

• 研究方法(Research Method) •

## 基于动态评估的取样法：经验取样法

段锦云 陈文平

(苏州大学心理学系; 教育部人文社科重点研究基地-苏州大学中国特色城镇化研究中心, 苏州 215123)

**摘要** 研究结果的普适性是心理学研究的黄金标准之一。为达到此目标, 研究者不断创造生态效度更高, 也更具现实意义的研究方法。通过在现实情境下对被试多次重复测量进行数据收集的研究方法——经验取样法(experience sampling method), 一经提出就备受研究者的青睐。文章在综合介绍经验取样法的概念、实施过程和数据处理的基础上, 比较和阐述了该方法的优劣势, 以及该方法的主要应用领域。文章最后对经验取样法的应用领域和未来发展作了展望。

**关键词** 经验取样法; 一般问卷法; 日记法; 生态效度; 个体内变异

**分类号** B841

### 1 引言

近年来, 心理学领域的研究越来越重视研究结果的一般化, 研究结果的普适性因此成为心理学研究的黄金标准。这必然要求心理学研究方法不断创新, 开发出更具现实意义的研究方法和手段。因此, 诸多学者在原有的问卷法、观察法、访谈法和实验研究等基本方法的基础上开发出了更具创新性、生态效度更高的方法, 如基于准实验设计的现场研究(Quasi-Experimentation Design based Field Research; QEDFR), 动态评估(Ambulatory Assessment)、经验取样法(experience sampling method; ESM)、生态瞬间评估(ecological momentary assessment)、即时数据获取(real-time data capture)等。对于后四种方法, 一般没有严格区分, 除了部分学者将其统称为动态评估(Klumb, Elfering, & Herre, 2009)或即时数据获取(Stone & Broderick, 2007)之外, 大部分学者将这些方法视为同一种(Ebner-Priemer, Eid, Kleindienst, Stabenow, & Trull, 2009), 或同一大类(Scollon, Kim-Prieto, & Diener, 2003; Swendsen et al., 2000;

Brandstätter, 1983; Ebner-Priemer & Trull, 2009)。通过广泛查阅文献, 笔者赞同大多数学者的观点, 认为这些研究手段在本质上都是在事件发生的真实情境下, 对被试的感受、态度、情绪、认知或行为的真实取样。尽管叫法不一, 但经验抽样的使用率最高, 又由于其操作步骤规范、生态效度高, 而致其广泛的应用。

为了增强研究结果的普适性, 在真实情境下进行动态即时研究的思想早有体现, 但一般认为经验取样法是由 Csikszentmihalyi, Larson 和 Prescott (1977)正式开创。他们在自然环境下对青少年的活动和经验的质量进行研究, 了解青少年的精力分配情况以及他们从事的活动与环境对他们的情绪的共同作用, 作者指出, 经验取样法对这类研究具有很好的适用性。另外, 也有学者着眼于经验取样法的本质, 认为经验取样法的源头是 Brandstätter (1983), 作者运用生态瞬间抽样研究日常生活中个体对他人的情绪反应, 发现情绪反应受到个体特征和所处环境的共同作用。尽管尚存争议, 但一般认为, 经验取样法是指在一段时间内, 研究者借助于辅助工具提醒被试, 让被试在事件发生或者随机的诸多瞬间回答问题从而进行数据收集的一种方法。如在一段时间内(如一周)的每一天中, 通过钟表、寻呼机、电话或者掌上电脑随机发送信号或信息来提醒被试, 让被

收稿日期: 2011-12-25

\* 国家自然科学基金(70902056), 教育部人文社科规划项目(09YJCZH087)资助。

通讯作者: 段锦云, E-mail: mgjyduan@hotmail.com

试即刻进行项目测量(Hektner, Schmidt, & Csikszentmihalyi, 2007)。

经验取样法一经提出,便受到心理学研究者的推崇。该方法最早在社会心理学领域盛行传播(Csikszentmihalyi et al., 1977; Brandstätter, 1983),并迅速扩展到临床心理学和精神病学(Margraf, Taylor, Ehlers, Roth, & Agras, 1987; Russell, Moskowitz, Zuroff, Sookman, & Paris, 2007),以及组织行为学领域(Klumb et al., 2009; Miner & Glomb, 2010; Glomb, Bhawe, Miner, & Wall, 2011)。经验取样法得到研究者青睐,一方面是由于已有心理学研究方法的生态效度不够;另一方面,能够在真实而复杂的情境下研究变量的真实存在和关系,还原事物的本来面貌,降低回忆偏差,也一直是研究者的关切点。

## 2 经验取样法的内容

经验取样法在本质上属于自我报告法(Csikszentmihalyi et al., 1977),该法要求研究者通过被试即时的自我报告来获取数据。由于经验取样法通常需借助信号发射工具,让被试在接收到信号后尽快对测量项目做出反应。所以,运用经验抽样的研究在被试要求、工具配置、抽样方式、反应时间、数据处理方面都有特定的要求。

### 2.1 被试要求

经验取样法只要求被试具备一般行为能力,识字并能够操作相关仪器即可。综合已有研究发现,研究被试在年龄方面跨度很大,既有不满10岁的儿童(Rah, Walline, Lynn Mitchell, & Zadnik, 2006),也有94岁的老人(Carstensen et al., 2011)。而且,被试遍及诸多领域,包括学生、公司员工、家庭主妇和退休老人等,这跟研究主题有关。

不论采样何种研究方法,被试参与研究有一个基本前提:被试必须是愿意参与研究,并且保证完成整个研究过程。经验取样法需要在一定周期内多次重复抽样,被试任务相对繁重,其正常工作或生活易受干扰,尤其考验被试的耐性和注意力,所以被试的选择和保持是该类研究最关心的问题。被试可能选择性的决定是否参与,并且由于工作压力原因而退出(Klumb et al., 2009)。鉴于此,研究者通常会采取一定措施吸引和维持被试,主要包括精神和物质上的奖励。精神上的包括让被试了解研究的重要性,与被试建立稳固

