比统计检验从大规模语料里把两者分拣出来，以便不给结构零分配概率质量。

这一族直指第四章。第四章提出“禁转”，说一条从未被观察到的转换至少有四种来源：物理上跳不过去、成本上不划算跳、规则上不许这样记、跳过发生了但被记录规则做成了看不见。前两种逐字就是结构零与抽样零；而第四章提出的“分拣判据”，在这一族里二十一年前就有了实现。

分离线（两条，都可判定）。

其一，统计学那一支分拣的是“不可能”与“没抽到”，二分；本书要分的是四分，多出来的两种——规则禁止、以及被观察系统涂抹——在那一支里没有位置。因为在那一支的设置里，观察系统是透明的：格子是什么就记什么。判定形式：拿两套都合规的判读手册去判同一份原始信号；若两套给出的零格完全相同，本书多出来的两种不成立，第四章应并入这一族。

其二，本书给了一个那一支没有的读数：残余率——被压低但不归零，而残余的高度由例外条款的宽窄决定。结构零是绝对的零，抽样零随样本量增大而消失；规则禁转两者都不是，它稳定地停在一个非零的小数上。判定形式：加大样本，抽样零消失、结构零不变、规则禁转的残余率不变——三者由此可分。

这一族反过来削本书的一条。那一支明确指出：结构零的判定通常要靠领域知识，而不是靠数据——数据只能给出候选。本书第四章那张分拣表的第三列“换判读规则”，恰恰在观察上通常做不到，因为一份数据只按一套规则判读过。于是本书必须承认：分拣判据在只有一套规则的记录上原理上执行不了，它是一个提案，不是一件已验证的工具。第四章已按此收窄。

十八之三·第二族·偶然空缺与系统空缺（语言学，一九三五年起）

它说到哪一步。一个符合全部语音规则、却碰巧不存在的词形（英语的 /blɪk/），叫偶然空缺；一个被规则挡住、因而不可能存在的词形（/pfnk/），叫系统空缺。这个二分是生成音系学的标准术语，围绕它有九十年的文献与整套实验方法（造词测试）。

它同样直指第四章，而且比统计学那一支更贴：语言学处理的正是“从未被观察到”这件事，而且它的判定手段是去问母语者一个从未出现过的形式合不合法——即用被观察系统之外的证据来分拣。

分离线。语言学那一支有一个本书没有的东西：母语者的直觉可以充当独立于语料的第二个证据源。睡眠判读没有对应物——没有人能被问“这个转换合不合法”。这是本书的劣势，不是优势，须如实写。本书能主张的只剩一句：在只有语料、没有第二证据源的领域里，偶然空缺与系统空缺原理上不可分；而睡眠判读正是这样的领域。

它反过来削本书的一条。语言学在九十年里发展出了造词测试这类办法，去逼近那个第二证据源。本书没有做类似的努力，甚至没有提出该往哪个方向找。第四章那条“去读规则本身怎么写的”是一个可行的替代，但它只能分开“规则禁止”与“物理不可能”，分不开“不划算”。

十八之四·第三族·可变单元问题（地理学与时空分析，一九三〇年代起）

它说到哪一步。把同一批点数据按不同的空间单元聚合，会得到不同的相关系数、不同的显著性、甚至相反的结论——这是可修改面积单元问题；把它搬到时间轴上就是可变时间单元问题，聚合、分段、边界三种改动都会改变被数出来的单位与结论。

它触及第一、二、八章。第一章那个“合并两个分期之后段数掉三成”是教科书级的分段效应；第二章的粗化检验直接就是这一支的方法。

分离线。这一支说的是“换一个单元，结论会变”，处方是做灵敏度分析、报告多种单元下的结果。本书说的是“在同一个单元下，某些东西根本没有位置”。两者正交：把灵敏度分析做到完美，未追认段仍然不构成一样东西，判余仍然在取最大值那一步被扔掉。判定形式：任取一份做过完整单元灵敏度分析的睡眠研究，看它报不报第二候选与分差。

它反过来削本书的一条。这一支已经证明：只要单元可变，任何基于单元计数的读数都不能被当作对象的性质。本书的八个可算读数里有五个是比例、其分母都是某种单元计数（段数、转换数）。因此本书的这些读数，同样只是在给定单元下才有意义，本书没有为它们做过完整的单元灵敏度分析。这是欠账。

十八之五·第四族·数学耦合与初值定律（一九三一年起）

它说到哪一步。一个差值天生就与它自己的某一端相关：取两串互相独立、标准差相同的随机数 x 与 y ， x 与 $x - y$ 的相关约为 0.707。因此“把差值与基线做相关”是一个已知的陷阱，有近百年的文献，在临床生理学里被反复重申。

它直指第七章。第七章报告的“下潜幅度的变异百分之六十六点七由白天那一端解释、百分之零点二由夜间那一端解释”，由三个矩完全决定，是算术不是发现——第七章自己在第二版里认了这一条。

分离线（这一条最干净）。那一支是给分析者的禁令：别把差值与它自己的一端做相关。本书是给报告者的要求：既然差值被印出来并据以行动，就必须把两端的方差与协方差一并印出来。完全遵守那条禁令，本书的问题原样存在；而公布端权对，不需要做任何一次耦合回归。

