

因果·影响·相关与预测辨析*

温忠麟 马 鹏 孟 进 王一帆

(华南师范大学心理应用研究中心/心理学院, 广州 510631)

摘 要 针对温忠麟等人(2024b)提出的影响关系, 葛泉语(2025)撰文提出下面质疑: (1)“影响关系”缺乏明确定义; (2)影响关系与因果关系区分不开; (3)不能因为因果目标无法达到便创设新目标; (4)所谓的影响关系应当称为“预测”。本文对此做出回应: (1)影响关系定义严密, 用的是“属加种差”的定义方法; (2)因果关系一定是影响关系, 但影响关系不一定是因果关系; (3)建立影响关系可以是非实验研究的一种目标, 优于仅仅建立相关关系; (4)预测是对变量关系的应用, 但不是关系本身, 与影响关系不相称。

关键词 影响关系, 相关关系, 因果关系, 影响因素, 预测

分类号 B841

温忠麟等人(2024b)在《心理学报》发表文章(以下简称温文), 主要内容如下: (1)说明将变量之间的影响 (influence) 理解为因果 (causal) 或相关 (correlation) 有什么不妥, 进而给出“影响关系”和“影响因素”的定义; (2)为了提出影响关系假设, 有哪些途径可以为变量之间的方向性提供理据; (3)当影响因素不止一个时, 有哪些共同作用类型。该文发表后, 引起学界关注: 人大复印资料全文转载, 科学网做了报道, 百度百科建立了“影响关系”词条, 多个学术公众号进行介绍甚至全文转载。

葛泉语(2025)撰文(以下简称葛文), 对温文提出商榷。葛文篇幅很长、内容很多, 但真正对温文形成质疑的可以概括为以下几点: (1)温文只提推论准则, “影响关系”缺失明确定义。(2)影响关系与因果关系区分不开, “影响”一词无论汉英、无论大众或学界, 均视为因果语言。(3)由于在手段上无法为因果目标提供因果证据, 便创设第三种目标, 是奇怪的。(4)有方向性的相关关系不宜称为影响关系, 这种关系一直有明确且公认的叫法——“预测”(predict)。

然而, 葛文不少论点似是而非, 多处表述前后

矛盾, 上述四点质疑没有一点是符合逻辑的。下面就葛文对温文的质疑做出回应, 但忽略那些与温文没有多少关系的问题。

1 “影响关系”的定义符合逻辑

葛文认为温文“最显而易见的疑点是, ‘影响关系’缺失明确定义”, 并且一再强调这个论点。事实究竟如何呢?

1.1 “属加种差”法定义的“影响关系”

温文在摘要中非常明确“将影响关系定义为有方向性的相关关系”, 正文中则将影响关系定义为“有方向性的共变”。显然, 温文是将共变和相关当作同义词使用, 可以相互替换。这里有两个问题需要讨论清楚, 一个是这样的定义算不算明确的定义? 另一个问题是, 相关和共变是否可以当作同义词使用?

给一个概念下定义的方法有多种, 其中“属加种差”是最常用的、逻辑严密的定义方法。做法是, 先将一个概念归入更大的类别(属), 再指出其与同类其他概念相比的独特特征(种差), 就可以明确定义该概念(Copi et al., 2019; Smith, 1995)。说得具体

收稿日期: 2025-03-18

* 国家自然科学基金项目(32171091)和广东省脑认知与人的素质发展基础学科研究中心(2024B0303390003)资助。

并列第一作者: 马鹏

通信作者: 马鹏, E-mail: mapeng@m.scnu.edu.cn 温忠麟, E-mail: wenzl@scnu.edu.cn

点, 假设已经有概念 A, 要定义的概念 B 是概念 A 的真子集, 那么概念 A 是概念 B 的属概念, 概念 B 是概念 A 的种概念, 种差代表了种概念 B 具有而 A-B (除了 B 以外的 A) 不具有的本质属性。例如, 矩形定义为有一个角是直角的平行四边形。其中, “平行四边形”是“矩形”的属概念, “有一个角是直角”是种差, 相当于在平行四边形基础上加了一个条件。

温文将影响关系定义为有方向性的相关关系, 这里“相关关系”是“影响关系”的属概念, 而种差是“有方向性”。这是非常严格而明确的定义了, 为什么葛文还说没有明确定义呢?

原来, 葛文将影响关系类比到因果关系: “为什么推论准则不可视同定义? 一个很现成的案例是因果关系。”葛文将影响关系的两个条件“相关+方向性”类比到因果关系的三个条件“相关+方向性+排他解释”, 但排他解释不属于因果的本质属性, 而是一个“兜底条款”: “排除了一切使其不能成为因果关系之解释的关系。这种尴尬的循环, 提示我们不可将推论准则视同定义”。没错, 因果关系三个推论准则不能作为因果关系的定义, 问题出在第三个准则。但影响关系的定义中没有这个导致循环的“兜底条款”, 因而葛文的类比是错误的。就是说, 虽然因果关系的定义不明确, 但是影响关系的定义是明确的。

从上面分析可知, 葛文误判的原因有两个: 一是没能识别温文对影响关系的定义完全符合“属加种差”的定义方法; 二是错误类比, 类比推理不像演绎推理, 类比的结果不一定正确。虽然将因果关系定义为“相关+方向性+排他解释”是有问题的, 但将影响关系定义为“相关+方向性”却没有问题。

葛文甚至还嫌温文没有给因果关系、相关关系下定义。温文新定义的是影响关系, 而因果关系、相关关系都是学界熟知的概念, 文献上已经有定义。首先, 影响关系的定义不需要用到因果关系的概念, 温文前提因果关系的用意是让读者容易理解影响关系的定位, 至于因果关系三个准则是否适合作为因果关系的定义, 或者世界上有多少种因果关系的定义, 与影响关系的定义没有关系。就是说, 温文本来就不需要给因果关系下定义, 只要说明因果关系是影响关系的真子集, 便可理解两者之间关系。其次, 温文是将相关关系当作已有定义的属概念来用的, 就像定义矩形的时候将平行四边形当作

已有定义的属概念一样。因此, 温文也不需要给相关关系下定义。尽管如此, 本文还是想借此机会讨论一下相关关系, 顺便说明温文将相关与共变当作同义词使用的合理性。

1.2 “影响关系”的属概念“相关关系”

