班组情景意识测量实验方法的比较分析

李鹏程,金 肖[1]
(南华大学 人因研究所,湖南 衡阳 421001)

[摘 要] 为了对班组情景意识测量方法的优劣进行全面分析和评价,文章着重对四种主要的班组情景意识测量方法进行简要介绍,方法包括 CAST 班组协调情景意识、SACRI 控制室情景意识调查表、CARS 班组情景意识评价量表和 TSAGAT 班组情景意识整体评估技术。同时进一步从方法本身、方法测量过程与结果等方面设置多个评价标准,对 12 个班组情景意识测量方法进行评价。结果表明 12 种测量方法各具特色,但通过建立的评价标准可对方法进行优劣识别,便于实验设置人员根据测量需求、实验环境等实际情况选择合适的测量方法或组合实验,为开发新的核电厂班组情景意识测量方法提供思路与建议。

[关键词] 数字化核电厂; 班组情景意识; 测量方法; 评价标准
[中图分类号] X946 [文献标识码] A [文章编号] 1673-0755(2019)06-0011-08

人因研究中的情景意识(SA),就是用来解释复杂人-机系统中操纵员理解系统和环境正在发生什么和如何发生的,是保持有效决策和高绩效的先决条件[1]。在复杂工业系统事故的处理过程中,因 SA 失误而不能正确完成后续复杂的行为会带来灾难性后果,如三哩岛核事故[2]、各种航空飞行事故等[3]。Endsley[4]研究指出,由人因失误导致的商用航空事故中,88% 的事故原因都涉及 SA 失误。班组情景意识(TSA)是指为完成班组任务,各班组成员所理解的班组任务中各要素的程度之和,但不重复计算各成员所理解的重叠要素[5]。个体的 SA 固然重要,但 TSA 更为重要,因为在像核电厂这样复杂的工业系统中,系统正常运行状态的监控和异常状态的成功处置均是由班组成员共同完成的,并且如果个体产生 SA 失误,则可能被班组其他成员发觉、纠正和恢复[6]。在像核电厂、空中交通管制等复杂的动态系统中,其安全更加依赖于班组的精准绩效而非个人绩效[7],诸多研究表明,班组的绩效水平与 TSA 是正相关的[8-9]。因此,TSA 对于班组任务的成功执行至关重要,是影响班组绩效并最终可能导致灾难性事故的关键因素。

班组任务的成功执行需要以较高的 TSA 为前提,因此精确测量 SA 和 TSA 尤为重要。但是目前的研究大多针对的是个体 SA,而对 TSA 的研究较为缺乏。虽然 TSA 的概念很早就被提出,但是限于 TSA 本身的复杂性,诸多学者对其测量方法有不同的见解,如石玉生[12]对三个班组情景意识测量方法进行分别评价,但没有进行系统地对比分析;Breton R 等[19]分析评价了八个班组情景意识测量方法,但没有设置整体标准进行对比分析;并且当前还没有针对数字化工业系统情景环境特征发展的 TSA 测量技术。因此,本文将构建 TSA 测量方法的评价标准,对多个 TSA 测量技术进行评估,识别现今的 TSA 测量技术存在的优势与不足,进而提出当前 TSA 测量技术发展的需求、挑战以及未来发展趋势,为开发更为合理的新型数字化核电厂 TSA 测量技术提供理论指导。

一 典型的 TSA 测量方法的介绍

由于目前对于 TSA 理论的研究还存在着较大的分歧和局限性,使得合适的 TSA 测量方法较为欠缺,已有的测量方法也存在着较大的差异,或多或少存在各自的局限性,如测量过于抽象不具体,测量问题的重要性没有区分等,因此,需要对已有的各种

