第四章 模型估计与分析：结构方程模型

本章内容

第一节　结构方程模型概述
第二节　结构方程模型的组成
第三节　参数估计与识别问题
第四节　模型拟合评鉴
第五节　验证性因素分析
第六节　路径分析
第七节　结构方程模型：统合模型分析
第八节　中介与调节
第九节　结构方程模型的正确运用

第五节 验证性因素分析

一、验证性因素分析原理

1. 探索性与验证性因素分析

传统上，研究者在进行因素分析之前，并未对于数据的因素结构有任何预期与立场，而是借由统计量来分析因素的结构。此种因素分析策略带有浓厚的尝试错误的意味，因此称为**探索性因素分析（EFA）**。然而有时，研究者在研究之初就已提出了关于特定的结构关系的假设，例如，某个概念的测量问卷是由数个子量表组成的，此时因素分析可以用来确认数据的模型是否为研究者所预期的形式，此种因素分析称为**验证性因素分析（CFA）**，具有理论检验与确认的功能。在技术层次来说，CFA是SEM的一种次模型（submodel），除了做因素结构检验之用，还可以与其他次模型整合，构成完整的SEM分析。

EFA与CFA最大的不同在于测量的理论架构（因素结构）在分析过程中所扮演的角色与检验时机：

- 对EFA而言，测量变量的理论架构是因素分析的产物，因素结构是在一组独立的测量指标或题目间，以数学程序与研究者的主观判断决定的一个具有计量合理性与理论适切性的结构，并以该结构代表所测量的概念内容。换句话说，**理论架构的出现在EFA中是一个事后（posterior）的概念**。
- 相对而言，CFA的进行则必须有特定的理论观点或概念架构作为基础，然后借由数学程序来确认该理论观点所导出的计量模型是否确实、适当。换句话说，**理论架构对于CFA的影响是在分析之前发生的，计量模型具有理论的先验性，其作用是作为一种事前（priori）的概念**。

第五节 验证性因素分析

一、验证性因素分析原理

1. 探索性与验证性因素分析

在SEM架构下，CFA为一种子模型，CFA分析的数学原理与统计程序都是SEM的一种特殊应用。由于SEM的模型界定能够处理潜在变量的估计与分析，具有高度的理论先验性，因此如果研究者在测量之初即对于潜在变量的内容与性质非常明确，已详细地加以推演，或有具体的理论基础，提出了适当的测量变量组成测量模型，那么借由SEM的分析程序，可以对于潜在变量的结构或影响关系进行有效的分析。

SEM中对于潜在变量的估计程序，即是在考验研究者先期提出的因素结构**（测量模型）**的适切性。

一旦测量的基础确立了，潜在变量的因果关联就可以进一步通过多元回归、路径分析的策略**（结构模型）**来加以探究。

一般而言，CFA可以说是进行整合性SEM分析的一个前置步骤或基础架构，亦可以独立进行。

在SEM的分析架构中，CFA所检验的是测量变量与潜在变量的假设关系，可以说是SEM中最基础的测量部分。它不但是SEM中其他后续高等统计检验的基础，更可以独立地应用在对信效度的考验与理论有效性的确认上。

第五节 验证性因素分析

一、验证性因素分析原理

2. 潜在变量的因素分析

潜在变量反映的是研究者关心但无法直接观察与测量的现象或假设性构念。从外表来看，构念所反映的通常只是一组有相互关联的行为现象，经由研究者基于理论的推导、文献的支持或个人主观的分析与演绎而提出的理论上存在的假设。

为了使构念得以有效地测定，需要经过下列四个程序:

① 构念必须要有明确的操作性定义来界定其内容与范畴;

② 用以测量构念的指标能够被明确地指出;

③ 测量同一构念的指标必须具有相当的一致性，测量不同构念的指标则有相当的区辨性，多元指标的一致与区辨性应能从观察资料中检验得出;

④ 经由统计检验的程序，观察数据可以支持或推翻构念是否存在的假设。

上述程序称为构念有效化（construct validation），从SEM的术语来说，这个过程就是检验构念的测量模型的过程。当把从测量变量中搜集到的协方差矩阵与由测量模型推导、计算出来的协方差矩阵加以比对时，两者若相符且达到一定的程度，我们即无法推翻测量模型存在的可能性，也就是说，研究者所提出的潜在构念的概念建构获得了支持。

第五节 验证性因素分析

一、验证性因素分析原理

2. 潜在变量的因素分析

Kline (1998)指出，对测量模型的检验有两个值得注意的逻辑上的错误：

第一，当测量数据无法指出测量模型不适切时（即无法推翻测量模型存在的可能性），并不能说明测量模型所代表的抽象建构确实存在。因为同一组测量指标有多种不同的组合方式，也就是说，可能有其他替代模型存在。

第二种逻辑上的错误是抽象概念实质化（reification）。当SEM程序确认了特定的测量模型最能反映观察数据时，仅代表某种抽象概念的测量程序的有效性得到了支持，但是该抽象建构是否确实存在则无法确知。也就是说，假设性构念不一定反映了真实世界的现象。如果研究以测量模型的结果去描绘客观的事实，则犯了抽象概念实质化的逻辑错误。

第五节 验证性因素分析

一、验证性因素分析原理

3. 验证性因素分析的特性

在SEM的术语中，测量模型的检验程序称为验证性因素分析，图5.1即是一个典型的CFA测量模型。图5.1中有两个相关的潜在变量F1与F2，F1由V1~V3这3个指标来测量，F2由V4~V6这3个指标来测量，E1~E6分别代表6个测量变量的测量误差。从潜在变量指到测量变量的单箭头代表研究者所假设的潜在变量对于测量变量的直接因果关系。经由统计过程对于这些因果关系的估计数称为因素载荷，有标准化与未标准化两种形式，性质类似于回归系数。

在整个模型当中，研究者所能具体测量的是6个测量变量，其背后受到某些共同的潜在变量的影响，因此测量变量可以说是内生变量，潜在变量与测量误差则为外源变量。从变异量的拆解原理来分析，每个测量变量的变异量可以被拆解成两部分：**共同变异（common variance）**与独特（或误差）变异（unique variance）。以F1为例，V1、V2、V3是用来测量潜在特质F1的指标或测量题目，其背后受到**同一个潜在因素的影响**——从数学关系来说，即是3个变量共变的部分。而测量误差的部分就是3个测量变量无法被该潜在变量解释的独特变异量或残差，彼此相互独立。

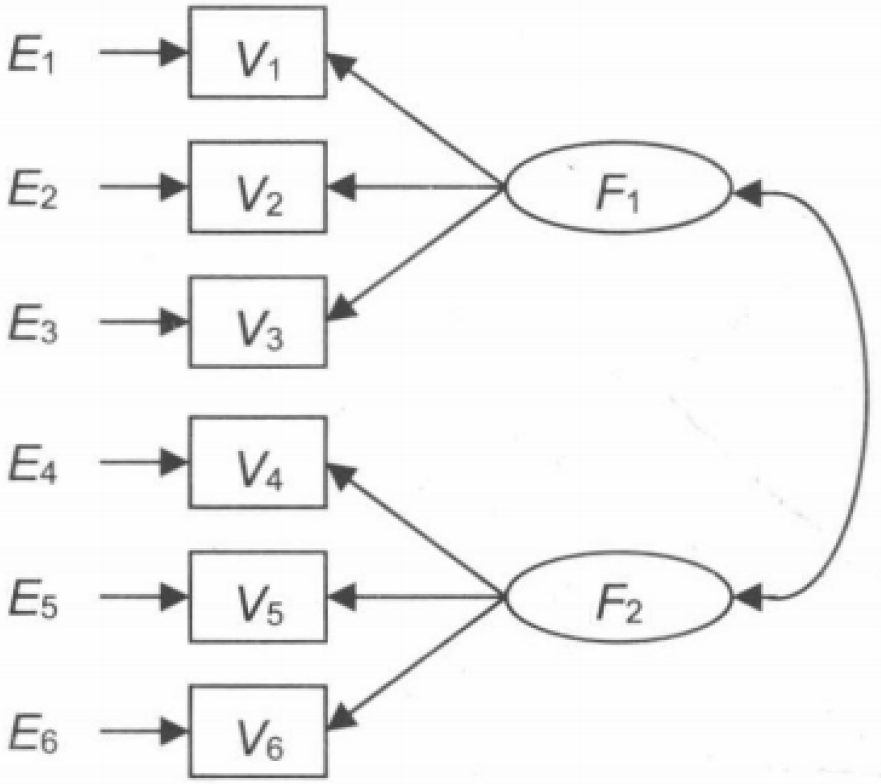


图 5.1 典型的 CFA 测量模型图

第五节 验证性因素分析

一、验证性因素分析原理

4. 测量误差与方法效应

CFA测量模型能够有效地处理路径分析当中变量具有一定测量误差的问题。CFA测量模型不但将每一个具有抽象特性的外源或内生变量以潜在变量的形式表示了出来，同时对于潜在变量的每一个测量变量进行了误差估计，即对独特变异量的分析。从测量变量拆解得出的独特变异量可能包含两种类型的测量误差。

第一种是随机误差（random error），也就是传统信度估计所处理的测量误差。造成随机误差的原因有很多，例如，测量的过程因素、受测者个人因素以及工具的因素等。但这些因素对于测量分数并无系统化的影响（例如系统性地高估或低估），因此称为随机性的误差来源。

第二种是系统误差（systematic error），对于测量分数会有系统化的影响，使测量分数以特定的模型偏离实际的真分数。理论上，系统化误差可以从测量变量中的独特变异量中抽离出来，因为系统性误差是在潜在变量之外仍有其他影响分数的变异来源导致的。最明显的一个例子是方法效应（method effect），也就是研究过程中对于变量的测量基于某一种特殊的方法（例如，纸笔测验），造成测量分数的系统性变化而无法反映真分数的一种现象。

针对这些方法/工具引发的系统误差，CFA的测量模型提供了技术上的解决策略，借由对共变关系的分析，配合多维测量的假设（见5.单维测量与多维测量），可以有效地估计存在于独特变异当中的系统性方法变异。 此外，研究者发展出了CFA的多重特质多重方法矩阵法（MTMM）程序来分析多重方法造成的多重方法效应问题。

第五节 验证性因素分析

一、验证性因素分析原理

5. 单维测量与多维测量

CFA测量模型与传统因素分析有一个相当大的不同，即测量变量与潜在因素之间的组合形态不但可以由研究者依其假设来指定，更可以突破单一变量只能反映单一潜在变量——单维测量（unidimensional measurement）——的限制。

以图5.2为例，每一个测量变量皆与特定的潜在变量相联结，但是V3与V5另有额外的联结（V3←F2、V5←F1），也就是说V3与V5同时与两个潜在变量有关，反映出对V3与V5两个变量的测量有其他变异来源的特殊测量模型，称为多维测量（multidimensional measurement）。

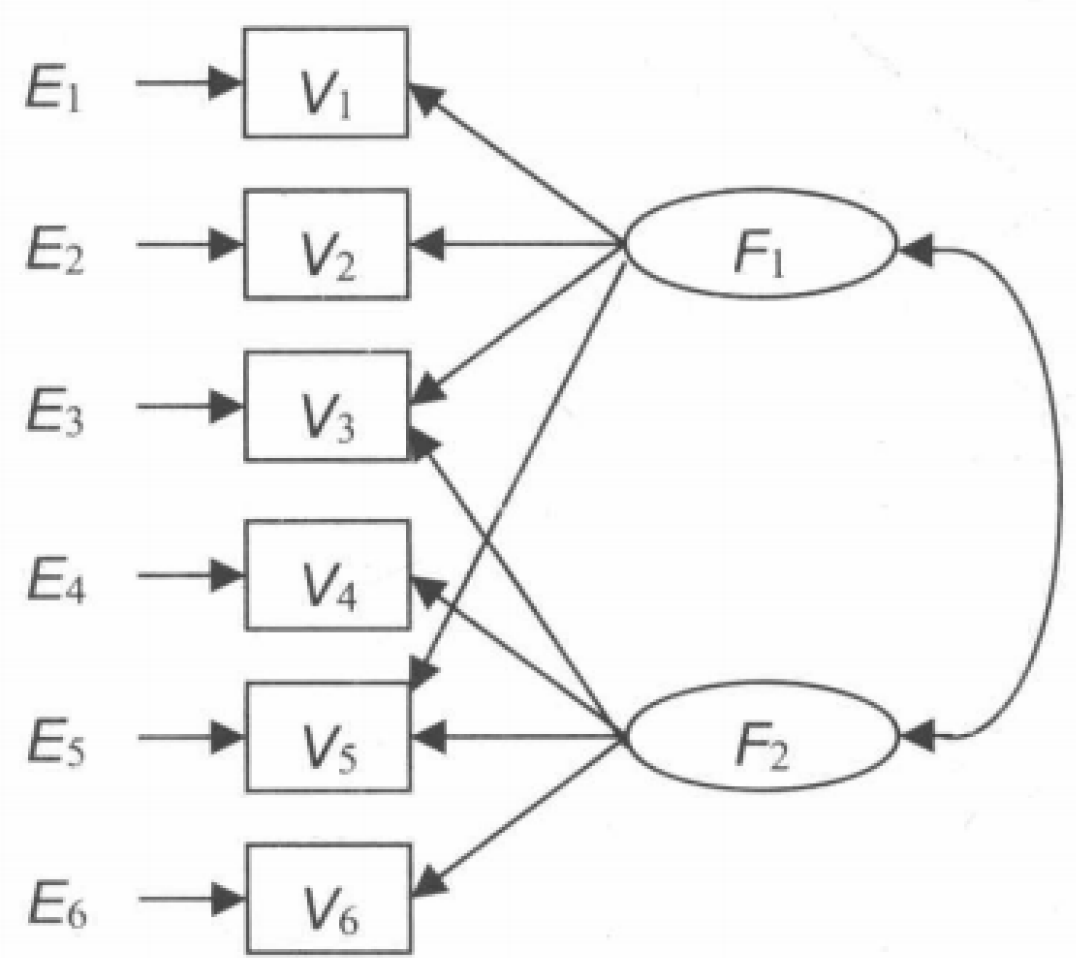


图 5.2 多维测量的 CFA 测量模型图

第五节 验证性因素分析

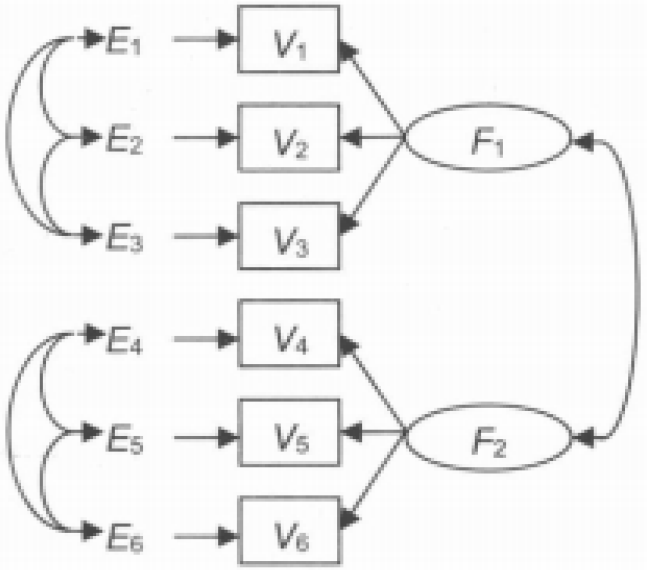
一、验证性因素分析原理

5. 单维测量与多维测量

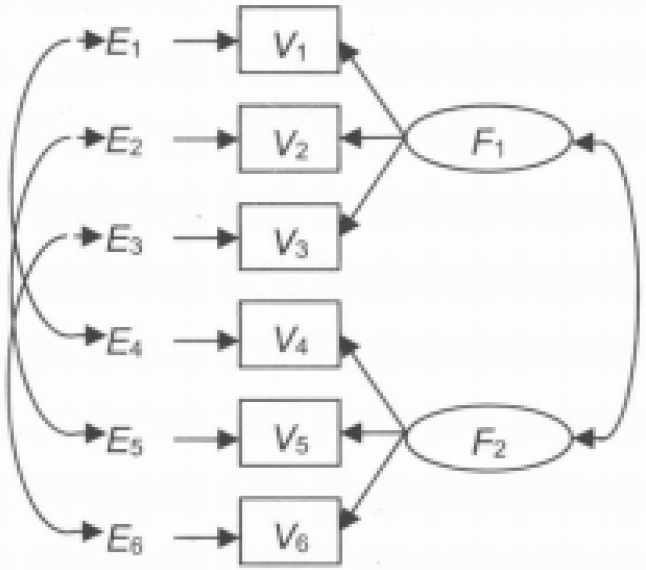
CFA测量模型除了允许测量变量与潜在变量具有多维关系之外，测量变量的误差项（独特变异）也可以与其他变量存在假设性的共变，也就是说，CFA测量模型在技术上允许测量变量的误差项为多维测量。最普遍的一种现象是使用相关误差（correlated measurement error）的测量模型，当误差项存在有意义的相关时，代表测量变量除了受特定潜在特质的影响之外，尚有其他未知的影响来源，须由对误差项的共变分析来估计。图5.3描绘出了两种不同的相关误差测量模型。在图5.3a与图5.3b中，误差项之间具有假设存在的共变，图5.3a中的关联误差发生在同一个潜在因素内，称为因素内关联误差模型，而图5.3b中的关联误差则跨越了不同的因素，称为因素间关联误差模型。

因素内关联误差模型最典型的例子是前面提及的方法效应，也就是说测量变量都是由同一种测量工具测量的。

因素间关联误差模型最典型的例子是对重测信度的测量。



（a）因素内关联误差的 CFA 模型



（b）因素间关联误差的 CFA 模型

图 5.3　两种关联误差的 CFA 测量模型图

第五节 验证性因素分析

二、测量模型的内部拟合检验

一个测量模型可否被接受以及参数估计的优劣，除了要从模型的整体拟合来看之外，还必须从模型的内在质量来衡量每一个潜在变量的适切性，又称内部拟合。

SEM虽然没有一个单一的效应量（R2）来进行模型的整体检验，但是可以就假设模型与观察模型之间的拟合程度进行整体检验。当模型拟合性被接受之后，我们得以针对个别的因素质量进行检测。在本质上，模型拟合性检验就是一种整体检验，个别因素的检验就是一种事后的评估程序。Hair等人(2006)认为，在CFA中，除了报告模型拟合指数之外，还必须进一步了解测量模型当中的个别参数是否理想（项目信效度），各潜在变量的组合情形是否稳定可靠（构念的信效度）。如果某些参数不甚理想，可以借由模型修饰程序来去除不良题目或增加参数，来提高模型的内在拟合。在具体做法上，较多被人采用的策略包括四项检验：项目质量、组合信度（ρc）、平均变异萃取量（ρv）以及构念区辨力。

第五节 验证性因素分析

二、测量模型的内部拟合检验

1. 项目质量检验

Bollen (1989)指出，构成一个有意义的潜在变量的最重要前提，是一组能够反映潜在构念意义的观察指标。换言之，构成潜在变量的题目必须具有相当的信效度，否则无法支撑一个潜在变量模型。

在古典测量理论中，信度（reliability）代表测量的可靠程度（trustworthiness），或对真分数的测量不受测量误差影响的程度。效度（validity）则反映了测量工具能够正确无误地测出潜在特质的程度，也就是研究者可以掌握抽象意义的程度。两者必须兼具，才能确保测量的质量。