它反过来削本书的一条。既然两端份额由三个矩完全决定，那么端权对不含任何超出这三个矩的经验内容。第七章把它称为“读数”是过强的，它是一个换算。本书在第十五章已把它归入“是装置不是读数”那一类。

十八之六·第五族·概化理论与非遍历性（心理测量学，一九七二年起）

它说到哪一步。概化理论把一次观测的变异分解到若干个侧面上——人、场合、评分者——并回答“要让某个用途足够可靠，需要在哪个侧面上多取几次”。而非遍历性的论证更硬：只有在极严格的条件下，才能把个体间变异的结构推广到个体内变异的结构，而真实的心理与行为过程几乎从不满足这些条件。

它触及第七、十、十一章。第十一章的“量重复度与位重复度”实质上是一次两侧面（人 $\times$ 场合）的分解；第十章的“一次事件被读成一种能力”，正是非遍历性论证的一个实例。

分离线。概化理论问的是“能不能推”——某个读数在多少次观测之后足够可靠；本书问的是“说话的是哪一端”以及“这个量有没有位置”。一个读数可以完全可靠（重测相关很高）而仍然

只在说它的一端；一个量可以有很高的组内相关而它的位置仍然完全不可重复——第十一章那对数正是如此：深睡百分比跨夜相关 +0.823，而窄口位置 -0.079。同一批夜、同一批人、同一次分解，两个方向。

它反过来削本书的两条。其一，第十一章自己承认：“低相关等于测不准”与“低相关等于没位置”在那批数据上不可分——这正是概化理论会指出的问题，而本书没有解决它。其二，本书大量读数只在每人一到两夜上算过，按概化理论的标准，本书的多数读数从未做过可靠性估计。

十八之七·第六族·特质与状态（人格与临床心理学，一九六一年起）

它说到哪一步。把一个观测值分解为稳定的特质成分与随场合变化的状态成分，是六十多年的标准做法，并有成套量表。

它直指第十一章，第十一章已把它排在自己分界节的第一位，并公开承认初稿完全没想到它。

分离线。特质与状态分的是同一个量的两个成分，二者相加等于观察值；本书分的是同一现象的两个属性——大小与位置——它们不相加，且位置那一侧没有特质成分。判定形式：对任何一个特质—状态分解，可以问“这个人的特质分是多少”；而对本书的位置，这个问题没有答案，因为位置在人身上不存在一个固定值。

它反过来削本书的一条。特质—状态那一支有一整套估计方法（多时点模型、潜状态—特质模型），可以把两个成分分开估出来。本书的位重复度只是一个跨次相关，它比那套方法粗得多，而本书没有说明为什么不用更成熟的工具。

十八之八·第七族·细窗判读（睡眠医学自身，二〇二〇年代活跃）

它说到哪一步。三十秒这一格的来历是纸速——每秒十毫米、一页三十厘米，于是一页三十秒——这是一九三〇年代的记录条件留下的。近年已有多组工作按五秒的迷你窗判读，并报告在三十秒判读下看不见的微结构不稳定；也有把睡眠—清醒判读做成时间上连续的深度学习工作。

这一族同时触及第一、二、四、八章，是本书在自己领域内最近的一位。

分离线（两条）。

其一，这一族做的是把刻度做细；本书说的是在给定刻度下，某些东西没有位置。刻度做到五秒，第一章的未追认段会变少；但只要每一格仍然必须有恰好一个标签，判余仍然在取最大值那一步被扔掉，第三章的缺口一点不动。判定形式：去查那些五秒判读的研究，看它们报不报第二候选与分差。

其二，也是更要紧的一条：这一族恰好是本书第四个“别处做了，这里没做”的正面对照，而且是本领域内部的。有人在做细窗判读，且已经报告出三十秒下看不见的东西；而临床报告格式仍然是三十秒一格、一格一个标签。技术已有而格式未改，这是本书的靶子，不是本书的功劳。

它反过来削本书的两条。其一，第八章说“判定它所缺的两样都不是技术难题——一把细于三十秒的刻度与一次标准化扰动”，而那把刻度正在被造出来。这既削弱了本书“没人做”的说法，又加强了本书“做了也没人拿去裁”的说法。第八章已按此改写。其二，第二章那条“十五秒对三十秒判读窗”的纯算术，在五秒判读下就不再成立——它是一条关于当前判读惯例的算术，不是一条结构性删除。第二章已按此降级。

十八之九·第八族·值守睡眠（职业健康，一九八八年起）

它说到哪一步。一九八八年就有脑电研究报告值守当夜的睡眠受损；此后三十余年积累了系统综述，并有实验证明仅在睡前给出值守指令即可改变随后的睡眠；也有研究报告过半数值守者称即使当夜零呼叫睡眠仍受影响。

它直指第十二章，第十二章已把它排在自己分界节的第一位并公开认账。

分离线。那一支研究的外部安排写在排班表上、当事人明确知道；第十二章关心的那一类不在任何表上。且那一支的变量是二分的（值守／不值守），第十二章的是一个连续的托管率。