“相关关系是两个变量之间存在相应变化的关系。一个变量变化时, 另一个变量也共变”(辛自强, 2017)。**相关和共变有一个共同的对立概念——“独立”**。变量之间不独立意味着相关或者共变。因为**可以将变量之间的相关定义为共变**, 所以温文将“相关”作为“共变”的同义词使用, 即把两个变量的影响关系定义为两个变量有方向性的相关或者有方向性的共变。

检验两个变量是否共变, 是统计分析可以做的事情。**单纯统计分析也只能证明共变, 无法证明因果或者影响**。如果两个变量都是连续变量, 通常计算皮尔逊相关系数并做统计检验。但更容易理解共变的统计方法是回归分析, 将其中任意一个变量作为自变量, 另一个变量作为因变量。如果自变量的系数显著不是 0 (回归显著), 则自变量的变化会引起因变量的变化, 即共变。如果两个变量有一个是连续变量, 另一个是类别变量, 可以使用方差分析做 F 检验(只有两个类别时也可以做 t 检验)。如果显著, 则各类别的均值不全相等, 即共变。如果两个变量都是类别变量, 可以做列联表分析。如果卡方检验结果是拒绝“两个类别变量独立”的假设, 说明一个变量的变化会引起另一个变量的变化, 即共变。

值得注意的是, 这里讨论的变量是随机变量, **相关关系是变量之间非确定性的关系**, 是一种集体规律, 不同于数学上的函数关系。这样, 两个变量之间不仅要考虑有无相关的问题, 还要考虑相关的程度。这方面涉及统计检验和效应量等内容(温忠麟 等, 2022), 因为与本文议题关系不大, 不拟展开。

相关既可以指线性相关, 也可以指非线性相关(戴步云, 简小珠, 2022; 甘怡群 等, 2019; Gravetter et al., 2021; 温忠麟, 2016; 辛自强, 2017; 张厚粲, 徐建平, 2020; 张敏强, 2010)。例如, 皮尔逊相关系数衡量的是线性相关, 而基于列联表分析的多个相关系数(如 Cramer 的 V 系数)、基于同序对和异序对的相关系数(如 Gamma 系数)、基于方差分析的相关比都是衡量非线性相关。但如果两个变量都是连续变量, 通常所说的相关指的是线性相关, 用皮尔

逊相关系数反映两个变量直线关系的程度。根据上下文不难区分所说的相关指的是线性相关还是一般的相关。两个变量不独立则相关,但不一定线性相关。然而,对于正态分布,“独立”与“皮尔逊相关系数为 0”是等价的,不独立一定是线性相关(温忠麟, 2016)。

国外有的文献将相关视为线性相关,而将包括非线性相关的一般情形称为关联 (association, Altman & Krzywinski, 2015)。关联是更一般的关系,一个变量提供另一个变量的信息,这种信息可以是线性、非线性,也可以不用量化变量间的关系,通过图形方式说明这种关系。

2 影响关系不等同于因果关系

葛文认为“影响关系等同于因果关系”,从下面两句话可以看出来。第一句是“从概念本身来看,该文(指温文)提出的推论准则并不能有效区分影响关系与因果关系”;第二句是“从概念向读者传递的意涵来看,‘影响’一词无论汉英、无论大众或学界,均视为因果语言”。第一句的意思是,温文定义的“影响”,与“因果”区别不开。第二句的意思是,说到“影响”就会被视为“因果”。为了拆解这两种说法,本节先简介文献上对因果关系是如何定义和理解的,然后给出影响关系不等同于因果关系的证据,最后举出属于影响关系但不属于因果关系的例子。

2.1 因果关系的不同界定

葛文说,因果关系与影响关系之界限暧昧。注意到因果关系是影响关系的真子集,对于某个具体的例子,如果难以区分是因果还是(因果以外的)影响,那不是影响关系的定义有问题。这就好比,如果区分不了一个生物是人还是(人以外的)动物,那不是动物的定义有问题。

定义因果关系不是一件简单的事, Hernán 和 Robins (2020) 讨论因果推断的专著《Causal Inference: What If》也未明确定义什么是因果关系。事实上,因果关系的定义因学科背景和理论框架的不同而呈现多样性:哲学探讨本质与逻辑,科学强调可验证性,法律侧重责任归属。

就心理学而言,学界比较认同的是 Rubin (1974, 2005) 在随机对照实验(后文简称实验)背景下的因果定义。葛文引用的新近文献也是类似定义:因果效应是指在干预 X 的不同条件下,变量 Y 的假设性差异(Bailey et al., 2024)。这个定

义相应的因果研究法是:在学科理论与逻辑基础上提出研究假设,在控制其他条件下对 X 做出干预,通过统计方法分析 Y 与 X 共变,从而做出因果推论。

在社科领域,研究者经常提到的是因果(推论)三准则:(1)因和果共同变化(共变), (2)因在果之前发生(方向性), (3)排除因果联系外的其他解释(排除他解释) (Morling, 2018; Shadish, et al., 2002; 辛自强, 2017)。使用统计分析的时候,上面第三个准则简化为:两个变量之间观察到的因果关系不能够被任何第三个变量解释,即两个变量之间不能够是虚假(spurious)效应(侯杰泰 等, 2021; 李文钊, 2018)。

对于第三个准则,葛文提出下面问题:“需要排除什么‘其它解释’才能推论因果呢? X 和 Y 之间的关系存在多种解释可能性,其中一种符合因果关系之定义、其它均不符合,因此,凡是使得 X 和 Y 之关系不符合因果关系定义的解释,均须排除。”

其实,对于实验研究,通过被试随机分组或者匹配,使得实验组和对照组可以看成两个“等组”,加上无关变量的控制,使得实验组与对照组唯一的差异就是实验者的干预,结果变量的差异就是干预导致的,因而不存在其他解释。所以,实验研究的结果是符合因果三准则的。至于说随机分组产生的抽样误差、实验引起的副作用、变量控制不够严密等问题,是实验信度或效度的问题(温忠麟, 1998),与非实验研究的问题不是同一层级。

对于非实验研究,如果想要证实因果关系才有葛文所说的问题,即不知道排除什么才能算是因果关系成立;但如果想要证伪因果关系,那就容易多了,只要找到因果关系外的一个其他解释,因果关系就不成立了。可以借用 Popper (1959) 的“科学理论不能被证实、但可以被证伪”的观点来说:在实验之外,因果关系不能被证实,只能被证伪。