就组成一个因素的个别题目来说，测量误差越小，表示测量题目受到误差的影响越小，能够测到真分数的程度越高。在SEM的测量模型中，测量残差由Θ_δ与Θ_ε两个矩阵的对角线的参数（δ与ε系数）反映，而这两组参数与因素载荷具有函数关系，因素载荷越高，测量残差越小。**Bagozzi与Yi (1988)认为，测量模型当中的测量残差必须具有统计显著性，才能确立一个潜在变量由一组带有测量误差的观察变量形成的这个前提。相反，如果测量误差太微弱而未达统计显著水平（或因素载荷太高，超过0.95），则意味该题足以完全反映该潜在构念的内容，测量模型的合理性即不复存在。除此之外，因素载荷系数的正负号也应符合理论预期，更不应出现超过-1~+1的数值。这些条件是测量模型的基本拟合指标（preliminary fit criteria）。**

第五节 验证性因素分析

二、测量模型的内部拟合检验

1. 项目质量检验

因素载荷除了反映测量误差的影响之外，也同时反映了个别题目能够用来反映潜在变量的程度。Hair等人(2006)认为，一个足够大的因素载荷代表题目具有良好的构念效度（construct validity）。一般而言，当因素载荷大于0.71时，即可以宣称项目具有理想质量，因为此时的潜在变量能够解释观察变量将近50%的变异，**这个λ≥0.71指标可以说是基本拟合指标当中最明确的一个判断标准**。

事实上，λ≥0.71原则来自传统因素分析当中对共同性（communality）的估计，这是个别题目反映潜在变量的能力的指标。Tabachnica与Fidell (2007)具体提出了下列标准（见表5.1）：当因素载荷大于0.71，也就是该因素可以解释观察变量50%的变异量之时，是非常理想的状况；当因素载荷大于0.63，也就是该因素可以解释观察变量40%的变异量之时，是很好的状况；但若载荷小于0.32，也就是该因素解释了不到10%的观察变量的变异量时，是非常不理想的状况——通常，这类题目虽然是形成某个因素的题目，但是贡献非常小，可以考虑删除该题，以提高整个因素的一致性。

表5.1 因素载荷的判断标准

λ	λ^2	状况		λ	λ^2	状况
0.71	50%	优秀		0.45	20%	普通
0.63	40%	很好		0.32	10%	不好
0.55	30%	好		0.32以下		不及格

第五节 验证性因素分析

二、测量模型的内部拟合检验

1. 项目质量检验

SEM分析软件中的多元相关平方（squared multiple correlation, SMC）数据反映了个别测量变量受到潜在变量影响的程度。SMC越高，表示真分数所占的比重越高；SMC越低，表示真分数所占的比重越低，信度越低。SMC的计算原理如下：

$$SMC_{var_i} = \frac{\lambda_i^2}{\lambda_i^2 + \Theta_{ii}}$$

其中，λ_i为个别测量变量的因素载荷，取平方后除以总变异量（解释变异量加误差变异量），即为个别题目的信度估计数。值得注意的是，SMC是以单维假设为基础的信度估计数，也就是一个测量变量仅受到单一潜在变量的影响（单一真分数变异来源）。如果一个测量变量受到两个或两个以上潜在变量影响，该式即不适用。

第五节 验证性因素分析

二、测量模型的内部拟合检验

2. 组合信度（ρ_c）

①测量信度的概念

观察分数=真实分数+误差分数

当误差为0时，观察分数可以完全反映真实分数。若以方差的概念来表示，观察分数的变异可以σ^2_{total}表示；而真实分数的方差可以σ^2_{true}表示；误差分数的方差则反映了测量误差波动情形，以σ^2_{error} 表示。三者关系如下：

$$\sigma^2_{total} = \sigma^2_{true} + \sigma^2_{error} \quad (5.2) \text{ 或 } 1 = \frac{\sigma^2_{true}}{\sigma^2_{total}} + \frac{\sigma^2_{error}}{\sigma^2_{total}} \quad (5.3)$$

公式5.3中，真实分数的变异除以变量的总变异量，代表一个测量分数的信度指标，即为信度系数（coefficient of reliability），以ρ表示。如公式5.4所示：

$$\rho = \frac{\sigma^2_{true}}{\sigma^2_{total}} = 1 - \frac{\sigma^2_{error}}{\sigma^2_{total}} \quad (5.4)$$

第五节 验证性因素分析

二、测量模型的内部拟合检验

2. 组合信度（ρ_c）

②SEM的组合信度估计

SEM测量模型的信度估计基本上延续了古典测量理论的观点，将信度视为真实分数所占的比例，而测量误差的变异即为观察分数当中无法反映真实分数的残差变异量。对于个别测量题目来说，由于测量变量分数的变动会受到潜在因素与测量误差的影响，而潜在因素影响产生的变异即代表真实分数的存在，因此，信度可以用测量变量的变异量能够被潜在变量解释的百分比（proportion of variance of a measured variable）来表示。Fornell与Larker (1981)基于前述SMC的概念，提出了一个非常类似内部一致性信度系数（Cronbach's α）的潜在变量的组合信度（composite reliability; CR或ρ_c）：

$$CR = \rho_c = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \Theta_{ii}} \quad (5.5)$$

其中，$(\sum \lambda_i)^2$为因素载荷相加求和后取平方的数值，$\sum \Theta_{ii}$为各观察变量残差方差的总和。

第五节 验证性因素分析

二、测量模型的内部拟合检验

2. 组合信度（ρ_c）

②SEM的组合信度估计

当测量模型中带有残差相关时，残差变异量估计数会因为残差间的相关而降低，因此ρ_c的估计必须将残差相关纳入计算，公式如下：

$$\rho_c = \frac{(\sum \lambda_i)^2}{(\sum \lambda_i)^2 + \sum \Theta_{ii} + 2\sum \Theta_{ij}} \quad (5.6)$$

其中，$\sum \Theta_{ij}$为第i与j题残差共变的总和。换言之，在利用SEM来估计模型之余，可以进行对测量工具的信度的估计。而且，除了可以估计整体量表的信度，也可以计算个别测量题目的信度。唯一的缺点是在计算整个因素（量表）的信度时，必须以人为的方式来计算，LISREL尚无模块来自动产生。

依据古典测量理论的观点，量表信度需达0.70才属比较稳定的测量，SEM的测量模型多沿用这一标准。但此标准的达成必须使各题的因素载荷平均达0.70以上，社会科学领域的量表不易达到此水平，因此，**Bagozzi与Yi (1988)建议ρ_c达0.60即可。 Raine-Eudy (2000)的研究指出，组合信度达0.50时，测量工具在反映真分数时即可获得基本的稳定性。**

第五节 验证性因素分析

二、测量模型的内部拟合检验

2. 组合信度（ρ_c）

③组合信度的区间估计

前述的ρ_c估计是一种点估计，如果考虑模型估计的随机性可能造成的误差影响，信度估计也应把样本波动因素纳入考虑，采取区间估计的做法来报告因素的信度，说明测量模型估计得到的ρ_c可以反映真正的总体信度的95%或99%置信区间。进行区间估计最重要的步骤是计算估计数的抽样标准误（standard error of sampling），Raykov (2002)提出CFA的因素信度的标准误计算公式如下：

$$SE(\rho_c) = \sqrt{D_1^2 Var(\mu) + D_2^2 Var(v) + 2D_1 D_2 Cov(\mu, v)} \quad (5.7)$$

其中，μ是标准化因素载荷估计数的总和，v是误差方差估计数的总和，Cov(μ,v)则是两者的协方差，D1与D2是量表信度系数分别对μ与v做偏微分，得到估计数：

$$D_1 = \frac{2\mu v}{(\mu^2+v)^2} \quad (5.8) \quad D_2 = \frac{\mu^2}{(\mu^2+v)^2} \quad (5.9)$$

一旦获得标准误的估计数后，即可获得信度的95%或99%置信区间：

$$95\%CI(\rho_c) = \rho \pm 1.96SE(\rho_c) \quad (5.10) \quad 99\%CI(\rho_c) = \rho \pm 2.59SE(\rho_c) \quad (5.11)$$

第五节 验证性因素分析

二、测量模型的内部拟合检验

3. 平均变异萃取量（ρ_v）

先前已经提及，测量题目的因素载荷越高，表示题目能够反映潜在变量的能力越高，因素能够解释各观察变量的变异的程度越大。因而可以计算出一个平均变异萃取量（Average Variance Extracted; AVE或ρ_v），来反映一个潜在变量能被一组观察变量有效估计的聚敛程度指标。公式如下：

$$AVE = \rho_v = \frac{\sum \lambda_i^2}{\sum \lambda_i^2 + \sum \Theta_{ii}} \quad (5.12)$$

其中，分母为各题的因素载荷平方加上误差变异，相加为1。因此，分母即为题数n：

$$\rho_v = \frac{\sum \lambda_i^2}{n} \quad (5.13)$$

换言之，ρ_v指标就是各因素的各题因素载荷平方的平均值。如果配合前述的λ≥0.71原则，那么ρ_v的判断标准即是0.50。当ρ_v大于0.50，表示潜在变量的聚敛能力十分理想，具有良好的操作性定义（operationalization）。从数学过程来看，ρ_v的概念其实就是传统EFA当中的特征值（eigenvalue）。也就是说，当各观察变量提供了一个单位变异量时，各因素的解释变异量也就是潜在变量变异量占总变异的百分比。换言之，CFA当中的每一个因素就是执行一次单因素的EFA的结果，ρ_v即为该单一因素的特征值。因此，ρ_v宜以概念本身来解释，而不宜解释成聚敛效度。

第五节 验证性因素分析

二、测量模型的内部拟合检验

4. 因素区辨力

Hair等人 (2006)除了引用ρ_v作为聚敛能力的指标，也指出了CFA估计结果所得到的潜在变量必须具有区分效度（discriminant validity），即不同的构念之间必须能够有效分离。

在具体的CFA操作技术上，有三种方式可以用来检验潜在变量的区辨力。第一种是相关系数的区间估计法，如果两个潜在变量的相关系数的95%的置信区间涵盖了1.00，表示构念缺乏区辨力。

第二种方法是竞争模型比较法，利用两个CFA模型来进行竞争比较。一个CFA模型是令两个构念之间相关的自由估计（效度模型），另一个CFA模型则是将相关设为1.00（完全相关模型；此模型等同于单一因素模型），完全相关模型由于少一个有待估计的参数，自由度多1，模型的拟合度也较低。如果效度模型没有显著地优于完全相关模型，即代表两个构念间缺乏区辨力。

第三种方法是平均变异萃取量比较法，比较两个潜在变量的ρ_v平均值是否大于两个潜在变量的相关系数平方。

Hair等人 (2006)将这些测量模型的内在质量的各种要求加以整理后，认为在这些检测都符合的情况下，测量的构念效度即可获得确保。但是对此说法应审慎，因为理想的测量模型的内在拟合或许可以提供聚敛与区分效度的证据，但非构念效度的充分条件。

第五节 验证性因素分析

二、测量模型的内部拟合检验

4. 因素区辨力

在心理计量领域，效度是一个建构的过程，而非可以由单一统计量数来充分支持。以上述程序来检验所获得的证据，或许可以作为测量模型的质量评估的部分证据与参考，但是要能作为测量工具反映构念效度的充分证据，还有一段距离。

更进一步地，测量工具的构念效度无法借由待测量表本身来证实。换言之，抽象特质是否能被正确地测量的计量证据必须超越研究者所关心的测量工具本身，以其他相近的测量工具来求得合理的关系，或以实验手段来证实测量分数的有效性。

另一个值得注意的问题是，Hair等人所谓的聚敛效度与传统的定义有所不同。Campbell与Fiske (1959)认为，聚敛效度是指以“不同方法”来测量相同特质的相关要高于所有的相关。此时，不同方法是指不同的量表或不同的测量方式（例如，自评与他评），而非一个分量表当中的不同题目（不同题目测量相同特质时的相关或因素载荷反映的是测量信度）。进一步地，区分效度的达成是不同方法或相同方法测量不同特质的相关高于不同方法测量相同特质的相关，此观点也无法单纯从单一量表的检验中获得。

第五节 验证性因素分析

三、LISREL的验证性因素分析

1. 验证性因素分析的操作步骤

① 发展假设模型。也就是针对测量的题目的潜在结构关系，基于特定的理论基础或是先期的假设，提出一个有待检验的因素结构模型。用SEM的术语来说，就是要建立一套假设的测量模型。

② 进行模型的识别，也就是将研究者所欲检验的测量模型转换成符合SEM分析的模型，以便利用统计软件来进行分析。

③ 执行SEM分析，进行参数估计与模型检验。

④ 进行结果分析，也就是分析SEM分析的报表结果，检验各项数据的正确性。

⑤ 模型的修正，以获得较佳的数据与结果。

⑥ 完成SEM分析，并做出报告。

第五节 验证性因素

三、LISREL的验证性因素分析

2. 验证性因素分析的操作

①模型界定

该量表题目的编写是根据研究者执行的前导研究所发现的影响组织创新气氛知觉的因素，包括“组织价值”“工作方式”“团队合作”“领导风格”“学习成长”“环境气氛”等六个因素。研究者针对这些因素编写题目，发展出评定量表。在本范例中，每一个因素仅取出3个题目作为代表，因此共有18个题目（代号A1~E3）。受测者在这些题目上的得分越高，代表所知觉到的组织气氛越有利于组织成员的创新表现。由于部分成员在部分题目上表示无法作答，所以实际应用于SEM分析的样本为完全作答的350位受测者。

表 5.2 组织创新气氛量表 18 题短版本的描述统计量

题目	平均数	标准差
1. 我们公司重视人力资产、鼓励创新思考。	4.42	0.98
2. 我们公司下情上达、意见交流沟通顺畅。	4.31	1.02
3. 我们公司能够提供诱因鼓励创新的构想。	4.07	0.97
4. 当我有需要，我可以不受干扰地独立工作。	4.02	1.16
5. 我的工作内容有我可以自由发挥与挥洒的空间。	4.25	1.16
6. 我可以自由地设定我的工作目标与进度。	4.24	1.09
7. 我的工作伙伴与团队成员具有良好的共识。	4.37	0.98
8. 我的工作伙伴与团队成员能够相互支持与协助。	4.34	1.03
9. 我的工作伙伴与团队成员能以沟通协调来化解问题与冲突。	4.31	1.05
10. 我的上司主管能够尊重与支持我在工作上的创意。	4.83	0.94
11. 我的上司主管拥有良好的沟通协调能力。	4.95	0.84
12. 我的上司主管能够信任下属、适当地授权。	4.83	0.91
13. 我的公司提供充分的进修机会、鼓励参与学习活动。	4.63	0.97
14. 人员的教育训练是我们公司的重要工作。	4.73	1.01
15. 我的公司重视信息收集与新知的获得与交流。	4.70	0.98
16. 我的工作空间气氛和谐良好，令人心情愉快。	4.23	1.17
17. 我有一个舒适自由、令我感到满意的工作空间。	4.63	1.09
18. 我的工作环境可以使我更有创意的灵感与启发。	4.49	0.94

第五节 验证性因素分析

三、LISREL的验证性因素分析

2. 验证性因素分析的操作

①模型界定

这个18题的评定量表基于研究者所提出的先期结构（六因素测量模型），6个因素与18个测量变量的关系可以利用图5.4的假设模型来表示。对组织气氛的知觉是一个多维度的个人知觉现象，以18个测量变量来测量受测者的知觉强度，同时不同的知觉维度（6个因素）之间可能具有相关。根据这些条件，对模型界定情形说明如下，并以路径图标示各参数（见图5.4）。

1) 模型中有18个测量变量（X1~X18被标签为A1~E3）与6个潜在变量（ζ_1~ζ_6）。

2) 模型中有18个测量残差（δ_1~δ_{18}），其变异量被自由估计。

3) 为了使6个潜在变量（因素）的量尺得以确立，每一个因素的方差被设定为1.00。

4) 每一个测量变量仅受单一潜在变量影响（单维假设），产生18个因素载荷参数（λ_1~λ_{18}）。

5) 因素共变允许自由估计，产生15个相关系数参数（ϕ_{21}~ϕ_{65}）。

6) 测量残差被视为彼此独立的，没有共变。

第五节 验证性因素分析

三、LISREL的验证性因素分析

2. 验证性因素分析的操作

②代码

LISREL、MPLUS、SIMPLIS、AMOS、R代码：略

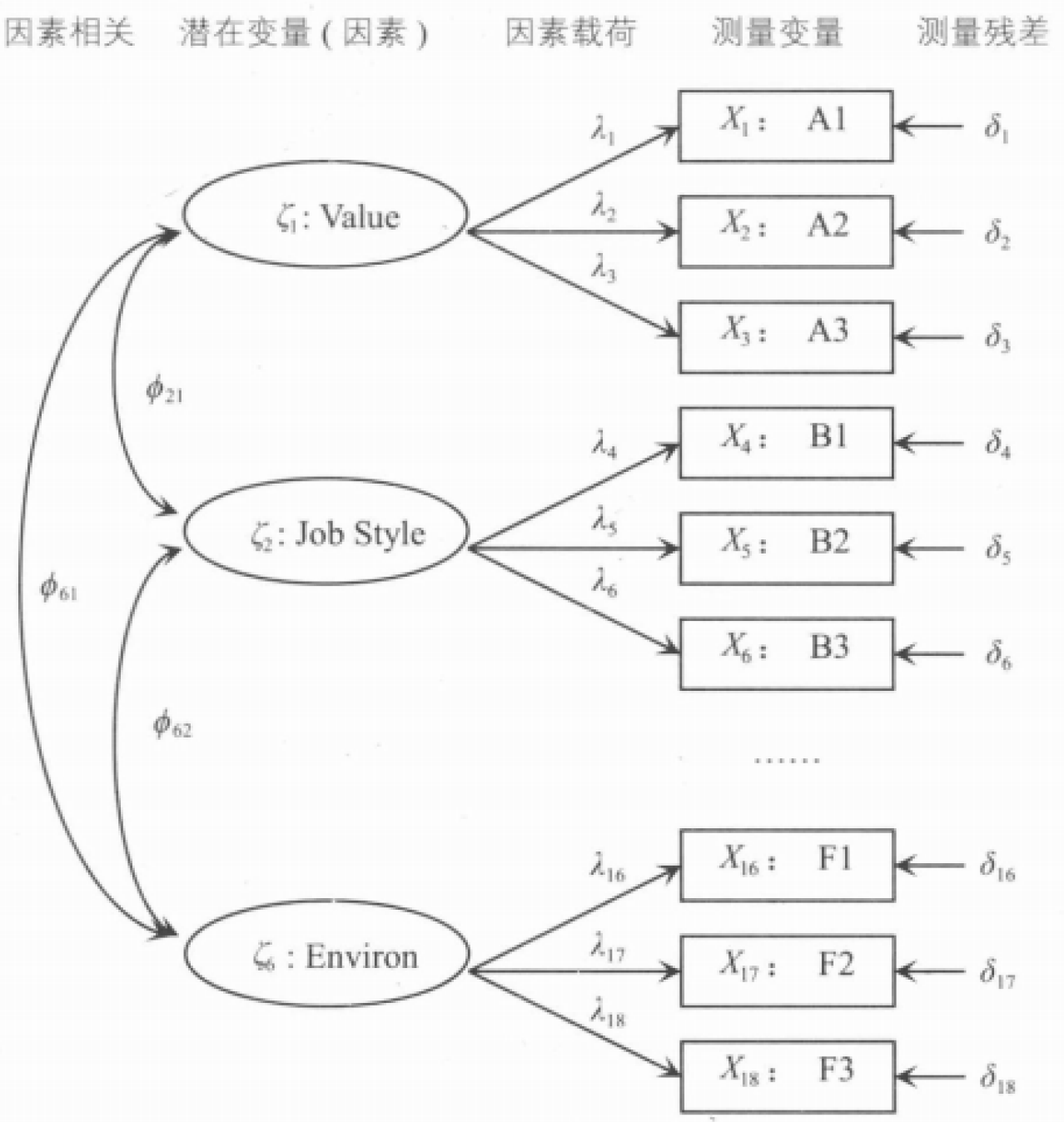
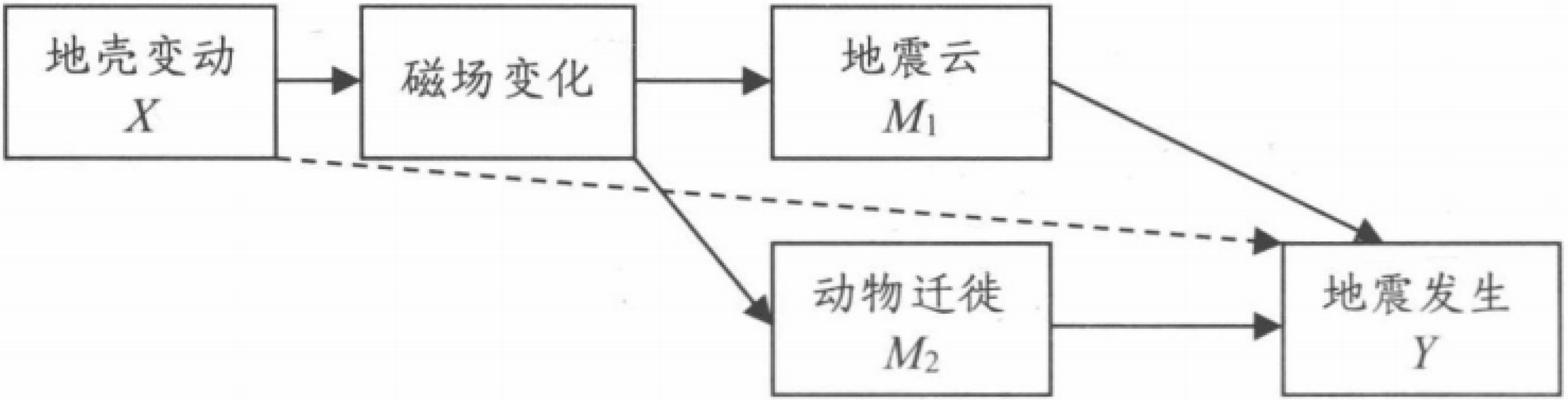