它反过来削本书的一条。那一支已有一项研究报告：在一百零五分贝闹钟、被试知道漏掉概率极低的条件下，值守夜与非值守夜的睡眠效率、总时长、各期时长无差异。这直接说明：无人接手时，仅把返还通道做可靠也可能足以让身体内转。第十二章据此把“三环闭合”由定义性主张降为经验问题。

十八之十·第九族·睡眠努力与反常意向（临床心理学，一九六〇年代起）

它说到哪一步。“越想睡越睡不着”是一个已命名、已量表化、已有疗法的临床构念：注意—意图—努力通路描述了从选择性注意到刻意入睡再到努力本身抬高唤醒的路径；有专门的睡眠努力量表；而反常意向疗法是一种被推荐的认知行为治疗成分。

这一族的位置很特别：它不是本书任何一章的对手，因为本书十四章一处也没有踩它。本书处理的全部是记录与归属，不处理意志与努力。

为什么要列它。因为它是本书唯一系统性回避的那一格，而回避的理由必须写出来：作者另有一篇面向普通读者的文章处理它，那篇文章讲的正是这一支的内容，且明确标明了出处，不主张原创。本书不重复它，也不与它竞争。

它反过来削本书的一条，而且很硬。第十一章建议读者做一次自我校准（连续若干天为自己的睡眠打分），第五章也有一条类似的动作。这一支会说：持续的自我监测本身就是派出一个守夜人。本书接受这条限制，并据此把所有自查动作限定为一次性的、做完就停的，写进了使用声明第四条。

十八之十一·第十族·通约化、残余范畴与社会授时

三支合为一族，因为它们共同处理“记录制度与它所记录的东西的关系”。

通约化（量化社会学，一九九八年）：把不同质的东西化为可比的量，是一种社会过程，也是一种权力形态；宣称某物不可通约是有代价的。它触及第一章那句“凡是一样东西，都可以按它占了多少来结算”。**分离线**：那一支处理的是通约的后果，本书处理的是通约之后账上还剩不剩第二本账。

残余范畴（分类研究）：落在分类之外、不留痕迹的那一类。它触及第一、三、四章。**分离线**：残余类可以通过增设类别来消化；而第三章的判余不能——增设类别只会制造新的判余，第三章的数据正是如此。

社会授时因子（情感障碍研究，一九八八年）：规律的社会作息——起床、吃饭、工作、与人接触的时点——稳住内部昼夜节律，扰乱它可以诱发情绪发作；配套有量化作息规律度的量表，并已长成一整套疗法。它直指第十三章，而第十三章成稿时零命中。**分离线**：那一支要求扰动是患者本人经历并可报出的生活事件，而且它的量表记录的是患者自己的作息事件，“是否有他人在场”只是那个事件的一个属性；第十三章要求的是对面那个人自己的日常安排变更，且患者未察觉。这条分离线反过来加固第十三章：那个“是否有他人在场”的字段，正是第十三章所说“问过第二个人的那一次，问的仍然是这个人”的又一个实例。

这一族反过来削本书的一条。社会授时那一支已经把“另一个人的作息影响你的节律”做成了一套疗法并有随机对照证据；第十三章的默借至今零采集，第一层主张是一个假设，不是一个发现。第十三章自己写了这一条。

十八之十二·十族之下有没有同一块地板

有一部分，而且必须把范围说准。

以人为报告单位的那些学科——临床评估、行为遗传学、社会支持研究、精神诊断——共同把一件事当作给定：一个原因要算得上原因，得能出现在某一方的自陈范围里。承认有一类原因两侧都自陈不出来，这些学科的采集工具就同时失去分母。

但这块地板不是普遍的，本书不打算把它说得比实情更大。环境流行病学从来不靠自陈：空气监测站、行政数据、气象记录都不问任何人。社会流行病学处理收入、住房、政策，也不靠当事人报

告。这块地板只在以人为报告单位的学科里成立——而睡眠评估的全套工具，恰好一件不落地属于这一类。

十八之十三·本章自己的三处限制

其一，十族里有八族是十四章写完之后才检索出来的。本书的检索半径长期只到“方法学加睡眠学”。这一条已在十八之一写明，此处重复一次，因为它是本书方法上最实的一处失败。

其二，本章的分离线全部是论证性的，没有一条被实际检验过。每一条都给了判定形式，但十条判定形式本书一条也没有执行。

其三，本章没有找过反例。第十六章那条三签名判据可以被两种方式证伪——找到满足三条而什么也没扔掉的装置，或找到扔掉了东西而不满足三条的位置——本书两种都没有系统地找过。

第十九章 对话

前一章处理的是对手。这一章处理的是三个不必被反驳、而应当被请到桌前的对话者。它们与本书的关系不是谁对谁错，而是：本书要说的事，它们各自已经走到了哪一步，为什么走到那里就停了。

十九之一·与判读手册对话

本书十四章里有九章的材料直接来自按判读手册标注的记录。手册是本书最大的债主，也是本书最主要的靶子。这一节把话当面说清。

手册做对的那件事，比本书批评的那些要大得多。一九六八年之前，不同实验室对同一段脑电给出的分期无法互相比较；手册使得全世界的睡眠记录第一次可以放在一起。可比性是一切后续研究的前提，而可比性正是靠“每一格必须给恰好一个标签”换来的。本书全书赖以立论的那批公开数据，本身就是这条规则的产物——没有这条规则，本书连数都数不成。