除了上面讨论的因果三准则和 Rubin (2005) 定义外,葛文提到了其他因果定义或标准,说明因果关系有多个不同的定义或者标准,不妨记为因果 A、因果 B、因果 C……。有鉴于此,可以将温文的图 1 修改为本文的图 1,以反映因果、影响和相关的关系。为简便起见,图 1 仅用三个因果进行示例。

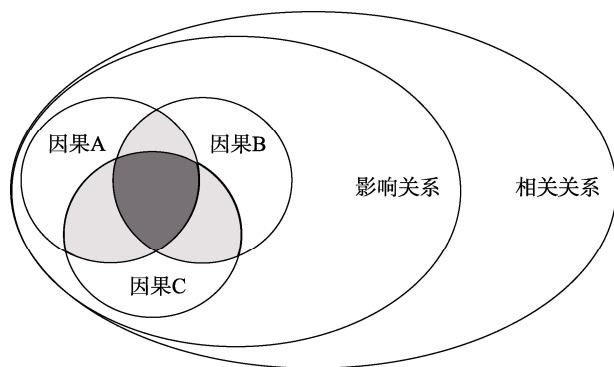


图1 因果、影响和相关三种关系示意图

2.2 影响关系不等于因果关系的证据

葛文有一句话比较让人费解：“无论大众或学界难以被温文作者说服相信不传达因果意涵。”在温文的概念框架之下，所有因果都是影响，即**有些影响是因果，所以在某些场合说影响其实说的就是因果**。例如，实验研究的结果，提到影响就是因果。温文从来就没有说过影响不传达因果意涵。温文的真实意思应当是这样的：影响关系有可能是因果关系，但不一定都是因果关系。这就好比说到动物的时候，有可能是人，但不一定是人；如果要明确指人而不包括别的动物，那就不要说动物，而是直接说人。类似地，像“作用”、“效应”、“导致”、“制约”等都可以传达因果意涵，但不能肯定是因果关系。葛文引用了《现代汉语词典》对动词“影响”的释义(含有“起作用”字眼)，除非白纸黑字说影响和因果是同义词或者两者有相同的定义，否则引用再多文献都不足以证明影响关系就是因果关系。

葛文认为的“影响关系等同于因果关系”这个命题，是一个**全称命题**，相当于说，所有影响关系都是因果关系。要证明这个命题成立，相当于要证明因果第三个准则是多余的，完全被前两个准则所覆盖。葛文以为只要证明因果第三个准则与前两个准则不独立就能证明其命题了，但不独立并不等于会被完全覆盖。虽然葛文试图举例说明其命题是对的，但全称命题是无法通过举例来证明的，再多的例子都徒然。葛文没有也永远不可能证明这个命题。

而要证伪“影响关系等同于因果关系”，就容易得多。可以用反证法从理论上证明，也可以通过一个反例否定其“全称”。下面两个证明都可以证伪“影响关系等同于因果关系”，后文将给出反例。

证明 1：如果因果关系与温文定义的影响关系没有区别，那么因果关系也可以使用温文的影响关

系定义。这样，因果关系就可以定义为“有方向性的相关”，行不通了吧。

证明 2：正如葛文所说，因果的定义或者标准不止一种，目前没有见到有一种因果可以包含所有的因果，说明任何一种因果之外，还存在别的因果，即存在别的影响关系(因为任何因果关系都是影响关系)。所以，任何一种因果关系，都只是影响关系的真子集，即影响关系不等同于任何一种因果关系。

顺便说一下，当今时代，有疑虑还可以问问“万能”的人工智能工具。虽然查询结果不能作为权威证据使用，但在人类“深耕”且不受资本利益影响的领域，查询结果还是值得参考的。向 DeepSeek 提问“影响关系和因果关系有何区别”或者向 ChatGPT 提问“What is the difference between causality and influence”得到的答复大同小异，下面是 DeepSeek 回答的主要内容(想了解细节的读者可以自己去查询一下)：这两个术语在实践中经常被混淆使用，但两者存在重要区别。因果关系是 X 直接导致 Y ，这种关系强调：(a) X 在时间上先于 Y ，(b) X 和 Y 以系统的方式一起变化，以及(c)其他解释已被排除或最小化。因果关系的确立需要更强的证据，通常需要对照实验、严格的统计方法或明确的逻辑推理。而影响关系是 X 会改变 Y 的发生方式或行为表现， X 可能仅是塑造(促成) Y 的因素之一。影响关系可以是一种间接的效果，有相关性以及合理性，但不一定构成因果关系。

2.3 影响关系不等于因果关系的案例

葛文多次表示想不通为何两个变量是影响关系却不是因果关系。“既然一个事件都已经对另一个事件起了作用，所谓‘起作用’必然是第二个事件发生了某些变化，那么请问，这种关系和因果关系有何不同？……就我个人而言，不易想出明明 X “影响”着 Y 、但 X 无论怎么改变都不会使 Y 产生丝毫差异的情形。”葛文先入为主，看到有“起作用”的字眼就认为是因果了。最后一句话葛文表达的意思是“如果不是因果， X 无论怎么改变都不会使 Y 产生丝毫差异”。这种推理错误太明显了，因为两个相关的变量就会共变。下面的例子，满足了影响关系定义，却不是因果关系。

大多数地方的公务员和事业单位员工的工资包括职务工资、工龄工资和绩效工资，所以职务、工龄和绩效都是工资的影响因素。按温文定义可以得出“工龄影响工资”的结果，方向性和相关性都满足。大众的日常用语也会说“工龄影响工资”。但工

龄与工资不是因果关系,因为可以给出别的解释:**薪酬政策才是原因**。在 1978 年改革开放之前的很长一段时间,我国公务员和事业单位的工资只与级别和地区有关,与工龄无关,所以说工龄与工资不是因果关系。

影响关系无需第三个(因果)准则,因而谁都没有机会来抬杠说:如果薪酬政策改变,工龄也不一定影响工资。影响关系是否成立只根据现实的情况,不需要排除其他解释,因而也就不需要考虑政策的事情。而因果关系有第三个准则,只要给出一个别的解释,因果关系就被否定了。