图 5.4 CFA 模型各参数路径图示

第六节 路径分析

一、路径分析的基本概念

1921年，Wright首度提出了“路径分析”的概念，利用回归方程来联结图中的X（自变量，independent variable，IV）、M（中介变量，mediator，Me）和Y（因变量，dependent variable，DV），建立一个统计模型来解释一组有关联的变量之间的因果关系，以推得经济上的因果结论。



第六节 路径分析

二、路径分析的模型界定与识别

1. 理论先行

路径分析的目的是建立因果解释模型，这也是社会科学领域用来检测因果模型的重要策略。但是因为因果关系所倚赖的基本前提多半难以确立，所以路径分析的结论也经常遭到质疑。

另一个重要的问题是如何排除与控制其他变量的影响，这使得因果关系的存在具有相当的稳定性与内部关系的有效性。在传统上，实验研究的优点即是可以精确地控制其他变量，借以观察X变量对于Y变量的影响。但是多数的社会与行为科学研究所探讨的变量数目多、关系复杂，许多变量因为基本性质或伦理考虑无法在实验室中操作，使得实验研究无法执行，大量使用统计控制程序的回归分析、路径分析或SEM即成为复杂的共变关系分析的重要替代方案，但同时这也是这类技术共同的问题根源。统计控制的程序往往因为不同的处理方式造成了数据的变异，在分析过程当中常常因为控制变量的调整造成了估计结果的改变，甚至扭转了研究的结论。因此，要从技术层面来追求一个稳定的、具有统计检验力（power）的参数估计程序的路径分析来达成对真相的发掘，往往收效有限。根本的解决之道仍是建立适切的理论基础与严谨的假设建构过程，并时时注意统计技术本身的限制与问题。

第六节 路径分析

二、路径分析的模型界定与识别

2. 模型的建立

路径分析的首要步骤是建立一个有待检验的路径模型（path model）。模型的建立除了基于研究者所关心的变量与关系，也必须斟酌相关研究与过去文献的观点，提出一个适当的模型以待验证。但是要如何选择适当的变量与安排适当的假设关系，是路径模型建立的两大挑战。以变量的选择来说，某一个因变量Y（例如学业表现）的影响因素除了研究者所关心的变量（例如动机因素）之外，可能存在其他重要的影响变量（例如，智力）。

从技术的观点来看，纳入越多的对于因变量具有解释力的自变量，可以有效提升模型的拟合度，但是从研究的观点，纳入过多的变量对于现象的解释不但没有帮助，反而容易造成概念上的混淆。因此，研究者无不希望提出一个精简的模型而能解释最大的变异，太多的讯息反而会造成判断的困扰。相反的，有时候研究者对于哪些变量对因变量具有影响力缺乏足够的讯息，尤其对于一些较为冷门或新兴的议题，过去的文献与既有的理论可能相当缺乏，造成研究者不知如何选择变量，或是无法提出有力的论证以支持模型的意义。过多或过少的信息都对模型的建立造成了困扰，最后的解决方案还有赖研究者自行通过归纳、推理与主观的分析来决定。

第六节 路径分析

二、路径分析的模型界定与识别

2. 模型的建立

其次是变量关系的决定，下表列举了路径模型中可能存在的变量关系，除了↔代表没有方向性的相关之外，其他的关系类型都与因果关系有关。当研究者获得观察数据之后，任何两个变量的共变都可以直接计算出来，所以相关可以说是具体存在而可以估算的变量关系，是整个路径分析的基础。其他所有关系的检验则建立在研究者的假设之上，需要适当的理论依据作为基础。

变量与符号	代表意义	关系类型
$X \leftrightarrow Y$	相关 Correlation	X 与 Y 为共变关系
$X \rightarrow Y$	单向因果关系 direct causal effect	X 对 Y 为直接效应
$X \rightarrow Y_1 \rightarrow Y_2$	单向因果关系 direct causal effect	X 对 Y_1 为直接效应，X 对 Y_2 为间接效应，Y_1 为中介变量
$X \rightleftarrows Y$	回溯因果关系 reciprocal causal effect	X 与 Y 互为直接效应，X 与 Y 具有回馈循环效果
$Y_1 \rightarrow Y_2 \rightarrow Y_3 \rightarrow Y_1$	循环因果关系 indirect loop effect	Y_1 对 Y_2、Y_2 对 Y_3、Y_3 对 Y_1 均为直接效应，Y_1、Y_2 与 Y_3 为间接回馈循环效果

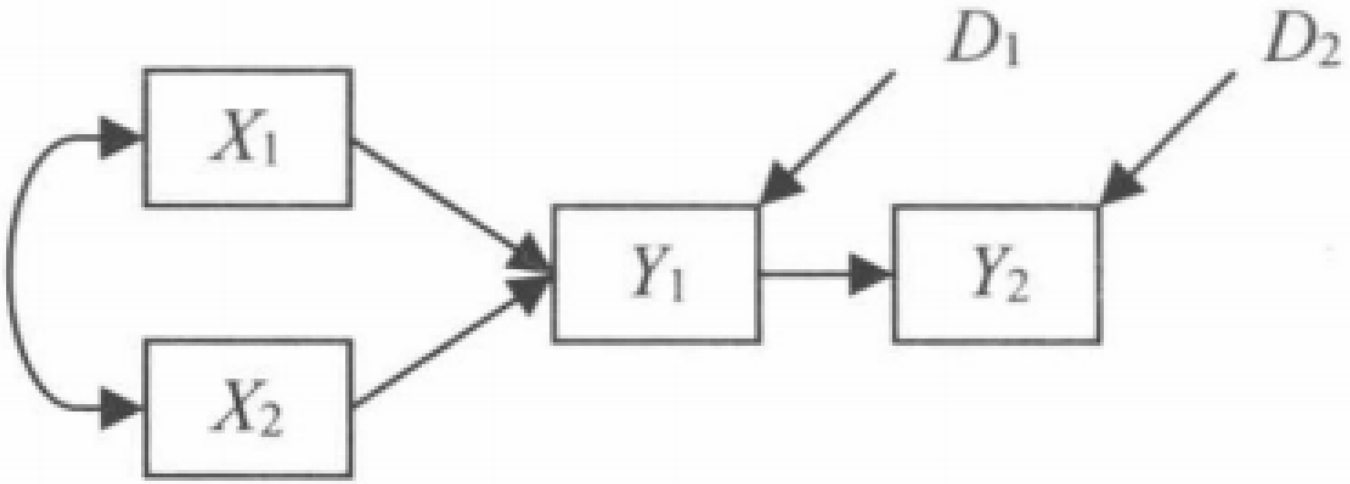
第六节 路径分析

二、路径分析的模型界定与识别

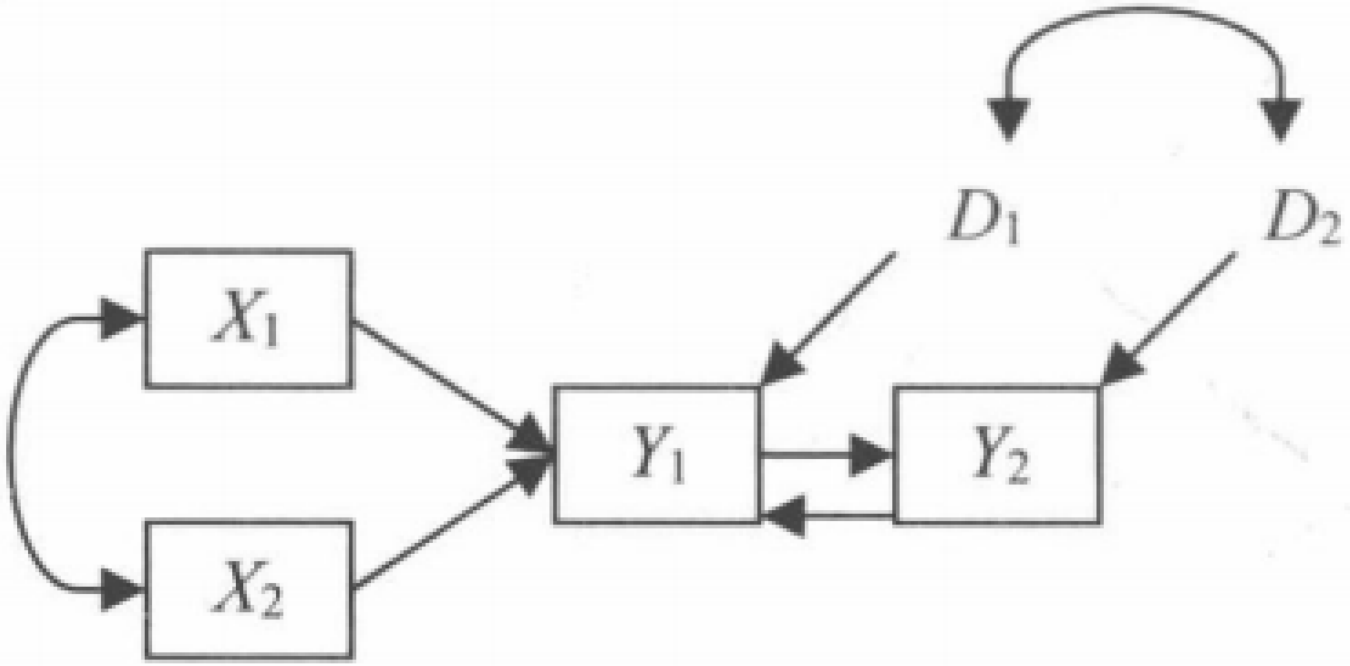
3. 递归模型与非递归模型

①模型的定义

借由变量关系的安排，路径分析有两种不同的基本类型：递归（不可逆）模型与非递归（可逆）模型（见图）。这两种模型的区别主要在于是否具有回溯性或循环因果关系（见上一页的表）。



（a）递归模型



（b）非递归模型

第六节 路径分析

二、路径分析的模型界定与识别

3. 递归模型与非递归模型

①模型的定义

在回归分析中，残差方差是因变量不能被自变量解释的部分，计算方法是以1减去 R^2 再乘以因变量的方差，得到非标准化残差方差（若不乘以内生变量方差则是标准化的方差）：

$$Var_D = \left(1 - R^2\right) \times S_y^2$$

值得注意的是，虽然Y1与Y2一定会被其他变量解释，但这里所谓的“其他变量”除了指外源变量（X1与X2）之外，也有可能是内生变量。换言之，Y1除了被X1与X2解释之外，也可以去解释Y2；同样的，Y2除了被X1与X2解释之外，也可以去解释Y1。然而，为了维系模型的稳定性与可解释性，Y1若作为Y2的解释变量，不宜“同时”作为被Y2解释的内生变量，此时即是路径分析所谓的递归模型（如图a所示）。在递归模型中，任何自变量与因变量的配对是不可逆的。

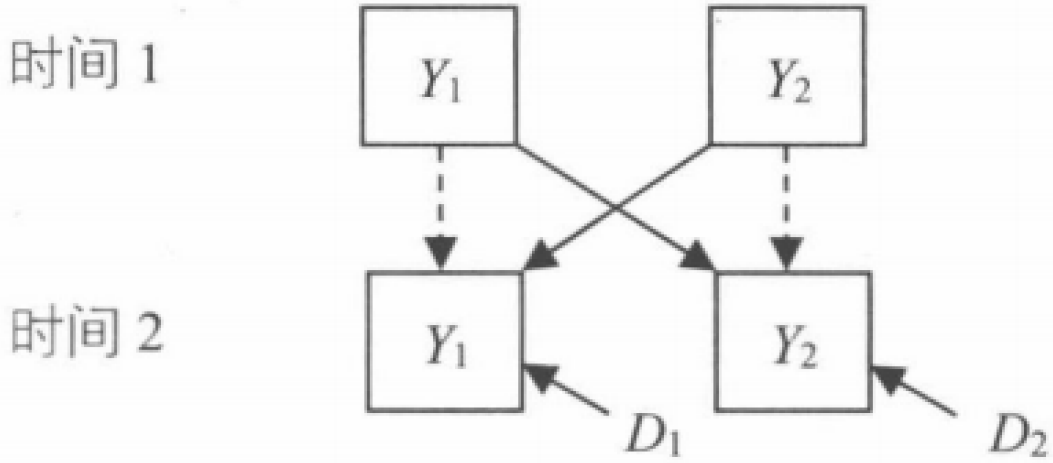
有时，有些研究无可避免地会碰到Y1作为Y2的解释变量，Y2又返回作为Y1的解释变量的情况（例如，经济学中的供给与需求之间的关系，或是管理学中的成就与动机之间的关系）。此时即会形成路径分析中所谓的非递归模型（如图b所示）。在非递归模型中，自变量与因变量的配对是可逆的。

第六节 路径分析

二、路径分析的模型界定与识别

3. 递归模型与非递归模型

②延宕模型



在具有时间落差的纵贯性或重复测量（repeated measures）的研究中，因为存在有限因果性延宕（finite casual lag），两个变量互为因果的假设模型比较可能被接受。因为互为因果的两个变量在经过一段时间的延宕后成为彼此的果与因，两个时间点的距离不论长短都提供了“因”“果”两变量成为他人的“果”“因”所需的时间落差，如图中的实线所示。

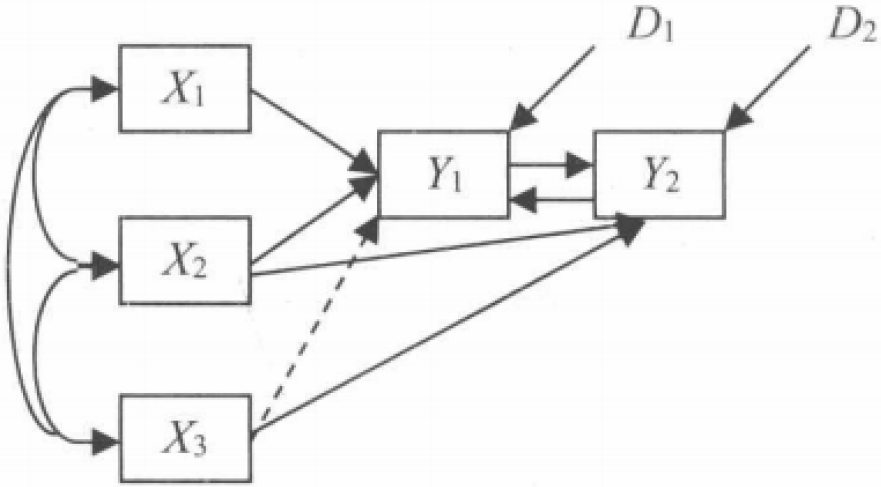
Maruyama (1997)指出，可逆模型经常会有无法收敛的问题，即因为干扰项的关系太强，参数不容易得到唯一解，甚至因为低度识别而无解。从前2页图b也可以看出，干扰项D1与D2如果不设定相关，也存在一种可能的影响关系：D1→Y1→Y2与D2→Y2→Y1，造成参数估计的不稳定。相对而言，不可逆的递归模型则容易被有效识别、获得终解，亦具有解释明确与理论清晰的优势。

第六节 路径分析

二、路径分析的模型界定与识别

3. 递归模型与非递归模型

③工具变量模型



Maruyama (1997)曾介绍一种如何在非递归模型中借由纳入工具变量（instrumental variable）来协助模型得以识别的做法。也就是纳入一个变量（工具）来解释某些内生变量但不影响其他内生变量，来协助解释内生变量间的复杂关系。

如前所述，互为因果的回溯变量关系在概念上不易解释，在统计上更有识别不足而无法得解的窘境。为了让模型得以识别，可以增加一些条件来协助参数的有效估计。例如，在模型中增加变量来解释互为因果的两个变量。