的关系等,物质上的主要是给予金钱奖励(Carstensen et al., 2011; Klumb et al., 2009; Russell et al., 2007),金钱的数量要仔细权衡,过多过少都会挫伤被试对研究本身的积极性。除了进行抽样前控制,也有少许研究在研究过程中进行流失被试的补充,以保证研究的样本量(Carstensen et al., 2011)。

总之,经验取样法对被试的关键要求是参与意愿和持续性。另外,在进行正式研究前需要对被试培训,这些将在“经验取样法实施过程”部分进行具体阐述。

### 2.2 配置工具

经验取样法需借助一定的工具,通常给被试配置的工具主要包括两种:信号发射工具和数据收集工具,有些研究中两者是一致的。

对被试当前情境下的即时评估需要能够立即与被试取得联系的工具,即信号发射工具。常用工具有电子分页装置、电话、寻呼机、钟表、掌上电脑、个人数字助理设备等。电子分页装置通过接收研究者发送的无线电信号,然后发出“哔哔声”提醒被试(Csikszentmihalyi et al., 1977),配置电话或者寻呼机可以以电话或者信息的方式通知被试(Hogarth, Portell, & Cuxart, 2007; Carstensen et al., 2011; Rah et al., 2006),也可以在钟表(Csikszentmihalyi & Hunter, 2003; Fisher, 2003; Snir & Zohar, 2008)、个人数字助理(PDAs)(Nielsen & Cleal, 2010)或者电脑(Rah et al., 2006)内编程,使之在特定时间发出声响,提醒被试作答。通常,研究者还需另外给被试配置移动通信工具,并开设热线方便研究者和被试联系,被试可以与研究者交流研究过程中的问题,研究者也可以督查被试(Carstensen et al., 2011; Rah et al., 2006)。

被试接收到信号后,需要立即回答测量问题,即对数据收集工具反应,数据收集的工具主要是电子问卷或纸质问卷。由于需要多次大量抽样,易导致被试厌烦或者疲劳,不认真答题,所以问题数目不宜过多,问卷长度一般控制在1~2分钟能完成即可。因此,在一般问卷法中所用问卷需要经过项目和代表性的处理才可应用于经验取样法的研究。通常每个变量只有几个代表性的项目,问卷还包括对当下情境的开放性测量。利用个人数字处理或者掌上电脑收集数据时,问卷完成当

下数据就自动收集,并且直接进入研究者的数据存储系统。其他数据一般要求当天反馈,如纸质问卷在当天结束后立即邮寄给研究者(Côté & Moskowitz, 1998; Carstensen et al, 2011; Russell et al., 2007)。研究者获得数据后,需要经过一定的筛选和加工,方可进行数据处理。

### 2.3 抽样方式

经验取样法的抽样方式可以分为两大类:第一大类是时间抽样(time-contingent),它又可分为两个子类,一种是任意时间抽样,即在每天的规定时间内(如早上7:30到晚上10:30),研究者任意发送一定次数的信号。由于短时间内的连续抽样易引起被试烦感和疲劳,且不易把握事件的动态变化,所以两次抽样最少间隔10分钟(Csikszentmihalyi & Hunter, 2003; Carstensen et al., 2011; Dimotakis, Scott, & Koopman, 2010);另一种是时间段抽样,即先将一天的时间划分为几个大时间段,或者细化为更多的时间段,研究者在每个或者随机的时间段内随机发送一次信号(Swendsen et al., 2000; Hogarth et al., 2007; Dimotakis et al., 2010)。如对情绪和酒精消费关系的研究中,先把一天分成上午、下午和晚上三个时间段,然后在每个时间段内抽样一次(Swendsen et al., 2000)。而在对日常生活中风险感知的研究中则先把时间分成很多以15分钟为单位的时间段,然后选取任意六个时间段发送信号(Hogarth et al., 2007)。时间划分根据科学界的原则——基于时间类的设计必须符合所研究事物的动态发展过程,必须关注研究对象的时间特征,防止周期性的影响,这样才能使抽样频率与所研究事物的特征匹配(Ebner-Priemer & Trull, 2009)。

另一类是事件抽样(event-contingent),即在目标事件发生的当下即时反应。事件抽样的研究对象主要是一些离散型的事件,如社会互动或者物质使用(Ebner-Priemer et al., 2009)等。Côté和Moskowitz (1998)研究了不同个性的个体在人际互动发生时的人际行为与情绪体验,Russell等(2007)则研究了边缘型人格障碍个体在社会互动中的人际行为和情绪体验。除了社会互动的研究,Margraf等(1987)研究了惊恐发作(panic attacks)时的症状表现和强度,并且让被试对自己惊恐发作进行分类。事件抽样主要依靠被试在目标事件发生后主动反应,一般不需要研究者发送信号提

醒,但是个别研究中仍然有声响提醒,如Côté和Moskowitz (1998)的研究,“哔哔声”是为了提醒被试记住自己有即时完成测量的责任,督促被试配合。

时间和事件抽样并非完全对立,有时需要结合二者进行事件与连续现象交互作用的研究。如Steiger等(2005)既将时间从早上10点到晚上10点划分为三个大时间段,在每个时间段内随机发送信号,要求被试接收到信号后即刻完成问卷,也让被试在每次暴饮暴食后完成相应问卷,将时间取样和事件取样结合进行暴饮暴食前因变量的研究。

### 2.4 实施过程

在正式取样前,研究者需要对被试进行简短培训,包括讲解研究的程序,具体如即时信号发送的时间设置、研究周期等。另外,还要让被试熟悉研究工具的使用,并且说明被试要求、注意事项,以及参与研究被试的报酬问题等。

经验取样法研究通常持续1~2周。大部分研究需要每天进行数据收集,也有仅仅限制在工作日进行(如工作情境中员工的情绪研究)。在整个研究周期内,被试正常工作和休闲,只是在收到提示信号后即时回答问题。研究者每天的抽样次数相同,操作过程重复。但是,工作中情绪与其他变量关系的研究中,往往需要在每天早上先进行一次情绪基线值的测量(Glomb et al., 2011; Miner, Glomb & Hulin, 2005)。