三十秒那一格的来历值得记住。它不是从生理学推出来的，是从纸速来的：每秒十毫米的走纸，一页三十厘米，于是一页三十秒。一个记录设备的物理尺寸，变成了一门学科的时间单位，并在纸张退场之后又存在了半个世纪。本书不认为这是错误；本书认为这是一件应当被记住的事——当一个单位的来历与它所测的对象无关时，凡是小于这个单位的东西，都需要被专门问一次它去了哪里。

本书对手册的全部要求只有两条，都是加法：其一，凡按某一版规则判读的记录，标上版本号（第三章）；其二，凡有规则明文规定某种情形下默认为某一期的，把该条款单列，以便使用者知道哪些转换的稀有来自生理、哪些来自条款（第四章）。

两条都不动分类，不改界线，不要求任何新设备。

十九之二·与概率化睡眠图对话

二〇一八年发表的一项工作，不给每个窗一个标签，而给每一期的概率——一整夜由此成为一条概率带而不是一串字母。它是第三章那条“保存第二候选与分差”的建议在技术上的实现，而且早了七年。

本书与它的关系不是竞争，是引用。第三章有一次预注册的检验触发了自己的撤销规则：同一份特征按两套体例分别训练，折叠之后硬标签不一致率只有百分之零点五九——软输出确实吸收了体例差异，第三章原先“两套手册之间”那条分离线因此整个撤回。撤回之后靶子挪了，而挪对了地方：判余不是在两套规则之间被扔掉的，是在取最大值写下标签那一步被扔掉的。

于是本书要向这一支问的，不是技术问题，是格式问题。概率化睡眠图存在了七年，而临床报告至今仍要求每个窗一个标签。能保住而没保住，就不是技术问题。

第三章顺带做了判余的第一次直接测量：判余低于零点二的窗占百分之九点零九，这些窗与人工标注的一致率是百分之四十一.三，而高判余窗是百分之八十八.六——出错概率高五点二倍，而报告对这两类窗的记录完全相同。

十九之三·与“别处已经做了”的三个领域对话

本书全书倚重三个正面对照，这一节把它们请到一起，因为它们合起来说的那句话，比任何一个单独说都清楚。

天文学。日出与日落不是一条线，是三段——民用、航海、天文曙暮光，各有明确的太阳高度角定义，各自进历书、各自有用途。天文学没有把边界当成边界，它给了边界三个名字和三组数字。

生态学。交错带一九〇五年就被命名。它有边缘物种——两侧都没有、只在交错带里出现的物种；它有锐边界与渐变带之分；它有专门测宽度的方法。生态学不但承认边界有宽度，还把“这条边界有多宽”做成了一个可以立项去测的量。

同一门医学的内部。药物滴定速率——升到目标剂量用多久——是一个必须写进方案、必须被伦理委员会看到、必须报告的量。而入睡的下降速率，可以从任何一份记录算出来，没有一份报告有这一

栏，也从来没有被裁决过。这是本书最喜欢的一个对照，因为它连领域都不用跨：同一门医学，两种做法。

三个对照合起来说的那句话是：把一条被画成分界的线，打开成一段有厚度、可命名、可清点、可立项研究的东西——这个动作已经被完成过三次，在三个互不相干的地方。它不需要新的哲学，不需要新的仪器，只需要有人问一次：这条线后面是不是有地方。

第二章那份清单最能说明问题。那是一份六项的核对表——这个东西有没有自己的名字？有没有自己的量纲？有没有人测过它的宽度？有没有专门的方法？有没有进过教科书？有没有独立的研究计划？拿它去核对生态交错带，六项全有；拿它去核对睡与醒之间那片区域，六项五无。

同一张清单，两个相反的结果。差别不在对象，在账本。

十九之四·一个不在桌前的对话者

最后要提一个本书请不到的对话者：睡不着的那个人。

本书十四章，没有一章看过一位失眠患者的记录。第九、第十、第十一、第十三这四章处理失眠，用的全部是公开的健康队列，和一批被库说明书载明为“轻度入睡困难、其余健康”的人。第十三章那句最不客气的话，是对本书自己说的：它的信息源结构，与它所批评的那些工具完全相同——那些工具只问患者，而本书只问文献。

这一条不能靠写作补上，只能靠采集补上。本书能做的，是把它写在能被看见的地方，而不是写在附注里。

全书结语 那一格里本来有什么

一·十四章合起来说了一句话

一个装置被要求给出恰好一个答案。它给出了，而给出的那一刻，别的候选被扔掉；扔掉这个动作不产生任何读数；于是记录看上去是完整的，而完整正是它读不出缺口的原因。

这句话在本书里发作了十四次，落在四个层上：每一格得有一个标签、每一次改善得有一个分数、每个人得归进一类、每笔账得记在一个人名下。

二·三种能被扔掉的东西

回头看这十四处，被扔掉的东西可以归成三类，而三类的处置完全不同。

第一类：本来就该被结出来的，只是没有那一栏。判读体例的版本号、每条建议的时点、试验基线期的第二夜、评估工具里问对面那个人的一条——这四样加起来不到十个字段，全部是加法，全部可撤。这一类是本书唯一有把握的处方。