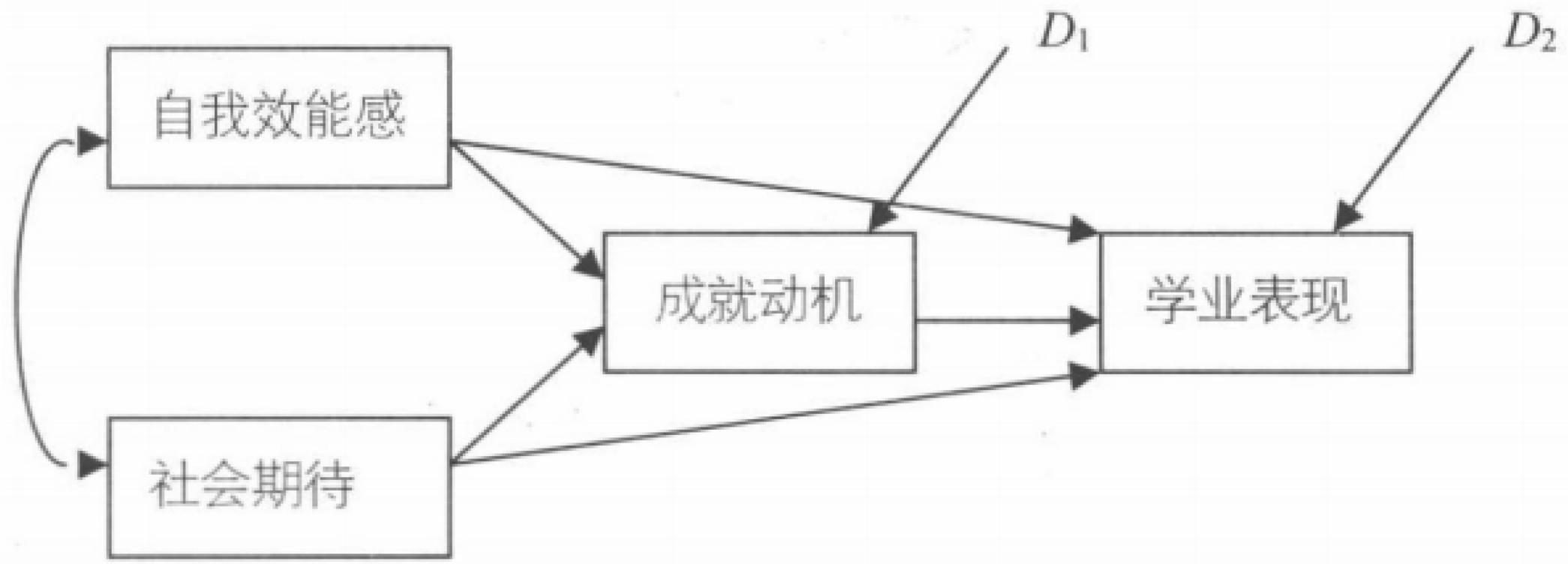
以上图为例，模型中外源变量共有X1、X2与X3，互为因果的变量仍为Y1与Y2，此时共有5个变量，共能产生(5+6)/2=15个数据点。模型中有12个待估计参数（包含实线与虚线以及D1与D2），因此模型得以识别。

第六节 路径分析

二、路径分析的模型界定与识别

4. 路径图与结构方程

现以学生学业表现的理论模型建构为例，说明一个假设性的路径模型。如下图所示，研究者认为学生的求学动机会影响其学业表现，在参酌相关文献后，认为学生受到他人影响的因素（社会期待）与个人对自己能力的判断（自我效能感）构成动机的基础，进而影响学业表现。整个模型可用下图的路径图来描述。



第六节 路径分析

二、路径分析的模型界定与识别

5. 直接效应与间接效应

经由上述的描述，变量之间的关系得以结构的方式呈现，每一个独立的回归方程都可用回归分析的原理与技术进行分析。对于每一个回归模型，自变量对于因变量的解释力可以由R2及F检验值来表示，而每一个箭头则可以获自回归系数，无论是显著还是不显著的回归系数，均需填注于路径模型箭头两侧。

箭头的回归系数若达显著，代表该因果变量间具有直接效应（direct effect），未显著的箭头回归系数则代表无直接效应存在。

两个变量之间除了可能具有直接效应，亦可能存在间接效应（indirect effect），也就是说，两个变量间具有一个或多个中介变量（mediated variable），变量与变量之间的直接效应均为显著，若有任何一个直接效应不显著，间接效应无法成立。【这是经典的三步法的论证，目前尽管已被推翻，但是仍有大量研究者使用】

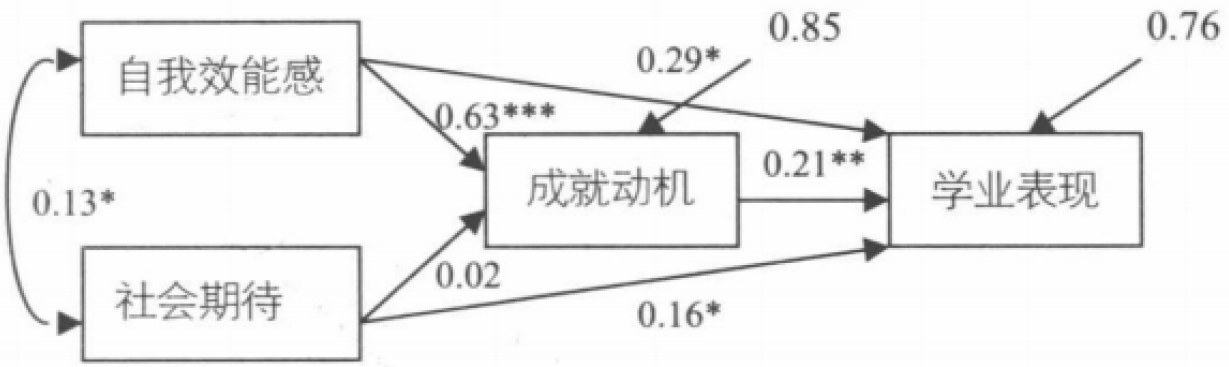
第六节 路径分析

二、路径分析的模型界定与识别

5. 直接效应与间接效应

①效应分解

注意！根据更新的规则，社会期待→成就动机→学业表现的路径中，尽管社会期待→成就动机的系数不显著，间接效应也可能存在，应该直接检验0.02×0.21是否显著!



	内生变量	
自变量	成就动机	学业成绩
自我效能感		
直接效应	0.63***	0.29**
间接效应	-	0.13*
整体效应	0.63***	0.42**
社会期待		
直接效应	0.02	0.16*
间接效应	-	0.00
整体效应	0.02	0.16*
成就动机		
直接效应		0.21**
间接效应		-
整体效应		0.21**

注：* 表示 $p<0.05$；** 表示 $p<0.01$；*** 表示 $p<0.001$。

第六节 路径分析

二、路径分析的模型界定与识别

5. 直接效应与间接效应

②模型衍生相关与模型拟合

从统计的角度来看，任何两变量之间都可以计算出一个协方差，反映两个变量的关系，称为观察相关（observed correlation），但表所列出的是经由路径模型推导并估计所得到的回归效果。Pedhazur (1997)详细说明了在路径分析中，如何将间接效应和直接效应加以合并，并纳入未被分析的拟似相关（spurious correlation，例如上一页图当中的自我效能感与社会期待的相关系数），计算出路径模型中任意两变量的模型衍生相关（model-implied or predicated correlation）。计算的原理称为轨迹法则（tracing rule）。

所谓的“轨迹”主要是寻找由于相关系数所导致的间接效应。每一个自变量对于内生变量的模型衍生相关是由先前所求得的整体效应再加上尚未被计算的相关导致的间接效应。以上一页的图为例，自我效能感对于学业成绩的模型衍生相关共牵涉四条轨迹：

1) 直接效应：自我效能感→学业表现=0.29

2) 间接效应：自我效能感→成就动机→学业表现=0.13

3) 相关间接效应（拟似效应）I：自我效能感↔社会期待→学业表现=0.13×0.16=0.02

4) 相关间接效应（拟似效应）II：自我效能感↔社会期待→成就动机→学业表现=0.13×0.02×0.21=0.00

因此，自我效能感对于学业成绩的模型衍生相关如下：0.29+0.13+0.02+0.00=0.44

第六节 路径分析

二、路径分析的模型界定与识别

5. 直接效应与间接效应

③间接效应与路径系数的检验

如果一个路径模型当中只有三个变量X、Me、Y，此时X→Me→Y的间接效应其实就是中介效应。换言之，X对Y的影响力除了可以从直接效应来看，更重要的是通过Me的中介效应来解释，甚至只有当X→Y的效应从原来不考虑Me的情况下为显著，而考虑Me之后变成不显著时，才被称为完全中介效应（两个变量间只有间接效应而无直接效应）(Baron & Kenny, 1986)。显然，间接效应可以说是路径分析中最重要的焦点。

间接效应的强度简单来说就是取X→Me→Y的两个标准化回归系数相乘，又称为路径系数 (Wright, 1960)。间接效应的强度是否具有统计的显著性，可以利用Sobel (1982)所推导的标准误来计算t检验值。但由于两个正态化的回归系数相乘后并不服从正态分布（呈现峰度为6的高狭峰分布），如果变量的平均数不为零，还有偏态问题，因此采用Sobel (1982)的公式导出的标准误为偏估计值（biased estimator）。Sampson与Breuning (1971)导出了路径系数的不偏估计数，得以进行路径系数的显著性检验。

虽然标准误的公式相继被提出，但经过模拟研究发现，Sobel (1982)所提出的公式仍是效率最佳的路径系数标准误 (Mackinnon, 2008)，这也是为何多数的SEM软件（例如，LISREL、EQS、Mplus）仍以Sobel (1982)作为间接效应的显著性检验方法。

第六节 路径分析

二、路径分析的模型界定与识别

5. 直接效应与间接效应

③间接效应与路径系数的检验

除了显著性检验之外，间接效应的强度也可以利用区间估计来描述，并得以进行不同的中介效应的相互比较。

近年来，由于统计仿真技术的进步与计算机指令周期的提升，对于间接效果标准误的估计得以利用重复取样技术来建立参数分布，求得参数的拔靴标准误（bootstrapping standard error），用以建立0.95置信区间（0.95CI）。

除了拔靴置信区间（0.95CI），贝叶斯估计标准误也逐渐受到重视，并认为可以取代拔靴标准误来检验间接效果的统计意义。事实上，贝叶斯估计法所建立的0.95CI也如同拔靴法的估计程序，利用重抽技术来反复取样借以获得参数分布。贝叶斯方法的不同之处，是基于先验分布信息的导入，结合实际样本的重抽分布，两者加以整合之后所得到的参数后验分布的标准偏差即贝叶斯标准误，进而建立贝叶斯置信区间（Bayesian-based credibility interval）。在Mplus软件中提供贝叶斯估计功能，可以便捷地执行间接效果的贝叶斯估计。

第六节 路径分析

二、路径分析的模型界定与识别

6. 结构方程模型的路径分析

先前提及第三变量将使X→Y的关系产生相当复杂的变量，如果一个模型存在第四或第五个变量，那么整个模型就相当复杂了。而路径分析就是一种利用回归模型来整合一连串复杂的变量关系所形成的完整模型。模型中所有变量的关系虽然未必全然是中介关系，也可能带有调节变量与非线性关系，但在本节的架构下，主要是以中介模型来分析变量关系。

其次，一般当研究者提及路径分析一词时，主要是指一组外显观察变量的关系，而不涉及潜在变量的定义与分析。近年来，由于SEM盛行，路径分析一词也逐渐与SEM当中的结构模型分析混淆并用。一般而言，当一个模型称为SEM时，是指模型中除了带有诸多变量间的因果解释关系之外，各研究变量还牵涉潜在变量的定义，因而有学者将其称为带有潜在变量的SEM（structural equation modeling with latent variables）。至于路径分析则是指仅包含外显变量的多重因变量的复杂回归模型，不涉及潜在变量的定义。近来，由于SEM的软件逐渐普及以及以SEM软件进行路径分析有诸多优点，因此绝大多数的路径分析已采用SEM软件以完整的共变结构来进行分析，为了有别于带有潜在变量的路径分析一词，Tabachnica与Fidell (2007)将其称为带有观察变量的SEM（structural equation modeling with observed variables）。

第六节 路径分析

二、路径分析的模型界定与识别

6. 结构方程模型的路径分析

以SEM来进行路径分析可以说是SEM最重要的应用价值之一。如果单从取代性来看，SEM的分析软件与技术原理可以完全取代过去路径分析的分析任务，更可以超越过去路径分析只能以外显变量作为分析变量的限制，通过统合模型（hybrid modeling）分析的运用，巧妙地将因素分析的概念与技术融合到路径分析的检验中，堪称计量技术的一大革命。

在传统的路径分析中，所处理的变量若以SEM的术语来表示，都应该以方块来表示，其性质属于测量变量，而无椭圆形的潜在变量。也就是说，路径分析当中的变量是假设没有测量误差的，即使误差明显（例如，信度很低），在路径分析过程中，变量的误差也将被忽略。SEM取向的路径分析可以说是SEM的一种应用特例，也就是没有包含任何潜在变量的结构模型分析。在SEM路径图中，变量皆以方块来表示。联结后方程通式如下：

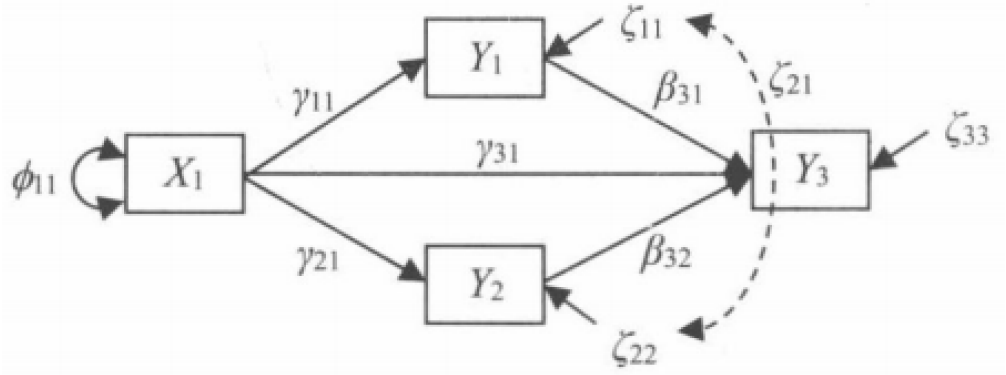
$$Y = \alpha + BY + \Gamma X + \zeta$$

其中，B为内生变量之间的回归系数矩阵，Γ为外源变量与内生变量间的回归系数矩阵，α为截距，ζ为回归残差。当SEM分析共变或相关矩阵的关系时，α为0，因此被忽略。此外，外源变量间的相关并没有反映在上述通式当中，但会在Φ矩阵中加以估计。

第六节 路径分析

二、路径分析的模型界定与识别

6. 结构方程模型的路径分析



以双中介变量模型为例，SEM取向的共变结构分析的参数如图所示。以结构方程来表示如下：

$$Y_1 = \gamma_{11}X_1 + \zeta_{11}$$
$$Y_2 = \gamma_{21}X_1 + \zeta_{22}$$
$$Y_3 = \gamma_{31}X_1 + \beta_{31}Y_1 + \beta_{32}Y_2 + \zeta_{33}$$

以矩阵表示如下：

$$\begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ \beta_{31} & \beta_{32} & 0 \end{bmatrix} \begin{bmatrix} Y_1 \\ Y_2 \\ Y_3 \end{bmatrix} + \begin{bmatrix} \gamma_{11} \\ \gamma_{21} \\ \gamma_{31} \end{bmatrix} [X_1] + \begin{bmatrix} \zeta_{11} \\ \zeta_{22} \\ \zeta_{33} \end{bmatrix}$$

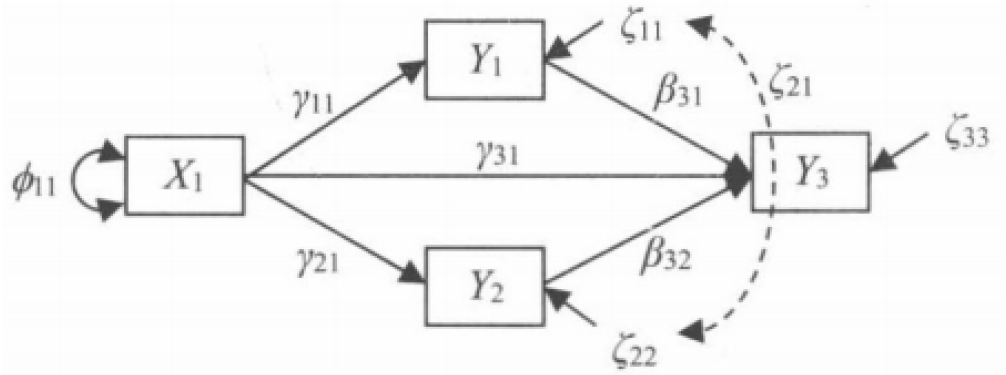
由于两个中介变量（Y1与Y2）之间可能存在相关，因此图中的解释残差ζ_{11}与ζ_{22}之间可以设定一个自由估计的参数ζ_{12}（虚线所示）。此时，残差变异矩阵（Ψ）与外源变量方差矩阵（Φ）如下：

$$\psi = \begin{bmatrix} \psi_{11} & & \\ \psi_{21} & \psi_{22} & \\ 0 & 0 & \psi_{33} \end{bmatrix}; \quad \Phi = [\phi_{11}]$$

第六节 路径分析

二、路径分析的模型界定与识别

6. 结构方程模型的路径分析



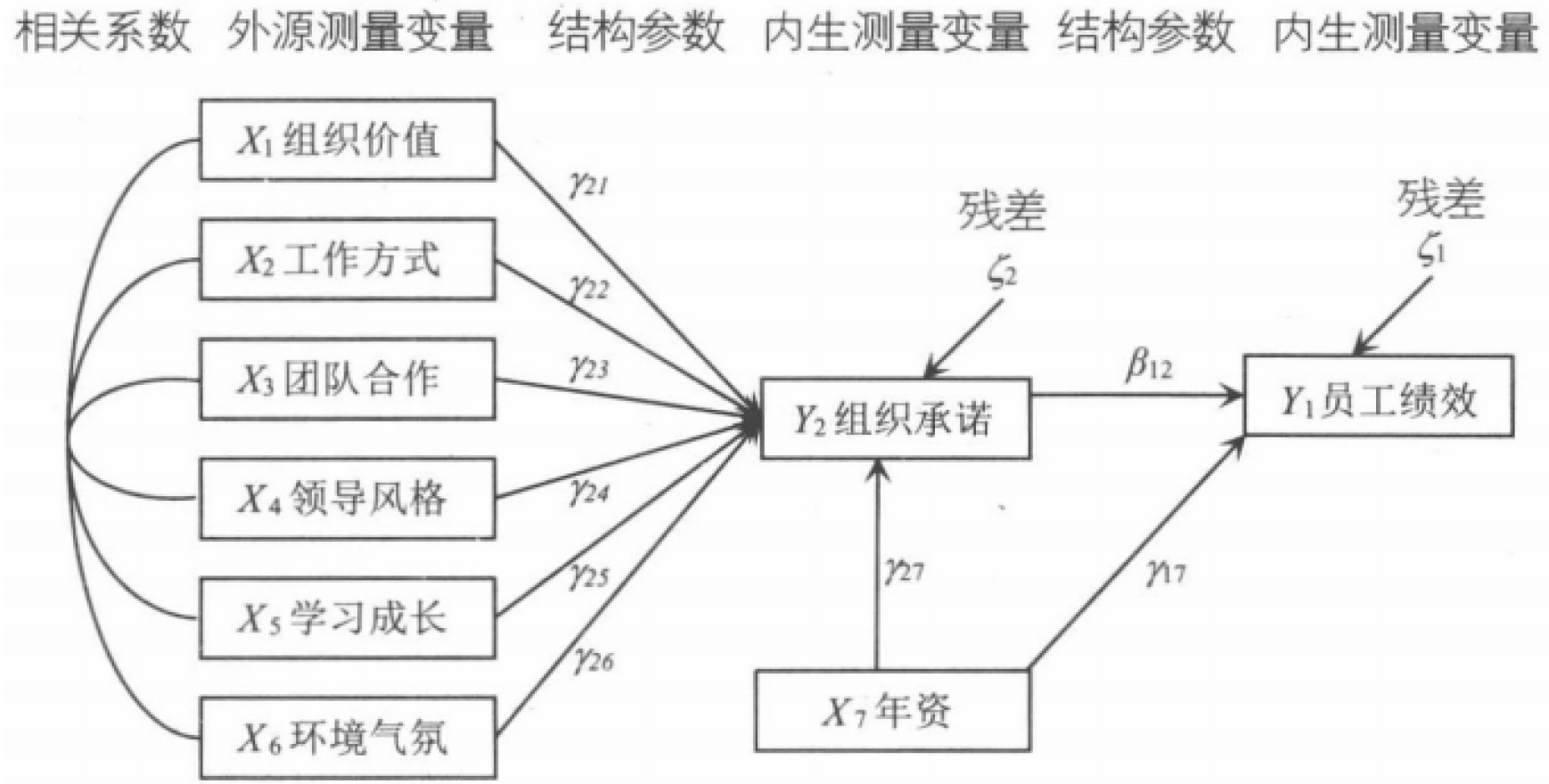
此模型的待估参数数目为10，数据点为(4+5)/2=10，两者相同，为一个充分识别的饱和模型，df=0，$\chi^2 = 0$，参数有唯一解，估计值的数据与传统普通最小二乘法回归分析的结果相近。由于模型拟合指数为0，因此无法进行模型优劣的检验（除非移除模型中的某些待估参数，例如ζ_{12}或γ_{31}。此时，路径模型背后的理论观点就略有不同了）。

