

• 研究构想(Conceptual Framework) •

密集追踪数据的收集方案优化和分析方法改进*

刘红云¹ 窦佳宁¹ 徐永泽²

(¹ 北京师范大学心理学部, 应用实验心理北京市重点实验室, 心理学国家级实验教学示范中心(北京师范大学), 北京 100875) (² 北京师范大学珠海校区文理学院心理系, 广东 珠海 519085)

摘要 密集追踪研究(Intensive Longitudinal Study, ILS)为探究心理状态的动态变化及相互作用机制提供了有效工具,但其高测量强度易导致参与者负担重、依从性低及数据质量不佳。目前尚缺乏能有效结合 ILS 特点的数据收集方案和分析方法。本研究将计划缺失设计(Planned Missing Design, PMD)的思想拓展到 ILS,在动态结构方程模型框架下,探究不同类型 PMD 的比较及其选择建议、含 PMD 的样本量规划及其优化方法,以及缺失数据处理方法的改进,同时将开发应用程序,结合心理学研究场景验证方案的可行性与适用性。

关键词 纵向数据分析, 结构方程模型, 缺失数据分析, 计划缺失设计, 样本量规划

分类号 B841.2

1 研究背景

近年来,随着心理学研究问题的不断深入和对个体特异性的重视,密集追踪研究(intensive longitudinal study, ILS)已经成为心理学、生命科学、医学和健康等领域普遍采用的研究范式之一。ILS 常采用经验取样法(experience sampling method)、生态瞬时评估(ecological momentary assessment)、动态评估(ambulatory assessment)和每日日记(daily diary)等在自然状态下评估个体日常的生活经历(Bolger et al., 2003)。其特点是可以对个体内(intra-individual)的动态关系以及个体间(inter-individual)差异进行研究(郑舒方 等, 2021; Hamaker & Wichers, 2017; Xiao et al., 2023)。相比于调查研究和实验室研究, ILS 可以有效降低回忆偏差,提高生态效度(罗晓慧, 刘红云, 2024; Trull & Ebner-Priemer, 2014)。更重要的是, ILS 为探讨个体内波动和个体间差异提供新视角,为深入理解心理状态的发生、发展和变化过程提供有效方法,为个性化干预方案的制定提供可靠依据

(König & Strahler, 2023; Luo, Hu, & Liu, 2024)。

但 ILS 中较高的测量强度(主要包括问卷长度和测量次数)会对所收集数据的数量和质量,以及研究结果的可靠性和有效性产生消极影响。在 ILS 中,参与者每天需要填写一份或多份问卷,通常持续数天甚至数月。这样频繁的评估可能会引起参与者作答反应的变化,导致数据质量下降(Eisele et al., 2023)。已有的实验研究和元分析表明, ILS 中的测量强度是影响数据质量的主要因素。增加测量强度可能会导致参与者的负担增加(Fuller-Tyszkiewicz et al., 2013; Santangelo et al., 2013)和依从性(compliance)降低(Eisele et al., 2022; Vachon et al., 2019),并引起研究中状态变量的个体内变异降低(Eisele et al., 2023; Fuller-Tyszkiewicz et al., 2013; Hasselhorn et al., 2022)和变量间动态关系的结果改变(Hasselhorn et al., 2022)。

虽然有研究者提出可以采用事后识别的方法检测无效作答以提升数据质量(Ward & Meade, 2023),并尝试将其拓展到 ILS 中(Hasselhorn et al., 2025; Jaso et al., 2022),但研究者普遍认为在研究设计阶段尽可能降低参与者负担是减小测量强度消极影响的首选策略(Eisele et al., 2023; Ward & Meade, 2023)。目前,在 ILS 中,仍缺乏从研究设计的角度探讨如何通过优化数据收集方案,有效

收稿日期: 2024-03-26

* 国家自然科学基金(32471145)资助。

通信作者: 徐永泽, E-mail: yz xu@bnu.edu.cn

降低测量强度和研究成本,以提升数据质量并降低结果偏差风险的研究。因此,有必要基于 ILS,探究数据收集的优化方案以及相应的数据分析改进方法,以直接有效地减小 ILS 中测量强度的消极影响。

在传统的横断研究和追踪研究(即面板研究)中,研究者常采用计划缺失设计(planned missing design, PMD)来降低测量强度。PMD 通过特定的研究设计产生缺失以降低测量强度,再运用合适的缺失数据处理方法得到无偏的估计量,可以从设计源头上解决 ILS 中测量强度有潜在消极影响的问题(Black et al., 2012; Eisele et al., 2022)。随着缺失数据处理方法的发展,大量研究表明这一设计思路非常有前景(Feng & Hancock, 2021; Little & Rhemtulla, 2013)。虽然 PMD 为有效降低问卷长度、减少测量次数和降低研究费用等一系列问题提供了解决方案,但仍然缺乏针对 ILS 及其常用统计模型,如动态结构方程模型(dynamic structural equation modeling, DSEM),探讨如何制定高效的 PMD 方案、合理规划研究样本量并有效处理缺失数据的研究。

综上,ILS 能够从个体内动态变化的视角揭示心理状态的发生机制、发展规律及变化特征,为探索个体内波动与个体间差异提供了新的研究视角。然而,ILS 较高的测量强度可能影响数据的数量与质量和研究结果的可靠性与有效性。目前,针对 ILS 及其常用统计模型,仍缺乏从研究设计角度优化数据收集方案,以降低测量强度和研究成本、提升数据质量并减少结果偏差的系统性研究。因此,本研究旨在结合行为科学研究的实际需要,重点探讨 ILS 中不同 PMD 方案的比较分析和选择建议、含计划缺失数据的样本量规划及其优化,以及缺失数据处理方法的改进和选择这三大问题,从设计源头探究 ILS 中经济高效的数据收集方案和更有效的数据分析方法。此外,研究还将开发样本量规划和缺失数据处理的应用程序,并围绕真实心理学研究问题,探讨测量方案设计、样本量规划和数据处理分析在应用研究中的可行性和适用性,为推动新兴技术和前沿方法的应用提供实践依据和实操工具。

2 国内外研究现状

2.1 计划缺失设计的研究进展和局限性

计划缺失设计允许研究者随机分配参与者,

并基于事先的计划收集含有预知缺失数据模式的设计(Little & Rhemtulla, 2013)。该方法可降低数据收集成本,减轻参与者负担,降低不认真作答和计划外的数据缺失比例,从而提高数据质量(Feng & Hancock, 2021; Little & Rhemtulla, 2013; Rhemtulla & Hancock, 2016)。更重要的是,由于预先设计数据的缺失机制,研究者可以借助先进的缺失数据处理技术,得到可靠有效的统计推断结果。PMD 被广泛应用于教育评价、心理测量和医学等研究领域(Johnson, 1992; Raghunathan & Grizzle, 1995)。

大量实证研究表明,在收集的有效数据量相等的情况下,PMD 在参数估计精度和检验力方面优于完整设计(Jorgensen et al., 2014; Little & Rhemtulla, 2013)。虽然有研究担心相比于完整数据的设计,PMD 可利用信息量减少,检验力可能会下降(Graham, 2012; Rhemtulla & Hancock, 2016)。但几乎所有研究结果均表明,在设计合理且缺失数据处理方法恰当的情况下,PMD 可带来明显的收益(Graham et al., 2006; Rhemtulla & Hancock, 2016; Silvia et al., 2014),即在有效降低参与者负担和降低研究成本的同时,不会降低参数估计的精度和检验力。因此,研究者普遍推崇在数据收集方案的设计阶段通过减轻参与者负担,提高其依从性和数据质量(王孟成,叶浩生,2014; McKnight et al., 2007)。但在重复测量次数较多(如 20 次以上)的密集追踪研究(ILS)中,如何基于 PMD 高效地收集优质的数据仍需更多关注与探索(Black et al., 2012; Eisele et al., 2022; Silvia et al., 2014)。

常用的 PMD 有 3 种基本类型(Little & Rhemtulla, 2013):多形式计划缺失设计(planned multiform missing design)、测量时间点计划缺失设计(planned wave missing design)和两方法测量计划缺失设计(two-method measurement planned missing design)。以下结合追踪研究简述 3 种常用的 PMD。

2.1.1 多形式计划缺失设计

多形式计划缺失设计,又称拆分问卷设计,该设计将整个问卷分为多种形式的较短子问卷,一次向参与者发放其中一种形式的子问卷,从而缩短问卷长度和调查时间(Graham et al., 2006)。多形式 PMD 在教育评价和心理测量中非常常见,常

用的题目组合形式有矩阵抽样设计和锚设计两种。两种设计共同点在于均可通过缩短测验长度,降低被试完成测验所需时间,从而降低作答负担,进而提高测验回收率与作答质量。但它们各自也存在一定的局限性(Graham et al., 2006)。矩阵抽样设计中,不同被试作答的题目组合差异较大,导致题目之间的协变量信息不完整,因而不利于准确估计题目间的相关性或结构模型参数(如协方差等)。锚设计则通过在各子卷中引入相同的“锚题”,使所有被试均作答该部分题目,从而支持题目间的相关分析及更为复杂的模型参数估计。然而,锚设计也存在一定局限性。一方面,由于每个子卷需包含相同的锚题,占据了问卷长度,在子卷长度相同时,其他题目提供的信息相较于矩阵抽样设计相对更少;另一方面,锚题的选择需具备较高的代表性和测量属性,这对测验编排与后续数据分析提出了更高的要求。图1以4个题目子集为例,呈现矩阵抽样设计和锚设计的测验组合模式,以及相应的个体内和个体间均平衡的施测方案。

以往研究表明,多形式PMD中不同形式测验题目的构成特点和测验组合方式会对研究结果产生影响,但若选择合适的题目层面的缺失数据处理方法,则各种测验组合形式对检验力影响很小(Gottschall et al., 2012; Rhemtulla et al., 2014)。然而,聚焦 ILS 的相关探究很少, Silvia 等人(2014)的研究是目前唯一基于 ILS 探讨多形式 PMD 的研究,结果表明题目层面的 PMD 可以得到无偏的参数估计结果和略微高估的标准误。但该研究基于