为了确保数据的真实性,需要控制被试的反应时间。对于时间抽样,被试接收到提示信号后应立即完成测量。通常情况下,为避免被试反应真实性的降低,被试的反应时间有限制,如组织行为学研究中人际相互作用的特点与情绪状态和员工幸福感关系的研究中,当信号发送超过10分钟被试仍未反应的,则将该次抽样作废(Dimotakis et al., 2010)。上述方式适合个人数字助理设备或者掌上电脑答题,对于纸质问卷则不好控制。为了降低被试的延迟答题,采用纸质问卷的研究一般要求被试每日回馈。单从时间控制这方面来说,采用电脑等工具进行数据收集更有利于被试控制。

### 2.5 数据分析

数据筛选是数据分析的前提,根据被试的答题态度、空缺率等对被试进行剔除和选择。答题

质量高、空缺次数少的问卷是处理的对象,有异常值、信息遗漏或者不可分辨的需要剔除(Miner et al., 2005; Goetz, Frenzel, Stoeger & Hall, 2010)。运用经验取样法取得的数据通常包括两个层次,第一层是每个个体每次抽样的数据,第二层是每个个体的数据。因此,经验取样法研究的数据分析通常是多层次分析,在此选取多层线性模型(Hierarchical linear model; 简称 HLM)作为数据处理的代表性方法进行介绍。

笔者结合 Goetz 等 (2010)的研究简单介绍,该研究的目的是检验关于控制和价值的认知评估是否与日常生活中的积极情绪相关,以及这两种认知评估在预测情绪时是否具有交互作用,并且这种关系是否受到成就情境的影响。该研究中的数据具有两层级结构,即层级一的测量水平嵌套于层级二的个体水平。在初步分析(preliminary analyses, 主要是描述性统计)的基础上,进行主效应检验(main analyses)。主效应分析运用 HLM 工具,它以每一次的情绪测量(层级一的变量)为因变量,以中心化了的控制、价值和成就情境为自变量,层级二水平的变量是由每个个体的测量水平聚合得到的。在 HLM 模型中,除了被试间和被试内变异的分析以外,还包括所有可能的自变量间的交互作用。利用 HLM 模型,研究者从个体内和个体外两个层面把握了控制和价值的认知评估对于状态情绪经验(瞬间情绪)的作用,以及二者对情绪的交互作用。

根据个体对自身群体的代表程度,可以将研究结果适度推广,经验取样法可以让学者更生态的验证理论构想(Scollon et al., 2003)。个体数据可推广的程度依赖于该个体的代表性,虽然样本可能存在方便取样,但是经验取样法所收集的个体内数据质量高,我们仍可以根据个体水平的反应,适度推断群体水平的表现(Hogarth et al., 2007)。

### 3 经验取样法的特点

在真实情境下纵向动态地研究变量间的关系,这大大地降低了回顾性偏差,尤其与传统横截面的单次评估(如传统一般问卷法)相比,优势更为突出。但是,任何方法都存在自身的局限性,经验取样法也不例外。

#### 3.1 与一般问卷法的比较

时至今日,问卷法由于具有标准化的设计和

评分,简易方便,仍旧是心理学研究中数据收集的主要方法之一。毫无疑问,对于主观的代表性的经验、态度和行为等的测量,问卷法具有很高的适用性,但是这种基于回忆的单次自我评估对于日常生活中真实行为数据的收集是不充分的(Fahrenberg, Myrtek, Pawlik & Perrez, 2007)。

问卷法的测量内容既可以是对过去代表性的经验、行为和态度的直接回忆,也可以是立足过去对未来的一般倾向性评估。Rah 等 (2006)同时运用一般问卷法和经验取样法对儿童的视觉活动进行研究,结果发现二者在对谈话时间的估计上存在显著差异。另外, Schwartz, Neale, Marco, Shiffman 和 Stone (1999)比较了关于应对方式风格的一般倾向性问卷调查和具体情境下应对方式的瞬时测量,结果表明应对风格的一般倾向性很难预测具体情境下个体的应对方式,倾向性测量与瞬间测量的特质的聚合水平之间共变性很少。这些结果表明,问卷法的测量结果与具体情境下的真实表现差距很大,利用经验取样法的研究结果更具真实性。Ross (1989)提出的稳定与变化的内隐理论(implicit theories of stability and change)解释了记忆偏差的原因,该理论认为,人们围绕他们认为最可能发生的事情这个隐含理论来组织记忆碎片,并以此建构以前的经验。

实质上,问卷法、访谈法、聚合实验数据法以及自传性研究等都属于回顾性方法,这些方法易受诸多因素影响而扭曲真实,比如情感效应、情绪记忆的一致性效应、持续忽略、峰值和谷值等(Ebner-Priemer & Trull, 2009)。经验取样法则是在经验发生的当下研究事件发生的动态过程,即通过测量真实情境下不断变化的现象,缩小了回顾偏差,进而提高了生态效度(Scollon et al., 2003; Swendsen et al., 2000; Dimotakis et al., 2010)。

总之,与一般问卷法相比,经验取样法可以降低回顾偏差,探究情境因素的作用,提高研究的生态效度。该法对于描述情境与变量之间的关系,个体内情感、认知和行为的连续变化以及个体内变量间的关系探索均具有较高的适用性。

#### 3.2 与日记法的比较

日记法(Daily Diary Studies)是指在每天结束时,被试记录下自己一天内的行为、社会交互或者情绪(Bylsma, Taylor-Clift, & Rottenberg, 2011)。日记法和经验取样法有一致之处,如 Ohly,

Sonnentag, Niessen 和 Zapf (2010)将日记法、经验取样法和事件取样(event-sampling)统称为日记研究(diary Studies)。这些抽样方法都是在自然情境下研究被试的想法、感受和行为,取样更接近真实状态,回顾偏差更少,如 Ptacek, Smith, Espe 和 Raffety (1994)运用日记法和回顾性报告来研究与考试相关的应付方式和策略,结果发现二者的一致性只有中等水平等。