第二类：只能被群体审计，不能被个人使用的。未追认段比例、成簇承载比、残余率、双向漏出率——这些量在一个人身上算出来没有意义，它们的用处是拿一批记录去问制度：“你这套规则在多大范围内制造了看不见的结构？”把它们写进个人报告，只会制造新的误读。第十五章那条引用纪律正是为此。

第三类，也是最要紧的一类：必须被允许继续扔掉的。不是每一样被扔掉的东西都该被捡回来。一个不允许留白的判读格式，换来的是全世界记录的可比性；一个综合分，换来的是评价的可操作性。这一类的判据是：当把一样东西结出来的动作，会改变或损害产生这样东西的那个动作时，它属于第三类。

第十一章给出了这一类最清楚的一个例子：本书建议读者做一次自我校准，而持续的自我监测本身会损害睡眠。于是那个校准只能是一次性的、做完就停的。一件必须靠不被持续测量才能保持完整的事，它的读数只能间断地取，且必须由承担那个动作的人自己取，并且随时可以撤回。

三·三个洞

其一，一位失眠患者的记录也没有看过。上一章已写。

其二，本书最想成立的那条主张，检验的风险度是零。第十七章已写。“十四个读数是十四样东西”这句话，本书实际检验过的部分只有八个，而那次检验不可能失败。

其三，说不出该由谁牵头。本书的处方要动的是判读手册、报告格式、量表条目、试验方案模板——这四样分属四个互不引用的共同体，而本书自己指出的第四层缺口正是“账按主体划”。一件跨在四个主体上的事，本书的处方落在了自己指认的那一格里。

四·最后一句

这本书从一个外行的问题开始：一件只持续十五秒的事，在一串三十秒的字母里是哪一个？

答案是不是哪一个。而十四章之后，这个答案有了一个更准确的形状：它不是没被记下来，它是被记成了别的东西，而记录不会因此报错。

这本书没有把它捡回来。这本书做的只是把那一格指出来，并且说明：那一格里本来有东西。它现在仍然在那里。

参考文献

本书按章列出文献。为免装书时制造虚假的统一，各章原有的著录方式保留原样：前四章与第十二章为线索体（在正文中给出出处），其余各章为著录体（章末列表）。这一差异是有意保留的。

跨章共用、且在多章承重的线索，另列如下，以便读者一次找齐：

- 睡眠判读手册（现行版与旧版）：第一、二、三、四、十七章共用。
- 公开判读记录数据集（老化队列七十八人一百五十三夜；药物队列二十二人四十四夜）：第一、二、三、四、六、八、九、十、十一、十四、十七章共用。本书十四章中有十一章用的是同一份记录，这是第十七章得以成立的前提。
- 周期性交替模式（睡眠微结构，四十年文献）：第一、二、十七章。
- 可变面积单元问题与可变时间单元问题：第一、二、八、十八章。
- 概率化睡眠图：第三、十八、十九章。
- 生态交错带（含边缘物种、锐边界与渐变带之分、移动分割窗）：第二、十八、十九章。
- 天文曙暮光分级：第二、十九章。
- 值守睡眠系列（职业健康，一九八八年起）：第十二、十八章。
- 数学耦合与初值定律：第七、十七、十八章。
- 注意—意图—努力通路 with 反常意向疗法：第十八章（本书正文不使用，仅作分界）。

 **本书未做的一件事：**全部承重文献未逐条回到原文核验。第十八章十族里有八族是十四章写完之后才检索到的。

附录一 十四项赌注、撤回条件与三条共同条款

每一章各下一注。本表把它们并列，并在末尾给出三条对全部十四项一体适用的“不算命中”条款。

表 A1-1·十四项赌注

章	押的是什么	使它为假的条件	处方的成本
一	判读记录里会出现“追认规则”这一栏	同批记录上周周期性交替模式的发生率与未追认段比例相关超过 0.6，则本构念被前者吸收	手册加一条条款说明，零新增测量
二	睡与醒之间那片区域会取得自己的字段	亚窗判读普及后，该区域的宽度收敛到零	另账登记，不进入现行份额加总
三	报告格式会保存第二候选与分差	判读者的实际犹豫与模型判余不相关	每窗多存一个数
四	判读手册的例外条款会被单列	提高编码器一致性后，规则型残余率显著下降	手册目录加一节，零新增测量
五	会有人测一次使用者的验收分辨率	三个改善量在使用者评分上可分辨	一次十四天的小样本校准
六	试验基线期会多记一夜	按样本层方法算出真实个体反应变异为零	每人多记一夜
七	差值读数会与它的两端一同公布	两端方差比普遍落在 0.8 至 1.25 之间	报告多两列，不需新数据
八	下降速率会被裁决一次	八十夜以上的样本仍全部落在零附近	一把已存在的细刻度+一次标准化扰动
九	会有同一批人同时具备细致主诉量表与多导记录	潜类别分析在同一批人上与主诉划分的调整兰德指数超过 0.5	一次并行采集
十	会有同一人多夜的记录用来判“一次”还是“一种”	前后两次中断长度的相关超过 0.3	每人多记两夜
十一	报告里会出现一个位置类字段	换一个更稳的位置定义后跨夜相关超过 +0.3	报告加一列
十二	会有指南把事件类别清单、替代持有人与睡眠记录绑定到同一区间	该读数与既有指标的相关始终超过 0.90	记录加三栏
十三	会有评估指南问患者以外的人“他自己的日常安排变更”	控制患者已报告的生活事件后，第三方报出的他人变更与失眠起始无关联	量表加一条条目
十四	每条睡眠建议会被标上时点	换一份带光照与活动的多日记录后，剖面变成单调远端更强	清单加一列