传统的两水平模型,没有结合 ILS 感兴趣的研究问题(如,考察变量的自回归和交叉滞后效应,以及这些效应的个体间差异);此外,该研究对于缺失数据的处理采用传统多水平模型默认的全息极大似然估计,没有考虑对题目水平缺失值处理效果更好的多重插补法。由于这项研究开展时,动态结构方程模型(dynamic structural equation modeling, DSEM; Asparouhov et al., 2018)尚未提出,适用于 ILS 的缺失数据处理新方法(如卡尔曼滤波法)也未得到合理应用,此研究基于多形式 PMD 得到的结果难以适应当前 ILS 的需求。因此,有必要结合 ILS 常用统计模型,采用前沿的缺失数据处理方法,探讨多形式 PMD 的有效设计方案。

2.1.2 测量时间点计划缺失设计

测量时间点计划缺失设计是研究者在追踪研究中事先分配缺失测量时间点的设计(王孟成,叶浩生, 2014; Mistler & Enders, 2012)。测量时间点 PMD 让参与者少参加一次或多次测试,以降低参与者的负担,提高其依从性,从而提升数据的质量,具有高效、省时和经济的特点。在长时追踪研究中,加速纵向设计(accelerated longitudinal design) (刘红云 等, 2025; 唐文清 等, 2014; Duncan et al., 1996)是一种典型的测量时间点 PMD。它结合了横断研究和纵向研究的特点,同时纳入群组(年龄组)和重复测量,考察某一较长年龄段特质的发展变化过程和特质间的相互影响。McArdle 和 Woodcock (1997)采用测量时间点 PMD 探讨了时间间隔对重测信度的影响,在其设计中所有参与者均参加第一次测试,而只参加其

	子集X	子集A	子集B	子集C		时间点1	时间点2	时间点3	时间点4	时间点5	时间点6	时间点7	时间点8	...
形式1	●	●	●	○	G1 (1/3)	形式1	形式2	形式3	形式1	形式2	形式3	形式1	形式2	...
形式2	●	●	○	●	G2 (1/3)	形式2	形式3	形式1	形式2	形式3	形式1	形式2	形式3	...
形式3	●	○	●	●	G3 (1/3)	形式3	形式1	形式2	形式3	形式1	形式2	形式3	形式1	...

(a) 锚设计以及相应的三形式个体内和个体间均平衡的施测方案

	子集A	子集B	子集C	子集D		时间点1	时间点2	时间点3	时间点4	时间点5	时间点6	时间点7	时间点8	...
形式1	●	●	●	○	G1 (1/4)	形式1	形式2	形式3	形式4	形式1	形式2	形式3	形式4	...
形式2	●	●	○	●	G2 (1/4)	形式2	形式3	形式4	形式1	形式2	形式3	形式4	形式1	...
形式3	●	○	●	●	G3 (1/4)	形式3	形式4	形式1	形式2	形式3	形式4	形式1	形式2	...
形式4	○	●	●	●	G4 (1/4)	形式4	形式1	形式2	形式3	形式4	形式1	形式2	形式3	...

(b) 矩阵抽样设计以及相应的四形式个体内和个体间均平衡的施测方案

图1 多形式计划缺失设计的测验组合形式和相应施测方案示例(以4个题目子集为例)

注: 实心圆点表示施测, 空心圆点表示未施测。G1(1/3)、G2(1/3)和 G3(1/3)表示将被试随机平均分为3组; G1(1/4)、G2(1/4)、G3(1/4)和 G4(1/4)表示将被试随机平均分为4组。

余 7 个时间点中的 1 次测试, 有计划地缺失了一部分参与者的数据。

然而, 很少有研究探讨测量时间点 PMD 中检验力的损失程度, 仅有几项关于长时追踪研究中增长曲线模型的研究关注了该问题。Mistler 和 Enders (2012) 的研究发现, 采用测量时间点 PMD 时, 潜在截距和斜率的平均水平的检验力下降很小。但 Rhemtulla 等人(2014)的研究表明, 与完整设计相比, 测量时间点 PMD 中截距和斜率个体变异的检验力明显下降; 某个时间点内题目水平的缺失对检验力的不利影响小于测量时间点的缺失。对于 ILS, 测量时间点 PMD 可能是最直接有效降低测量次数的方法, 但具体应如何制定有效可行的设计方案, 如何合理分析数据并解释结果, 迄今仍没有可供参考的研究。

2.1.3 两方法测量计划缺失设计

两方法测量计划缺失设计由 Graham 等人(2006)提出, 是指在研究设计阶段, 通过平衡“精贵”与“低廉”两种测量方法的 PMD。两方法测量 PMD 适用于同一个构念有两种测量方法的情境, 其中一种成本昂贵但精度高(简称“精贵”测量), 另一种简便易行、成本低廉但偏差相对较大(简称“低廉”测量)。例如, 吸烟行为可采用检验唾液中可替宁含量的“精贵”测量, 也可采用参与者自陈量表的“低廉”测量。当研究者从“低廉”测量转向“精贵”测量时, 往往需要在时间、材料或设备、专家技术人员或助手、经济成本和侵入性等方面进行权衡(Peri, 2022; Rhemtulla & Hancock, 2016), 经常面临从多种测量方法中进行选择的问题。两方法测量 PMD 通过构造两种或两种以上有效性

和成本不同的测量方法的组合, 可产生比任何一种单独测量方法都更高的效益(Graham et al., 2006)。两方法测量 PMD 的基本思想是对多数参与者采用“低廉”方法以提升统计检验力, 而对少数参与者同时采用“精贵”和“低廉”方法来校正估计偏差, 两者综合以得到更有效的结果。最重要的是, 以往研究发现即使只有小部分随机子集同时采用“精贵”和“低廉”方法测量, 也基本不影响其纠偏能力(Garnier-Villarreal et al., 2014; Graham, 2012; Peri, 2022; Rhemtulla & Hancock, 2016)。随着潜变量建模方法的发展, 很容易在结构方程模型的框架下构建基于两方法测量 PMD 的偏差校正模型(见图 2), 并进行缺失数据的处理和分析。

目前关于两方法测量 PMD 的研究主要集中在横断研究和长时追踪研究, 实证研究和模拟研究的结果一致表明这种 PMD 可以有效提高检验力并降低研究成本。Graham 等人(2006)基于回归模型探讨了不同情况下两方法测量 PMD 的优势及其影响因素, 研究发现在总费用相同时, 相比于完整设计, 所有情况下两方法测量 PMD 都能得到更高的检验力和精度更高的参数估计; “低廉”测试的信度越高, 设计所需“精贵”测试的样本量越少; 两种测试的花费差异越大, 两方法测量 PMD 的效果越好; 回归效应越大, 所需“精贵”测试的样本量越大。此外, 基于回归模型的实证研究也表明, 两方法测量 PMD 能有效提升检验力, 降低研究成本, 且不会降低预测变量对结果变量的预测效度(Drake et al., 2013; Zawadzki et al., 2012)。Garnier-Villarreal 等人(2014)将两方法测量 PMD 推广到长时追踪研究并发现, 当反应偏差由

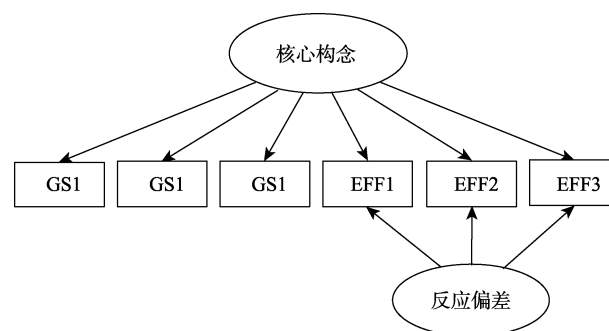


图2 偏差校正模型示意图

注: “精贵”测量方法(GS)包括 GS1、GS2 和 GS3 三个测量指标; “低廉”测量方法(EFF)包括 EFF1、EFF2 和 EFF3 三个测量指标。

随时间变化的因素表示时,每次测量都应该包含“精贵”方法,以实现反应偏差载荷、核心构念载荷和路径系数的精确估计。Peri (2022)基于潜变量增长模型探讨了两方法测量PMD的表现,结果表明对于平均变化趋势的估计,即使在小样本和高缺失率的情况下,两方法测量PMD也能得到无偏的、高检验力的结果,但个体间变异的方差协方差估计仍需更高的样本量以保证参数估计的精度。由此可见,以往研究大多证实两方法测量PMD具有明显的优势。然而,目前在ILS中对这种PMD的探讨尚属空白。考虑到ILS中重复测量次数更多、成本更高、信度分析更复杂等特点(Luo, Hu, & Liu, 2024; Xiao et al., 2023),探讨如何合理分配不同测量方法,以有效地提高检验力和测量精度是更具有挑战性的研究问题。

综上,以往基于横断研究和长时追踪研究的结果证实,PMD具有降低参与者负担、提升数据质量、节约研究成本,乃至提高研究效度等多方面的优势。然而,目前在ILS中对PMD的研究非常少。由于密集追踪与长时追踪关心的问题不同,常用的模型不同,且ILS的重复测量次数远多于长时追踪研究,时间间隔更短,基于长时追踪研究的结果难以推广到ILS。因此,有必要在ILS的框架下,探讨如何设计有效的PMD方案,以解决ILS中测量强度大所带来的参与者负担重、研究成本高等问题。此外,还需基于ILS中数据分析常用的DSEM,分析各种PMD的有效性和适用条件,为应用研究者提供方法支持和实践指导。

2.2 计划缺失设计在数据收集前、后面临的挑战

虽然计划缺失设计能从事先设计的角度有效降低参与者负担,节约研究成本,提升收集到的数据的可靠性和有效性,但是在ILS中,相比于完整设计,PMD收集的数据具有以下两个研究难点。首先,基于PMD的密集追踪数据的缺失比例较高。由于设计上参与者没有完成部分题目或部分时间点的测量,收集的数据中含有较高比例的缺失。其次,基于PMD的密集追踪数据的缺失机制复杂。虽然事先设计可保证PMD中计划缺失的数据满足完全随机缺失(missing completely at random, MCAR)机制或随机缺失(missing at random, MAR)机制,但实际研究中参与者常常还会由于设计以外的各种外部或内部因素产生非计划数据缺失(陈楠,刘红云,2015;唐文清等,

2014; Schafer & Graham, 2002),这可能导致实际收集数据中混杂了其他缺失机制的数据。因此,相比于完整设计,基于PMD的ILS在数据收集前、后都面临挑战。一方面,在数据收集前的方案制定和预注册阶段,需要综合考虑PMD的缺失比例和缺失机制进行样本量规划,并对拟采用的PMD的效率进行评价。另一方面,在数据收集后的数据分析阶段,需要基于适宜ILS的统计分析模型,结合PMD的缺失比例和复杂的混合缺失机制,选择和发展合理高效的缺失数据处理方法。下面主要从含计划缺失数据的检验力分析和样本量规划,以及缺失数据处理方法两个方面综述该领域的研究进展。