[收稿日期] 2019-10-30
[基金项目] 国家自然科学基金资助(编号:51674145);博士后科学基金项目资助(编号:2016M600633);湖南省自科基金项目资助(编号:2017JJ2222);湖南省社科基金项目资助(编号:16YBA314);2019 南华大学研究生科研创新项目(编号:193YXC051)。
[作者简介] 李鹏程(1978-),男,湖南邵阳人,南华大学人因研究所教授,博士。
1 南华大学经济管理与法学学院硕士研究生

TSA 测量方法构建评估标准来进行评价,以识别他们各自的优劣,为发展更为客观合理的方法提供指导。由于篇幅所限,主要选择目前较为常见、认可度较高的 4 种 TSA 测量方法进行介绍,其他具体方法见各自的参考文献。

(一)CAST 班组协调情景意识

方法简介:班组协调情景意识(Coordinated Awareness of Situation by Teams,CAST)是一种基于协调的班组情景意识测量方法。Gorman J C[10]认为在不熟悉的领域,其情景状态空间会对任务性能产生影响。在任务状态空间的这一区域,单个班组成员对情景失误有自己的局部感知,但是需要班组感知的协调,以便及时充分感知失误并对其采取行动。每个班组成员都能感知到失误的不同方面,并且必须协调这些视角,以便充分地感知到失误并对其采取行动。因此,TSA 不能被理解为单个班组成员感知的总和,而是作为一个班组在通过班组协调进行的失误转换下保持不变的各个部分的综合。

实验过程:Gorman J C 与 Cooke N J[11]将该方法应用于一个模拟无人机侦察实验中,班组对通信通道故障的感知。该实验设置一半的班组成员(实验组 1)在一个共同的环境中完成任务,他们可以通过耳机交流,并且可以离开控制台去见其他班组成员。另一半班组成员(实验组 2)在分布式环境中执行任务,使用的通信模式与实验组 1 的条件相同,但是位于不同的房间或小隔间中。假设这种协同位置的微妙操作可能会由于远距离协调的困难而对分布式班组的 TSA 产生负面影响。

方法总结:这种测量方法的核心观点是认为班组情景意识的核心特征是针对环境的动态变化进行适应以完成班组目标。那么在班组完成特定目标的过程中,如果人为引入意外变化与障碍,然后记录班组为了应对这种突发状态而产生的一系列行为,包括信息的收集、沟通与交流以及最后的处理方法,对这一系列行为进行评价与分析,就能够对该班组的班组情景意识进行评价[12]。

CAST 呈现了 TSA 的研究人员简化实验过程,利用简单的二组目标实验组来理解班组是如何以及为什么要这样做。班组协调并不独立于单个班组成员的输入,CAST 代表了这些班组协调之间的综合。因此,无论班组的规模或专业知识的分布如何,简化人为因素的变量对于研究 TSA 非常有用。然而,需要更多的研究来确定归因于共享心智模型的人为因素与在复杂班组中解释 TSA 时的协调过程。

(二)SACRI 控制室情景意识调查表

方法简介:控制室情景意识调查表(Situation Awareness Control Room Inventory,SACRI)以调查表的形式直接调查操纵员对信息收集、过程状态以及过程预测方面的情景意识水平[13]。SACRI 是在情景意识整体评估技术(the Situation Awareness Global Assessment Technique,SAGAT)的基础上加以改进的。SACRI 研究的目的是研究 SAGAT 应如何应用于过程控制研究。研究的内容包括量表的问题清单、操作人员能力方面的要求、场景的设计、回答的评分和比较备选系统设计的程序,其中问题清单是该方法的独特之处。

问题清单是通过与专家的讨论和 HAMMLAB 文件的检查而编制,以调查操作人员对压水反应堆模拟器的情景意识。该清单是通过在研究中使用它所获得的经验而修订的。可以看出,问题与流程参数相关,如流程、级别、温度、压力、泵的状态等。