三条共同的“不算命中”条款

一、由本书作者或本书推动写入的，不算。一条预言若靠预言者自己去实现，它测的是影响力，不是判断力。

二、已经存在的条款不算。例如向同床者询问患者的鼾声与肢体动作，一九八九年就有；询问照护者对患者困难的观察，二〇一四年就有。问的是第二个人，说的仍然是患者本人——这一类一律不算命中。

三、只在研究工具里出现、未进入正式记录或正式判定的，不算。一个字段若不进入计分、不进入报告、不进入诊断，它的存在不改变本书所描述的那件事。

撤回由谁执行：以上十四项，凡被判为假者，由作者在本书的后续版次中删除该章的第一层主张，保留第二、三层，并在修订说明中写明判假的来源。

附录二 十四个读数的编码手册

本附录供他人直接照做。手册内不出现本书的章名与构念名，以便使用者不必先接受本书的论述即可执行。凡需要判断的地方，一律给出边界例。

通则六条

一、**分析窗必须声明**。写明起点（记录起点／熄灯时刻／首个睡眠窗）与终点（记录终点／熄灯后固定时长／最后一次入睡）。同一批数据换一种窗，多项读数的显著性会翻面。

二、**入睡后清醒的口径必须声明**：最后一次醒来之后的时间算不算。改这一处，超过四分之一的夜会换类别。

三、**判读体例版本必须声明**（旧体例分深睡两期／现行体例合并为一期）。

四、**凡报告比例，必须同时报告逐个体的中位数与汇总比例**。二者可以相差极大。

五、**凡报告比例或差值，必须同时报告它两端各自的方差与协方差**。

六、**凡按时间戳切窗，起止索引若不构成正区间，必须显式跳过**，不得依赖负索引的默认行为。

逐读数定义（八个可从判读标注直接算出的）

读数	分子	分母	缺失规则
最小段比例	长度为一窗的段数	总段数	无
成簇承载比	落在含两次以上转换的粗窗中的转换数	总转换数	转换数为零记缺失
体例折叠损失	旧体例转换数减现行体例转换数	旧体例转换数	旧体例转换数为零记缺失
残余率	深睡直接转入第一期的次数	该次数加深睡转入第二期的次数	分母为零记缺失
下降时长	入睡点到首个深睡窗的窗数	—	整夜无深睡记缺失
命中界线数	三条界线中命中的条数	—	无熄灯时刻则不可算
长断占比	一夜最长的一次中断时长	入睡后清醒总量	清醒总量为零记缺失
窄口位置	最长的连续“无深睡无快速眼动”区间中点距睡眠期起点	睡眠期总长	并列取最早

四个边界例（照抄可用）

- 一段被单个非五分期窗（体动／未判）打断的同分期序列：**按打断处断开**，不合并。断与不断会使总段数相差约百分之一。
- 睡眠期的界定：**首个第一期／第二期／深睡／快速眼动窗到末个**。清醒不参与界定。
- 剪裁：剪去首尾连续的清醒与非五分期窗后，**首尾各回留三十分钟**。只剪不留会使总窗数减少一成以上。
- 分期转换：**跨越非五分期窗处不计为一次转换**。计与不计会使转换总数相差约百分之零点六。

两个不进入本手册的读数，及其原因

- 验收分辨率**：需要同一个人对同一批夜的重测评分，至少七对。判读标注给不出。
- 他人自身行为条目占比**：分母是评估工具的计分条目数，不是睡眠记录的量。以一个字母条款为一条（诊断标准）、以一个官方编号的可评分应答格为一条（量表）；**信息源是他人、且所问内容是他人自身行为、且进入正式计分**，三条同时满足才计入分子。

附录三 四十二个学科入口、消融总表与叠计比

一·四十二个入口

表 A3-1·十四章各自的三门学问及其在该章的位置

章	立论位（给结构）	推翻位（给反例）	约束位（给让步）
一	化学	音乐学	社会学
二	市场营销学	有机化学	行星科学
三	法律学	房地产估价学	细胞学
四	经济学	影视学	心理学
五	工程学（测量系统分析）	信息学	戏剧学
六	交通学	逻辑学	气候学
七	年老学	结构学	舞蹈学
八	生理学	物理学	哲学
九	群论	量子力学	佛学
十	西医学	戏剧学	易经学
十一	认知学	化学工程	体育学
十二	社会学	音乐学	化学
十三	遗传学	园林学	道德经
十四	时间生物学	控制工程	照明工程