2.2.1 计划缺失设计的检验力分析和样本量规划

随着心理学研究者对学术不端和研究可重复性问题的讨论日趋激烈,国内外越来越多的学术期刊推行预注册(pre-registration)制度,以促进科研过程和结果的公开透明,提高研究的可重复性(Liu et al., 2025; Nosek et al., 2022)。预注册时,检验力分析被认为是合理规划样本量的金标准(Schoemann et al., 2014)。传统样本量规划主要是基于虚无假设显著性检验(null hypothesis significance test, NHST)的检验力分析,要求样本量必须使检验力达到预设标准。检验力分析主要包括公式推导方法和基于蒙特卡洛(Monte Carlo, MC)模拟的方法(刘玥等, 2023; Arend & Schäfer, 2019; Moerbeek, 2022)。公式推导方法含有关于分布的强假设,当数据不符合假设时可能得到有偏差的结果(Judd et al., 2017)。当设计中含有缺失数据时,需要在公式中对缺失数据的影响进行校正,校正也可能带来新的偏差。另外,对于嵌套结构的数据和含有多个随机参数的复杂模型,有时很难推导出检验力和样本量计算的公式(Arend & Schäfer, 2019; Lafit et al., 2021)。基于MC模拟的方法采用预设参数并基于特定模型重复生成数据,再基于重复生成的数据估计参数,统计多次重复中显著性结果的比例(Liu et al., 2025)。该方法不需要推导参数分布,适用范围广,容易处理多参数的复杂数据模型(Liu et al., 2019)。更重要的是,对于含缺失数据的设计,MC模拟法只需要在生成数据时按照计划采用的数据收集方案,设置缺失比例和缺失模式直接模拟即可(Schoemann et al., 2014)。考虑到计划缺失的密集追踪数据中,数据

结构复杂, 研究问题涉及的随机参数较多, 同时又有较高比例的缺失数据, 适合采用 MC 模拟法进行检验力分析和样本量规划。已有研究采用 MC 模拟法在密集追踪研究(ILS)中进行检验力分析和样本量规划(Lafit et al., 2021, 2022; Schultzberg & Muthén, 2018)。

然而, 以往关于 ILS 中检验力分析和样本量规划的研究仍存在以下局限: ①现有研究采用的统计分析模型不能完全回答 ILS 研究者感兴趣的研究问题。现有研究仅局限于只包含随机自回归效应的动态结构方程模型(DSEM)或多水平模型, 但在 ILS 中, 研究者往往对不同变量之间的动态双向关系, 即变量间交叉滞后效应(或称溢出效应)更感兴趣(Li et al., 2022; Li et al., 2024; Luo, Liu, & Hu, 2024); ②已有研究无法评价基于 PMD 的 ILS 的样本量规划和设计方案的优劣。目前研究所提供的分析软件等工具均没有考虑缺失数据对检验力的影响, 也就无法评估 ILS 中不同 PMD 方案的优劣。③随着模型复杂程度增加, 已有检验力分析和样本量规划方法的计算量呈几何级增长, 分析效率低。当统计模型涉及多维参数时, 其组合变得异常复杂, 为了尽可能考虑不同条件的组合, MC 模拟方法需要进行大量计算, 这导致计算量巨大且效率低下(Richter et al., 2022), 对于适用范围更广更灵活的 DSEM 建模存在现实可行性上的挑战。

为了解决上述第三点局限, 近期有研究者引入了基于机器学习预测的代理建模框架, 针对 t 检验、方差分析、含缺失数据多重插补的回归模型等 9 种模型优化多维参数情境下的检验力分析(Zimmer & Debelak, 2025), 并开发 R 软件包 *mlpwr* (Zimmer et al., 2024)。然而, 在其考虑的 9 种模型中, 没有适用于 ILS 分析的模型(如 DSEM)。而且, 关于含缺失数据的样本量规划, 他们的研究只考虑了 MCAR 的比例, 并没有结合不同类型 PMD 的系统探讨。因此, 针对 ILS 感兴趣的研究问题, 有必要在更一般的 DSEM 框架下, 结合 PMD 预设的缺失比例和缺失机制, 借鉴机器学习领域一些高效率的算法, 探讨检验力分析和样本量规划的问题, 以有效地评价不同的 PMD 方案。

2.2.2 计划缺失设计的缺失数据处理方法

关于缺失数据处理方法的研究大多基于具体的研究问题, 结合相适宜的统计分析模型, 通过

评估不同缺失处理方法在各种缺失比例和缺失机制下对模型关心参数的推断统计结果的影响, 来选择适用于当前问题和模型的缺失数据处理方法(刘红云, 张雷, 2005; Li et al., 2017)。设计上计划缺失满足 MCAR 的假设(Little & Rubin, 2002; Silvia et al., 2014; Vicente, 2023), 追踪研究中的计划缺失满足 MCAR 或 MAR 的假设(Peri, 2022; Xu & Logan, 2024)。但由于 ILS 中实际收集到的数据可能存在混合的缺失机制, 因此, 在实际数据分析阶段不仅需要考虑能处理较高比例缺失数据的方法, 还应该考虑更复杂多样的数据缺失机制。

Schafer 和 Graham (2002)以及 Enders (2025)对缺失数据处理方法的发展进行了详尽的回顾。首先, 传统的缺失数据处理方法主要有删除不完全个案法(如列删除和对删除)和单一插补法(如均值插补和回归插补)。这类方法适用于 MCAR 缺失机制, 但大量研究表明这类方法会导致标准误增大, 检验力下降(Enders, 2011)。因此, 即使 PMD 中缺失数据满足 MCAR, 也不推荐使用删除不完全个案法或单一插补法(Schafer & Graham, 2002)。这类方法更不适用于 ILS 中 PMD 收集到的较高比例的缺失数据的处理。

其次, 现代的缺失数据处理方法主要有全息极大似然(full information maximum likelihood)估计和多重插补(multiple imputation)法。两种方法对 MCAR 或 MAR 缺失机制都能得到无偏的估计结果, 且在 MCAR 机制下, 两种方法比传统的删除不完全个案法和单一插补法得到的参数估计更准确, 检验力也更高(Enders, 2011)。以往大量研究证明, 全息极大似然估计和多重插补法在处理缺失数据时存在优势(Ji et al., 2018; Xu & Logan, 2024; Yuan & Bentler, 2000)。在密集追踪的情景下, Ji 等人(2018)的研究发现多重插补法在处理 ILS 中的缺失数据时仍有较好的表现。但也有研究者指出, ILS 中重复测量次数较多, 而被试人数相对较少, 全息极大似然估计和多重插补法可能会遇到模型收敛的问题(Enders, 2025; Liu et al., 2024)。概括而言, 基于 PMD 的大量研究表明在横断研究和长时追踪研究常用的模型中, 全息极大似然估计和多重插补法表现较好, 但目前仍缺乏在 ILS 中结合其分析常用的 DSEM 建模框架对这两种方法的可行性和在各种 ILS 条件下的适用性的系统探究。

另外, 缺失数据处理方法近20年的新进展主要体现在因式回归设定 (factored regression specifications) 法的应用 (Ibrahim, 1990; Ibrahim et al., 2002; Keller, 2022; Lüdtke et al., 2020)。这种方法运用概率链规则将联合分布分解为一系列单变量的条件分布, 是一种基于模型的缺失数据插补方法, 可与全息极大似然估计、多重插补法和贝叶斯估计结合使用。它解决了以往方法依赖于多元正态分布假设的问题 (Enders, 2025)。传统极大似然估计在处理缺失协变量时通常依赖于对协变量的联合分布作多元正态分布假设。但当模型中包含非线性效应或交互项时, 这一假设往往不成立, 进而导致估计偏差。而因式回归设定则通过概率链规则将完整数据的联合分布因式分解为“目标模型”(即研究者关心的模型)与“协变量模型”(即缺失预测变量的分布模型)两个部分, 而协变量模型进一步被分解为一系列单变量的条件分布, 每一个条件分布可以根据变量类型灵活设定, 可与全息极大似然估计、多重插补法和贝叶斯估计结合使用, 从而使得因式回归设定能适应分类变量、非正态连续变量等复杂场景 (Enders, 2025)。研究者将其应用于心理学研究, 探讨了该方法在多水平模型中的应用, 结果表明在大多数情况下这种方法可以得到近似无偏的估计结果 (Du et al., 2022; Enders et al., 2020; Keller, 2022)。Alacam 等人 (2025) 将其用于题目水平的多形式 PMD 研究, 研究表明因式回归设定法具有高灵活性和可拓展性等优势。可见, 因式回归设定法的发展为复杂模型、类别数据、非正态数据以及多维题目层面的缺失数据处理提供了新思路和新方法。

此外, 结合 ILS 数据中时间相依性的特点, 近年来有研究者将时间序列分析中处理缺失数据的卡尔曼滤波法引入密集追踪数据的分析模型 (Hamaker & Grasman, 2012)。Asparouhov 等人 (2018) 采用卡尔曼滤波法处理 ILS 中时间间隔不均等的问题, 并探讨插补缺失值比例对参数估计精度的影响。结果表明在有效测量次数较高时 (如 60 次), 即使缺失比例很高 (如 80%), 仍能得到较为准确的参数估计结果。此外, 在采用 DSEM 分析密集追踪数据时, Mplus 软件因其已易操作、易理解等特点, 备受应用研究者青睐 (Muthén & Muthén, 1998-2017)。可见, 应用研究者对 ILS 中

缺失数据处理技术需求的增长速度远超过了这些方法本身的完善和发展的速度。目前亟需方法学研究对卡尔曼滤波法和其他更常用且适用范围更广的缺失数据处理方法 (即, 全息极大似然估计和多重插补法) 进行比较分析, 并提供更契合 ILS 应用研究需求的实践建议和实用工具。