再看一个类似的案例。公共汽车公司通常都有针对未成年人的优惠规定,这里看一个简化版的规定:(1)身高不超过 1.2 米儿童免费乘车(不单独占座);(2)身高 1.2~1.5 米的儿童半价优待(提供座位)。根据温文的定义,可以说身高影响票价(免费、半费和全费)。大众的日常用语也会说“身高影响票价”。但身高与票价之间不是因果关系,公交政策规定才是原因。如果没有了针对未成年人的优惠规定,身高与票价的相关为零;如果改为按年龄优惠,票价和身高还是相关,但同样年龄的儿童(即控制年龄),不会因为身高而改变票价。这时,身高是票价的(代理)影响因素,但与票价没有因果关系。

说到身高的影响,葛文也有一个例子,以为研究者根据温文可以得出“身高影响词汇量”的结果:“默许研究者将这种结果描述为‘身高影响一年后的词汇量’,令人不安。”举出这样的例子,暴露了其作者没有读懂温文。下面就这个话题,看看一个正常的研究者根据温文会得到什么样的结果。

如果一位研究者提出假设:身高影响词汇量,然后根据方向性和相关性,确实可以验证其假设。但一个正常的研究者,在提出“身高影响词汇量”这个假设之前,更早就会**根据常识和逻辑**提出“年龄影响词汇量”的假设。这就有两种可能性,一种是研究者认为有了年龄不用考虑身高了;另一种是研究者认为可以假设两者都是词汇量的影响因素。如果是前者,不会得到葛文说的结果。现在来看后者,根据温文“多重影响因素的共同作用类型”一节的方法,研究者将同时考虑年龄和身高对词汇量的影响(即做词汇量对年龄和身高的回归),身高的回归系数将不再显著(即同龄人中词汇量与身高相关不显著)。根据温文的分类和表述,研究结果是“身高是年龄影响词汇量的**代理影响因素**”。所以说,根据温文不会得到“年龄影响词汇量”这么粗浅的

研究结果。

如何理解代理影响因素呢?回到上面说的“身高影响票价”的例子,其实优惠的初衷是针对年龄:不超过 6 周岁免费,超过 14 周岁全价。但很多地方乘坐公共汽车不像乘坐飞机和火车那样需要查验身份证,因而方便查验的“身高”就成了“年龄”的代理,优惠政策直接按身高去规定。6 周岁儿童的正常身高在 1.101~1.244 米范围,因而 1.2 米就近似代表了 6 周岁。

再看一个例子,香港的公众泳池规定,超过 8 周岁或者超过 1.35 米的儿童,不准进入异性更衣室(简称为“进入许可”),无论按温文定义或者大众理解,都可以说年龄和身高影响进入许可,当然这种影响关系也不是因果关系。有的地方只按年龄规定,有的地方只按身高规定,可见泳池规定才是原因。这个例子中,进入许可是根据性意识觉醒水平制定的。年龄和发育水平都会影响性意识觉醒水平。控制了年龄,发育水平还是与性意识觉醒水平相关(即同龄人的性意识觉醒水平与发育水平相关)。发育水平不易直接观测,而身高是发育水平的一个外显指标。这样,年龄和身高都是进入许可的影响因素,此时身高不是年龄的代理,而是发育水平的指标。

许多基于问卷的研究中的影响关系,却不像上面的案例那样可以明确不是因果关系,但又无法设计随机对照实验进行研究,因而使用“影响”来描述结果比使用“因果”更加恰当。

回到葛文的例子,得到“身高影响词汇量”这样的研究结果固然令人不安,但如果将“影响”研究换成“相关”研究,得到“词汇量与身高相关”这样的研究结果就能心安了吗?如果还心安,那么“词汇量与拇指长度相关”、“词汇量与饭量相关”的结果呢?可见,葛文试图用那个例子来否定温文定义的影响关系是经不起推敲的。

无独有偶,葛文误读温文可不止一处。如果说上面的误读还比较隐秘,那么下面这一段误读却是一目了然:“他们举例,研究者在横向研究中让一些学生报告一下自己的情绪和学习成绩(GPA),只要回归系数显著,他们主张研究者无须排除任何替代解释就可以将这个结果描述为‘情绪影响 GPA’”。这不是温文的例子,而是引用教育心理学顶刊 *Journal of Educational Psychology* 上的一篇文章,该文是横向问卷调查类研究,使用了 impact、influence 和 affect,但没有说是 causal。温文有一节专门讨论如何为影响关系的方向性假设找到理据,

这是统计以外需要做的事情,哪里说过回归系数显著就是“影响”了?

葛文还引用了 Haber 等人(2022)的文章,其中让被试对 affect、impact、influence 的因果意涵打分(无、弱、中、强)。这里问卷的设计非常关键。因为所有因果都是影响,所以有的影响就是因果。如果问“影响是否含传达因果意涵”,无论结果是什么,都证明不了影响是否与因果等同。正确做法是让被试判断“如果 A 影响 B, 则 A 与 B 有因果关系”,选择“正确”选项的被试,才是认同“影响等同于因果”。我们通过问卷星收到了 1998 份大学生或以上学历的被试对此判断的回答,只有 11.3% 的被试认同“影响等同于因果”。

3 目标往往被手段所决定

葛文认为,温文“创设一个不同于因果关系和相关关系的第三种目标,是奇怪的,因为这类研究本质上仍抱持着因果性目标。”还说温文混淆目标与实现目标的手段,这是所有疑点的本质根源。这得问问那些非实验的研究人员,他们的目标是不是冲着因果结论去的,至少温文作者没有葛文说的这个意思。

正如温文指出的,主流的心理学研究方法教科书,都认为只有实验研究才可能得出因果关系。上面一节的分析也说明,按照因果三准则的话,实验之外的因果难以令人信服。因此就有人提出条件较低的冠名因果,其中格兰杰因果(Granger causality)是一种基于“预测”的因果关系(Granger, 1969)。

葛文的话用来批评包括格兰杰在内的冠名因果的作者比批评温文要合适得多,因为那些作者的目标都是因果,手段上无法提供因果证据,创设了一个较低的第三目标(如格兰杰因果)。既然在某些领域严苛的因果标准达不到,降低标准定义一个新的因果没有什么好奇怪的。格兰杰因果在经济学上已经被普遍接受。而且,今天用来检验格兰杰因果的方法也与当初格兰杰提出的定义相去甚远。这个例子说明,科学研究不会死抱一个目标不变。