二·消融总表：删掉哪一家，哪些结论会变

表 A3-2

章	删掉推翻位会怎样	删掉约束位会怎样
一	三家共同踩着的那句话无法被推翻，全章不成立	失去“段落账与时长账”这对账，主张退回单账
二	界带无从提出	失去“不得主张多待在界带里更好”这条限制
三	“同时成立的多个正确答案”无来源	失去分化路径这条对照
四	四签名塌回三签名，第四种无从命名	顺序由计划维持这一层丢失
五	可分辨性无形式定义	失去“验收端分不开导致争论无法解决”这个两百五十年的实例
六	“进度不可加总”无来源，全章不成立	失去“信号可长期低于内部变率而仍真实”这条限制
七	边界条件可在一轮内改写这条丢失	失去“弧覆盖率不是越高越好”这条限制
八	路径速率决定终态这条丢失，全章不成立	失去“本质上的副产品”这个类别
九	“同一密度矩阵多种系综”这条丢失	失去“身份相待而有”这条呼应（可替换，见下）
十	判词结构无来源	救场惯例这条丢失（只引行业惯例，见下）
十一	位置由规定而非测量决定这条丢失	失去“位置随机不是好事”这条限制
十二	时间语法的命名丢失（不影响任何预测，见下）	失去麻醉那条硬限制
十三	借景五类丢失	失去“属性不为任何一物所持有”这条呼应（可替换，见下）
十四	“输出端不是合法干预点”丢失，全章不成立	失去“白天那端被建筑锁死”这条辩护

⚠ 表中标“可替换”或“不影响任何预测”的四处，本书如实认定为修辞承重，不作材料承重：第九章的佛学、第十二章的音乐学、第十三章的道德经、第十章的戏剧学（只引行业惯例，未引理论著作）。这四章因此实际是“两门半”承重，本书不主张三门等重。

三·叠计比

同一门学问在本书中出现多次，须公开计数。

学问（含分支合并计）	出现章	次数
化学族（化学／有机化学／化学工程）	一、二、十一、十二	4
工程族（测量系统／交通／结构／控制／照明）	五、六、七、十四	4（五个分支）
社会学	一、十二	2
音乐学	一、十二	2
戏剧学	五、十	2
心理学族（心理学／认知学）	四、十一	2
物理学族（物理学／量子力学）	八、九	2

本书的叠计比至少是四。

四·一组三学科被问了两遍

第一章与第十二章用的是同一组：**社会学 × 化学 × 音乐学**，而且落在同一道题（睡眠是什么）上。这是本书最该被质疑的一处编排，必须正面处理。

两章的用法确实零重叠：

- 第一章的音乐学取的是**形式功能可被后续事件回溯改判**；第十二章取的是**休止与节拍期待**。
- 第一章的化学取的是一组**突触蛋白的磷酸化随清醒累积**；第十二章取的是**腺苷、线粒体与清除机制之争**。
- 第一章的社会学取的是**作息制度是被产业时间表安排出来的**；第十二章取的是**社会豁免与值守责任**。

两章交出的量形状相同而内容不重叠：第一章给出的是“一段时间能不能成段”，第十二章给出的是“一次回应由谁承担”。第十七章在同一份记录上算过第一章的读数，而第十二章的读数在那份记录上根本算不出来——这两章在验证表上落在完全不同的类别里。

代价也要写出来。同一组学科被问两遍，意味着**这两章继承了同一批盲区**：三门学问都不处理“记录制度本身是谁定的”这个问题，因此两章都看不见判读手册与评估量表分属两个互不引用的共同体这件事——而那正是第十五章第四条互不蕴含的内容。这个盲区在两章里重复了一次。

五·三条本书无法自检的判断，各指名下一门学问

- “从来没见过转换有四种来源”这条分类是否穷尽——应交给**分类学与档案学**（分类系统如何处理落在分类之外的东西）。
- “扔掉的那一份不留痕迹”是否为记录这件事的普遍性质——应交给**会计学与审计学**（复式记账正是为了让某一类损失留下痕迹而发明的）。
- 本书的四层是否穷尽了“强制作答”发生的位置——应交给**测量理论与形式认识论**。

附录四 预注册文件、脚本与哈希

本书全部经验结论可独立复现。凡涉及统计的章，预注册文件在跑任何统计之前写定并封存，哈希列于下表；脚本随书提供。

表 A4-1·预注册与脚本

章	预注册	哈希 (SHA-256, 前十二位)	主要脚本	数据
一	有	—	段落清点与追认判定	老化队列前四十夜
二	有	—	转换间隔、粗化、两种零模型	老化队列前四十夜
三	有 (三次)	—	体例折叠计数、抹线补跑、谱特征与交叉验证	老化队列，另取六人原始信号
四	有	—	越格计数、残余率、配对检验	老化队列+药物队列
五	有	—	配对复算、地板效应分层、两窗对照、零分布对照	药物队列二十二人
六	有 (两次)	—	四条预测、噪声底对照、窗敏感性	药物队列+老化队列无干预两夜
七	有	8df05c6173d0	五条预注册、拆混淆、方差分解、四问补跑	二十四小时连续记录，十九人
八	有	—	四条预测、自助区间、粗化稳健性	老化队列前四十夜
九	有	—	三条界线命中、潜类别、风险度	老化队列+药物队列安慰剂夜
十	有	—	长断、零分布对照、清醒时长口径更正	老化队列一百五十三夜
十一	有	b671800c039b	五条预注册、跨夜检验、置换零分布	老化队列一百五十三夜
十二	有	附录列数据与预注册双哈希	公开歌单数据复算	十三万余首独立曲目
十三	有	—	八份工具七十八条计分条目清点，四种编码口径	公开量表与诊断标准
十四	有	—	预测力—滞后剖面、年龄分层、风险度置换	老化队列一百五十三夜
十五	有	df9f203446ff	run15.py	老化队列一百五十三夜，四种粗细
十七	有	5455ec98871d	edfio.py／ run17.py／ run17b.py	老化队列一百五十三夜