值得说明的是, 与 Ibrahim 等 (2001) 提出的用于处理 MNAR 数据的联合建模方法 (joint multinomial model) 相比, 因式回归设定法在适用范围与建模策略上存在差异。Ibrahim 等人 (2001) 所提出的方法基于广义线性混合模型框架, 通过 Monte Carlo EM 算法应对复杂的非单调缺失情况, 并构建灵活的多项式分布形式以刻画缺失机制, 从而使模型能够处理缺失依赖于自身值的 MNAR 情形。该方法在理论上具有高度的通用性和严谨性, 但其实现依赖于高维积分与 Gibbs 采样, 计算成本较高, 尤其在样本量大或模型中包含多个随机效应时, 常面临收敛速度慢与估计效率低的问题。相较而言, 因式回归设定法虽然建立在相对较弱的 MAR 假设之上, 但具有计算效率高、模型设定灵活等优势, 能够适配多种变量类型 (如分类变量、偏态连续变量) 及复杂模型结构 (如交互项或非线性效应)。在本研究所涉及的包含交互项的回归模型中引入因式回归设定法, 为提升协变量缺失情形下的参数估计精度提供了切实可行的解决方案。

通过对缺失数据处理方法的系统回顾发现, 全息极大似然估计和多重插补法是被验证效果较好且运用较广的缺失数据处理方法, 因式回归设定法则是近年来缺失数据处理中极具应用潜力的新发展方法, 而卡尔曼滤波法是 ILS 中缺失数据处理常用的默认方法。针对密集追踪数据分析中越来越广受运用的 DSEM, 尚无研究系统比较不同方法的表现及其适用条件, 也无研究探讨因式回归设定法在此建模框架下的拓展与运用。因此, 从缺失数据处理方法发展和改进的角度, 亟待将因式回归设定法拓展运用于 DSEM 分析, 探讨适用性更强、更灵活的缺失数据分析新方法。从缺失数据处理方法选择的角度, 应在 DSEM 的框架下对全息极大似然估计、多重插补法、Mplus 默认的卡尔曼滤波法和拓展的因式回归设定法的表现进行系统比较, 为基于 PMD 的密集追踪应用研究提供更实用高效的缺失数据处理方法。

3 问题提出

以往研究表明,密集追踪研究(ILS)中较高的测量强度可能会对参与者的主观感受、收集数据的质量和研究结果的准确性等产生消极影响。计划缺失设计(PMD)能从研究设计的源头有效降低被试负担和结果偏差,是减少测量强度消极影响的首选策略(Ward & Meade, 2023)。虽然 PMD 在传统的横断和长时追踪研究中已有较多的运用,但仍缺乏在密集追踪情境下对各种 PMD 方案的适用性和有效性的系统探析。

由于基于 PMD 的收集数据存在较高的缺失比例和复杂的缺失机制,基于 PMD 的 ILS 在数据收集前、后都面临挑战。对于数据收集前样本量的确定,尚无研究在 ILS 常用的动态结构方程模型(DSEM)建模框架下,探究有缺失数据的样本量规划问题。因此,仍需要对基于 PMD 的 ILS 开发合理规划样本量的实用方法,并优化解决传统方法计算量大、难以得到最优解等问题。对于数据收集后缺失数据的处理,仍没有研究探讨如何将常用的全息极大似然估计和多重插补法以及基于因式回归设定法拓展运用于 ILS 中 DSEM 框架下的缺失数据处理,也缺乏对这些方法和 Mplus 中默认设置且实际应用较多的卡尔曼滤波法的比较分析。因此,有必要切实考察基于 PMD 的 ILS 中缺失数据的复杂性,为其缺失数据处理开发和选用更实用高效的分析方法。

综上,本研究旨在通过对基于多种 PMD 的 ILS 中数据收集前、后全流程的综合探究,完善基于 DSEM 的检验力分析和样本量规划方法和工具,拓展复杂模型中缺失数据的处理方法和实践应用,最终有效解决 ILS 中参与者负担重、依从性差、数据质量低等问题。此外,本研究还将通过实用程序开发和实证案例演示,推动 PMD 在心理学研究中的实践应用,丰富国内外有关 ILS 中数据收集优化方案的方法与实践指导,助力心理学问题的细化深入研究和科学系统考察,为心理学的生态系统论提供更丰富的方法技术支撑和应用研究支持。

4 研究构想

本研究旨在探讨如何将计划缺失设计(PMD)思想应用于密集追踪研究(ILS),通过优化密集追

踪数据收集方案(包括测量设计方案和样本量规划)和改进缺失数据处理方法,解决如何更合理、高效收集数据,更科学、有效分析数据的问题。本研究将在 ILS 常用的 DSEM 框架下,采用贝叶斯参数估计方法探究数据收集前的设计方案优化和数据收集后的缺失数据处理方法改进,并提供数据收集前、后全流程的方法指导和实操工具。研究整体上采用理论和应用相结合的方法,研究内容框架如图 3 所示。

研究 1~2 聚焦于数据收集前阶段,旨在比较各类 PMD 中的有效测量方案,并基于检验力分析提出样本量规划方法。研究 3 则关注数据收集后阶段,重点探讨基于 PMD 的 ILS 中缺失数据处理方法的改进。研究 4 则基于实证研究,系统演示从数据收集方案设计、数据处理分析到 DSEM 结果解读的完整流程。

4.1 研究 1: 密集追踪研究中计划缺失设计方案的比较及其选择建议

研究 1 主要围绕测量方案设计,关注密集追踪数据收集中 PMD 方案的制定、有效性分析、不同方案的比较和选择建议的问题。本研究将从 DSEM 参数估计精度和检验力分析等方面,评价不同 PMD 方案的适用条件和影响因素。基于常见的 3 种 PMD 类型和预期服务于应用实践的目的,研究 1 分为以下 4 个子研究。

4.1.1 研究 1-1: 多形式计划缺失设计方案及其有效性比较

研究 1-1 的主要目的是基于多形式 PMD,探讨有效缩短问卷长度的密集追踪数据收集方案,以及不同方案的适用条件和优劣比较。首先,给出 ILS 数据收集中的多形式 PMD 方案。PMD 方案的制定主要考虑 3 个设计因素:问卷长度、多形式子测验的组合方式和子测验的施测方式。然后,分析不同方案的有效性和适用条件。考虑的有效性影响因素主要有重复测量次数、被试人数、题目的因素载荷、自回归系数和交叉滞后系数的均值和方差。在每种模拟条件下,研究先基于双潜变量动态结构方程模型(DSEM)生成满足模型参数设定的完整数据,再根据相应的 PMD 方案生成每种方案对应的含计划缺失的数据,重复生成 500 次。研究将基于双潜变量 DSEM 分析数据,采用贝叶斯估计对模型中的参数进行估计,并使用默认的卡尔曼滤波法处理缺失数据,同时以完整

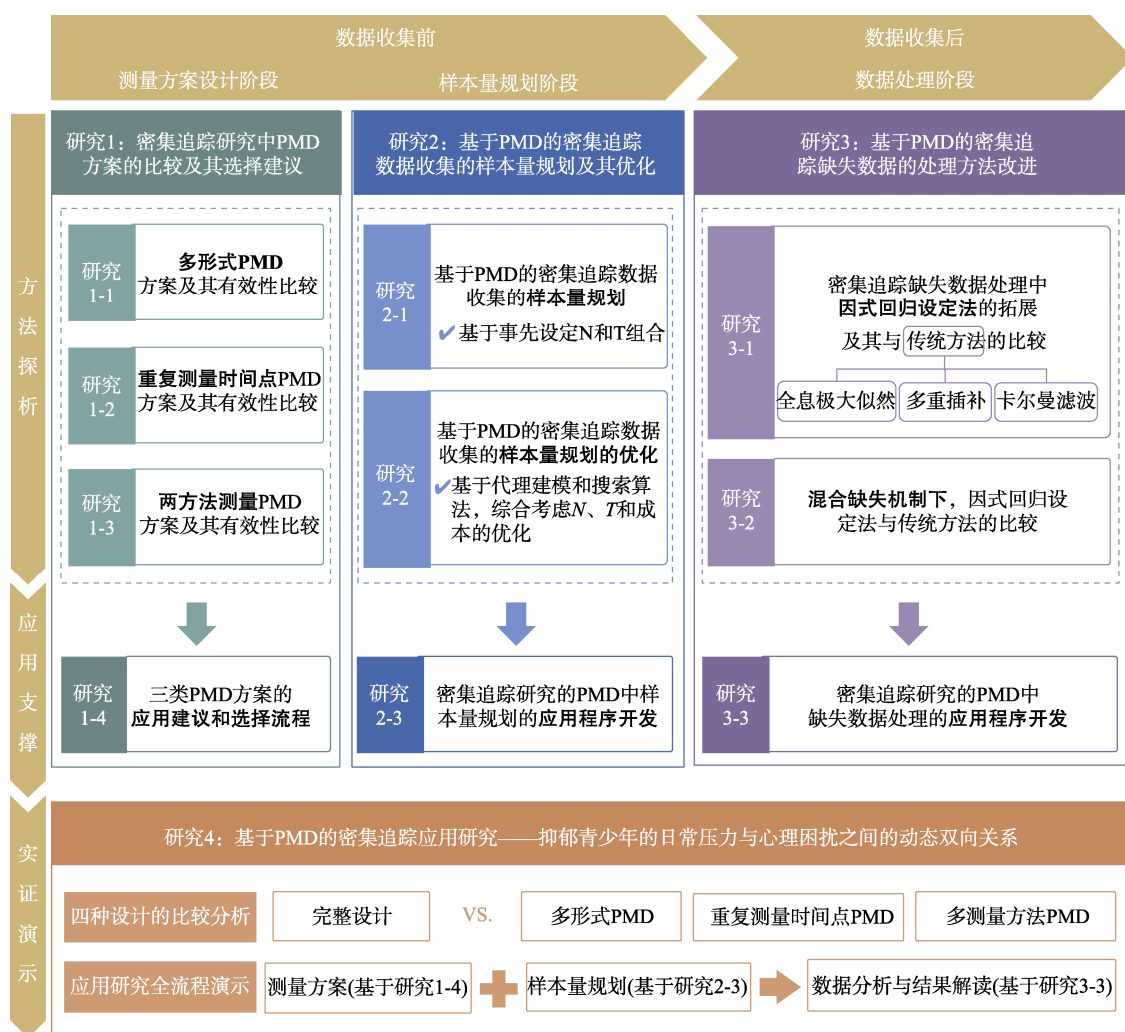


图3 研究内容框架图

数据分析结果作为比较基准。研究主要从参数估计的相对偏差、标准误、均方根误差、置信区间覆盖比率、第 I 类错误、统计检验力等方面评价模型参数估计的准确度以及不同多形式 PMD 方案的有效性。研究 1-1 将为 ILS 的应用研究者提供多形式 PMD 方案的选择依据。