受过心理学主流课程训练的研究者,基于问卷做研究的时候,没有多少人会以因果为目标的。不少文章还会在讨论部分明确说明:本研究使用问卷数据,无法得到因果结论。多数时候,得到的变量之间关系都被视为相关关系。但对有方向性的相关,部分人会使用“影响”的说法,并不认为影响就是因果,研究者方面如此,审稿人方面也如此。如果研

究者运气好,碰到这样的审稿人(以及主编),非实验研究论文写着影响关系也能被接受。

例如,即使 *Social Indicators Research* 主编 Bartram 等人(2024)明确将“influence”列入“只能传达因果意涵的语言”(葛文有引用此例),并警告投稿人在横向研究中使用将贬损手稿价值,但在该文发布后,该刊依然录用了在结果表述时使用了“influence/impact”的论文(Minkande, 2025; Song, 2025),而且还是横向研究。在一些重要期刊(例如, *Journal of Educational Psychology*, *Journal of Applied Psychology*),非实验研究使用“influence/impact/affect”等词表达影响关系的论文也不少。

上述现象不仅可以反驳葛文“影响等同于因果”,而且可以反驳葛文所谓“混淆目标与实现目标的手段”。对于非实验研究,即使是纵向数据也比研究者想象中更难做出因果推断(Bailey et al., 2024)。如果研究者总是以因果推断为目标,以因果推断的有效性作为评判研究有无价值的标准,那么容易得到因果推断的问题就会迅猛发展,而不易做出因果推断的问题则无人问津(许琪, 2024)。温文将有方向性的相关关系明确定义为影响关系,可以帮助非实验研究者摆脱因果证据不足带来的种种困扰,更加关注研究问题本身,如实报告研究的是影响关系。此外,温文也鼓励非实验研究者尽量控制第三变量,使得影响关系尽量接近因果关系,可以提供更多的信息。

4 “预测”代替不了“影响”

葛文认为,有方向性的共变关系,一直有明确且公认的叫法——“预测”。说公认有点言过其实,但确实有一部分审稿人或者期刊编辑,不同意研究者在非实验尤其是横向研究中使用“影响”一词,而是建议使用“预测”来替代。葛文引用了《现代汉语词典》对预测的解释:“预先推测或测定”。从词义上不难看出,发出预测的主语不应当是一个变量。一个变量可以影响另一个变量,但说“一个变量预测(预先推测或者测定)另一个变量”,算不算病句呢?

例如,工作负担影响教师工作满意度,用预测的说法,正确的应当是:“研究者预测,工作负担会降低教师工作满意度”,“工作负担对教师工作满意度有预测作用”,“工作负担可以用来预测教师工作满意度”,等等。目前文献上经常会见到类似“工作负担预测教师工作满意度”的说法,确实比前面的说法简单,如果能正确理解其含义,似乎没有必要

反对这种用法。但葛文建议用预测指代“有方向性的相关”，可以合理推测其受到这种句式的误导了。

本节要论述的是，使用预测指代“有方向性的相关”有两个根本性问题：其一，现有预测的应用范围超出相关关系的外延，因而“预测”与“影响”不匹配；其二，预测方向性是人为确定的，反映不了变量之间的真实方向性。

4.1 统计上凡是“相关”皆可用来“预测”

综合字典释义、常识和统计知识，可以将“预测”理解为：根据现有的资料、数据、信息和特定的方法对未知事件或趋势进行推断的过程。

按预测方法分类，可以分为定性方法、定量方法和定性定量相结合的混合方法。定性方法包括德尔菲法、专家会议法、情景分析法、历史类比法等。定量方法包括时间序列分析、回归分析、机器学习、蒙特卡罗模拟等。可见，预测应用范围很广。下面聚焦大家比较熟悉的回归分析来讨论。

“预测”一词作为统计上的概念，许多读者是在回归分析中见到的。回归分析本质上是一种相关分析，只要两个变量相关(无论有无方向性)，就可以建立回归模型用一个变量预测另一个变量，许多教科书对回归分析都持这样的观点(例如，黄希庭，张志杰，2010；盛骤等，2008；王天夫，2006；温忠麟，2016；施雨，2005；辛自强，2017；张厚粲，徐建平，2020；张敏强，2010)。许多国外的教科书更是将相关和回归放在一起论述，甚至在章节标题上出现 regression/correlation (例如，Cohen et al., 2002；Gravetter et al., 2021)。

葛文也举了一个例子：“研究者利用银行交易记录(所反映出的人类行为)来预测一个人的人格特质(Gladstone et al., 2019)，无须关心是花销影响人格、还是人格影响花销、又或是第三变量同时影响二者，只要已知花销、能在可接受的精确度上猜中人格，预测性目标就达成了。”可见葛文是认可预测不需要方向性，只要相关就可以预测，怎么能说“有方向性的共变关系，一直有明确且公认的叫法——‘预测’”呢？

因果、影响或者相关，在量化的时候都是依靠相关显著性和效应量。对预测而言，除了显著性和效应量，还多了“个体的预测值及其置信区间”，“均值的预测值及其置信区间”。**是否可用来预测，用显著性来衡量，预测的准确性用效应量和置信区间来衡量。**效应量越高(如回归分析中的 R^2)，准确性越高；置信区间越小，准确性越高。

影响关系可以用来做“预测”，但反过来，能用预测指代影响关系吗？当然不行，两者不等同。影响关系是相关关系的真子集，而相关关系都可以用来“预测”。当一个研究结果表述为“A 可以预测 B”，或者更严格地说，“A 对 B 有预测作用”，并没有比“A 与 B 相关”提供更多的变量之间方向性的信息，这是因为“预测”的方向性随研究目的而定。

4.2 “预测”的方向性随研究目的而定

只要两个变量之间有相关关系，就可以在形式上“确定”自变量和因变量，建立回归模型进行预测(张敏强，2010)。谁是自变量，谁是因变量，这个方向性根据研究目的而定。温文在其讨论部分已经说明：影响关系的方向性必能成为回归模型的方向性；但反过来，一个回归模型的方向性，未必能成为影响关系的方向性。

Cohen 等人(2002)下面这句话很值得应用统计工作者牢记：“科学研究中的因果推断，取决于数据是怎么产生的，而非数据是怎么分析的”。就算是以因果冠名的模型，也要看数据是怎样产生的才能推断因果是否成立。同理，**数据是怎么分析的，改变不了变量之间本来的方向性。**