如果图7.7再增加任何一个参数（例如，ζ_{23}），将造成识别不足的问题，模型无法收敛估计。换言之，只有在模型自由度大于0的非饱和模型（non-saturated model）中，拟合指数才得以计算。在模型可识别的情况下（自由度大于0），研究者可采用模型竞争策略，借由增减参数来比较不同的嵌套模型（nested model；参数少的较小模型嵌套在参数较多的较大模型当中），此时模型的优劣可以利用两个模型的卡方差异量（$\Delta\chi_2$）是否达到自由度为两个模型自由度差异量（Δdf）下的统计显著水平，来检验两者拟合是否有别（额外增加的参数是否能够改善模型拟合性），称为卡方差异检验（chi-square difference test）。

第六节 路径分析

三、LISREL的路径分析

本研究的主要假设是组织气氛的知觉影响员工的承诺感，进而影响员工的工作表现。此时组织承诺扮演中介变量的角色，年资变量则是个干扰变量或控制变量，其角色不但影响了组织只承诺，也影响工作产出，它调节了承诺感与员工绩效的预测力。路径图如图所示。



第六节 路径分析

三、LISREL的路径分析

图中说明了研究变量的假设关系，也列出了路径分析的主要参数。当以组织承诺为因变量时（Y1），6个组织气氛变量（X1~X6）与年资变量（X7）为自变量，当以员工绩效为因变量（Y2）时，仅有年资与组织承诺为自变量，整个模型的回归方程如下：

$$Y1=b1x1+b2x2+b3x3+b4x4+b5x5+b6x6+b7x7+a1$$

$$Y2=b7X7+b8X8+a2$$

在方程中，b与a分别为回归系数（斜率）与截距。在SEM的路径分析中，截距设定为0，回归系数则视其性质属于不同的参数矩阵。在本范例中，整个模型的观察变量共有9个，因此测量数据数为(9×10)/2=45（DP=45）。7个外源变量之间都可能具有相关，但是图中显示年资变量与其他6个组织气氛的测量并未假设具有相关，因此在模型界定时，需予以说明。

模型中有0个测量残差，但两个内生变量有2个解释残差（ζ_1和ζ_2），其变异量被自由估计。第一个内生变量（员工绩效）被2个外源变量解释，故B矩阵中有2个结构参数（β1、β2）。第二个内生变量（组织承诺）被7个外源变量解释，故Γ矩阵中有7个结构参数（γ1~γ7）。6个外源变量之间的相关允许自由估计，表示Φ矩阵将产生30个相关系数。

第六节 路径分析

三、LISREL的路径分析

LISREL、SIMPLIS、MPLUS、AMOS、R代码：略。

第七节 结构方程模型：统合模型分析

一、统合模型的基本概念

1. 路径分析与因素分析模型的整合

同时带有测量模型与结构模型的SEM分析称为统合模型分析，可以说是SEM各种基本模型的整合运用，也可以视为路径分析与验证性因素分析的综合分析，又称为结构回归模型（structural regression models）。

从路径分析的原理来看，统合模型的结构化假设可以假定各种直接或间接的因果关系，并利用模型拟合度分析技术来进行参数估计，找出路径系数并检验其显著性；从因素分析的角度来看，统合模型的测量模型可以用来检测一群测量变量背后的潜在因素，验证研究者提出的假设模型，并以竞争模型的比较来寻找最佳的因素结构。

在测量的内在基础与外部的结构关系皆能兼顾的优势下，统合模型分析一举整合了路径分析与因素分析的核心概念，回避了这些传统多变量统计技术的缺失与限制，并提供各种统整性指数与统计量数，来检测各项参数与模型特性，例如，路径分析对整体效应、间接效应的估计在SEM模型中可以快速有效地完成，并提供显著性检验来检验效果的统计意义，此为SEM最有价值的贡献。

第七节 结构方程模型：统合模型分析

一、统合模型的基本概念

2. 统合模型的构成

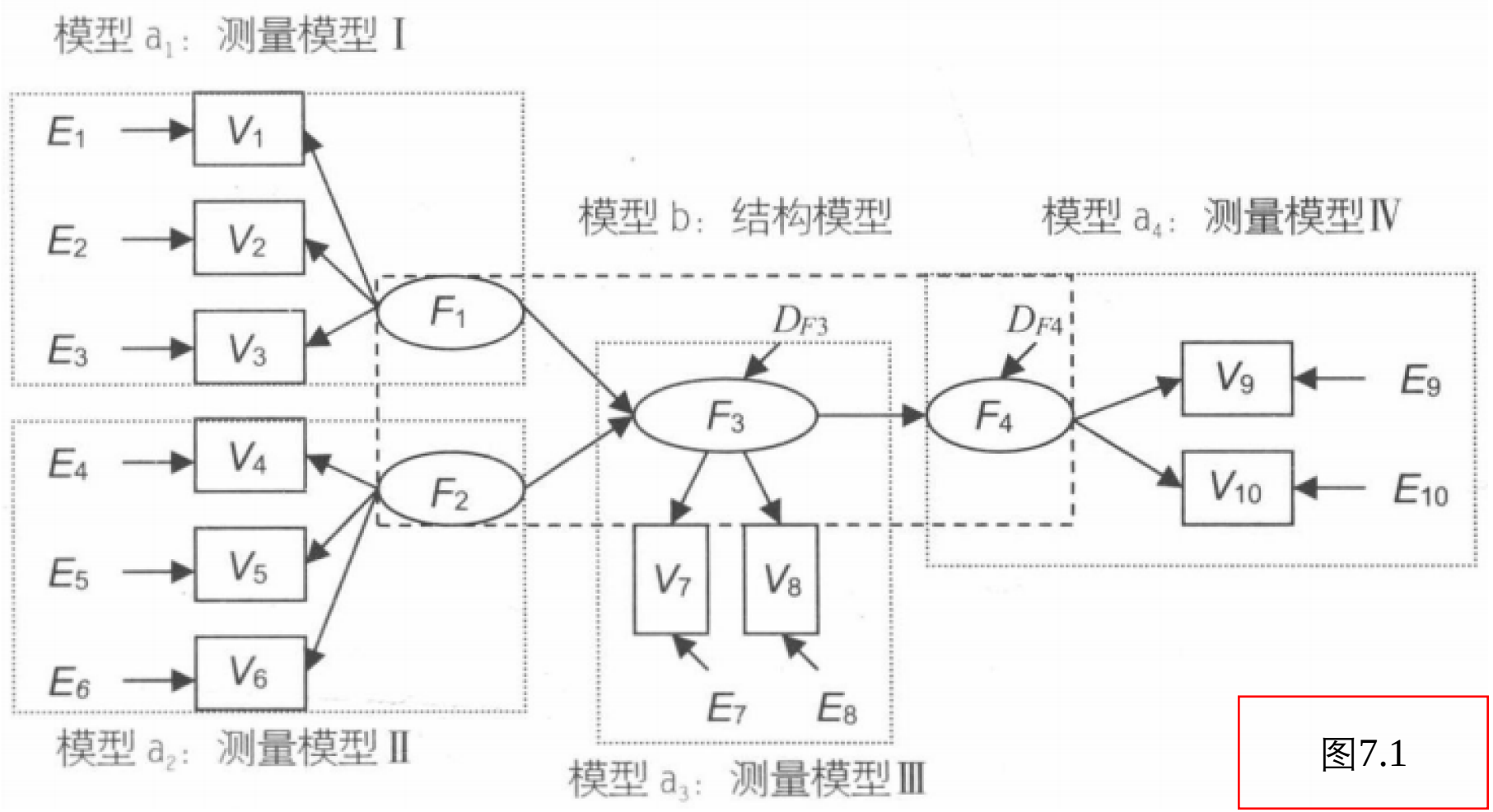


图7.1

第七节 结构方程模型：统合模型分析

一、统合模型的基本概念

2. 统合模型的构成

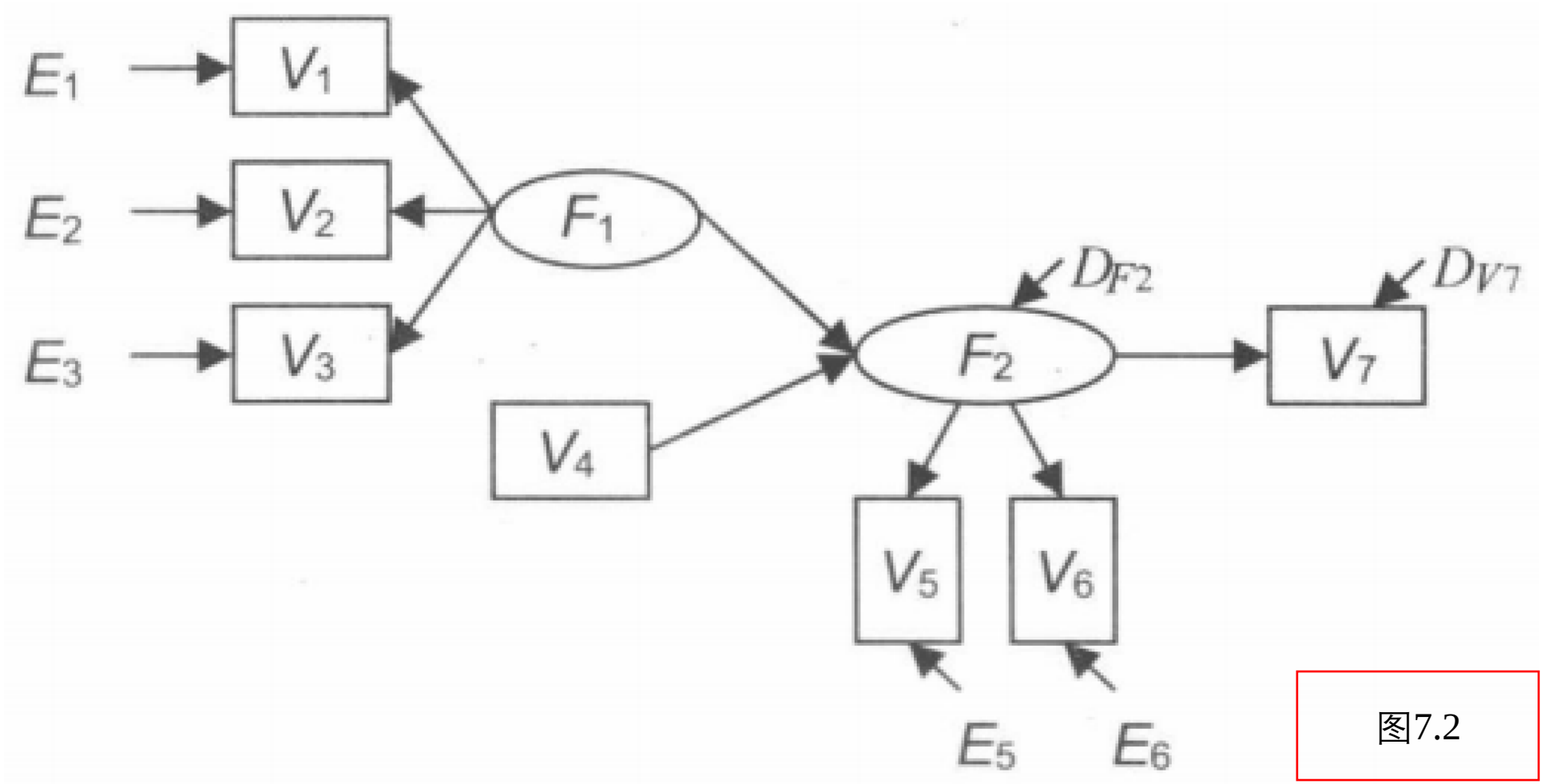


图7.2

第七节 结构方程模型：统合模型分析

一、统合模型的基本概念

3. 统合模型方程

①测量模型的方程

图7.1当中有四个独立的测量模型，每一个模型代表了一个特定的潜在变量与两个以上的观察变量之间的组合关系，每一个观察变量可以被视为一个内生变量，它的分数受到潜在变量（外源变量）的影响以及测量误差（ε）的影响，可以用一个单独的方程来表示其间的关系：

$$x = \Lambda_x \xi + \delta \quad (\text{模型} a_1 \text{与} a_2)$$
$$y = \Lambda_y \eta + \varepsilon \quad (\text{模型} a_3 \text{与} a_4)$$

在上述两个通式当中，ξ表示潜在外源变量（F1与F2），η表示潜在内生变量（F3与F4），x与y是各潜在变量所影响的观察变量，在概念上属于内生变量。也就是说，对于每一个个别的观察变量，不论是x或y，它的分数都可以被切割为由共同的潜在变量影响的部分，以及无法被解释的部分（称为测量误差，由ε与δ表示），Λ_x与Λ_y即表示潜在变量各自的因素载荷系数矩阵（matrices of factor loadings）。

以测量模型I（a1）为例，潜在变量F1与三个测量变量的模型关系由三个联立方程组成，Λ_x表示F1潜在变量所影响的三个观察变量V1、V2、V3在F1上的因素载荷，三个观察变量无法被潜在变量解释的部分为测量误差，分别为E1、E2、E3。整个模型的组成就是一个典型的单一因素的因素分析模型，其他三个测量模型的方程是类似的概念。

第七节 结构方程模型：统合模型分析

一、统合模型的基本概念

3. 统合模型方程

②结构模型的方程

上述潜在变量与相对应的测量变量的模型关系显示出了测量模型的组成，而结构模型的组成方程如下：

$$\eta = B\eta + \Gamma\xi + \zeta$$

η是m×1个潜在内生变量向量（vector of endogenous latent variables），也就是研究者所关心的潜在因变量，即图7.1中的F3与F4；ξ是k×1个潜在外源变量向量（vector of exogenous latent variables），是研究者拿来解释或预测因变量的独立变量；B是m个内生变量的m×m回归系数矩阵，Γ是m个内生变量与k个外源变量的m×k回归系数矩阵，ζ为m×1个内生变量无法被解释的干扰向量（vector of disturbance）。B与Γ两个矩阵所代表的是方程的标准化回归系数矩阵，性质等同于传统多元回归的回归系数，多个内生变量的组合关系即为传统的路径分析模型，这些回归系数也就自动成了路径系数，这些系数都是研究者最后所要报告标示出来的重要参数。

第七节 结构方程模型：统合模型分析

一、统合模型的基本概念

4. 统合模型的识别

在统合模型当中，由于合并了测量模型与结构模型，牵涉的参数相当众多且复杂，因此在识别上的处理也就必须格外的小心。

首先，在测量模型部分，为了使潜在变量与观察变量之间的关系被顺利地估计出来，不论是潜在内生变量还是潜在外源变量，在各潜在变量之下所属的各观察变量中，挑选其中一个观察变量与潜在变量的参数，也就是因素载荷（Λ_x或Λ_y）固定为1.00，以便作为其他参数的参照，并可以让潜在变量的方差自由估计。此时，由这一特定的潜在外源变量组成的测量模型就是一个典型的单一因素模型。

一旦测量模型的参照设定完成，整体模型的识别性是否足够可以利用t法则来计算。Bollen (1989)提出了一个二阶段法则，将统合模型的分析区分为两个步骤：第一步是以因素分析的方式单纯进行测量模型的参数估计；第二步则将第一步所建立的潜在因素视为观察变量，以路径分析的模型来进行结构模型分析。

第七节 结构方程模型：统合模型分析

二、统合模型的分析步骤

基本上，统合模型的执行与先前所介绍的验证性因素分析与路径分析相似，需遵循假设模型的提出、模型识别的计算、参数估计、模型拟合检验、最终解的报告等程序。但是由于统合模型的内容包括测量模型与结构模型两个部分，因此在实际估计时，必须考虑这两种模型的执行顺序。如果不区分测量模型与结构模型，而把整个SEM视为一个完整模型，进行一次估计而获得所有的参数估计数，报告一个模型拟合指数来评估模型的优劣，称为**一阶段模型（one-step modeling）**。如果模型拟合不佳，可进行模型修饰，直到模型拟合良好为止。

相对的，如果一个带有潜在变量的路径分析模型将测量模型与结构模型分成两个模型进行检测，则称为**两阶段模型（two-stage modeling）**。第一阶段确定因素结构的拟合性，第二阶段则是在不改变测量模型的前提下，增加结构模型的设定，并评估结构模型界定之后的拟合性。此种做法是目前广为学者建议的分析策略。Anderson与Gerbing (1988)主张两阶段模型对于潜在变量的效度评估具有重要的价值，因为测量模型的检测可以提供潜在变量的聚敛与区分效度的信息，结构模型则可提供预测效度的证据。

第七节 结构方程模型：统合模型分析

二、统合模型的分析步骤

Mulaik与Millsap (2000)提出了四阶段的策略，其操作方式是利用一系列的嵌套模型，借由模型竞争比较策略来评估各模型的拟合性。

阶段一：未限制测量模型（unrestricted measurement model）。这是一种探索性测量模型分析，以决定能让理论协方差矩阵与观察协方差矩阵达到最理想拟合下的潜在变量数目。（对每一个测量变量与每一个潜在变量的因素载荷均加以估计。）

阶段二：限制性测量模型（restricted measurement model）。执行研究者所决定的因素结构，确认模型拟合，并决定测量模型的最终参数估计数。（不属于某特定潜在变量的因素载荷设定为0，此举将使这一CFA模型较前一个阶段的模型拟合变差。）

阶段三：完整结构模型（structure model）。依据研究者所关心的理论假设关系估计潜在变量间的参数，无关联的结构参数设定为0。

阶段四：替代模型竞争比较（alternative model competition）。依据其他理论观点提出的结构模型，调整结构模型（包括增减结构参数，对结构参数加以设限，或是利用参数的置信区间来比较参数的差异），以比较不同结构模型的拟合性，推估最合理的结构模型。

第七节 结构方程模型：统合模型分析

二、统合模型的分析步骤

多阶段模型的优势是可以获得更丰富的模型拟合数据，借以厘清影响模型估计的原因为何。Joreskog与Sorbom (1983)以及Bollen (1989)均曾指出，测量模型的良好拟合提供了后续结构模型的重要基础，因而测量模型必须在结构模型之前进行确认。换言之，如果测量模型不佳，结构模型的分析即缺乏合理的测量基础。

从测量模型与结构模型中的参数数目也可以看出，测量模型中所带有的参数远多于结构模型，因此一个良好拟合的测量模型表示绝大部分的参数（包括因素载荷、测量残差、潜在变量的共变等）都获得了良好合宜的估计值，也确保了后续的结构模型分析的质量。