由于日记法是在每日结束时评估并记录,本质上是基于某一点的回顾性自我报告,因此同样具有回忆扭曲的可能(Russell et al., 2007)。Stone 和 Broderick (2007)指出回忆周期的长度影响回顾偏差,回忆周期越短,偏差越少。所以,日记法仍然具有回顾偏差,只是比一般回顾性的方法偏差小。与日记法不同,经验取样法关注经验发生的当下,而不是通过回忆进行评估(Hektner et al., 2007)。已有研究表明,将经验取样法与日记法,以及回顾性的访谈和问卷收集的数据进行比较,发现患者具有夸大回顾结果的倾向(Margraf et al., 1987)。而且,Bylsma 等 (2011)对抑郁症患者的研究也表明,日记法和经验抽样所得结果的相关很低,这可能是由于回忆偏差和记忆错误造成。不过,对于正常个体的研究发现,两种方法的结果相关性较高,这说明日记法依然存在有效性。但是,在上述研究中,运用能让个体对事件进行系统的再加工的日记重构法(Day Reconstruction Method)的研究结果,与经验取样法研究结果的相关性很高,这说明了日记重构法能够弥补日记法与经验取样法之间的差距,是二者间的桥梁。

总之,日记法和经验取样法都具有研究环境真实、生态效度较高的特点。但是,两者相比,经验取样法由于是在事件发生的当下收集数据,回顾偏差更少,而日记法的数据收集仍具有时滞性。

### 3.3 经验取样法的优劣总结

从经验取样法与其他研究方法的比较中,我们可以总结得出经验取样法的诸多优势,如生态效度高,可以观察个体内的变异,更好的把握环境的作用等。

第一,经验取样法具有较高的生态效度。

经验取样法最注重的是经验研究的生态性,注重在自然或者个体惯常所处的情境下研究事物的状态或者变量之间关系,期望所研究的思想、

情感和行为更真实可信,实质上这也是研究设计者的主要目的(Hormuth, 1986)。

运用经验取样法在动态过程中研究变量之间的关系,可以更好的把握情境的作用,理解行为的可变性。被试在反应时关注的是研究变量的瞬时值,而并非静态条件下变量的评估值,更好的捕捉到了变量的真实状态(Miner et al., 2005)。关于日常生活中风险评估的研究发现,运用经验取样法的风险评估要比运用回顾性方法评估的风险大,这是由于经验取样法研究的即时性,剔除了实验室条件的限制和记忆效应的影响(Hogarth et al., 2007)。另外,在情绪的前因变量的研究中,测试项目聚焦于当前的活动,而非一般性的情境或者事件,这使得对状态情绪经验的评估更加清晰具体,从而提高了测量的效度(Goetz et al., 2010)。

第二,经验取样法可以考察个体内变异。

经验取样法的另一个优点是能够在个体内研究各变量之间的关系,能够考察个体内的变异。Côté 和 Moskowitz (1998)利用事件抽样的研究表明,个性特质不仅与情绪和行为的平均水平存在共变,而且也与随着时间变化的行为和情感存在共变。这种动态评估是个体内设计与个体间设计的结合,既可以考察个体内的相关,也可以考察个体间差异,这两种方式的结合促使我们更好的了解个体是如何体验、感觉、思考和行为的(Klumb et al., 2009)。

已有情绪与绩效关系的研究表明,在绩效中组内变异和组间变异具有同等的重要性(Beal, Weiss, Barros, & MacDermid, 2005)。Miner 等 (2005)的研究也表明,个体内情绪的变异占工作中情绪总变异的一半以上,这说明情绪的被试间研究的局限性,被试间的设计只分析了工作情绪变异的一半,工作情境中情绪的研究需要关注被试内的变异。另外,对于幸福感(well-being)的研究也表明,个体内的测量是很重要的,它可以捕捉幸福感随着时间的动态变化(Dimotakis et al., 2010)。

用经验取样法进行研究,更精准具体的把握了变量之间的关系,尤其在控制了被试间差异后,各变量之间的关系更加明确,为实际应用提供了有力的证据。

第三,运用经验取样法可以进行更精确的纵向研究。

运用经验取样法的研究具有纵向性质, 都是在一定时期内(如 1~2 周)通过多次重复评估探讨诸如情绪与行为的关系、精神病理学中精神病症状等(Fischer, 2009)。方法本身由于是多次测量已使研究结果更真实可靠, 若运用经验取样法进行跨度更大的纵向研究(如两个月或以上)则结果更准确。Carstensen 等 (2011)为了研究情感体验随年龄增长的发展变化, 每五年进行一次为期一周的经验抽样, 结果表明, 随着年龄的增长幸福感增高, 且积极情绪较多者的死亡率低。因此, 用经验取样法进行纵向研究在减少回忆偏差, 提高生态效度的同时, 还可以更好地验证变量间的真实因果关系。

任何方法因其本身的特殊性, 除了突出的优势以外, 总有其不足之处。经验取样法由于在真实情境下的多次抽样, 可能存在被试的接受性不足和反应性改变、测量项目的代表性降低, 以及基于自我报告的固有的局限等。经验取样法的主要不足之处包括:

首先, 被试对研究方法的接受性以及研究过程中的服从性问题。

由于多次重复抽样对被试正常工作的干扰, 可能导致被试对经验取样法的接受性下降。被试拒绝参与研究, 或中途流失。一方面, 被试对研究方法的接受性不高, 可能导致被试选择性偏差。即个体可能因为某些原因参加或者不参加研究, 导致样本的代表性降低。如 Dimotakis 等 (2010)对人际相互作用的特点与情绪和幸福感关系的研究中, 一些处在高度相互作用中的个体就不愿意参与该研究。另一方面, 被试对研究方法的接受性也能影响他们对研究的积极性, 从而降低对测量的应答率和效度。