举一个例子来说明预测的方向性与影响的方向性可能不一致。利用鞋码与身高的相关性，刑侦人员经常使用留在罪案现场的脚印长度(鞋码长度)预测罪犯身高。但实际上身高是比鞋码更靠前的变量。理论上，身高是“脚长影响鞋码”的代理影响因素；现实中，一个人的身高比其所穿的鞋码更早出现，所以不能说鞋码影响身高。其实，上面提到的葛文例子“研究者利用银行交易记录来预测一个人的人格特质”，也是预测的方向性与实际方向性不一致，实际的方向性应当是人格影响银行交易记录。

说得更明白一点，预测分析可以是探索性的，可以没有假设。用回归分析做预测时，变量之间的方向性是人为确定的，实际情况可能没有方向性，也可能与回归分析的方向性刚好相反。但影响分析只能是验证性的，需要有假设，研究者要为方向性假设提供理据。值得一提的是，因为中介分析的是变量之间的影响机制，即变量 X 是如何影响 Y 的(温忠麟，叶宝娟，2014)，所以中介分析只能是验证性的，变量之间的方向性需要理据(温忠麟等，2024a, 2024b)。如果研究者利用相关关系做预测和中介分析，违背了中介分析的作用和功能，因为在相关关系中变量之间没有方向性，用来预测的时候变量之间的方向性是人为确定的，讨论中介变量的作用就

失去了意义。

5 葛文中的矛盾表述

前面分析说明了葛文的四点质疑缺乏逻辑, 现在看看葛文中的多处矛盾表述。

例 1. 葛文认为温文定义的影响与因果区分不开, “因果关系与影响关系之界限暧昧”(1.3 节第一段); 但又质疑: “他们认为, 只要将某些原本被称为相关关系的研究重命名为影响关系, 就能使这些相关性研究不会违反中介模型是因果模型的本质要求、可以名正言顺地实施中介分析。”(1.4 节最后一段)

温文说有人将影响理解为因果, 有人将影响理解为相关。没想到葛文同时扮演了两种角色, 前面质疑温文定义的影响关系与因果关系区分不开, 后面却质疑温文将某些原本被称为相关关系的重命名为影响关系。

按温文的图一, 三种关系从概念上说有明显区分, 并且有包含关系。有些相关关系是影响关系甚至是因果关系, 有些影响关系是因果关系。但对一个具体例子, 如果不知道有没有方向性, 也就区分不了是相关还是影响。如果没有通过实验去排除其他解释, 也难以区分是影响还是因果。但葛文将概念的区分与具体例子中变量关系的判别混为一谈, 导致前后矛盾。

例 2. 一方面葛文对温文在相关关系与因果关系之外创设影响关系这个第三目标感到奇怪(1.4 节第二段); 另一方面大篇幅谈论因果目标之外的预测目标: “举几个例子就可以快速了解为什么有的研究者只关心预测目标、不关心因果目标。”(2.1 节第二段)。这个预测目标是第几个目标呢?

例 3. 葛文谈到委婉用语的时候, 一方面引用文献来表达不赞成因果推断中使用委婉用语(1.4 节最后一段); 另一方面却认为, 若研究者担心无法表明方向, 例句“来自祖先多样性较高地区的人们更愿意公开表达情感”似乎是一个非常不错的表达方式(第 2 节第四段)。这种说法, 如果不是自相矛盾, 至少算得上双重标准。

例 4. 前面说有方向性的相关一直有明确且公认的叫法——“预测”(第 2 节第一段); 后面却举例说: “……无须关心是花销影响人格、还是人格影响花销、又或是第三变量同时影响二者, 只要已知花销、能在可接受的精确度上猜中人格, 预测性目标就达成了”(2.1 节第二段)。这样说来, “预测”还能

取代“影响”吗?

6 总结

本文回应了葛文对温文的质疑。针对“影响关系”没有明确定义的说法, 本文说明其定义用的是“属加种差”的定义方法, 完全符合逻辑, 其严密性可以与数学概念的定义媲美。本文还解释了影响关系的属概念相关关系, 相关即共变。除了因果关系与影响关系的概念是否等同的问题外, 葛文涉及因果的各种论述, 包括因果三准则不能作为因果定义、因果有多种定义等等, 都不能构成对温文影响关系定义的有效质疑, 因为影响关系的定义不需要使用因果的概念。

本文明确指出葛文的“影响关系等同于因果关系”是一个全称命题, 从而说明其使用再多的例子也无法证明该命题。在简介因果的定义和标准后, 本文从理论上证明影响关系不等同于因果关系, 并给出多个容易理解的案例, 其中影响关系成立但因果关系不成立。还顺便在案例中呈现和解释了代理影响因素的地位和作用。

对于葛文“不能因为手段上无法为因果目标提供因果证据, 便创设新目标”的说法, 本文认为目标不能独立于手段, 骑着自行车上不了天, “与其盯着天上飞的白日梦, 不如抓住手里的热馒头”。非实验研究达不到证实因果的目标, 那就踏实做好影响关系研究, 也优于仅仅得到相关关系。

在统计上, 凡是“相关”皆可用来“预测”, 说明葛文建议的用“预测”代替“影响”是行不通的。预测本质上是利用变量间的关系进行统计推断, 但不等同于关系本身。变量关系是预测的基础, 而预测是关系的应用。就是说, 预测和因果、影响、相关不是同一个范畴的概念, 后面三个都是变量之间关系, 但预测是应用变量关系的一种研究方法, 葛文的一位审稿专家就很中肯地指出了这一点。

查询各种人工智能工具, 可以帮助理解影响关系和因果关系的区别。向 ChatGPT 提问“What is the difference between causality and influence”, 不同版本的 ChatGPT 或者不同时间的回答可能有出入, 但关键信息是一样的: 因果关系中, 一事物对另一事物的作用是确定且直接的, 即一事物直接导致另一事物的发生; 影响关系中的作用可以较弱、间接、具有或然性, 即一事物会塑造或改变另一事物, 但无法完全决定它。再看看 ChatGPT 给的一个影响关系的例子: 父母的习惯会影响孩子的吸烟行为, 但

