

相关不等于因果

第 4 讲 · 用“屏幕时间与睡眠”判断结论能说多强

贯穿问题

如果屏幕时间与睡眠质量一起变化，我们已经知道了什么？还缺什么，才可以谈“导致”？

学习目标

能区分关联与因果；识别反向因果和混杂；比较横断面、纵向与随机实验；改写过强标题。

本讲只记住一条总原则

先看研究怎样收集证据，再决定结论能说多强

同一组数据，可以被写成多强的句子？

先圈动词，再看证据是否撑得住

结论层级	常见动词	当前横断面案例能否直接支持
描述	有多少、分布怎样	可以，但只限当前样本与测量
关联	相关、一起变化、较高者通常较低	可以，需说明样本、方向和测量
预测	可帮助预测	需要额外验证预测表现
因果	导致、使得、造成	不能只凭本次横断面关联直接下结论

本讲的保守改写

不写“手机使用导致大学生失眠”。写：“在本次调查样本中，较长的睡前屏幕使用时长与较差的自评睡眠质量相关。”

横断面：像在同一时刻拍一张照片

能看见谁和谁一起出现，但往往看不出谁先发生



它能做什么

描述某个时间窗口的分布；发现变量之间是否有关联；为后续追踪或实验提出问题。

它不能自动做什么

不能凭一次同步测量确定时间先后；不能自动排除反向因果；不能自动排除混杂因素。

稳妥句式

“在【具体样本】中，【暴露】与【结果】相关。” 不要因为“不能直接证明因果”就反过来说“横断面研究毫无价值”。

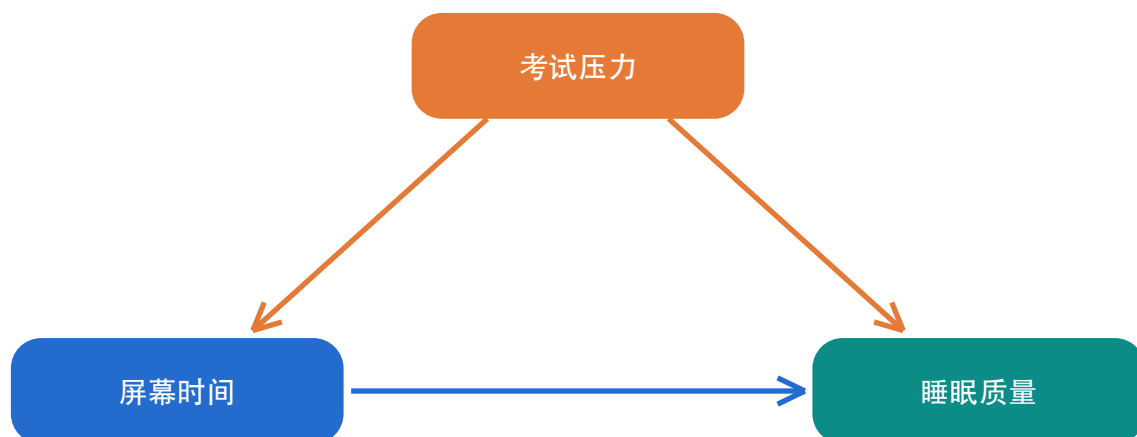
关联出现后，先检查两种常见替代解释

反向因果

不一定只是“刷手机 → 睡得差”；也可能“本来睡不着 → 刷手机更久”。

混杂因素

第三因素同时与暴露、结果有关，使两者看起来像存在更强的直接关系。



入门检查：第三因素是否与暴露有关？是否也与结果有关？

三所高校屏幕时间与睡眠质量调查

施工噪声遮住的样本量，应在本页核对

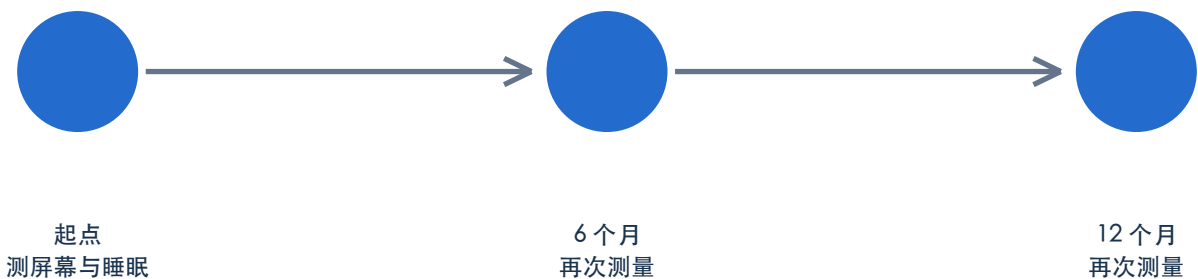
项目	案例记录	解释边界
样本	最终纳入 N=580，来自 3 所高校	不能自动代表所有大学生
暴露	手机系统记录过去 7 天睡前屏幕使用时长	相对客观记录，但仍有设备和定义边界
结果	自评睡眠质量量表；分数越高表示睡眠越好	不是临床失眠诊断
结果	$r=-0.28$ ， $P<0.001$	负号说明总体方向，不是“少睡 28 分钟”