研究过程中需要被试多次反应, 被试对研究的服从性也可能影响研究效果。已有研究表明, 在经验抽样过程中被试并不总是遵守研究者的指示(Jacobs et al., 2005)。由于每天多次被声响提醒, 被试开始觉得有趣, 后来就会很厌烦, 尤其是在特定的时间内, 可能导致数据收集工作的耽误、延迟甚至拒绝数据输入(Fahrenberg et al., 2007)。被试对研究要求的不服从, 导致研究数据质量下降, 研究结果效度降低。

其次, 经验取样法可能改变被试的反应性。基于经验取样法的研究往往需要被试自己记录采

样时间, 如“何时进行的测量”、“完成测量用的时间”等, 这会导致被试对下一次的测量形成预警(Hormuth, 1986)。如 Swendsen 等 (2000)在情绪和酒精依赖的关系的研究中, 多次重复测量可能使被试产生了期望效应, 混淆研究结果。

再次, 测量项目的数量和代表性问题。经验取样法的多次抽样, 易导致被试疲惫和敷衍, 所以为取得真实数据, 常常需要减少测量项目的数量, 降低变量描述的详细程度(Swendsen et al., 2000)。这些操作可能降低研究项目的代表性, 降低测量的信效度。所以, 如何选择更具代表性的项目, 在项目数量减少的同时, 保证变量的测量效度, 这是研究者面对的重要问题。

最后, 自我报告法的固有局限。虽然经验取样法与其他方法相比具有诸多优点, 但该方法仍是依赖于自我报告的。研究过程没有实验室研究的严格控制, 缺乏研究者对行为和环境的直接观察, 仅依靠个体内的重复评估, 所以可能导致共同方法偏差。因此, 研究结果在得到来自更具控制性和可观察的实验研究的支撑时往往更具说服力(Russell et al., 2007)。

#### 4 经验取样法的主要应用领域

经验取样法最先应用于社会心理学领域, 随后扩展到临床心理学和组织行为领域, 至此而备受研究者的青睐。经验取样法的风靡与其自身优势紧密相连, 较高的生态效度促使该法在更多领域更为广泛的应用。

##### 4.1 社会心理学领域

经验取样法最先应用于社会心理学领域(Csikszentmihalyi et al., 1977; Brandstätter, 1983), 该方法多用来研究个体的认知、情绪和行为, 以及具体情境因素的影响。

传统上把情绪当做一种特质(trait)来研究。实质上情绪是一种极易变动并且依赖于情境的状态(state), 越来越多的研究开始在真实具体的情境下研究情绪, 经验取样法对这类动态研究非常适用。Csikszentmihalyi 和 Hunter (2003)运用经验取样法研究了小学生日常生活中的幸福感, 研究揭示了情境、当下活动以及习惯与个体幸福感的关系, 该方法很好的将幸福感产生的具体情境从长期性的条件中分离出来。Goetz 等 (2010)通过一周内每天六次的随机抽样进行收集数据, 验证了

认知因素(控制和价值)与积极情绪关系密切,以及二者在预测积极情绪中的交互作用。Carstensen等(2011)纵向的研究了积极情绪与年龄的关系,发现随着年龄增长情绪变得更加积极,并且日趋稳定和复杂,而且拥有较多情绪体验的人更长寿。

经验取样法的应用日益广泛, Graham (2008)研究了动态条件下夫妻关系的自我扩展(self-expansion),验证了夫妻共同活动中的激活水平与夫妻关系的质量积极相关。Hogarth 等(2007)还运用经验取样法和回顾性问卷法共同评估了日常生活中的风险,结果发现当下的风险评估要比回顾性的风险评估大,这表明经验取样法更好的捕捉到了当下情绪等具体因素的影响。

#### 4.2 临床心理学领域

传统的临床心理学研究方法是基于回顾性的自我报告,如结构化/半结构化访谈、问卷法、日记法等。这些方法都是基于被试的回忆,因此得到的信息含混不清,往往存在扭曲(Stone & Broderick, 2007)。由于回顾性的方法可能有记忆消退、情感效应、情绪记忆的一致性、持续忽略等如此多的缺陷(Ebner-Priemer & Trull, 2009),研究者更青睐能够即时抽样的研究方法。

在临床心理学领域,将被试在正常生活中借助计算机获得的自我报告、行为记录以及生理测量等统称为动态评估(ambulatory assessment)。具体方法包括持续监控、时间和事件抽样法、现场测验、现场实验、症状监控等,在该领域中,经验取样法、生态瞬间评估、即时数据获取属于术语不同的同类方法(Fahrenberg et al., 2007; Ebner-Priemer & Trull, 2009)。这类方法足够创新,对环境敏感度高,能够考察环境和症状的交互作用,有利于研究结果的实际应用。

经验取样法在临床心理学领域应用广泛。Swendsen 等(2000)利用经验取样法研究了情绪和酒精消费之间的关系,有效捕捉到了紧张不安(nervousness)导致酒精消费的增多,饮酒之后紧张感下降等临床现象,验证了自我医疗(self-medication)理论。Steiger 等(2005)研究了神经症性贪食症患者暴饮暴食前后情绪和自我概念的变化,以及相关生物因素的影响。Bylsma 等(2011)研究了重度和轻度抑郁患者对日常事件的反应,验证了情绪闪亮效应(mood-brightening

effect),即与健康者相比,积极事件可以使抑郁者的情绪得到更多改善。

运用经验取样法还可以研究精神病理学症状随时间的兴衰起伏,以及环境、个体经验和精神病理学症状的相互作用(Ebner-Priemer et al., 2009; Oorschot, Kwapil, Kwapil, & Myin-Germeyns, 2009)。如 Margraf 等(1987)研究了惊恐发作的症状,并将经验取样法和问卷法及访谈法的研究结果进行了比较,发现传统方法具有回顾性夸大倾向。由于回顾性的自我报告和临床观察等研究方法的结果效度的模糊性, Russell 等(2007)运用基于事件的经验取样法研究了自然状态下边缘性人格障碍患者的情绪体验和人际行为的稳定性和可变性。不过,用经验取样法研究精神病的行为模式、情绪、阳性症状是可取的,但是研究精神病者的阴性症状则不可行(Oorschot et al., 2009)。van Winkel 等(2008)运用经验取样法研究日常生活中对压力、情绪和精神病症状的评估,证明了基因和环境共同对精神病症状的形成起作用。