4.1.2 研究 1-2: 重复测量时间点计划缺失设计方案及其有效性比较

研究 1-2 的主要目的是基于重复测量时间点 PMD, 探讨有效降低测量次数的测量设计方案, 并比较不同方案的优劣及其适用条件。首先, 制定重复测量时间点 PMD 方案。借鉴已有时间序列分析中缺失数据处理的相关研究结论 (Liu &

Molenaar, 2014) 和预研究结果, ILS 中重复测量时间点 PMD 方案拟采用两阶段设计, 其中数据收集的初始阶段 (第一阶段) 采用完整设计, 数据收集的中后期阶段 (第二阶段) 采用测量时间点 PMD。对于第二阶段的测量时间点 PMD, 拟考虑平衡设计和个体一致设计两种方案。此外, 拟考虑的设计因素有第一阶段完整数据所占比例和第二阶段缺失数据所占比例。然后, 分析不同方案的有效性和适用条件。关心的结果评价因素包括被试人数、重复测量次数、自回归系数和交叉滞后系数的均值和方差。考虑的有效性影响因素、数据生成方式、分析方法和评价指标与研究 1-1 相同。研究 1-2 将为 ILS 的应用研究者提供重复测量时

间点 PMD 方案的选择依据。

4.1.3 研究 1-3: 两方法测量计划缺失设计方案及其有效性比较

研究 1-3 的主要目的是基于两方法测量 PMD, 探讨研究总成本固定的情况下, 如何最优分配两种测量方法的样本量, 探讨提高参数估计精度和检验力的数据收集方案, 并评价不同方案的优劣及其适用条件。首先, 借鉴横断研究和长时追踪研究的潜变量偏差校正模型(Graham et al., 2006; Xu & Logan, 2024), 将其扩展到 ILS 的两方法测量 PMD 中。其次, 基于固定的总成本制定多种两方法测量 PMD 方案, 考虑两种测量方法费用的差异、缺失数据比例和重复测量次数三个设计因素。然后, 分析不同方案的有效性和适用条件。考虑的影响因素有两种测量方法的信度、测量指标个数、自回归系数和交叉滞后系数的均值和方差。关心的影响因素包括重复测量次数、“精贵”方法和“低廉”方法的测量信度、两种方法测量指标的个数以及自回归系数和交叉滞后系数的均值和方差。数据生成、数据分析和评价指标与研究 1-1 相同。研究 1-3 将为 ILS 的应用研究者提供两方法测量 PMD 方案的选择依据。

4.1.4 研究 1-4: 三类计划缺失设计方案的应用建议和选择流程

研究 1-4 的主要目的是基于研究 1-1、研究 1-2 和研究 1-3 的结果, 从应用研究的角度, 提供 ILS 中采用 PMD 方案的建议, 以及制定测量方案时应考虑的影响因素和一般流程。首先, 整合分析研究者如何根据研究目的、测量工具特点和研究成本, 选择适合的 PMD 类型。其次, 对于每种类型的 PMD, 基于模拟研究中有效性比较和影响因素的分析结果, 给出 PMD 方案制定和选择的建议, 以及研究实践中需要注意的问题。最后, 结合应用研究的实际情况, 从研究者和研究参与者两个角度评价各种测量方案的可行性。研究 1-4 将为 ILS 的应用研究者提供如何制定 PMD 方案的操作化流程和选择建议。

4.2 研究 2: 基于计划缺失设计的密集追踪数据收集的样本量规划及其优化

当测量方案确定后(研究 1 的内容), 研究者还需确定研究的样本量(在 ILS 中, 包括被试人数(N)和重复测量次数(T)。研究 2 重点关注密集追踪数据收集前, 如何基于检验力分析, 结合不同类

型 PMD 方案, 确定被试人数和重复测量次数。研究将在动态结构方程模型(DSEM)的框架下, 采用蒙特卡洛(MC)模拟的方法, 解决样本量的规划、优化及其实现的问题。研究 2 分为以下 3 个子研究。

4.2.1 研究 2-1: 基于计划缺失设计的密集追踪数据收集的样本量规划

研究 2-1 的目的是在 3 类 PMD 中, 基于研究 1 推荐使用的设计方案, 探究在 DSEM 框架下解决含计划缺失数据的样本量规划问题。以往关于 ILS 中样本量规划的研究都未考虑缺失数据的影响, 但含缺失数据是 PMD 的特点。因此, 研究 2-1 的关键问题在于如何在样本量规划中纳入计划缺失数据对参数估计和检验力的影响。此外, 考虑到 ILS 对个体内变量间动态关系及其个体差异的关注, 如何将样本量规划的方法拓展到应用研究普遍采用的 DSEM 也是研究 2-1 拟重点考察的问题。研究 2-1 将在更灵活的统计模型框架下, 为应用研究者提供含计划缺失数据的 ILS 的样本量规划方法。

研究 2-1 拟采用 MC 模拟法, 将检验力分析和样本量规划拓展到密集追踪 PMD 中, 并基于 DSEM 及其简化模型实现样本量的规划。研究借鉴 Arend 和 Schäfer (2019)以及 Schultzberg 和 Muthén (2018)的方法, 基于检验力分析确定 ILS 中的样本量, 其不同之处主要在于需考虑设计阶段计划缺失数据对模拟数据生成和处理的影响, 以及 ILS 中应用研究更普遍采用的 DSEM。研究的基本思路和步骤如下: (1)选择模型并设定参数; (2)确定计划缺失设计方案; (3)选择缺失数据处理方法; (4)模拟生成重复样本; (5)分析重复样本并计算检验力; (6)确定样本量。

4.2.2 研究 2-2: 基于计划缺失设计的密集追踪数据收集的样本量规划的优化

研究 2-1 虽然可以解决 PMD 中的样本量规划问题, 但由于 DSEM 常常包括较多的参数且其估计速度较慢, 随着研究希望考察的 N 和 T 组合数的增加, 样本量规划所需的计算量将非常大。更重要的是, 传统的样本量规划思路基于事先设定的 N 和 T 的组合计算检验力, 可能会由于设定的局限而无法得到最优化的样本量规划方案。另外, 在考虑研究经费的设计中, 随着参数的增多, 样本量规划的实现将更加困难。近年来, 机器学习

方法的快速发展以及在各个领域的拓展应用,为解决多变量和多参数问题提供了全新且有效的思路。研究 2-2 拟借鉴 Zimmer 和 Debelak (2025)的研究,将机器学习中基于代理模型预测的思想和搜索算法应用于检验力和样本量规划问题,提出基于代理建模和搜索算法的样本量规划的优化思路,探讨高效的预测代理模型,重点解决研究 2-1 遇到的计算量大、可能难以获取最优解的问题,最终实现 ILS 中 PMD 数据收集方案的优化设计。研究 2-2 不仅可以为应用研究者提供更高效率的样本量规划新算法,还能进一步优化基于 PMD 的数据收集方案。

4.2.3 研究 2-3: 密集追踪研究的计划缺失设计中样本量规划的应用程序开发

研究 2-3 旨在从实际应用层面解决 ILS 中复杂模型的检验力分析和样本量规划问题,同时解决数据收集方案优化设计和应用程序操作实现的问题。研究 2-3 拟基于研究 2-2,针对含计划缺失的密集追踪数据收集中样本量规划及其优化的实际需求,借鉴 Zimmer 等人(2024)开发 *mlpwr* 软件包的思路和 *simsem* 包(Pornprasertmanit et al., 2013)模拟生成缺失数据的思路,将其拓展到动态结构方程模型(DSEM)及其系列简化模型和含有缺失数据设计的情境中。借助 Shiny 框架,开发简单易用、界面友好、交互性和可视化程度高、适用范围广的动态交互式的检验力和样本量规划的应用程序,提供使用手册和案例演示,以帮助应用研究者合理规划密集追踪研究的数据收集方案。

4.3 研究 3: 基于计划缺失设计的密集追踪缺失数据的处理方法改进

研究 3 首先将近年来新发展的具有潜力的因式回归设定法(Enders, 2025)拓展到 DSEM 中,并将其与已有常用方法进行比较。此外,研究 3 还将切实考虑 ILS 中可能存在的混合缺失数据机制,以进一步检验新拓展方法的表现,分析其与已有方法的比较。最后,从应用研究的角度,开发适用于密集追踪缺失数据处理的应用程序。研究 3 分为以下 3 个子研究。

4.3.1 研究 3-1: 密集追踪缺失数据处理中因式回归设定法的拓展及其与传统方法的比较

研究 3-1 的目的是将因式回归设定法拓展到 DSEM,提供含计划缺失的 ILS 中更优的缺失数据处理新方法。研究 3-1 的主要内容包括针对 3

种类型 PMD,将因式回归设定法在题目层面(多形式 PMD)、测量时间点层面(重复测量时间点 PMD)和被试层面(两方法测量 PMD)进行拓展;并将其与传统的全息极大似然估计、多重插补法、以及线性卡尔曼滤波法进行比较,考察计划缺失比例、重复测量时间点数量和被试人数的影响,考察新方法的优势及其适用条件。

为了将因式回归设定法拓展应用于 DSEM,本研究根据 DSEM 中随机参数较多的特点,拟将因式回归设定法与贝叶斯估计相结合。基于马尔科夫蒙特卡洛(MCMC)算法,采用 Gibbs (或 Metropolis-Hastings) 抽样实现参数和缺失数据的迭代抽取。以双潜变量 DSEM 为例,基于因式回归设定法的缺失数据插补的思路和流程如下:

①设定因式回归分解模型:基于概率链规则将 DSEM 中多变量的联合分布分解为一系列单变量条件分布的乘积。每个条件分布可视为实现迭代插补过程所需关注的回归模型。例如,条件分布 $f(Y|X_1, X_2)$ 可通过以 Y 为因变量、 X_1 和 X_2 为自变量的回归得到 Y 的分布。②基于 MCMC 算法,采用 Gibbs (或 Metropolis-Hastings) 抽样抽取后验分布并插补缺失数据。③输出插补后的数据集:检查模型收敛情况和缺失变量的后验分布,基于后验分布,抽取并导出插补后的数据集。

采用 MC 模拟研究法,结合 3 种不同类型 PMD,验证因式回归设定法的有效性,并与已有表现较好的方法(如全息极大似然估计、传统的基于模型的多重插补、线性卡尔曼滤波和因式回归设定法)进行比较。基于的模型、完整数据的生成思路以及所考虑的 DSEM 模型的真值条件与对应的研究 1 相同,评价指标也与研究 1 相同。