⚠ 前十四章的预注册哈希未在本书成书时统一收齐，仅第七、十一、十五、十七章留有完整哈希记录。这是本书的一处程序缺陷，如实列出。后续版次应补齐。

三条复现纪律

一、复现本书任何一个数之前，先按附录二通则一至三声明三项口径（分析窗、清醒时长、体例版本）。口径不同，数字不同，而两套都对。

二、本书报告的相关系数一律为斯皮尔曼相关，成对删除，各对的有效样本量随读数缺失而不同，已逐处标出。

三、复现出与本书不同的数字时，先查口径，再查实现，最后才怀疑数据。本书自己因口径与实现问题更正过至少四次，全部写在相应各章的正文而非附注。

后记

这本书写完之后，我数了一下它自己的账。

十四章里，有九章报告了自己至少一条预注册预测的失败，其中三章触发了自己写死的撤销规则，把一条已经写好的主张整个撤回：第二章撤回了成簇那条读数的全部证据力，第三章撤回了“两套手册之间”那条分离线，第十一章撤回了整个“闸期”的构念。第五章四条预测全线不成立，第八章四条预测全线不成立。

这些不是本书的瑕疵，是本书唯一可能有的力气来源。一本讲“失效不产生信号”的书，如果自己的失效不产生信号，它就什么都没说。

也要数另一半的账。

第一，本书最想成立的那条主张，检验的风险度是零。第十七章那条预测在设计上就不可能失败。我在写那份预注册时，专门把风险度这个装置写了进去，因为第十四章刚刚用它查出过同样的病；然后我仍然写出了一条不可能失败的预测。知道一个陷阱存在，与不掉进去，是两件事。

第二，本书的检索半径长期只到“方法学加睡眠学”。第十八章那十族里有八族，是十四章全部写完之后才查出来的——结构零与抽样零、偶然空缺与系统空缺、概化理论与非遍历性、社会授时因子。这些二分在统计学、语言学、心理测量学、情感障碍研究里各有几十年到九十年的文献，而十四次检索一次也没有碰到它们。一个缺口没有在任何一次检索的记录上留下条目，不是因为它不在，是因为检索规则没有为它写下条款。

第三，本书一位失眠患者的记录也没有看过。

第四，本书的写成本身依赖着一批算不出来的东西。用来复算的公开数据库、期刊的开放摘要页、量表的公开评分说明、把这些挂在网上二十年而无人道谢的那些维护者——他们不知道本书在依赖他们，本书在动笔时也没有把他们记为条件。取消其中任何一项，本书都写不出来，而参考文献表一行不会变。文献标记的是被引用的作品，不是使这次引用成为可能的那些安排。

关于人机分工，写在这里：文献检索、公开数据的复算、结构检查与语言编辑使用了人工智能工具；全部脚本经我复核并逐项与已发表数字对表（第十七章附有对表）。判断、取舍与全部结论由我负责，包括其中的错误。

最后要谢的，是那些把三十秒定下来的人。

一九六八年之前，两个实验室对同一段脑电给出的分期无法互相比。是那一代人用一本手册换来了全世界睡眠记录的可比性——而可比性正是靠“每一格必须给恰好一个标签”换来的。本书赖以立论的那批数据，本身就是这条规则的产物。没有这条规则，我连数都数不成。

指出代价不等于否定它。一个能在如此稀薄的记录条件下持续有效半个多世纪的体系，值得被认真地清点一次。

王德生

二〇二六年八月

一件只持续十五秒的事，在一串三十秒的字母里是哪一个？

答案是：不是哪一个。判读规则要求每一格必须落一个分期，占不满半格的事被并进相邻的那一期，在整夜九百六十个字母里不留一行——而记录不会因此报错。

本书十四章，是同一个形状的十四次发作。它们分别落在四层上：每一格得有一个标签、每一次改善得有一个分数、每个人得归进一类、每笔账得记在一个人名下。四层的共同机制只有一条——某个装置被要求给出恰好一个答案，而给出答案这一步扔掉了别的候选，且扔掉不留痕迹。

书中没有一处指出违规或疏漏。每一次损失都发生在规则被完整遵守的时候。十四章里有九章报告了自己预注册预测的失败，三章触发了自己写死的撤销规则，把已经写好的主张整个撤回。第十七章把十四个读数放到同一份公开记录上一次算齐，并查出本书最想成立的那条预测在设计上就不可能失败。

-
- 十四章跨学科论证，四十二个学科入口，全部消融检验列表公开
 - 全部经验材料取自公开数据集，预注册与脚本随书列出，可独立复现
 - 三百页，二十八万五千字，四百三十九张表

定价 US\$25.00

ISBN 979-8-90690-001-2

德麦国际出版社（新加坡）