经验取样法在临床心理学的应用可以明确各种行为的临床表现,进行及时的监控和治疗。尤其对于精神病学领域,明确各种症状不仅为精神病的治疗提供依据,也为疗效的评估提供参照。

#### 4.3 组织行为学领域

当今时代,竞争环境日益复杂,整个组织处在不断变化的动态过程中。利用传统的问卷法和访谈法所得的研究结果难以保证其效度。因此,能够捕获具体情境下变量瞬时状态的经验取样法受到组织行为学研究者的追捧。Klumb 等(2009)总结得出运用动态评估类研究方法的研究主要聚焦于工作紧张和应对工作压力,工作和非工作的不同,工作场所中的交互作用,工作态度和与工作相关的情绪的研究等。

当情绪事件理论(affective events theory)提出后,研究者认识到员工经历的工作事件对员工情绪状态以及随后的态度、行为产生影响。利用经验取样法的个体内研究迅速增多,并且大多聚焦在人际互动领域。早期, Miner 等(2005)研究了积极、消极工作事件与快乐度(hedonic tone)的关系,以及快乐度与组织公民行为(organizational citizenship behaviour)、工作撤出(work withdrawal)的关系, Miner 和 Glomb (2010)又研究了客服人员的情绪状态与任务绩效、工作撤出行为的关

系, 以及元情绪特质(trait meta-mood)对这种关系的调节作用。Bono, Foldes, Vinson 和 Muros (2007)研究了社会互动中领导行为对员工工作情绪的影响, 与领导互动比与同事互动体验的积极情绪更低, 但是变革型领导行为(transformational leadership)可以起调节作用。Dimotakis 等 (2010)利用经验取样法研究了工作场所中人际互动的特点对工作中的情绪和日常幸福感的影响。另外, 基于情绪调节(mood regulation)理论, Glomb 等 (2011)运用经验取样法研究了员工做出组织公民行为前后情绪状态的变化, 验证了利他行为具有调节情绪的作用, 即消极情绪可引发利他行为, 而利他行为又会引起积极情绪。

经验取样法可以结合自然观察和标准化问卷的优势, 更好的研究组织内的现象(Snir & Zohar, 2008)。Fisher (2003)用经验取样法研究了工作绩效和工作满意度的关系, 解释了外行的信念与学者们的研究结果不一致的原因。外行由于将个体内水平上绩效和满意度的相关推广到个体间水平, 主观的加强了二者之间的关系, 从而得出与研究结果不同的结论。Snir 和 Zohar (2008)则利用经验取样法研究了工作狂(workaholism)认知和情绪方面的特点, 完善了关于工作狂的定义。Ohly 和 Fritz (2010)验证了长期的工作特点和利用经验取样法测得的日常的工作特点(时间压力和工作控制)都与日常的创新和主动行为积极相关, 并且长期的工作特点以日常测量的工作特点为中介, 个体对工作特点的评估在整个关系中起部分中介作用。

经验取样法可以与一般问卷法组合进行研究。Ilies, Scott 和 Judge (2006)用一般问卷法测量员工的个人特质, 用经验取样法收集情绪和工作满意感、及组织公民行为数据, 探索个人特质和情绪状态对当下行为的影响。类似地, Nielsen 和 Cleal (2010)为研究工作中“流体验”(flow)的前因变量, 用经验取样法将流体验从流体验条件中分离, 获取流体验的瞬间状态以及流体验当下的活动, 并用问卷调查来研究稳定的工作特点对流体验的预测作用(Nielsen & Cleal, 2010)。

随着组织情境变动性的增强, 运用经验取样法可以更真实的反映变量间关系, 将来类似的动态评估方法的将更加普遍。

## 5 未来研究展望

心理学研究越来越重视研究结果的推广性, 研究者因而一直致力于高生态效度的研究方法的创新和开发。经验取样法作为在真实情境下把握事物动态变化的研究方法, 必将得到越来越广泛的应用。对于经验取样法研究未来的发展, 笔者认为主要包括两个方面, 一是经验取样法本身, 经验取样法需要克服自身的不足, 不断改进程序达到研究效果的最优; 另一个是应用领域的扩展, 即经验取样法的研究应该在更广泛的领域中应用。

第一, 方法本身的完善。

样本良好的代表性是研究结果推广的前提, 所以未来研究需要进一步控制被试的选择性偏差。而且, 未来研究还应该增强被试对该研究方法的接受性, 以及对研究要求的服从性, 保证数据的质量, 提高研究的效度。具体来说, 在未来研究中要力求抽样方式与被试的日常生活情境相匹配, 保证应答率。在某些情况下, 被试由于工作的特殊性不能及时反应的, 研究者应该采取相应措施, 如 Nielsen 和 Cleal (2010)研究的被试由于是管理人员不能及时应答, 所以研究者将抽样的次数设置的比同类研究多。另外, 研究收集数据的期望必须与被试的要求保持均衡, 为了促进被试参与和保持, 过频的长时间测量可能需要更多的金钱或者精神上的激励(Miner et al., 2005)。

测量项目的筛选和编制、研究周期的设置也是未来研究需要注意的地方。由于经验取样法所用的测量项目多来源于已有问卷, 测量项目的减少和项目描述程度的降低势必对研究效度造成威胁, 所以如何选取高质量有代表性的题目是研究者面临的挑战。未来研究应该更注重测量项目的选择方法, 根据变量的维度、研究的对象和研究情境选出最有代表性的测量项目。另外, 测量周期的长度也要与被试的工作相匹配, 周期过长可能降低被试的合作性和服从性, 过短则不利于研究结果的效度。概而言之, 测量周期既要符合被试的现实状况, 又要满足研究的需要。

由于多次抽样, 基于自我报告而产生的共同方法偏差是比较少的, 为使研究结果更真实有效, 仍需要研究者给予关注。已有学者认识到这点并进行了控制, 如 Glomb 等 (2011)在采用经验取样