从参数的性质来看，结构模型的参数多为单方向的回归系数，不论是A→B或B→A，其实都是A↔B的一种特例，而测量模型是把潜在变量的相关（即A↔B）设定为自由估计，因此如果测量模型获得理想拟合，那么结构模型的单方向参数也应有一定的适切性。Anderson与Gerbing (1992)主张，测量模型所提供的估计是测量的计量基础，而结构模型所检测的是理论关系，因此结构模型所看重的并不是测量的合理性，而是参数的方向性背后所代表的理论合理性。

发生测量模型的参数合理但结构参数不理想的可能性之一，是当相关系数过高时，可能产生共线性问题。在多元回归当中，多元共线性问题原本就是威胁参数估计与解释的重要议题，而非测量问题。

第七节 结构方程模型：统合模型分析

三、变量组合与聚合

SEM分析经常以变量组合策略（item parceling）来简化测量模型，从而在比较简化的情形下估计结构模型。例如，当某个潜在变量有6个观察变量时，可以将6个变量聚合（aggregate）成3个变量（每两题加在一起或求平均值），来降低模型的复杂度，聚合层次的变量分数称为组合分数（parcel scores）。另外，当研究样本太少（例如低于200）而统合模型很大、所需估计的参数很多时，一种变通的做法是将潜在变量改为观察变量来处理（把构成各潜在变量的题目组合成单一变量），来进行路径分析。

1. 变量组合的方法

在SEM分析中，变量组合并没有一致的方式，但有一个基本前提是构念具有单维性。也就是说，只有测量同一个构念的测量题目才能够进行组合。其他的组合方法各有目的，研究者必须了解各种方法的特殊性与使用时机，才能发挥变量组合的优势。例如，当观察变量的正态性不理想时，变量组合的方式应能抵消各题的偏态，以提高组合后变量的正态性。此外，也有学者主张就题目的内容来组合题意相近的观察变量，使得组合后的观察变量拥有可解释的意义。

第七节 结构方程模型：统合模型分析

三、变量组合与聚合

1. 变量组合的方法

①辐射组合法

较早的一种方法是Cattell (1974)和Cattell & Bursdal (1975)所提出的辐射组合法（radial parceling），它是利用两阶段的EFA，基于观察变量在各萃取因素当中的因素载荷高低，依次将若干题目组合成组合变量，再进行变量较少的EFA，确认组合变量是否仍可以得到类似的因素结构。此法适用于多因素的因素模型（例如，对性格的测量），但此法费时费力，而且可能会把不同构念的观察题目组合到不当的群组中。

②随机分布法

最直观的方法是随机分布法（random assignment），其做法是将测量同一个构念的题目以随机方式或系统性方式加以组合。例如，一个原有9个随机排列的观察变量的构念，可将第1~3题、第4~6题、第7~9题依次求和并除以题数3形成三个带有原始测量尺度的新变量，简化成一个带有3个观察变量的测量模型。Little等人 (1999; 2002)认为这种策略在多数情况下会比依据题目内容来组合更理想，但是未必会得到效度最好的组合变量。

第七节 结构方程模型：统合模型分析

三、变量组合与聚合

1. 变量组合的方法

③项目构念平衡法

为了改善随机法的缺点，Little等人 (2002)提出了一种更能找到反映同一个构念的一组最相似的组合变量的方法，称为项目构念平衡法（item-to-construct balance approach），是利用CFA得到的因素载荷数值，将同一个构念中的题目的载荷依高低顺序将各题平均分布到各个组合变量中。例如，若要将6个题目组合成3个组合变量，那么因素载荷第1、2、3高的3个题目依次成了3个组合变量的第一题[定锚题（anchor item）],然后载荷次高的另外3题则以相反的次序（第6、5、4题）归入三个组合变量中成为第二题，使最高载荷的三题与低载荷的三题可以平衡组合在一起，使组合后的新变量有一致的变异量。此种方法适用于当研究者希望组合后的观察变量有一致的因素载荷时，即平行指标（Parallel indicators）。如果研究者同时关心各变量的平均数，则可套用相似的原则，将各题依照截距的高低排列，将各题组合成平均数相当的新组合变量。

第七节 结构方程模型：统合模型分析

三、变量组合与聚合

1. 变量组合的方法

④同源法

项目构念平衡法所组合的变量是方差或平均数最相当的一组平行测量，相对而言，另一种策略则是令组合后的因素结构与组合前的结构最相似的同源法（congeneric parcels）。具体做法是将最相似的完全标准化因素载荷者组合成一个组合变量。例如，将6个题目组合成3个组合变量，于是因素载荷第1、2高的题目组合成第一个观察变量，第3、4高者组成第二个组合变量，依此类推。如此一来，因为组合变量下的观察变量同构性最高，所形成的组合变量的重要性差异即是按原始因素结构内的因素载荷权重排序的。因而，对组合后的新变量进行简化的因素分析可以得到与组合前最相似的因素结构。事实上，这一做法类似于辐射组合法，以及Hall等人 (1999)提出的误差分离法（isolated uniqueness procedure），所不同的是，它采用CFA分析数据。Fletcher与Perry (2007)的模拟研究发现，此法的效果比项目构念平衡法好。

第七节 结构方程模型：统合模型分析

三、变量组合与聚合

2. 变量组合的优缺点

如果第一阶段所获得的测量模型的拟合性理想，后续利用组合分数来进行组合分数的统合分析会有诸多优点。除了模型可大幅简化，提高模型拟合度，获得较理想的估计解之外，也可以让测量模型得到较佳的测量信度、较高的变量解释力或共同性、观察变量的尺度更具有等距性与正态性、更理想的检验效力（样本量与题数之比较高），以及可避免特殊题目的干扰。

Bandalos与Finney (2001)指出，组合分数之所以会有较佳的拟合，主要是因为当利用组合分数时，观察变量矩阵中的元素数目会大幅度降低，同时也让观察矩阵与导出矩阵的差异（残差）减少，进而大幅提升模型拟合的数据。另外，各潜在变量的信度会提高，因为潜在变量对观察变量的解释力提高了（共同性提高），测量误差的比例也随之下降。因此，根据这一原理，在进行两阶段的统合模型估计时，如果测量模型并不十分理想，可以在第一阶段尝试将测量模型简化，以利第二阶段的结构模型的估计。Bandalos与Finney (2001)整理了7个期刊当中的论文，发现在317篇文章中有62篇（占20%）涉及组合变量的运用。

第七节 结构方程模型：统合模型分析

三、变量组合与聚合

2. 变量组合的优缺点

组合变量的处理虽然有诸多优点，但是也有一些限制。最重要的一点是研究者不应为了简化而简化，而需提出简化的理由与统计数据上的支持。例如，能够加以组合的题目，必定具有相同的量尺（例如，都是二分变量或是同点数的利克特量尺）。此外，能够组合成简化题目的一组观察变量应具有单维性（uni-dimensionality），即它们是测量同一个构念的变量，受到同一个潜在变量的影响。如果一群观察变量的相关并不高，或是其性质是在测量不同的构念，且具有多维性（multidimensionality），即不宜组合成组合变量。因为观察变量之间的关系可能非常复杂，如果径自组合变量，可能会遮蔽有意义的因素结构，造成不当的测量模型。

一般在进行变量组合之前，必须经过验证性因素分析的检验，以确保构念的单维性。观察变量的因素载荷越高，表示单维性越稳固，越适合进行题目组合。进一步地，在高阶验证性因素分析当中，如果第一阶潜在变量能够形成更高阶的二阶潜在因素，表示这些低阶因素具有高阶的单维性，因此也可以利用组合变量策略，改以测量变量处理第一阶的潜在变量，用来定义高阶的潜在变量。

第七节 结构方程模型：统合模型分析

三、变量组合与聚合

2. 变量组合的优缺点

Meade与Kroustalis (2006)引入了测量恒等性（measurement invariance）的概念，强调组合变量的使用前提是这些变量必须在不同情境（例如，前后测）或不同团体（例如，性别）中具有测量恒等性，以确保组合过程不致产生团体偏误。在进行恒等性评估时，最重要的是去检测因素载荷。因为变量的组合直接影响变量的方差与协方差的数值，所以组合后的观察变量的因素载荷在跨样本间都应具有恒等性。如果研究者关心潜在变量的平均数，那么组合变量的截距的恒等性就必须纳入考虑。Meade与Kroustalis (2006)以仿真数据获得的研究建议是，为了确保组合变量的恒等性，建议研究者在组合之前逐一检查个别观察变量的恒等性，找出导致不恒等的变量，这样一来比较有可能得到理想的恒等性的结论。

第七节 结构方程模型：统合模型分析

四、LISREL的统合模型分析

本范例所采用的数据库的目的是探讨教师的创意教学行为是否受到教师自身对于执行创意教学的胜任感的自我评价（也就是创造行为的自我效能感）的影响。也就是说，教师的创意教学自我效能感越强，越有可能展现出创意教学的行为。然而，教师创意教学自我效能感的高低可能受到教师个人性格特质与组织生活经验（组织社会化程度）的影响，因此，研究假设可写为：

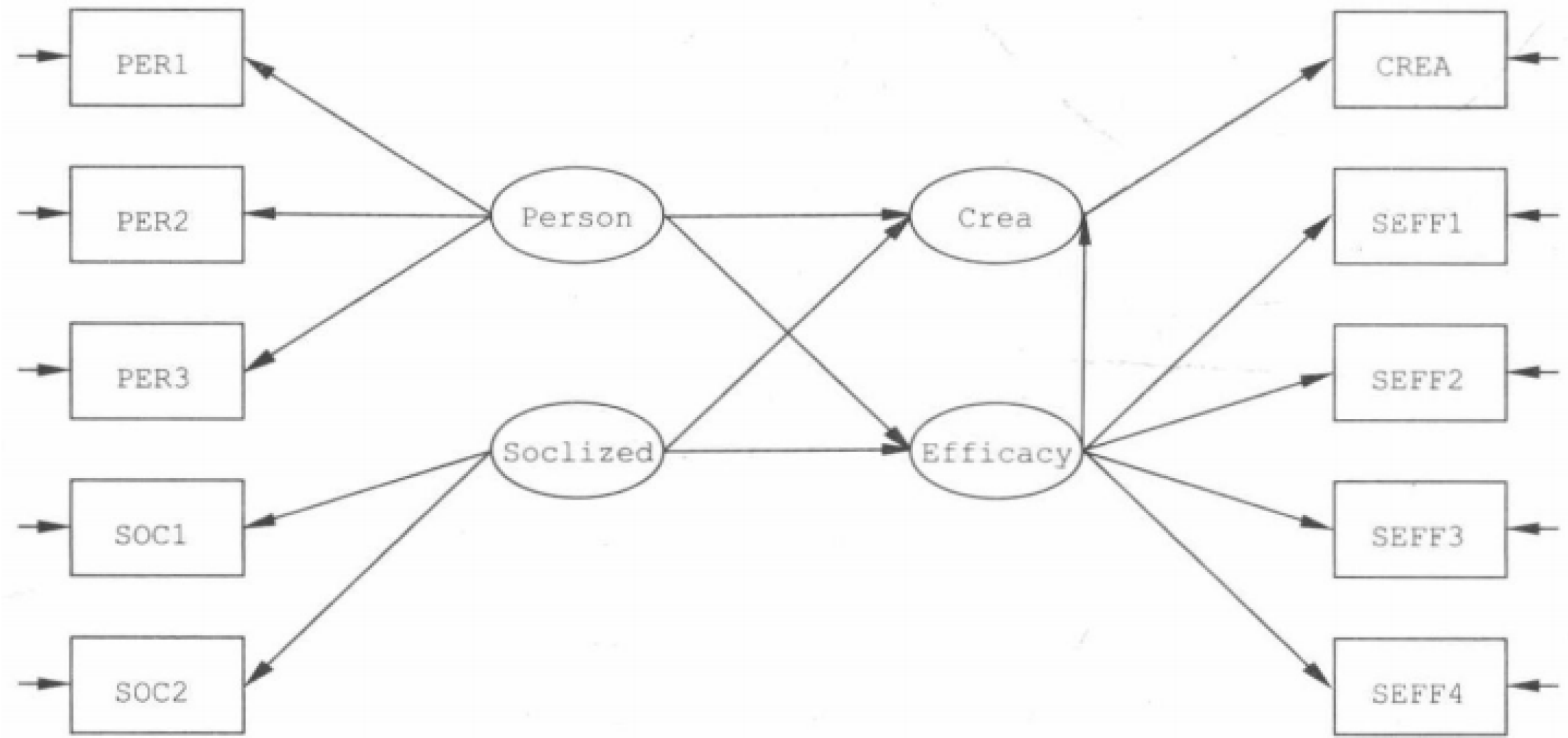
1) 教师创意教学自我效能感越高，越能够表现出创意教学行为。

2) 教师个人创造性格越强、组织社会化程度越高，创意教学自我效能越高。

3) 教师个人创造性格与组织社会化程度两项特质会通过自我效能感的中介作用，间接影响教师的创意教学行为。

第七节 结构方程模型：统合模型分析

四、LISREL的统合模型分析



第七节 结构方程模型：统合模型分析

四、LISREL的统合模型分析

LISREL、SIMPLIS、MPLUS、AMOS、R代码：略。

第八节 中介与调节

一、中介与调节的基本概念

中介与调节是社会科学研究中重要的方法学概念，近年来受到研究者相当程度的重视。主要原因是研究者经常遇到第三变量的混淆与干扰而影响变量间的解释关系。假设今天要以x解释y，可利用简单回归原理取y对x做回归（regress y on x），得到x的斜率系数表x对y的影响力。

在真实世界中，除了x与y之外，还可能存在一个或多个可能对于x→y的关系发生影响的第三变量，例如z变量。在研究实务上，如果研究者认为两变量的关系x→y可能受到第三变量z的影响时，最简单的处理方式是采取多元回归策略，将x与z皆作为解释变量，一并投入回归模型，通过x与z之间的统计控制来去除z对于x→y的干扰。此时，第三变量z被称为控制变量（control variable）或干扰变量（confunding variable）。如果x→y的关系受到z的干扰而呈现减弱或弱化，原始所观察到x→y的关系就称为虚假关系（spurious relationship）。相反的，x→y的关系可能因z的存在而提升，于是干扰变量被称为压抑项（suppressor），表示残差变异因为纳入z被“压抑”而减少。

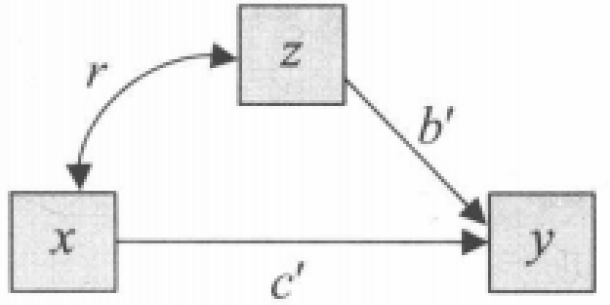
将z纳入模型将使x→y的关系发生变化，主要是因为x与z之间存在一定的共变关系。一般在回归分析中，如果多个解释变量具有相关，称为共线性问题。多元回归模型把共线性视为威胁而欲予以去除。但事实上，x与z之间所存在的共变很可能反映了特殊的影响关系，如果加以正视，可能成为重要的发现。

第八节 中介与调节

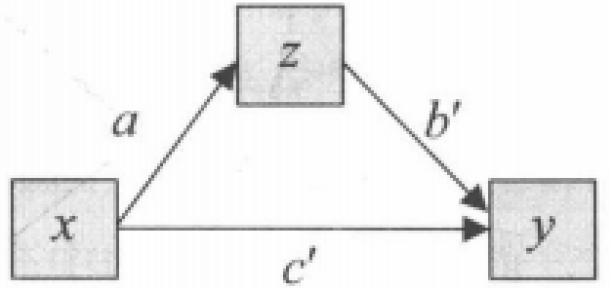
一、中介与调节的基本概念

例如，组织气氛良好提升了员工对公司的认同（x→z），然后，认同感提升造成了工作满意的提高（z→y），形成了一个x→z→y的影响路径，又称为中介关系。此时，第三变量z称为中介者，扮演x与y之间的中继角色。

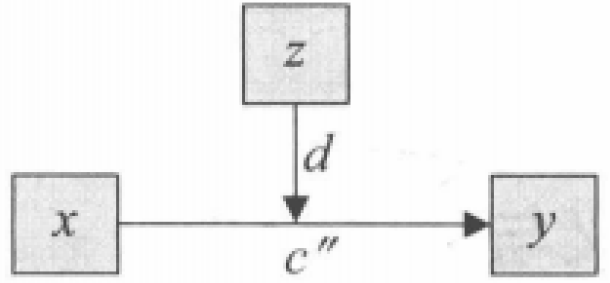
然而，员工的认同感（z）被视为中介者只是组织气氛（x）与认同感（z）这两个变量之间存在关系的一种解释与处理方式。另一种可能性则是认同感（z）扮演调节者的角色，即x→y关系在不同的认同感状况下会有不同的效应。换言之，组织气氛与认同感可能一起对于员工满意度发生作用，即x与z对y存在交互作用，x→y的效应必须就z的不同状态发生条件性的变化。



（a）第三变更 z 为控制变更



（b）第三变更 z 为中介变更



（c）第三变更 z 为调节变更

第八节 中介与调节

二、中介效应的统计原理

若以线性回归模型来整理三种第三变量角色，可由预测方程来表述如下：

$$\hat{y}_c = \beta_{0c} + cx \quad (1)$$

$$\hat{y}_b = \beta_{0b} + bz \quad (2)$$

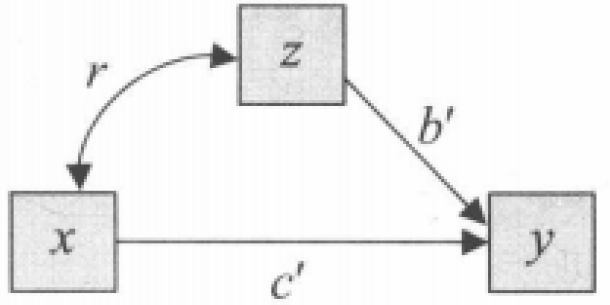
$$\hat{z}_a = \beta_{0a} + ax \quad (3)$$

$$\hat{y}_{bc} = \beta_{0bc} + c'x + b'z \quad (4)$$

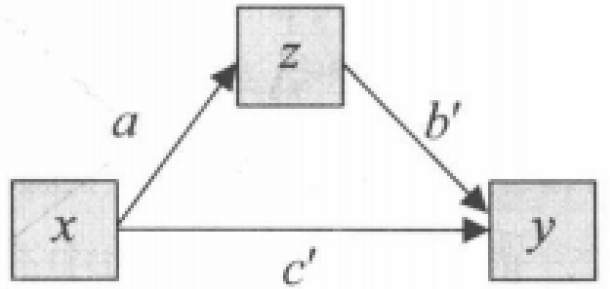
$$\hat{y}_{bcd} = \beta_{0bcd} + c''x + b''z + dxz \quad (5)$$

首先，公式1~3是三个简单回归方程，借以表现x→y、z→y、x→z三种关系。方程当中的回归系数都是单独只有一个解释变量的零阶系数，回归系数分别为a、b、c。公式4则将x与z一起纳入作为解释变量，(x+z)→y当中的两个回归系数b’与c’是控制了另一个解释变量下的净效应。