4.3.2 研究 3-2: 混合缺失机制下,因式回归设定法与传统方法的比较

研究 3-2 切实考虑实际 ILS 中更复杂的数据缺失机制,拟探讨既有计划缺失又有非计划缺失的混合缺失机制下,新拓展的因式回归设定法的表现及其适用条件,并进一步将其与传统的全息极大似然估计、多重插补法和线性卡尔曼滤波法进行比较。研究 3-2 在研究 3-1 的基础上,采用模拟研究方法,在更复杂、更接近真实 ILS 的条件下,即将在计划缺失的基础上加入非计划缺失机制(MCAR、MAR、MNAR 和混合机制)和各种非计划缺失数据所占比例(10%、20%和 30%)两个影

响因素时,拓展检验因式回归设定法在参数估计和检验力等方面的表现,以及与其他缺失数据处理方法的比较,重点探讨缺失机制和缺失比例对不同处理方法的影响。研究 3-2 将进一步为应用研究者提供实际 ILS 中缺失数据处理方法的选择依据和建议。

4.3.3 研究 3-3: 密集追踪研究的计划缺失设计中缺失数据处理的应用程序开发

研究 3-3 旨在结合应用研究的现实情况,为以往常用和新拓展的缺失数据处理方法提供选择依据,为各种方法的应用实现提供操作工具。本研究拟从应用研究的角度,基于研究 3-1 和研究 3-2,借鉴 Alacam 等人(2025)利用 R 软件包 *Blimp* (Keller & Enders, 2021)和 *Mplus* 8.10,在 DSEM 的建模框架下,采用 R 语言实现缺失数据插补和分析的流程,开发适用于 ILS 中缺失数据处理分析的应用程序,并提供相关应用案例和操作手册。同时,基于不同缺失数据处理方法的比较结果,完善研究 2-3 应用程序中缺失数据处理方法的相关内容。

4.4 研究 4: 基于计划缺失设计的密集追踪应用研究——抑郁青少年的日常压力与心理困扰之间的动态双向关系

为了帮助应用研究者更深入地理解如何基于 PMD 开展 ILS,研究 4 将选取密集追踪情境下有代表性的心理学研究问题,收集相关实证数据并完成数据分析,对完整设计和 3 种 PMD (多形式 PMD、重复测量时间点 PMD 和多测量方法 PMD)进行比较分析,对测量方案设计、样本量规划到数据分析与结果解读的应用研究全流程进行操作演示。通过具体案例分析,研究 4 将展现如何将研究 1~研究 3 中的方法结论和程序软件应用于实际研究,促进 PMD 在 ILS 中的运用与实践。

研究 4 针对至少有轻微抑郁症状的青少年,选取日常压力与心理困扰这两个密集追踪情境中有较多实证基础并广受关注的研究变量,采用基于 PMD 的 ILS 探究目标群体中两者的动态双向关系,为青少年抑郁筛查预警与干预实践提供实证依据。随着心理学研究问题的不断深入和新兴技术方法的快速发展,越来越多的研究者开始探讨压力与个体的行为状态之间的动态交互过程。心理困扰(psychological distress)作为抑郁与焦虑状态在日常生活中的重要体现,它与个体日常压力

之间的相互影响也受到许多研究的关注与探讨(Goldring & Bolger, 2021)。

(1)设计方案

研究 4 拟招募 4 组至少有轻微抑郁症状的青少年被试,分别参加①完整设计(即无计划缺失)、②多形式 PMD、③重复测量时间点 PMD 和④多测量方法 PMD。根据研究 1 的方法选择与应用建议,确定多形式 PMD、重复测量时间点 PMD 和多测量方法 PMD 的设计方案。

(2)研究对象

根据研究 2 中样本量规划的方法指导和实用程序确定各种设计所需的样本量和重复测量时间点数量(包括测量天数和每天的重复测量次数)。在被试招募阶段,要求青少年被试完成 PHQ-9 (nine-item Patient Health Questionnaire; Kroenke et al., 2001)。该问卷包括 9 道测量抑郁症状严重程度的题项,要求被试根据过去两周的感受,从 0 (“完全没有”)到 3 (“几乎每天”)进行评价。问卷总分 ≥ 5 表示至少有轻微抑郁症状。

(3)测量工具

在完整设计、多形式 PMD 和重复测量时间点 PMD 中,采用自陈式量表测量青少年的日常压力和心理困扰。对于日常压力,采用改编适用于生态瞬时评估的压力感知量表(EMA-adapted Perceived Stress Scale; Murray et al., 2023)进行测量。对于心理困扰,采用适用于密集追踪情境的 PHQ-4 (four-item Patient Health Questionnaire)进行测量(Kroenke et al., 2009)。在多测量方法 PMD 中,除了上述自陈式量表(“低廉”测量),还将通过心率带记录参与者在研究期间的心率变化,并基于心率变异性(heart rate variability, HRV; Kim et al., 2018; Thayer et al., 2012)对个体日常压力感受进行生理测量(“精贵”测量)。本研究基于心率变异性的时域特征和频域特征,将心率变异性具体操作化为以下两个指标:RMSSD (相邻 R-R 间隔差值的均方根信号)和 HF-HRV (心率变异性的高频信号,即 0.15~0.4 Hz),每个指标首先按 5 分钟分段计算,再基于每种设计中两次连续自陈量表的测量时间间隔,分别计算所有 5 分钟指标结果的平均值,作为对该时间间隔内,个体压力程度的生理测量指标结果。

(4)数据分析计划

在数据分析阶段,首先采用研究 3-1 和研究

3-2 推荐的缺失数据处理方法,运用研究 3-3 开发的实用程序处理各种设计中的缺失数据,再采用双变量 DSEM (McNeish & Hamaker, 2020)估计日常压力和心理困扰的自回归效应和两者间的交叉滞后效应。对 4 种设计的结果分别进行合理的解读,并对 3 种 PMD 的结果与完整设计的结果进行比较分析。

5 模型建构与创新

合理高效的研究设计方案是提升数据质量、保证研究结果可靠和有效的前提。密集追踪研究 (ILS)可推动心理学研究问题从个体间视角向个体内视角的转变,为从个体内层面深入理解心理的发生、发展和变化过程提供有效的方法,为科学实践中个性化干预方案的制定提供可靠的依据。然而,ILS 中较高的测量强度可能会对其收集数据的质量等产生诸多负面影响。如何优化密集追踪数据的收集和分析方案,从源头上解决测量强度过高的潜在风险,目前尚缺乏系统的研究。本研究可以从研究方法及其应用的层面弥补这方面的不足。

本研究的主要特色在于,研究选题的提出密切结合心理学研究的实际需求和面临的实际问题。本研究从 ILS 实施中遇到的测量强度过高和数据质量下降等实际问题出发,结合 ILS 设计的特点,旨在基于计划缺失设计(PMD),发展出更经济实用且能有效降低参与者负担的数据收集方案。同时,本研究还将重点突破 PMD 方案中的样本量规划和缺失数据处理等关键技术问题。这不仅可以促进心理学研究的设计方法和分析技术的发展,更能帮助心理学应用研究者在更加科学、合理、高效的 ILS 设计方案下,探讨个体特质发生、发展的动态规律,为心理学理论的不断发展和深入探讨提供方法支持。

本研究的创新点主要体现在四个方面。

(1)在密集追踪数据收集方案的设计思路上具有创新性。结合 ILS 设计的特点和数据收集过程中遇到的主要问题,提出 PMD 的思路,拟从研究设计的源头提出有效减小测量强度并提升数据质量的数据收集方案。这不仅可以丰富数据收集方案设计的理论方法研究,还可以有效提高实证研究的效率。

(2)在推动交叉学科领域方法的拓展迁移与整

合应用方面具有创新性。围绕 ILS 的测量方案设计中样本量规划遇到的计算量巨大且难以得到最优规划方案的问题,探索将机器学习预测模型、搜索算法与动态结构方程模型(DSEM)整合应用的技术路线。这不仅可为优化样本量规划提供新方法,还可以促进不同学科前沿技术的交叉融合,发展拓宽传统心理学量化研究的学科边界。

(3)在缺失数据处理新方法的改进和应用上具有创新性。本研究拟拓展的因式回归设定法可有效解决数据分布非正态、类型不统一,以及 ILS 统计分析模型中随机参数多、不容易收敛等问题。该方法与贝叶斯估计的有效整合,为缺失比例高、缺失机制复杂、样本量小、变量个数多的计划缺失密集追踪设计提供了更切合实际的解决方法和技术手段。

(4)在实践应用领域具有创新性。关于 PMD 方案的有效性及其影响因素的研究分析,可为应用研究中密集追踪数据收集方案的选择提供可靠依据;优化的样本量规划方法和应用程序可为应用研究者提供样本量规划和检验力计算的新方法;真实案例研究的设计和分析中所使用的方法、工具、以及所探讨的问题,都属于心理学实践应用领域的创新性探索。

参考文献

- 陈楠,刘红云.(2015).基于增长模型的非随机缺失数据处理:选择模型和极大似然方法.《心理科学》,38(2),446-451.
- 刘红云,王丽娟,李若璇.(2025).追踪数据分析方法及其应用.北京师范大学出版社.
- 刘红云,张雷.(2005).追踪数据分析方法及其应用.教育科学出版社.
- 刘玥,方梵,刘红云,雷怡.(2023).混合效应均值-方差模型的建构和样本量规划探索.《心理科学进展》,31(6),958-969.
- 罗晓慧,刘红云.(2024).密集追踪研究中测验信度的估计:多层结构和动态特性的视角.《心理科学进展》,32(4),700-714.
- 唐文清,张敏强,黄宪,张嘉志,王旭.(2014).加速追踪设计的方法和应用.《心理科学进展》,22(2),369-380.
- 王孟成,叶浩生.(2014).计划缺失设计——通过有意缺失让研究更高效.《心理科学进展》,22(6),1025-1035.
- 郑舒方,张沥今,乔欣宇,潘俊豪.(2021).密集追踪数据分析:模型及其应用.《心理科学进展》,29(11),1948-1969.
- Alacam, E., Enders, C. K., Du, H., & Keller, B. T. (2025). A factored regression model for composite scores with