法研究组织中的情绪时,每次测验的题目并非完全相同,且测验时间随机,有效避免了共同方法偏差。另外,未来研究在抽样方式的设置上应注意与研究对象的匹配性。

第二,应用领域的扩展。

经验取样法自提出之日起,就受到了各领域学者的青睐,该法在社会心理学、临床心理学和组织行为学领域都得到广泛应用。在追求高效度的研究范式的要求下,未来该方法在各领域还会得到广泛应用。

经验取样法在社会心理学的应用从最初的情绪、行为方面向外扩展,开始研究更深层次的认知因素,并进行理论的验证。未来研究可以在现实生活中更多的检验和完善已有理论,结合情境的复杂性,挖掘深层次的发生机制,指出理论适用的具体条件。

在精神病理学领域,我们希望未来研究在关注精神病理学症状和特点的同时,更加注重个性或者症状特点、从事活动和个体行为间的相互作用。另外,利用经验取样法研究社会互动下的人际行为与情绪,可以同时研究情境评估等认知因素的作用,也可以将社会互动划分为更具体的事件进行研究。经验取样法除了研究精神病的模式、症状外,还可以用来评估精神病的疗效。

在组织行为学领域,由于组织变革加剧,组织活动和员工情绪及行为都处在时刻变化之中,为此该领域对经验取样法的需求也将更多。在未来研究中可以将经验取样法与问卷法、访谈法等结合使用,更准确的把握组织现象,如 Côté 和 Moskowitz (1998)利用经验抽样收集人际行为与情绪的数据,通过一般问卷收集个体外向性、宜人性和神经质等个性特征,研究了个性特征及人际行为对情绪的影响。

除了上述领域,经验取样法在其他研究领域也应该推广使用。如在发展和教育心理学领域,运用经验取样法可以更准确掌握青少年儿童的认知、情绪和行为发展特点。在咨询心理学领域中,笔者认为经验取样法可以监控来访者的问题行为,评估咨询效果。

总之,经验取样法作为一种高效度的动态研究方法,在国外受到研究者的重视,并在诸多领域开展了大量研究。在日益重视研究结果普适性的今天,运用能够在具体情境下进行动态研究的

方法是学术界的一大趋势。因此,在国内经验取样法应用相对匮乏的情况下,介绍和综述该研究方法的实施过程、优缺点和应用领域,希望能够为该研究方法在国内的应用提供参照。

## 参考文献

- Beal, D. J., Weiss, H. M., Barros, E., & MacDermid, S. M. (2005). An episodic process model of affective influences on performance. *Journal of Applied Psychology, 90*, 1054–1068.
- Brandstätter, H. (1983). Emotional responses to other persons in everyday life situations. *Journal of Personality and Social Psychology, 45*, 871–883.
- Bylsma, L. M., Taylor-Clift, A., & Rottenberg, J. (2011). Emotional reactivity to daily events in major and minor depression. *Journal of Abnormal Psychology, 120*(1), 155–167.
- Bono, J. E., Foldes, H. J., Vinson, G., & Muros, J. P. (2007). Workplace emotions: The role of supervision and leadership. *Journal of Applied Psychology, 92*, 1357–1367.
- Carstensen, L. L., Turan, B., Scheibe, S., Ram, N., Ersner-Hershfield, H., Samanez-Larkin, G. R., et al. (2011). Emotional experience improves with age: Evidence based on over 10 years of experience sampling. *Psychology and Aging, 26*(1), 21–33.
- Côté, S., & Moscovitz, D. S. (1998). On the dynamic covariation between interpersonal behavior and affect: Prediction from neuroticism, extraversion, and agreeableness. *Journal of Personality and Social Psychology, 75*, 1032–1046.
- Csikszentmihalyi, M., & Hunter, J. (2003). Happiness in everyday life: The uses of experience sampling. *Journal of Happiness Studies, 4*, 185–199.
- Csikszentmihalyi, M., Larson, R., & Prescott, S. (1997). The ecology of adolescent activity and experience. *Journal of Youth and Adolescence, 6*, 281–294.
- Dimotakis, N., Scott, B. A., & Koopman, J. (2010). An experience sampling investigation of workplace interactions, affective states, and employee well-being. *Journal of Organizational Behavior, 32*, 572–588.
- Ebner-Priemer, U. W., Eid, M., Kleindienst, N., Stabenow, S., & Trull, T. J. (2009). Analytic strategies for understanding affective (in)stability and other dynamic processes in psychopathology. *Journal of Abnormal Psychology, 118*(1), 195–202.
- Ebner-Priemer, U. W., & Trull, T. J. (2009). Ambulatory assessment: An innovative and promising approach for clinical psychology. *European Psychologist, 14*, 109–119.