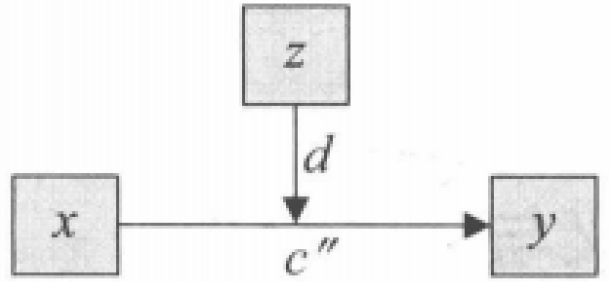
由于图b当中z与y均作为结果变量，因此可说是从公式1、2、3延伸到公式4的串联，而得到图b的中介关系；相对而言，图c当中仅有y作为结果变量，但x与z之间具有交互作用，因此可视为由公式1+公式2→公式4→公式5的复杂化，最后得以讨论变量如何发生调节作用。



（a）第三变更 z 为控制变更



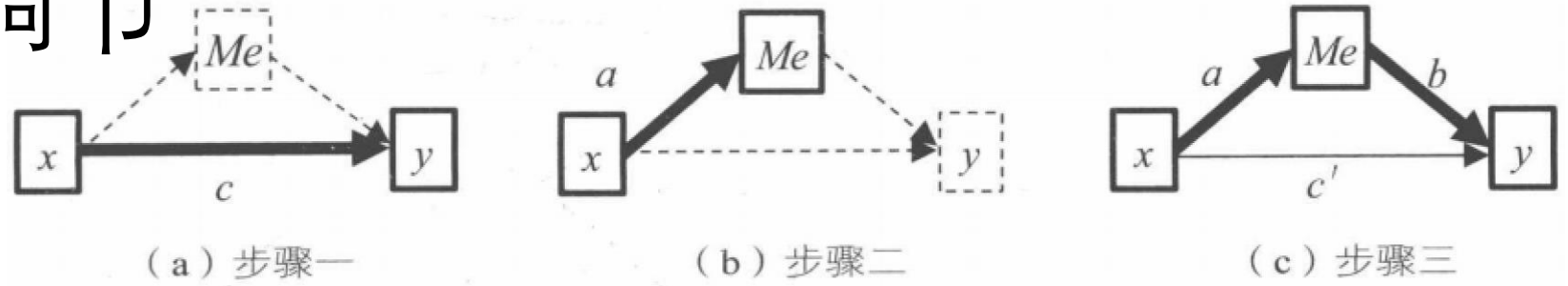
（b）第三变更 z 为中介变更



（c）第三变更 z 为调节变更

第八节 中介与调节

二、中介效应的统计原理



1. 中介效应的定义与估计

中介效应（mediation effect）可定义成第三变量z在x与y两变量当中所存在的传递效应。此时，第三变量z作为中介者，标示为Me。Me的中介作用有两种途径：x→Me→y与y→Me→x，但一般均以y作为最后的结果，x作为最前面的前因，因此在没有特别说明的情况下，中介效应指的是x→Me→y影响过程。

①多阶段因果关系拆解法

由于中介效应涉及多个变量之间的影响关系，因此最务实的做法就是逐一检查各系数的状况，来判定中介效应是否存在。最经典的做法是依据Baron与Kenny (1986)所建议的三步骤四条件原则。【已被认为是不正确的做法，但是仍有很大的影响力】

第一个步骤是检验x→y解释力，即回归系数c必须具有统计显著性，如图a所示。第二个步骤是检验x→Me的解释力，即a必须具有统计显著性，如图b所示。第三个步骤是若同时考虑x与Me对结果变量y的影响，Me→y的效应b必须具有统计显著性，但是x对y的解释力c消失，即证实x→y的关系是经由Me传递造成的，如图c所示，称为完全中介效应（completed mediation effect）；如果x对y的解释力c没有完全消失，即c虽有明显下降成c’，但c’仍具有统计显著性，则称为部分中介效应（partial mediation effect）。

第八节 中介与调节

二、中介效应的统计原理

1. 中介效应的定义与估计

②间接效应估计法

另一种评估中介效应的策略，是计算间接效应（indirect effect）并进行显著性检验。在仅有x、y、Me三个变量存在的情况下，x→Me→y的间接效应可由x→Me与Me→y两个回归系数的乘积，或是由x→y的c到c’的递减量求得，如公式6所示。

$$间接效应=c-c'=a\times b \quad (6)$$

从效应拆解的关系来看，公式1当中x→y的零阶回归系数c称为总效应（total effect）。当把中介变量纳入模型后，总效应被拆解成两个部分：公式4当中x→y的影响力c’是排除z的影响力后的净效应，x→z与z→y两个直接效应回归系数的乘积反映了中介变量z的间接作用，这些效应具有加成性而可表述如下：

$$总效应=直接效应+间接效应=c'+a\times b \quad (7)$$

如果间接效应的显著性检验达到显著水平，或是c到c’递减量具有统计意义，即可作为中介效应的证据。

第八节 中介与调节

二、中介效应的统计原理

1. 中介效应的定义与估计

②间接效应估计法

对于x→y纳入单一中介变量Me的间接效应可由a与b的乘积来估计，如果可以找出a×b的抽样分布，估计其标准误，即可进行中介效应是否显著不为0的考验或0.95置信区间的建立。最常用的检验公式基于Sobel (1982)所导出的a×b样本估计数的标准误，利用z检验或t检验来评估，称为Sobel检验，公式如下。

$$t = \frac{ab}{\sqrt{s_b^2 a^2 + s_a^2 b^2}} \quad (8)$$

公式8的分母为联合标准误，可由a与b的标准误求得，并利用传统普通最小二乘法回归分析或ML估计。但由于a与b这两个非标准化回归系数的抽样分布虽符合正态分布，a与b的联合概率分布（回归系数相乘）并不服从正态分布，而是峰度为6的高狭峰分布，同时如果解释变量的平均数不为零，还有非对称的偏态问题，使得Sobel (1982)所导出的标准误为偏估计值且不符合正态要求，因此Sampson与Breunig (1971)对标准误进行了修正：

$$s'_{ab} = \sqrt{s_b^2 a^2 + s_a^2 b^2 - s_b^2 s_a^2} \quad (9)$$

第八节 中介与调节

二、中介效应的统计原理

1. 中介效应的定义与估计

②间接效应估计法

虽然公式9修正了非正态问题，但是当样本量太小时（N<200）常会发生估计数为负值的非正定问题而无法有效估计。Bobko与Rieck (1980)建议在进行中介效应分析前，先将x、y、Me标准化，并利用三者的相关系数来计算标准误。

除了检验ab的显著性，c-c’的标准误也可以由个别系数标准误求得，借以进行t检验或0.95的置信区间的建立：

$$t_{(N-2)} = \frac{c-c'}{\sqrt{s_c^2+s_{c'}^2-2s_c s_{c'}\sqrt{1-r_{xz}^2}}} \quad (10)$$

虽然检验方法与标准误公式相继被提出，但经过模拟研究发现，Sobel (1982)所提出的原始公式仍是效率最佳的间接效应标准误。这也是为何多数软件（例如，LISREL、EQS、Mplus）仍以Sobel (1982)作为间接效应的显著性检验方法。

第八节 中介与调节

二、中介效应的统计原理

1. 中介效应的定义与估计

②间接效应估计法

近年来，由于统计模拟技术的进步与计算机指令周期的提升，对于间接效应标准误的估计得以利用重复取样技术来建立参数分布，求得参数的拔靴标准误，用以建立0.95置信区间（0.95CI）。其计算原理是以研究者所获得的观察数目为N的数据为总体，从中反复进行k次（例如，k=5000次）的置回取样（sampling with replacement），来获得k次间接效应的参数分布。这一分布的标准偏差所反映的即是间接效应的抽样误差，也就是标准误。由于拔靴标准误反映了“真实的”参数抽样分布，因此不需受限于正态概率与对称分布的要求，因此可以利用拔靴标准误所建立的间接效应参数分布，取其第2.5百分位数与第97.5百分位数数值为上下界，检查0.95CI是否涵盖0，借以判定间接效应是否显著不为0。

基于拔靴标准误所建立的0.95CI虽不受分布形态的限制，但随着所估计的参数类型不同，拔靴重抽得到的参数分布不对称情况也各有不同。为了改善拔靴0.95CI上下界的对称性，可进行偏误校正（bias correction）。但是最近的一些模拟研究发现，用未经偏误校正的拔靴置信区间来评估间接效应的结果未必较差，有时反而有更理想的表现，例如，在小样本量（N<200）时，保留正确的虚无假设的能力较佳（犯一类错误的可能性较低）。

不论是否经过偏误校正，以拔靴法建立的0.95CI称为拔靴本位标准误（bootstrapped-based confidence interval），Mplus等主流软件皆已纳入拔靴功能来建立偏误校正的拔靴置信区间，来评估间接效应的统计意义，逐渐成为检验间接效应的常规技术。

第八节 中介与调节

二、中介效应的统计原理

1. 中介效应的定义与估计

②间接效应估计法

除了拔靴本位0.95CI，贝叶斯估计标准误也逐渐受到重视，并认为可以取代拔靴标准误来检验间接效应的统计意义。事实上，贝叶斯估计法所建立的0.95CI也如同拔靴法的估计程序，利用重抽技术来反复取样借以获得参数分布。贝叶斯方法的不同之处是基于先验分布信息的导入，结合实际样本的重抽分布，将两者加以整合之后所得到的参数后验分布的标准偏差，即贝叶斯标准误，进而建立贝叶斯置信区间。

贝叶斯标准误的应用越来越容易进行，主要受惠于计算机科技效能的提升，以及马可夫链蒙特卡洛（Markov Chain Monte Carlo; MCMC）算法的应用。例如，Mplus将贝叶斯估计纳入分析模块，大大简化了贝叶斯估计的程序撰写与模型设定工作。

第八节 中介与调节

三、调节效应的统计原理

1. 调节效应的定义与估计

调节效应是指某一个自变量对于因变量的解释因调节变量（Mo）的不同状态而发生改变。如果调节变量是类别变量，那么自变量对于因变量的解释或预测在调节变量的不同组别之下会有所不同。如果调节变量是连续变量，自变量对于因变量的解释或预测则会随着调节变量的高低而发生变化。

调节效应的概念源于二因子实验设计中A与B两个因子对于因变量的交互效应。当A与B两个自变量“联合”对于因变量y发生作用时，两个自变量在因变量上所造成的效应称为主效应（main effect），两个自变量联合对y产生的A×B影响力称为交互作用。

在回归分析中，如果取x来对y进行解释，其直接效应可由简单回归估计所得到的β1系数表示。如果今天要将Mo作为x→y关系的调节者，必须增加两项来进行非线性回归分析：第一项是Mo→y的直接效应，影响力以β2表示；第二项是x×Mo→y的交互作用，影响力以βint表示，此时称为带有交互作用的回归（regression with interaction）。当多元回归包含交互作用项时，可进行调节效应分析，因此又称为调节回归（moderated multiple regression; MMR）：

$$y = \beta_0 + \beta_1 x + \beta_2 Mo + \beta_{int} xMo + [\varepsilon] \quad (11)$$

若βint显著不为0，表示交互作用存在，可据此进行x→y关系的调节效应分析。但如果βint不显著，表示x与Mo无交互作用，可将交乘项移除，仅保留Mo变量在模型中作为控制变量之用。

第八节 中介与调节

三、调节效应的统计原理

2. 调节效应的解释方法

交互作用回归的调节效应分析可利用公式11的移项重整来进行说明。如果令Mo为x→y的调节变量，公式11重整如下：

$$y = \beta_0 + \beta_1 x + \beta_2 Mo + \beta_{int} xMo + [\varepsilon^y] = (\beta_0 + \beta_2 Mo) + (\beta_1 + \beta_{int} Mo)x + [\varepsilon^y] = \beta_0^* + \beta_1^* x + [\varepsilon^y] \quad (12)$$

公式12当中，$(\beta_0 + \beta_2 Mo)$与$(\beta_1 + \beta_{int} Mo)$为x→y简单回归在纳入调节变量后的新截距（$\beta_0^*$）与新斜率（$\beta_1^*$），这两者均为随Mo变化的变动系数（varying coefficient）而非固定值。换言之，在调节回归当中，Mo作为调节变量，将通过两个途径影响x→y的回归方程：β2调整x→y的截距，βint调整x→y的斜率。

x→y的截距会因为Mo的改变而变动（截距的调节），截距调节的幅度由β2反映。但是由于截距高低并非反映x→y的解释力的强弱，使得β2的显著性并不受重视。

x→y的斜率会随着Mo的变动而改变（斜率的调节），斜率变动程度由βint反映。从统计的观点来看，只有当βint具有统计意义（显著不为零，βint≠0）时，斜率的调节才有意义；如果βint没有统计意义（βint=0），x→y的影响力不会随Mo的变动而变化。

当βint系数具有统计意义时，x→y的斜率将被Mo调节，调节的方向由βint系数的正负向来称呼：如果βint为正值，β_1^*将随Mo的放大而放大，称为正向调节（positive moderation）；如果βint为负值，β_1^*将随Mo的放大而降低，称为负向调节（negative moderation）。

第八节 中介与调节

三、调节效应的统计原理

2. 调节效应的解释方法

估计值	交互作用系数的影响		
	$\beta_{int}>0$ 正向调节	$\beta_{int}=0$	$\beta_{int}<0$ 负向调节
$\beta_1>0$	β_1^*随 Mo 放大，正值更加趋正（正效应正调节）	$\beta_1^* = \beta_1$	β_1^*随 Mo 放大，正值趋缓转负（正效应负调节）
$\beta_1<0$	β_1^*随 Mo 放大，负值趋缓转正（负效应正调节）	$\beta_1^* = \beta_1$	β_1^*随 Mo 放大，负值更加趋负（负效应负调节）

第八节 中介与调节

三、调节效应的统计原理

2. 调节效应的解释方法

更直观的解释方式：

主效应方程： $y = \beta_0 + \beta_1 x + \beta_2 Mo + [\varepsilon^y]$

调节效应方程：$y = \beta_0 + \beta_1 x + \beta_2 Mo + \beta_{int} xMo + [\varepsilon^y]$

如果调节效应方程中的β_{int}与主效应方程中的β_1符号相同，则意味着Mo增强了主效应（x）的影响，反之则削弱了主效应的影响。

第八节 中介与调节

三、调节效应的统计原理

3. 调节效应的分析程序

从分析的先后次序来看，调节效应是当交互作用具有统计显著性之后的事后检验。换言之，交互作用是一个整体考验，检验两个自变量是否会一起联合对因变量产生影响。调节效应则是交互作用下的一种特例，反映在某一个自变量作为调节变量的情况下，在不同水平或程度高低下的另一个自变量对因变量的影响力是否发生改变。从操作的角度来看，调节效应分析会因为调节变量是类别变量或连续变量而有所不同，但原理相同。

基本上，类别调节变量由于变量数值相对较少（类别数为k），所以对调节效应的分析相对单纯。首先，研究者必须先将类别调节变量转换成k-1个虚拟变量，并创造k-1个交互作用项，然后就带有交互作用项的多元回归进行整体考验，检查k-1个交互作用项是否显著。如果交互作用显著，研究者仅需分别就调节变量的不同类别进行k次x→y的简单回归分析，即可得知在调节变量的不同水平下，x→y的影响力分别为何，称为分组回归（separated regression）。例如，当调节变量Mo有两个类别时，需进行两个分组回归；如果有k组，则进行k次分组回归。如果调节变量编码为{0,1}的二分变量，更直接的做法是估计含交互作用项的调节回归方程，令Mo=0与Mo=1，即可导出两组简单回归方程。作图时，仅需绘制出各水平下的回归线，x→y的解释力的差异将反映在回归线的斜率差异上。

第八节 中介与调节

三、调节效应的统计原理

3. 调节效应的分析程序

如果调节变量与x都是连续变量，调节效应分析就相对复杂。这是由于连续调节变量的数值范围为连续光谱，并没有特定的固定状态，研究者必须自行定义调节变量的不同“水平”，来讨论x→y关系的变化。

如果Mo与x都是连续变量，且交互作用显著不为0，研究者必须指定其中一个为调节变量，另一个为被调节的变量。然后寻找反映调节变量高低水平的分割点（例如，在调节变量平均数以上及以下一个或两个标准差的位置），此时将可得到调节变量（以z表示）的平均数（z_M）、加一个（或两个）标准差得到高限（upper limit; z_H）与减一个（或两个）标准差得到低限（lower limit; z_H）等三个条件值（conditional value; CV），据以进行x→y简单回归。

为了便于结果解释与交互作用图的绘制，一般在进行交互作用回归之前，多会先将连续变量以总平均数进行中心化，使连续变量的平均数设定为0，即进行总平减，利用离均差来进行主效应或交互作用的估计。至于交互作用项，亦是利用平减后的变量进行相乘，使交乘项也具有离均差的性质。

第八节 中介与调节

四、调节式中介与中介式调节

虽然中介与调节是两个截然不同的方法学概念，但是在研究实务上，两者可能会一起发生。在Baron和Kenny (1986)及James和Brett (1984)关于中介与调节效应分析的经典论文中，他们除了指出了中介与调节效应的意义与检测原理，还提出了中介与调节的组合效应的概念雏形，即调节式中介（moderated mediation; MoMe）与中介式调节（mediated moderation; MeMo）。后来的学者则详细解释了这两个名词的异同与统计原理，甚至发展出了专属的分析模块（例如，PROCESS, Hayes, 2013）来简化其分析流程，使得调节与中介的组合效应更容易应用于实际研究当中。

顾名思义，Mo**Me**是指某一中介效应被特定调节变量Mo所调节，主要名词是**Me**（因此加着重号），Mo则是形容词，因此中介效应是研究分析的重点，中介效应必须先于调节效果成立，再进行中介效应当中的路径分析，例如，直接效应与间接效应如何受到Mo影响而发生条件化的变动。

相对而言，Me**Mo**一词是以调节效应为核心，**Mo**先于Me存在，Me用来形容**Mo**，表示调节效应在中介关系中扮演特定的角色。

第八节 中介与调节

四、调节式中介与中介式调节

1. 调节式中介（Mo**Me**）

如果研究者所关心的是中介效应如何被Mo调节，首先必须检查中介效应是否成立，即间接效应当中的各项效应（间接与直接效应）是否具有统计意义，进一步才能检验间接效应是否会随Mo的不同状况而变化，因此MoMe又可称为条件化间接效应（conditional indirect effect）。如果从传统回归路径分析或SEM观点来看，多个相连的直接效应的乘积可称为路径系数，因此间接效应a×b只是路径系数的一种特例，因而MoMe也被称为调节性路径分析（moderated path analysis）。