- item-level missing data. *Psychological Methods*, 30(3), 462–481. <https://doi.org/10.1037/met0000584>
- Arend, M. G., & Schäfer, T. (2019). Statistical power in two-level models: A tutorial based on Monte Carlo simulation. *Psychological Methods*, 24(1), 1–19. <https://doi.org/10.1037/met0000195>
- Asparouhov, T., Hamaker, E. L., & Muthén, B. (2018). Dynamic structural equation models. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(3), 359–388. <https://doi.org/10.1080/10705511.2017.1406803>
- Black, A. C., Harel, O., & Matthews, G. (2012). *Techniques for analyzing intensive longitudinal data with missing values*. In M. R. Mehl & T. S. Conner (Eds.), *Handbook of research methods for studying daily life* (pp. 339–356). The Guilford Press.
- Bolger, N., Davis, A., & Rafaeli, E. (2003). Diary methods: Capturing life as it is lived. *Annual Review of Psychology*, 54(1), 579–616. <https://doi.org/10.1146/annurev.psych.54.101601.145030>
- Drake, K. M., Longacre, M. R., Dalton, M. A., Langeloh, G., Peterson, K. E., Titus, L. J., & Beach, M. L. (2013). Two-method measurement for adolescent obesity epidemiology: Reducing the bias in self-report of height and weight. *Journal of Adolescent Health*, 53(3), 322–327. <https://doi.org/10.1016/j.jadohealth.2013.03.026>
- Du, H., Alacam, E., Mena, S., & Keller, B. T. (2022). Compatibility in imputation specification. *Behavior Research Methods*, 54(6), 2962–2980. <https://doi.org/10.3758/s13428-021-01749-5>
- Duncan, S. C., Duncan, T. E., & Hops, H. (1996). Analysis of longitudinal data within accelerated longitudinal designs. *Psychological Methods*, 1(3), 236–248.
- Eisele, G., Vachon, H., Lafit, G., Kuppens, P., Houben, M., Myin-Germeys, I., & Viechtbauer, W. (2022). The effects of sampling frequency and questionnaire length on perceived burden, compliance, and careless responding in experience sampling data in a student population. *Assessment*, 29(2), 136–151. <https://doi.org/10.1177/1073191120957102>
- Eisele, G., Vachon, H., Lafit, G., Tuyaerts, D., Houben, M., Kuppens, P., Myin-Germeys, I., & Viechtbauer, W. (2023). A mixed-method investigation into measurement reactivity to the experience sampling method: The role of sampling protocol and individual characteristics. *Psychological Assessment*, 35(1), 68–81. <https://doi.org/10.1037/pas0001177>
- Enders, C. K. (2011). Analyzing longitudinal data with missing values. *Rehabilitation Psychology*, 56(4), 267–288.
- Enders, C. K. (2025). Missing data: An update on the state of the art. *Psychological Methods*, 30(2), 322–339. <https://doi.org/10.1037/met0000563>
- Enders, C. K., Du, H., & Keller, B. T. (2020). A model-based imputation procedure for multilevel regression models with random coefficients, interaction effects, and non-linear terms. *Psychological Methods*, 25(1), 88–112.
- Feng, Y., & Hancock, G. R. (2021). Oh no! They cut my funding! Using “post hoc” planned missing data designs to salvage longitudinal research. *Child Development*, 92(3), 1199–1216. <https://doi.org/10.1111/cdev.13501>
- Fuller-Tyszkiewicz, M., Skouteris, H., Richardson, B., Blore, J., Holmes, M., & Mills, J. (2013). Does the burden of the experience sampling method undermine data quality in state body image research? *Body Image*, 10(4), 607–613. <https://doi.org/10.1016/j.bodyim.2013.06.003>
- Garnier-Villareal, M., Rhemtulla, M., & Little, T. D. (2014). Two-method planned missing designs for longitudinal research. *International Journal of Behavioral Development*, 38(5), 411–422. <https://doi.org/10.1177/0165025414542711>
- Goldring, M. R., & Bolger, N. (2021). Physical effects of daily stressors are psychologically mediated, heterogeneous, and bidirectional. *Journal of Personality and Social Psychology*, 121(3), 722–746.
- Gottschall, A. C., West, S. G., & Enders, C. K. (2012). A comparison of item-level and scale-level multiple imputation for questionnaire batteries. *Multivariate Behavioral Research*, 47(1), 1–25. <https://doi.org/10.1080/00273171.2012.640589>
- Graham, J. W. (2012). *Missing data: Analysis and design*. Springer Science & Business Media.
- Graham, J. W., Taylor, B. J., Olchowski, A. E., & Cumsille, P. E. (2006). Planned missing data designs in psychological research. *Psychological Methods*, 11(4), 323–343.
- Hamaker, E. L., & Grasman, R. P. P. (2012). Regime switching state-space models applied to psychological processes: Handling missing data and making inferences. *Psychometrika*, 77(2), 400–422. <https://doi.org/10.1007/s11336-012-9254-8>
- Hamaker, E. L., & Wichers, M. (2017). No time like the present: Discovering the hidden dynamics in intensive longitudinal data. *Current Directions in Psychological Science*, 26(1), 10–15. <https://doi.org/10.1177/0963721416666518>
- Hasselhorn, K., Ottenstein, C., & Lischetzke, T. (2022). The effects of assessment intensity on participant burden, compliance, within-person variance, and within-person relationships in ambulatory assessment. *Behavior Research Methods*, 54, 1541–1558. <https://doi.org/10.3758/s13428-021-01683-6>
- Hasselhorn, K., Ottenstein, C., & Lischetzke, T. (2025). Modeling careless responding in ambulatory assessment studies using multilevel latent class analysis: Factors influencing careless responding. *Psychological Methods*, 30(2), 374–392. <https://doi.org/10.1037/met0000580>
- Ibrahim, J. G. (1990). Incomplete data in generalized linear models. *Journal of the American Statistical Association*, 85(411), 765–769. <https://doi.org/10.1080/01621459.1990.10474938>

- Ibrahim, J. G., Chen, M. H., & Lipsitz, S. R. (2001). Missing responses in generalized linear mixed models when the missing data mechanism is nonignorable. *Biometrika*, 88(2), 551–564. <https://doi.org/10.1093/biomet/88.2.551>
- Ibrahim, J. G., Chen, M.-H., & Lipsitz, S. R. (2002). Bayesian methods for generalized linear models with covariates missing at random. *Canadian Journal of Statistics*, 30(1), 55–78. <https://doi.org/10.2307/3315865>
- Jaso, B. A., Kraus, N. I., & Heller, A. S. (2022). Identification of careless responding in ecological momentary assessment research: From posthoc analyses to real-time data monitoring. *Psychological Methods*, 27(6), 958–981. <https://doi.org/10.1037/met0000312>
- Ji, L., Chow, S. M., Schermerhorn, A. C., Jacobson, N. C., & Cummings, E. M. (2018). Handling missing data in the modeling of intensive longitudinal data. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(5), 715–736. <https://doi.org/10.1080/10705511.2017.1417046>
- Johnson, E. G. (1992). The design of the national assessment of educational progress. *Journal of Educational Measurement*, 29(2), 95–110. <https://doi.org/10.1111/j.1745-3984.1992.tb00369.x>
- Jorgensen, T. D., Rhemtulla, M., Schoemann, A., McPherson, B., Wu, W., & Little, T. D. (2014). Optimal assignment methods in three-form planned missing data designs for longitudinal panel studies. *International Journal of Behavioral Development*, 38(5), 397–410. <https://doi.org/10.1177/0165025414531094>
- Judd, C. M., Westfall, J., & Kenny, D. A. (2017). Experiments with more than one random factor: Designs, analytic models, and statistical power. *Annual Review of Psychology*, 68(1), 601–625. <https://doi.org/10.1146/annurev-psych-122414-033702>
- Keller, B. T. (2022). An introduction to factored regression models with Blimp. *Psych*, 4(1), 10–37. <https://doi.org/10.3390/psych4010002>
- Keller, B. T., & Enders, C. K. (2021). *Blimp user's guide (Version 3)*. <https://www.appliedmissingdata.com/blimp>
- Kim, H. G., Cheon, E. J., Bai, D. S., Lee, Y. H., & Koo, B. H. (2018). Stress and heart rate variability: A meta-analysis and review of the literature. *Psychiatry Investigation*, 15(3), 235–245. <https://doi.org/10.30773/pi.2017.08.17>
- König, L., & Strahler, J. (2023). Advancing health psychology through ecological bio-psycho-social assessments. *Zeitschrift für Psychologie*, 231(4), 241–242.
- Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. *Journal of General Internal Medicine*, 16(9), 606–613.
- Kroenke, K., Spitzer, R. L., Williams, J. B., & Löwe, B. (2009). An ultra-brief screening scale for anxiety and depression: The PHQ-4. *Psychosomatics*, 50(6), 613–621. [https://doi.org/10.1016/S0033-3182\(09\)70864-3](https://doi.org/10.1016/S0033-3182(09)70864-3)
- Lafit, G., Adolf, J. K., Dejonckheere, E., Myin-Germeys, I., Viechtbauer, W., & Ceulemans, E. (2021). Selection of the number of participants in intensive longitudinal studies: A user-friendly shiny app and tutorial for performing power analysis in multilevel regression models that account for temporal dependencies. *Advances in Methods and Practices in Psychological Science*, 4(1), 2515245920978738. <https://doi.org/10.1177/2515245920978738>
- Lafit, G., Sels, L., Adolf, J. K., Loeys, T., & Ceulemans, E. (2022). PowerLAPIM: An application to conduct power analysis for linear and quadratic longitudinal actor-partner interdependence models in intensive longitudinal dyadic designs. *Journal of Social and Personal Relationships*, 39(10), 3085–3115. <https://doi.org/10.1177/02654075221080128>
- Li, J., Luo, X., & Liu, H. (2024). Dynamic bidirectional relation between state mindfulness and suicidal ideation among female college students: The moderating effect of trait mindfulness. *Death Studies*, 49(4), 347–358. <https://doi.org/10.1080/07481187.2024.2329180>
- Li, M., Chen, N., Cui, Y., & Liu, H. (2017). Comparison of different LGM-based methods with MAR and MNAR dropout data. *Frontiers in Psychology*, 8, 722–722. <https://doi.org/10.3389/fpsyg.2017.00722>
- Li, R., Liu, H., Chen, Z., Wang, Y. (2022). Dynamic and cyclic relationships between employees' intrinsic and extrinsic motivation: Evidence from dynamic multilevel modeling analysis. *Journal of Vocational Behavior*, 140, 103813. <https://doi.org/10.1016/j.jvb.2022.103813>
- Little, R. J. A., & Rubin, D. B. (2002). *Statistical analysis with missing data*. Hoboken, John Wiley & Sons.
- Little, T. D., & Rhemtulla, M. (2013). Planned missing data designs for developmental researchers. *Child Development Perspectives*, 7(4), 199–204. <https://doi.org/10.1111/cdep.12043>
- Liu, H., Xie, Q. W., & Lou, V. W. Q. (2019). Everyday social interactions and intra-individual variability in affect: A systematic review and meta-analysis of ecological momentary assessment studies. *Motivation and Emotion*, 43, 339–353. <https://doi.org/10.1007/s11031-018-9735-x>
- Liu, H., Yuan, K.-H., & Li, H. (2025). A systematic framework for defining R-squared measures in mediation analysis. *Psychological Methods*, 30(2), 306–321. <https://doi.org/10.1037/met0000571>
- Liu, S., & Molenaar, P. C. (2014). iVAR: A program for imputing missing data in multivariate time series using vector autoregressive models. *Behavior Research Methods*, 46, 1138–1148. <https://doi.org/10.3758/s13428-014-0444-4>
- Liu, Y., Hau, K. T., & Liu, H. (2024). Linear mixed-effects models for dependent data: Power and accuracy in parameter estimation. *Multivariate Behavioral Research*, 59(5), 978–994. <https://doi.org/10.1080/00273171.2024.2350236>
- Lüdtke, O., Robitzsch, A., & West, S. G. (2020). Analysis of