- Fahrenberg, J., Myrtek, M., Pawlik, K., & Perez, M., (2007). Ambulatory assessment-monitoring behavior in daily life settings: A behavioral-scientific challenge for psychology. *European Journal of Psychological Assessment*, 23, 206–213.
- Fischer, J. E. (2009). Experience-sampling tools: A critical review. In *Mobile Living Labs*. Bonn, Germany, Working paper.
- Fisher, C. D. (2003). Why do lay people believe that satisfaction and performance are correlated? Possible sources of a commonsense theory. *Journal of Organizational Behavior*, 24, 753–777.
- Glomb, T. M., Bhave, D. P., Miner, A. G., & Wall, M. (2011). Doing good, feeling good: Examining the role of organizational citizenship behaviors in changing mood. *Personnel Psychology*, 64, 191–223.
- Goetz, T., Frenzel, A. C., Stoeger, H., & Hall, N. C. (2010). Antecedents of everyday positive emotions: An experience sampling analysis. *Motivation and Emotion*, 34, 49–62.
- Graham, J. M. (2008). Self-expansion and flow in couples' momentary experiences: An experience sampling study. *Journal of Personality and Social Psychology*, 95, 679–694.
- Hektner, J. M., Schmidt, J. A., & Csikszentmihalyi, M. (2007). *Experience Sampling Method: Measuring the Quality of Everyday Life*. Thousand Oaks, CA: Sage.
- Hogarth, R. M., Portell, M., & Cuxart, A. (2007). What risks do people perceive in everyday life? A perspective gained from the experience sampling method (ESM). *Risk Analysis*, 27, 1427–1439.
- Hormuth, S. E. (1986). The sampling of experiences *in situ*. *Journal of Personality*, 54, 262–293.
- Ilies, R., Scott, B. A., & Judge, T. A. (2006). The interactive effects of personal traits and experienced states on intraindividual patterns of citizenship behavior. *Academy of Management Journal*, 49, 561–575.
- Jacobs, N., Nicolson, N. A., Derom, C., Delespaul, P., van Os, J., & Myin-Germeys, I. (2005). Electronic monitoring of salivary cortisol sampling compliance in daily life. *Life Sciences*, 76, 3431–2443.
- Klumb, P., Elfering, A., & Herre, C. (2009). Ambulatory assessment in industrial/organizational psychology fruitful examples and methodological issues. *European Psychologist*, 14, 120–131.
- Margraf, J., Taylor, C. B., Ehlers, A., Roth, W. T., & Agras, W. S. (1987). Panic attacks in the natural environment. *The Journal of Nervous and Mental Disease*, 175, 558–565.
- Miner, A. G., & Glomb, T. M. (2010). State mood, task performance, and behavior at work: A within-persons approach. *Organizational Behavior and Human Decision Processes*, 112, 43–57.
- Miner, A. G., Glomb, T. M., & Hulin, C. (2005). Experience sampling mood and its correlates at work. *Journal of Occupational and Organizational Psychology*, 78, 171–193.
- Nielsen, K., & Cleal, B. (2010). Predicting flow at work: Investigating the activities and job characteristics that predict flow states at work. *Journal of Occupational Health Psychology*, 15, 180–190.
- Ohly, S., & Fritz, C. (2010). Work characteristics, challenge appraisal, creativity, and proactive behavior: A multi-level study. *Journal of Organizational Behavior*, 31, 543–565.
- Ohly, S., Sonnentag, S., Niessen, C., & Zapf, D. (2010). Diary studies in organizational research: An introduction and some practical recommendations. *Journal of Personnel Psychology*, 9, 79–93.
- Oorschot, M., Kwapi, T., Delespaul, P., & Myin-Germeys, I. (2009). Momentary assessment research in psychosis. *Psychological Assessment*, 21, 498–505.
- Ptacek, J. T., Smith, R. E., Espe, K., & Raffety, B. (1994). Limited correspondence between daily coping reports and retrospective coping recall. *Psychological Assessment*, 6, 41–49.
- Rah, M. J., Walline, J. J., Mitchell, G. L., & Zadnik, K. (2006). Comparison of the experience sampling method and questionnaires to assess visual activities in pre-teen and adolescent children. *Ophthalmic and Physiological Optics*, 26, 483–489.
- Ross, M. (1989). Relation of implicit theories to the construction of personal histories. *Psychological Review*, 96, 341–357.
- Russell, J. J., Moskowitz, D. S., Zuroff, D. C., Sookman, D., & Paris, J. (2007). Stability and variability of affective experience and interpersonal behavior in borderline personality disorder. *Journal of Abnormal Psychology*, 116, 578–588.
- Schwartz, J. E., Neale, J., Marco, C., Shiffman, S. S., & Stone, A. A. (1999). Does trait coping exist? A momentary assessment approach to the evaluation of traits. *Journal of Personality and Social Psychology*, 77, 360–369.
- Scollon, C. N., Kim-Prieto, C., & Diener, E. (2003). Experience sampling: Promises and pitfalls, strengths and weaknesses. *Journal of Happiness Studies*, 4, 5–34.
- Snir, R., & Zohar, D. (2008). Workaholism as discretionary time investment at work: An experience sampling study. *Journal of Applied Psychology: An International Review*, 57, 109–127.
- Steiger, H., Gauvin, L., Engelberg, M. J., Kin, N. M. K. N. Y., Israel, M., Wonderlich, S. A., et al. (2005). Mood- and restraint-based antecedents to binge episodes in bulimia

- nervosa: Possible influences of the serotonin system. *Psychological Medicine*, 35, 1553–1562.
- Stone, A. A., & Broderick, J. E. (2007). Real-time data collection for pain: Appraisal and current status. *Pain Medicine*, 8, 85–93.
- Swendsen, J. D., Tennen, H., Carney, M. A., Affleck, G., Willard, A., & Hromi, A. (2000). Mood and alcohol consumption: An experience sampling test of the self-medication hypothesis. *Journal of Abnormal Psychology*, 109, 198–204.
- van Winkel, R., Henquet, C., Rosa, A., Papiol, S., Fañanás, L., De Hert, M., et al. (2008). Evidence that the COM<sup>TV</sup>Val158Met polymorphism moderates sensitivity to stress in psychosis: An experience-sampling study. *American Journal of Medical Genetics Part B: Neuropsychiatric Genetics*, 147B, 10–17.

## Ambulatory-assessment Based Sampling Method: Experience Sampling Method

DUAN Jin-Yun; CHEN Wen-Ping

(Department of Psychology, Soochow University;

Key Research Institute of Education Ministry-Center for Chinese Urbanization Studies,

Soochow University, Suzhou 215123, China)

**Abstract:** The generalizability of research findings is the golden standard of psychological research. To achieve this goal, researchers constantly develop more realistic research methods with higher ecological validity. Experience sampling method (ESM), a method of data collection by which participants respond to repeated assessments at moments during the course of the time in real situation, is highly favored by various researchers since it was proposed. This article introduced this data collection method and its statistical analyses processes, compared and elaborated the strengths and weakness of this method, and illustrated its main application fields. Finally, the future studies of experience sampling method were discussed as well.

**Key words:** experience sampling method; general questionnaire survey; daily diary study; ecological validity; within-person variation