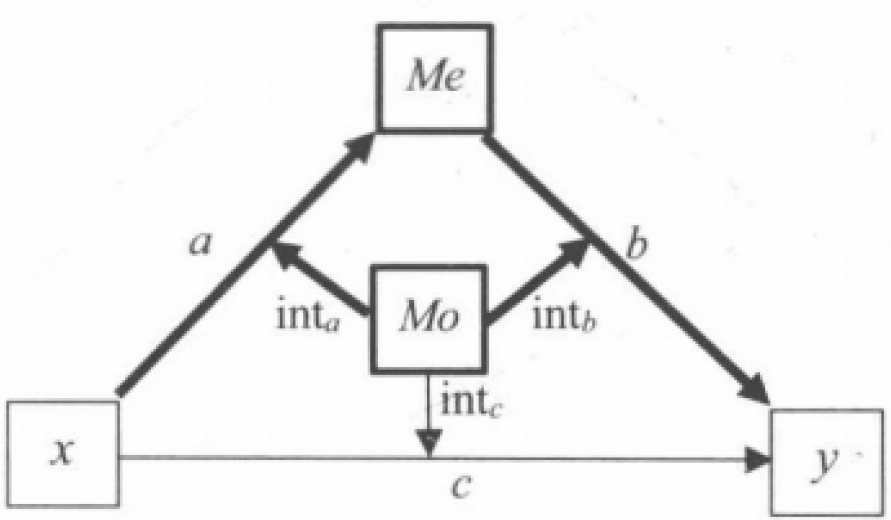
如果x→y的简单回归受到一个中介变量（Me）与一个调节变量（Mo）的影响，这一Mo可能对于3个直接效应产生调节作用，如图中的int_a、int_b、int_c。值得注意的是，图a是从研究设计的观点来陈述各变量的影响方向，也就是用来表述研究者所关心的“效应”或“机制”为何，因此称为“假设概念图”。图中的系数符号仅作为各效应标示之用，不重要的系数并未被标示出来。相对而言，图b则完整呈现了统计方程的成分拆解，因此可称为“参数效应图”。

第八节 中介与调节

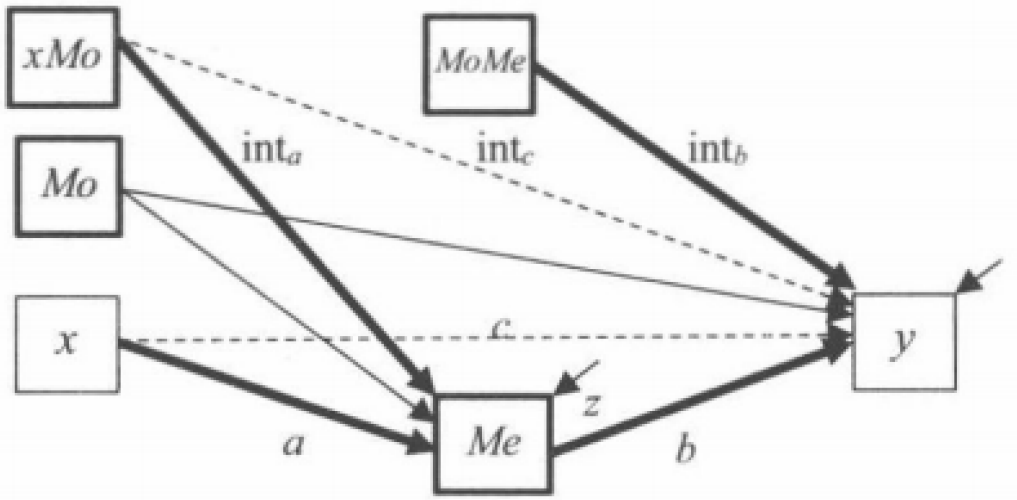
四、调节式中介与中介式调节

1. 调节式中介（MoMe）

由于本节讨论的是MoMe，因此图a当中的x→Me与Me→y两个直接效应皆以粗单箭头联系，目的在凸显中介效应必须先行成立。图中的Mo同时对x→Me、Me→y、x→y三个直接效应发生调节作用，所有的效应都被调节。从MoMe的观点来看，图中最重要的调节应该发生在x→Me与Me→y两个直接效应之上，因此图a当中的Mo对于这两个直接效应a与b均以粗的单箭头进行指向。相对而言，Mo对x→y的直接效应c的调节并非重点（因为中介效应并不在乎x→y直接效应的统计意义），因此图中对于int_a和int_b标以粗的单箭头，至于int_c并非关心的重点。在执行MoMe时，甚至可以把int_c项移除，进行较简化的MoMe分析。



（a）假设概念图



（b）参数效应图

第八节 中介与调节

四、调节式中介与中介式调节

2. 中介式调节（MeMo）

相对于MoMe把重点放在间接效应如何被调节，MeMo则是在检验xMo→y的交互作用是否会因为纳入Me而消失或显著减弱，可检测带有调节效应的xMo→Me→y的中介过程是否成立。换言之，MeMo先必须有一个显著Mo调节x→y的int_c交互作用，其次估计同时纳入其他直接效应与交互作用后的新的Mo对x→y影响的交互作用项int_c，最后检测$int_c - int_c'$差异是否具有显著意义，即图中的int_c是否从显著变成不显著，或是显著降低到int_c'。如公式15~17所示。

$$y = \beta_0 + \beta_c x + \beta_{d(y.Mo)} Mo + \beta_{int_c} xMo + [\varepsilon] = \left(\beta_0 + \beta_{d(y.Mo)} Mo\right) + \left(\beta_c + \beta_{int_c} Mo\right)x + [\varepsilon] \quad (15)$$

$$Me = \beta_0 + \beta_a x + \beta_{d(Me.Mo)} Mo + \beta_{int_a} xMo + [\varepsilon] = \left(\beta_0 + \beta_{d(Me.Mo)} Mo\right) + \left(\beta_a + \beta_{int_a} Mo\right)x + [\varepsilon] \quad (16)$$

$$y = \beta_0' + \beta_c' x + \beta_{d(y.Mo)}' Mo + \beta_{int_c}' xMo + \beta_b' Me + \beta_{int_b}' MoMe + [\varepsilon'] = \left(\beta_0' + \beta_{d(y.Mo)}' Mo\right) + \left(\beta_c' + \beta_{int_c}' Mo\right) x + \left(\beta_b' + \beta_{int_b}' Mo\right) Me + [\varepsilon'] \quad (17)$$

第八节 中介与调节

四、调节式中介与中介式调节

2. 中介式调节（MeMo）

第一个步骤是检验int_c系数的统计意义，因此(x+Mo+xMo)→y的交互作用回归（公式15）必须先行检测，β_{int_c}必须具有统计意义，此时x→y的总效应为a+int_c。第二个步骤则是以Me为结果变量，建立(x+Mo+xMo)→Me（公式16）来检测xMo→Me的β_{int_a}是否显著不为零。第三个步骤是建立(x+Mo+xMo+Me+MeMo)→y（公式17）来检测β'_{int_c}的作用是否消失。

为了简化MeMo的检验，Muller、Judd与Yzerbyt (2005)以系数降低法来检测先期存在的调节效应减少量是否具有统计意义，如公式18所示。

$$\beta_{int_c} - \beta'_{int_c} = \beta'_b\beta_{int_a} + \beta_a\beta'_{int_b} \quad (18)$$

事实上，MoMe或MeMo的检验虽然概念复杂，但是在统计上没有差异，Hayes (2013)甚至呼吁研究者完全抛弃MeMo的概念，仅就MoMe来检测同时带有Mo与Me的模型。从研究设计的角度来看，研究者最重要的工作是清楚说明研究假设为何，并提出理论与文献来支持，分析时则针对研究者所关心的效应进行检验，详细说明各项效应，应可充分说明MoMe与MeMo的现象是否存在。

第八节 中介与调节

五、LISREL的中介与调节分析

LISREL、SIMPLIS、MPLUS、AMOS、R代码：略。

第九节 结构方程模型的正确运用

一、正确运用SEM的相关议题

1960年可以视为统计技术发展的一个分水岭。之前的心理计量学的发展以古典的推论统计技术为主，例如，t检验、方差分析、皮尔逊相关以及最小二乘回归；之后的发展则着重于各种统计问题的校正技术，可称为现代统计学。1994年，《结构方程模型》期刊第一卷的出版掀起了统计技术的第三次革命。

尽管SEM的发展风起云涌，但同时也遭遇了许多困境与挑战。例如，美国南加州大学的心理系教授Norman Cliff在1983年以相当严厉的口吻质疑SEM的不当使用：

1. 研究者所获得的数据无法替我们完全确认或否认一个模型的正确性，因为模型是人为的，而且可以用各种方法重新定义；
2. 具有时间性的先后次序的证据并不代表因果；
3. 潜在变量的命名是一个主观的过程，而非客观的事实，对潜在变量的估计存在名义谬误（nominalistic fallacy）的陷阱；
4. 事后的解释与调整具有诚信与可信度的问题，也就是驳斥了部分研究者大量使用模型修饰程序来获得理想拟合度的不当做法。

Cliff的批评可以说是警世之语，事实上，Cliff也相当看重SEM的发展，他在1983年的文章中最后提道：最后，我必须再次强调，像LISREL这类的分析工具的确提供了一个空前的、史无前例的机会，使我们能够把这类研究做好。

第九节 结构方程模型的正确运用

一、正确运用SEM的相关议题

1. SEM运用的三个关键议题

Nesselroade (1991)指出，使用SEM这一类高阶的复变量分析技术，有三个层面的问题必须注意：个体（individuals）、测量（measures）和情境（occasions）。

首先，与个体有关的问题，就是与抽样、样本有关的问题。任何推论统计都涉及抽样的过程，因此，有关抽样与样本对于统计分析的影响，在任何一本统计著作中皆有详细说明。对于统计稍有概念的研究者，对抽样的原理与相关定理的重要性也应知之甚详。对于SEM而言，由于分析的核心原理是比较理论值与实际值的差异程度，因此，SEM深受样本与样本统计量的影响。尽管SEM发展了诸多的估计样本所带来的影响的分析策略（例如，复核效化）与统计指数（例如，ECVI），但是这些技术或指数也有自己的限制，更重要的是，无法改变抽样本身会带来各种问题的事实。因此，抽样程序与样本特性可以说是SEM成败的关键之一。

第九节 结构方程模型的正确运用

一、正确运用SEM的相关议题

1. SEM运用的三个关键议题

其次，测量的问题也深深影响着SEM的操作。所谓测量的问题主要在于对潜在变量的估计与分析。不论是古典测量理论对于心理测验的界定，还是因素分析统计技术对于因素的决定，都反映了心理计量学最重要的任务之一——能够有效地处理抽象心理构念的测量问题。对于潜在心理构念，我们可以视其为一个行为或心理特征的概念总体，我们所进行的测量是从这个概念总体抽取相当的样本指标来反映这个总体概念。因此，测量的成败取决于我们如何对于这个测量总体进行定义。一旦对于总体加以定义，我们才可能发展出相对应的一组样本指标来反映该总体的基本特征。结合此前谈到的抽样问题来看，从概念总体选取指标样本，就好比从一群人口总体中抽取样本来进行研究，有异曲同工之意。但是，不同的是人口总体通常并没有太困难的定义问题，因为人的集合是具体的事件。然而对于心理特质的定义则是一个相当抽象的程序。这些有关测量题目如何选取、借以反映潜在变量性质以及不同的选题程序对于SEM影响的选择效果问题，是SEM成败的第二个关键。

第九节 结构方程模型的正确运用

一、正确运用SEM的相关议题

1. SEM运用的三个关键议题

最后，SEM使用的情境也是SEM分析的重要影响因素。举例来说，当SEM应用于重复测量的研究数据分析时，不同测量时间点的重复测量的操作方式在前面提到的样本与测量问题之外，又增加了一个不同情境的干扰效果。在近年来相当蓬勃发展的成长曲线分析的应用上，对于不同时间点下的SEM测量应如何进行，有着诸多的讨论，反映了这个议题的重要性。此外，复核效化的应用也说明了SEM的结果在不同人口总体上的类化性，这也是一个必须加以探讨的重要议题。

进一步地，本门课并没有特别讨论对多重特质多重方法矩阵（MTMM design）研究的SEM分析，这类应用也涉及SEM模型在多元测量模型当中的应用问题。如果我们把多重方法视为SEM分析的不同情境（好比重复测量的不同时间点），MTMM设计的应用也反映了情境影响SEM分析这个议题的重要性，因为时间序列分析、复核效化以及MTMM设计的应用都是最近在期刊上往来讨论得非常热烈的课题。我们或许不应该说情境议题是SEM分析成败的另一个关键，而应将它视为SEM更高度发展的跳板因素。

第九节 结构方程模型的正确运用

一、正确运用SEM的相关议题

2. SEM的决策建议

从原理来看，SEM是一套非常复杂的统计技术，但是由于应用软件的便捷与强大功能，用户不需对于SEM的原理与内涵有深入了解，即可以快速熟练地操作SEM，得到大量的数据。窗口版的LISREL、EQS、Amos等SEM分析工具更可以结合绘图功能与文书作业软件，让使用者甚至不必绘制变量关系的路径图，直接从软件中制作。从SEM的推广来看，这些便捷的作业模型并不令人鼓舞。

综观统计技术的运用，从初等统计、多变量统计到SEM，每一种统计方法都难免掺杂主观的判断。尤其是当统计的原理越趋于繁复时，所涉及的人为判断往往更多且决策更为困难。基本上，操作SEM时有一些值得注意之处，虽然在之前已有讨论，下面还是列举了10点有关SEM分析决策的提醒与建议。这些提醒与建议综合了各个层面的概念，供大家参考。

① 对于模型的优劣，应该避免下武断与绝对的结论。记得在单一的SEM分析中，我们只能证实某一个模型可以拟合观察数据，也就是该模型没有错误地代表观察值，但是我们无法证明某一个模型是绝对正确的。

② 对于所检验的模型皆以理论为基础。在没有确切根据或没有逻辑推理的合理性之前，不要随意改动模型的界定。即使模型修饰指数告诉我们某些参数可以变动调整，也必须全盘考虑整个模型的设计原理，避免“为改善拟合度而修饰”的做法。

第九节 结构方程模型的正确运用

一、正确运用SEM的相关议题

2. SEM的决策建议

③ 尽可能地检验不同的模型，了解不同模型的计量特性。如果可以，研究者应该从不同的理论观点去提出竞争的对立假设来进行相互比较。必要时，可以进行复核效化的检验，利用不同的样本来重复检验，以确认模型估计的类化能力。

④ 当一个SEM模型当中兼含测量模型与结构模型时，宜先进行测量模型的检验，待测量模型具有相当的合理性之后，再进行结构模型的参数估计。使得SEM模型评估具有测量的“渐近合理性”。

⑤ 对于模型界定的适当性，不应只考虑统计量的意义，而需考虑概念的本质与意义。尤其应该注意潜在变量的多重指标是否具有理论与实务的适当性。

⑥ 对于模型拟合度的评估应同时采用各种指数，因为不同的指数反映不同的模型的计量特征。同时，不论指数数据的好坏，均应翔实报告，不应“报喜不报忧”或“粉饰太平”。

⑦ 尽可能地先行检验数据的分布特性，使其不受偏离值与遗漏值的影响。必要时应进行多变量的正态化或遗漏值检验，找出潜藏于变量之间的特殊数据。

第九节 结构方程模型的正确运用

一、正确运用SEM的相关议题

2. SEM的决策建议

⑧ 对于测量变量的选择，应尽可能地符合正态化假设的基本特性，例如，避免二分变量或顺序变量的虚拟化；单一测量变量的尺度尽可能增加，并扩大其变异量（例如，可以将多题集合成单一测量变量，提高单一指标的变异量，使其符合正态化的特质）。

⑨ 当不同模型具有类似的拟合度时，使用简效原则来决定最终接受的模型。也就是越简单的模型（估计的参数较少者），越具有较高的简约性。

⑩ 样本量越大越好。

第九节 结构方程模型的正确运用

二、SEM的解释与应用

1. 因果关系的论证

在SEM的使用上，最常被讨论的是对因果关系的质疑。SEM最早的名称是因果模型分析，到目前为止，仍有很多SEM研究者将回归分析或是SEM视为非实验研究可以得到因果结论的重要技术。然而，如果从研究设计的角度来看，只有经过严格实验处理的研究，将受试者随机分派，然后在严谨的实验控制下检验研究者所操纵的自变量变化对于因变量的影响，所得到的结果才符合因果关系的基本条件。其他任何形式的研究设计或统计推论，所得到的仅可能是对变量的预测或解释，而非因果关系。

在实验研究当中，研究者所操纵的自变量（因）是影响因变量（果）的直接原因，在时间顺序上有前因后果之分。但是，非实验研究的变量关系往往没有一个清楚的方向与单一轴线，变量的影响关系无法清楚区分。同时，在变量的选取上，通常是从因变量来回溯、探讨可能具有解释力的前置变量，再放入模型中来检验研究者的推理是否正确。因此，如果要以SEM分析的测量原理从在同一个时间点测量得到的变量关系上推导出因果关系，是不切实际的。SEM的统计原理与其他相关及回归分析的原理并无二致，在推导因果结论上的限制是相同的。除非SEM分析所进行的研究在测量上能够证明变量的前因后果（例如，纵向研究，变量的测量是在不同时间点上的测量），此时，关系的讨论或许可以被接受为一种因果的概念；否则，任何其他形式的SEM分析都应该避免做出因果结论。

从因果推断专业来看：①时间顺序上有先后不代表存在因果关系；②一般的SEM分析只能做出相关关系的结论；③因果中介分析，这一前沿方法可以用SEM推断因果关系。

第九节 结构方程模型的正确运用

二、SEM的解释与应用

2. SEM分析的推论限制

SEM是一套用以检验特定假设模型的统计方法学。因此，SEM最主要的一个目的是检验研究者所提出的理论或概念架构是否具有实证的意义。整个SEM的分析程序都离不开研究者所提出的SEM模型，因此，研究者是否可以在提出研究问题的第一时间，就通过理论推导与文献检阅过程，选择适当的研究变量，提出有意义的研究假设，去说明变量的关系，进而发展出适切理想的假设模型，即成为SEM模型的计量检验程序是否可以顺利完成的一个基本条件。换句话说，SEM分析技术只是一套统计的方法与分析的策略，SEM本身无法创造理论或知识，而是需要研究者以其智慧去整理前人所发展出来的理论或知识，建构一套适当的概念模型，然后再以SEM技术协助研究者完成模型的分析与讨论。

SEM分析最大的禁忌是过度倚赖技术指标数据与过度推论。SEM分析只能用来评估研究者所提出的假设模型是否适切，但是，究竟何者才是真正能够反映真实世界的变量之间关系的模型，这一个结论并不能够从模型拟合指数的数据中得到。因为除了研究者所提出的模型之外，同样的一组观察变量可能有许多不同的模型组合，这些基于同样的观察数据的假设模型可能都有理想的拟合度，SEM分析无法区辨这些计量特征类似的理论模型何者为真。SEM的使用者不但必须谨记统计方法学本身的限制，也必须避免陷入过度推论的陷阱当中。

第九节 结构方程模型的正确运用

二、SEM的解释与应用

3. SEM分析的解释

从先前的讨论中，我们可以意识到SEM的分析有其限制，但是这并不意味着SEM分析的结果缺乏说服力。只要依据SEM分析的原理，逐步完成各项设定与检验工作，我们还是可以从分析结果中得到许多具有启发性的研究发现的。

首先，使用者必须了解SEM分析的各项参数估计与模型拟合度数据是变量因果关系的必要但非充分的证据。SEM分析的证据可以说明某一个因果概念是可能存在的，但是不能据此排斥其他模型的存在。除非SEM分析直接检验了其他假设模型的计量特性，得到清楚的结论，证明该模型是不适当的，否则，SEM分析的结果应仅就该研究所使用的样本与所检验的变量关系来讨论。

此外，SEM分析结果受到诸多因素的影响，例如样本量的大小、变量正态性问题、估计方法的选择、潜在变量的量尺化设定等。因此，在撰写研究报告的同时，也应详细说明这些技术层面的特殊之处与处理原则，尽可能使分析过程透明化。此举除了有利于其他读者了解，也让其他研究者得以复制研究的结果，是维系SEM分析的公正性、客观性的一个重要手段。