- interactions and nonlinear effects with missing data: A factored regression modeling approach using maximum likelihood estimation. *Multivariate Behavioral Research*, 55(3), 361–381. <https://doi.org/10.1080/00273171.2019.1640104>
- Luo, X., Hu, Y., & Liu, H. (2024). Assessing between- and within-person reliabilities of items and scale for daily procrastination: A multilevel and dynamic approach. *Assessment*, 32(1), 61–76. <https://doi.org/10.1177/10731911241235467>
- Luo, X., Liu, H., & Hu, Y. (2024). From cross-lagged effects to feedback effects: Further insights into the estimation and interpretation of bidirectional relations. *Behavior Research Methods*, 56, 3685–3705. <https://doi.org/10.3758/s13428-023-02304-0>
- McArdle, J. J., & Woodcock, R. W. (1997). Expanding test–retest designs to include developmental time-lag components. *Psychological Methods*, 2(4), 403–435.
- McKnight, P. E., McKnight, K. M., Sidani, S., & Figueredo, A. J. (2007). *Missing data: A gentle introduction*. Guilford Press.
- McNeish, D., & Hamaker, E. L. (2020). A primer on two-level dynamic structural equation models for intensive longitudinal data in Mplus. *Psychological Methods*, 25(5), 610–635.
- Mistler, S. A., & Enders, C. K. (2012). Planned missing data designs for developmental research. In B. Laursen, T. D. Little, & N. A. Card (Eds.), *Handbook of developmental research methods* (pp. 742–754). The Guilford Press.
- Moerbeek, M. (2022). Power analysis of longitudinal studies with piecewise linear growth and attrition. *Behavior Research Methods*, 54, 2939–2948. <https://doi.org/10.3758/s13428-022-01791-x>
- Murray, A. L., Xiao, Z., Zhu, X., Speyer, L. G., Yang, Y., Brown, R. H., ... Ribeaud, D. (2023). Psychometric evaluation of an adapted version of the perceived stress scale for ecological momentary assessment research. *Stress and Health*, 39(4), 841–853. <https://doi.org/10.1002/smi.3229>
- Muthén, L. K., & Muthén, B. O. (1998–2017). *Mplus User's Guide* (8th ed.). Los Angeles, CA: Muthén & Muthén.
- Nosek, B. A., Hardwicke, T. E., Moshontz, H., Allard, A., Corker, K. S., Dreber, A., ... Vazire, S. (2022). Replicability, robustness, and reproducibility in psychological science. *Annual Review of Psychology*, 73, 719–748. <https://doi.org/10.1146/annurev-psych-020821-114157>
- Peri, J. I. (2022). *Growth modeling applications in two-method measurement planned missing designs* [Unpublished Doctoral dissertation]. The Ohio State University.
- Pornprasertmanit, S., Miller, P. J., Schoemann, A. M., & Jorgensen, T. D. (2013). *simsem: Simulated structural equation modeling version 0.5-0* [computer software]. Available at the Comprehensive R Archive Network.
- Raghunathan, T. E., & Grizzle, J. E. (1995). A split questionnaire survey design. *Journal of the American Statistical Association*, 90(429), 54–63. <https://doi.org/10.1080/01621459.1995.10476488>
- Rhemtulla, M., & Hancock, G. R. (2016). Planned missing data designs in educational psychology research. *Educational Psychologist*, 51(3–4), 305–316.
- Rhemtulla, M., Jia, F., Wu, W., & Little, T. D. (2014). Planned missing designs to optimize the efficiency of latent growth parameter estimates. *International Journal of Behavioral Development*, 38(5), 423–434. <https://doi.org/10.1177/0165025413514324>
- Richter, N. F., Hauff, S., Ringle, C. M., & Gudergan, S. P. (2022). The use of partial least squares structural equation modeling and complementary methods in international management research. *Management International Review*, 62, 449–470. <https://doi.org/10.1007/s11575-022-00475-0>
- Santangelo, P. S., Ebner-Priemer, U. W., & Trull, T. J. (2013). *Experience sampling methods in clinical psychology*. Oxford University Press.
- Schafer, J. L., & Graham, J. W. (2002). Missing data: Our view of the state of the art. *Psychological Methods*, 7(2), 147–177. <https://doi.org/10.1037/1082-989X.7.2.147>
- Schoemann, A. M., Miller, P., Pornprasertmanit, S., & Wu, W. (2014). Using Monte Carlo simulations to determine power and sample size for planned missing designs. *International Journal of Behavioral Development*, 38(5), 471–479. <https://doi.org/10.1177/0165025413515169>
- Schultzberg, M., & Muthén, B. (2018). Number of subjects and time points needed for multilevel time-series analysis: A simulation study of dynamic structural equation modeling. *Structural Equation Modeling: A Multidisciplinary Journal*, 25(4), 495–515. <https://doi.org/10.1080/10705511.2017.1392862>
- Silvia, P. J., Kwapil, T. R., Walsh, M. A., & Myin-Germeys, I. (2014). Planned missing-data designs in experience-sampling research: Monte Carlo simulations of efficient designs for assessing within-person constructs. *Behavior Research Methods*, 46, 41–54. <https://doi.org/10.3758/s13428-013-0353-y>
- Thayer, J. F., Åhs, F., Fredrikson, M., Sollers III, J. J., & Wager, T. D. (2012). A meta-analysis of heart rate variability and neuroimaging studies: Implications for heart rate variability as a marker of stress and health. *Neuroscience & Biobehavioral Reviews*, 36(2), 747–756. <https://doi.org/10.1016/j.neubiorev.2011.11.009>
- Trull, T. J., & Ebner-Priemer, U. (2014). The role of ambulatory assessment in psychological science. *Current Directions in Psychological Science*, 23(6), 466–470. <https://doi.org/10.1177/0963721414550706>
- Vachon, H., Viechtbauer, W., Rintala, A., & Myin-Germeys, I. (2019). Compliance and retention with the experience sampling method over the continuum of severe mental

- disorders: Meta-analysis and recommendations. *Journal of Medical Internet Research*, 21(12), e14475. <https://doi.org/10.2196/14475>
- Vicente, P. C. (2023). Evaluating the effect of planned missing designs in structural equation model fit measures. *Psych*, 5(3), 983–995. <https://doi.org/10.3390/psych5030064>
- Ward, M. K., & Meade, A. W. (2023). Dealing with careless responding in survey data: Prevention, identification, and recommended best practices. *Annual Review of Psychology*, 74, 577–596. <https://doi.org/10.1146/annurev-psych-040422-045007>
- Xiao, Y., Wang, P., & Liu, H. (2023). Assessing intra- and inter-individual reliabilities in intensive longitudinal studies: a two-level random dynamic model-based approach. *Psychological Methods*. Advance online publication. <https://doi.org/10.1037/met0000608>
- Xu, M., & Logan, J. A. R. (2024). Two-method measurement planned missing data with purposefully selected samples. *Educational and Psychological Measurement*, 84(6), 1232–1244. <https://doi.org/10.1177/00131644231222603>
- Yuan, K. H., & Bentler, P. M. (2000). Three likelihood-based methods for mean and covariance structure analysis with nonnormal missing data. *Sociological Methodology*, 30(1), 165–200. <https://doi.org/10.1111/0081-1750.00078>
- Zawadzki, M. J., Graham, J. W., & Gerin, W. (2012). Increasing the validity and efficiency of blood pressure estimates using ambulatory and clinic measurements and modern missing data methods. *American Journal of Hypertension*, 25(7), 764–769. <https://doi.org/10.1038/ajh.2012.40>
- Zimmer, F., & Debelak, R. (2025). Simulation-based design optimization for statistical power: Utilizing machine learning. *Psychological Methods*, 30(3), 513–536. <https://doi.org/10.1037/met0000611>
- Zimmer, F., Henninger, M., & Debelak, R. (2024). Sample size planning for complex study designs: A tutorial for the *mlpwr* package. *Behavior Research Methods*, 56, 5246–5263. <https://doi.org/10.3758/s13428-023-02269-0>

Optimization of data collection plans and improvement of data analysis methods for intensive longitudinal studies

LIU Hongyun¹, DOU Jianing¹, XU Yongze²

(¹ Beijing Key Laboratory of Applied Experimental Psychology, National Demonstration Center for Experimental Psychology Education (Beijing Normal University), Faculty of Psychology, Beijing Normal University, Beijing 100875, China)

(² Department of Psychology, Faculty of Arts and Sciences, Beijing Normal University at Zhuhai, Zhuhai 519085, China)

Abstract: Intensive longitudinal studies (ILS) provide a powerful tool for investigating the dynamic changes and interaction mechanisms of psychological states. However, the high measurement intensity in ILS often leads to increased participant burden, reduced compliance, and compromised data quality. Currently, there is a lack of optimized data collection designs and analytical methods that effectively account for the unique characteristics of ILS. This study extends the concept of planned missing design (PMD) to ILS and examines, within the dynamic structural equation modeling (DSEM) framework, the comparison of different PMD strategies and their selection criteria, sample size planning and optimization for studies incorporating PMD, and improved methods for handling missing data. Additionally, we will develop an application to validate the feasibility and applicability of these approaches in real-world psychological research scenarios.

Keywords: longitudinal data analysis, structural equation modeling, missing data analysis, planned missing design, sample size planning