并不会直接导致孩子吸烟。其他影响因素还包括同龄人压力或个人选择。**如果你相信 ChatGPT 的说法,会感觉到“影响”简直是为基于问卷的研究结果量身打造的一个术语!**

因为在实验研究中“影响就是因果”,不少学者有了这种思维定势。使用影响一词描述问卷研究结果能否被接受,在乎碰到什么样的审稿人,无论中文还是英文期刊都一样。这就可以解释为什么很多作者投稿时被要求避免使用“影响”的表述,但还是有基于问卷的文章使用“影响”的说法也能见刊。温文给影响关系明确定义,为基于问卷的研究工作者提供了有别于“相关”级别(包括关联、关系、联系、预测等)的成果表述方式。其实影响关系不是温文首创,非实验研究的文献上在结果中使用“影响”的表述并不少见,温文只不过试图为这种做法奠定基础。

在英文中,除了 influence, 还有 impact 和 affect 有“影响”的意思,其中 influence 与因果的区别最大。英文文章要用“影响”表述结果的话,使用 influence 比较妥当。如果碰到将影响当作因果的审稿人,找到在投期刊或者级别更高的期刊(行内口碑更好的期刊)上用了“影响”表述结果的同类研究文章,加上 ChatGPT 对影响与因果的理解,去回应审稿意见或许可以过关。否则,还得使用“相关”级别的术语去表述结果。

综上所述可以得出结论:(1)影响关系定义严密,通过方向性的逻辑理据加上统计上相关性检验,可以判断变量之间是否为影响关系;(2)因果关系是特殊的影响关系,在实验研究中两者是同义词,但在其他场合影响关系不一定是因果关系;(3)实验之外,不必以因果为目标,建立影响关系可以作为非实验研究的一种目标;(4)预测是对变量关系的应用,与影响关系不相称。陈述“一个变量可以用来预测另一个变量”,相当于根据变量相关性做了预测,但不能说明变量之间一定有影响关系。

参 考 文 献

- Altman, N., & Krzywinski, M. (2015). Association, correlation and causation. *Nature Methods*, 12(10), 899–900.
- Bailey, D. H., Jung, A. J., Beltz, A. M., Eronen, M. I., Gische, C., Hamaker, E. L., ... Murayama, K. (2024). Causal inference on human behaviour. *Nature Human Behaviour*, 8(8), 1448–1459.
- Bartram, D., Alaimo, L. S., Avery, E., Bardo, A., Bella, E. D., Binder, M., ... Tani, M. (2024). Towards the next fifty years of Social Indicators Research: Some guidance for authors. *Social Indicators Research*, 174(1), 1–17.
- Cohen, J., Cohen, P., West, S. G., & Aiken, L. S. (2002). *Applied multiple regression/correlation analysis for the behavioral sciences* (3th ed.). Routledge.
- Copi, I. M., Cohen, C., & Rodych, V. (2019). *Introduction to logic* (15th ed.). Routledge.
- Dai, B., & Jian, X. (2022). *Applications of SPSS26.0 statistical analysis in psychology and education*. Beijing: Beijing Normal University Press.
- [戴步云, 简小珠. (2022). *SPSS 26.0 统计分析在心理学与教育学中的应用*. 北京: 北京师范大学出版社.]
- Gan, Y., Zhang, Y., & Zheng, L. (2019). *Introduction to psychological and behavioral Statistics* (2nd ed.). Beijing: Peking University Press.
- [甘怡群, 张轶文, 郑磊. (2019). *心理与行为科学统计* (第二版). 北京: 北京大学出版社.]
- Ge, X. (2025). The so-called influence relationship requires caution: Commentary on Wen et al. (2024). *Acta Psychologica Sinica*, 57, 1098–1107.
- [葛泉语. (2025). 所谓影响关系有待商榷: 对温忠麟等人 (2024) 的评论. *心理学报*, 57, 1098–1107.]
- Gladstone, J. J., Matz, S. C., & Lemaire, A. (2019). Can psychological traits be inferred from spending? Evidence from transaction data. *Psychological Science*, 30(7), 1087–1096.
- Granger, C. W. J. (1969). Investigating causal relations by econometric models and cross-spectral methods. *Econometrica*, 37(3), 424–438.
- Gravetter, F. J., Wallnau, L. B., Forzano, L. B., & Witnauer, J. E. (2021). *Essentials of statistics for the behavioral sciences* (10th ed.). Cengage Learning.
- Haber, N. A., Wieten, S. E., Rohrer, J. M., Arah, O. A., Tennant, P. W. G., Stuart, E. A., ... Fox, M. P. (2022). Causal and associational language in observational health research: A systematic evaluation. *American Journal of Epidemiology*, 191(12), 2084–2097.
- Hernán, M. A., & Robins, J. M. (2020). *Causal inference: What if*. Boca Raton: Chapman & Hall/CRC.
- Hau, K-T., Wen, Z., & Cheng, Z. (2021). *Structural equation model and its applications (Edition with Mplus)*. Beijing: Educational Science Publishing House.
- [侯杰泰, 温忠麟, 成子娟. (2021). *结构方程模型及其应用 (Mplus 版)*. 北京: 教育科学出版社.]
- Huang, X., & Zhang, Z. (2010). *Research methods in psychology* (2nd ed.). Beijing: Higher Education Press.
- [黄希庭, 张志杰. (2010). *心理学研究方法* (第 2 版). 北京: 高等教育出版社.]
- Li, W. (2018). The potential outcomes model in the causal inference: Origins, logic, and implications. *Public Administration Review*, 11(1), 124–149.
- [李文钊. (2018). 因果推理中的潜在结果模型: 起源、逻辑与意蕴. *公共行政评论*, 11(1), 124–149.]
- Minkande, J. B. (2025). Farming households facing food insecurity in Cameroon: What indicators for resilience? *Social Indicators Research*, 177, 809–832.
- Morling, B. (2018). *Research methods in psychology: Evaluating a world of information* (3rd ed.). W.W. Norton & Company.
- Popper, K. R. (1959). *The logic of scientific discovery*. New York: Routledge.
- Rubin, D. B. (1974). Estimating causal effects of treatments in randomized and nonrandomized studies. *Journal of Educational Psychology*, 66(5), 688–701.
- Rubin, D. B. (2005). Causal inference using potential outcomes:

- Design, modeling, decisions. *Journal of the American Statistical Association*, 100(469), 322–331.
- Shadish, W. R., Cook, T. D., & Campbell, D. T. (2002). *Experimental and quasi-experimental designs for generalized causal inference*. Boston: Houghton-Mifflin.
- Sheng, Z., Xie, S., & Pan, C. (2008). *Probability theory and mathematical statistics* (4th ed.). Beijing: Higher Education Press.
- [盛骤, 谢式千, 潘承毅. (2008). *概率论与数理统计* (第四版). 北京: 高等教育出版社.]
- Shi, Y. (2005). *Applied mathematical statistics*. Xi'an: Xi'an Jiaotong University Press.
- [施雨. (2005). *应用数理统计*. 西安: 西安交通大学出版社.]
- Smith, R. (1995). Logic and the definition of terms: Aristotle's theory of definition. In W. Wians (Ed.), *Aristotle's philosophical development: Problems and prospects* (pp. 123–144). Rowman & Littlefield.
- Song, L. (2025). Unveiling the neighborhood effect on intergenerational mobility: Which scale is more relevant? *Social Indicators Research*, 176, 5–28.
- Wang, T. (2006). Causal analysis in social research. *Sociological Studies*, 7(4), 132–157.
- [王天夫. (2006). 社会研究中的因果分析. *社会学研究*, 7(4), 132–157.]
- Wen, Z. (1998). New approaches to some issues about statistical tests in educational experiments. *Education Research and Experiment*, (1), 63–66.
- [温忠麟. (1998). 对教育实验统计检验若干问题的再探讨. *教育研究与实验*, (1), 63–66.]
- Wen, Z. (2016). *Statistics in psychology and education* (2nd ed.). Guangzhou: Guangdong Higher Education Press.
- [温忠麟. (2016). *心理与教育统计* (第二版). 广州: 广东高等教育出版社.]
- Wen, Z., Liu, F., Zheng, Y., Liao, X., & Huang, Y. (2024a). Why are there so many empirical articles on moderating or mediating effects? *Chinese Journal of Applied Psychology*, 30(4), 291–297.
- [温忠麟, 刘方, 郑渊丹, 廖心怡, 黄亦南. (2024a). 为何调节效应或中介效应的实证文章那么多? *应用心理学*, 30(4), 291–297.]
- Wen, Z., Wang, Y., Ma, P., & Meng, J. (2024b). The influence relationship among variables and types of multiple influence factors working together. *Acta Psychologica Sinica*, 56(10), 1462–1470.
- [温忠麟, 王一帆, 马鹏, 孟进. (2024b). 变量之间的影响关系和多重影响因素的共同作用类型. *心理学报*, 56(10), 1462–1470.]
- Wen, Z., Xie, J., Fang, J., & Wang, Y. (2022). Methodological research on hypothesis test and related issues in China's mainland from 2001 to 2020. *Advances in Psychological Science*, 30(8), 1667–1681.
- [温忠麟, 谢晋艳, 方杰, 王一帆. (2022). 新世纪20年国内假设检验及其关联问题的方法学研究. *心理科学进展*, 30(8), 1667–1681.]
- Wen, Z., & Ye, B. (2014). Analyses of mediating effects: The development of methods and models. *Advances in Psychological Science*, 22(5), 731–745.
- [温忠麟, 叶宝娟. (2014). 中介效应分析: 方法和模型发展. *心理科学进展*, 22(5), 731–745.]
- Xin, Z. (2017). *Research methods in psychology* (2nd ed.). Beijing: Beijing Normal University Press.
- [辛自强. (2017). *心理学研究方法* (第2版). 北京: 北京师范大学出版社.]
- Xu, Q. (2024). Fifty years of causal inference: Achievements, challenges, and responses. *Academic Monthly*, 56(11), 131–144.
- [许琪. (2024). 因果推断五十年: 成就、挑战与应对. *学术月刊*, 56(11), 131–144.]
- Zhang, H., & Xu, J. (2020). *Modern psychological and educational statistics* (5th ed.). Beijing: Beijing Normal University Press.
- [张厚粲, 徐建平. (2020). *现代心理与教育统计学* (第5版). 北京: 北京师范大学出版社.]
- Zhang, M. (2010). *Educational and psychological statistics* (3rd ed.). Beijing: People's Education Press.
- [张敏强. (2010). *教育与心理统计学* (第三版). 北京: 人民教育出版社.]

Distinguishing between causality, influence, correlation, and prediction

WEN Zhonglin, MA Peng, MENG Jin, WANG Yifan

(Center for Studies of Psychological Application / School of Psychology, South China Normal University, Guangzhou 510631, China)

Abstract

Wen et al. (2024) discussed the following three issues: (1) explaining why it is inappropriate to understand influence relationship between variables as causal or correlation relationship, and then providing the definitions for two terms, influence relationship and influence factor; (2) summarizing several ways to find evidence for justifying the directionality when modeling the influence relationship. (3) categorizing multiple influence factors working together.

Ge (2025) questioned Wen et al.'s article as follows: (1) the concept of "influence relationship" is not clearly defined; (2) influence relationship and causal relationship are indistinguishable; (3) one cannot create a new goal just because the means cannot provide causal evidence for the goal of causality; (4) the so-called influence relationship should be called "prediction". In response to these concerns, the present article offers clarifications and justifications.

First, the influence relationship has been rigorously defined by using a “Genus and Differentia” approach in Wen et al.’s paper. The influence relationship can be determined by using logical reasoning and statistical correlation testing. We also explain “correlation”, the genus concept of the influence relationship.

Furthermore, any causal relationship is the influence relationship, and the two are equivalent in studies through randomized controlled experiments, whereas in other contexts influence relationship may not necessarily be causal relationship. We provided easily understandable cases where the influence relationship was established but the causal relationship was not. It also presented and explained the status and role of proxy effects in those cases.

Beyond experimentation, establishing influence relationships is a suitable goal for investigation and research, which is better than the goal of establishing correlation. In statistics, any “correlation” can be used for “prediction”, and the direction of prediction can differ from the actual direction of the variable relationship. Prediction is essentially statistical inference based on the relationships between variables, but it is not the relationships themselves. Causality, influence, and correlation are all relationships between variables, whereas prediction concerns the application of variable relationships. Therefore, prediction is not an appropriate substitute for influence relationship.

Keywords influence relationship, correlation relationship, causal relationship, influence factor, prediction