可以写

“在本次 580 人样本中，屏幕时间较长者的自评睡眠质量得分总体较低 ($r=-0.28$) 。”

纵向：在多个时间点继续跟踪

更容易说明先后，但“更强”不等于“自动证明”



比横断面多解决什么

更有机会确认暴露在后续结果之前；能观察变化轨迹；可比较不同起点的人后来怎样。

仍要检查什么

未测量混杂；测量误差；时间变化的共同原因；中途退出者与留下者不同造成的失访偏差。

推荐表述：纵向研究增强了时间顺序证据，但仍不能仅凭“跟踪过”就宣布因果成立。

随机抽样 ≠ 随机分组

两个词都含“随机”，解决的问题不同

随机抽样

从哪里选人？
从总体中按随机机制抽取样本。

主要帮助：样本更可能代表总体，关系到结论能推广多远。

随机分组

入组后怎样分条件？
把参与者随机分到实验组和对照组。

主要帮助：各组起点更可比，增强因果判断。

随机实验也有边界

还要看参与者是否真的按要求执行、测量是否可靠、样本是否适合推广。不能为了实验强迫参与者长期严重缺觉；伦理可行性决定了有些问题不能直接做随机实验。

显著、调整过，都不是“因果通行证”

P<0.05 能告诉什么

在特定统计假设下，当前或更极端数据有多意外。它不能替研究设计补上时间顺序，也不能自动排除混杂。

“已调整” 能告诉什么

模型处理了被测量、被正确放入模型的因素。未测量、测量很差或模型不合适的因素仍可能留下影响。

措辞对照

过强：“回归已经消除了全部混杂。”
稳妥：“分析调整了考试压力、咖啡摄入和课程负担等已测量因素，但仍可能存在残余混杂。”

统计显著回答“数据有多反常”；因果判断还要回答“证据从哪里来”。

老师明确会考：四步检查一条“研究发现”

1

研究设计是什么？

横断面、纵向、实验，还是其他设计？

2

实际测到了什么？

手机系统记录、自评量表，还是临床诊断？

3

作者声称了什么？

描述、关联、预测，还是“导致”？

4

声称超过证据了吗？

时间顺序、反向因果、混杂和推广边界是否处理？

一句话口诀

设计 → 测量 → 声称 → 是否越界。先完成这四问，再决定标题要保留还是降级。

把过强结论改成与证据相配的句子

A 一次问卷发现咖啡摄入与焦虑得分相关

原标题：咖啡导致大学生焦虑

改写：_____

B 跟踪一年发现起点压力与后续睡眠有关

原标题：压力已经被证明会造成失眠

改写：_____

C 随机分组减少睡前通知，依从性只有 54%

原标题：关闭通知对所有学生都有效

改写：_____

每次改写至少保留：样本 / 测量 / 关联或比较 / 关键限制。

一页作业：简化因果图 + 约 300 字说明

上半页：把三个节点和可能的箭头补完整

屏幕时间

考试压力

睡眠质量

下半页：约 300 字，回答下面三个问题

1. 当前研究设计和实际测量是什么？
2. 为什么时间顺序、反向因果和混杂仍需处理？
3. 怎样把“导致失眠”改写成更稳妥的结论？

最终截止：2026-08-16 22:00 | 1 页 PDF | 不要求 1000 字或额外 3 篇文献

合上讲义，先口答这 6 题

1. 为什么横断面关联不能直接推出因果？

2. 反向因果在屏幕与睡眠案例中是什么？

3. 考试压力为什么可能是混杂因素？

4. 纵向研究比横断面多解决了什么，又没有自动解决什么？

5. 随机抽样和随机分组分别主要解决什么？

下次课前

阅读第 8—10 页；带上本页口答中最不确定的一题。下次课：2026-08-20 14:00—15:30，博学楼 C103。