实验过程:实验中这些过程参数是与操纵员发生交互作用而得到的,而且通过问题的设计使得这些问题都涉及三层 SA 水平,第一层 SA 是需要操纵员回答“过去”或“正常”情况下操纵员可直接观察到的或能读到的参数,第二层是关于“过去”或“正常”情况下不能被直接观察到的参数,但可以基于可观察到的参数进行推理得到,第三层 SA 涉及所有的三类问题(过去、现在和未来),第三类问题主要是基于可直接观察到的参数和推理得到的事实来预测未来的情况[13]。在试验的过程中通过计算机来管理问题,操纵员在暂停时通过登录来对随机选择好的问题进行回答,一般包括 12 个问题,时间在 3-7 分钟,Endsley 研究表明模拟机暂停在 5 分钟之内,操纵员的 SA 水平不会衰减[15]。实验后,可通过回答问题的正确率来确定操纵员的 SA 水平[13]。

方法总结:DAVID N H 等[14]制定的 SACRI 方法适用于核电厂压水反应堆报警系统,它能够客观地评估情况意识的动态性质,评估操作员对核电厂情景状态的认识,以及在各种不同过程状态下的应用情况和界面配置。SACRI 的应用在系统测试和评估的水平上有些粗糙,且需要有经验的操作人员,才能有效地从 SACRI 获得可靠的态势感知度量,但不能否认它的可用性;它与操纵员绩效的相关性在于它基于建议的情境意识模型(Endsley 情景意识三层次模型)[16]。它可以与其他性能度量一起用于生成对竞争设计替代方案的更完整的评估。另外,SACRI 的使用并不局限于核电站控制室的先进报警系统只要确定了操作员在不同情况下的情景意识影

响因素,就可以应用。因此,它可用于其他过程工业,如石油平台和石化厂,既可用于过程状态信息系统的评价,也可作为操作员业绩的衡量标准,用于培训和评价。

(三)CARS 班组情景意识评价量表

方法简介:班组情景意识评价量表(Crew Awareness Rating Scale,CARS)包括 8 个部分的问卷量表,涉及操作员的心智内容与 SA 的心智处理过程[17-18]。CARS 是用来评估 4 个层次的 SA:(1)感知,新信息的同化/感知(The Assimilation of New Information);(2)理解,在情景中理解信息;(3)预测,预测未来可能的发展;(4)整合,通过行动对上述进行综合集成[18]。

实验过程:实验中,对于上述 4 个功能要求被试评价:(1)该方面的内容:可靠准确吗?(2)该方面处理:容易维持?每个条目的评级分为 4 个等级,对应 4 分,范围从理想情况到最坏的情况,分别为:明确的"+";明确的"-";可能的"+";可能的"-"[19]。实验人员将答案记录在 2×2 的网格里。

被试在完成试验之后,对涉及 4 个层次的 SA 内容和过程的 8 个调查(问题)进行自我评估和回答,收集各个条目的数据,涉及可证实的感到的 SA(正确的或不正确的)以及个人对该 SA 的确认(确定的或不确定的)[13]。表格中(2)和(4)是不协调的(Dissonant),特别是表格中(2)的危害比较大[20]。

方法总结:CARS 对军队里人员的 SA 提供了主观评价,并且它有时用作 TSA 测量,但是具体问题的适应性以及回答的确信度还需要进一步考虑,如对评估结果的置信度进行扩展,像 McGuinness 和 Systems[21]采用的 SA 定量分析(Quantitative Analysis of Situational Awareness, QUASA)方法中就对 SA 调查进行了扩展,如表 1 所示。

表 1　被试对回答问题的自我评价

测量描述(Probe Statement)	评价	置信度
一排敌方的坦克 现在正在离开这个城市	[　]真 [√]假	[　]非常高 [　]高 [　]中 [√]底 [　]非常低

McGuinness 和 Foy 报告说,这种方法已经成功地用于研究航空公司飞行员的眼球运动探测、商用飞机飞行甲板上的自动化信任以及军事指挥和控制[22]。

(四)TSAGAT 班组情景意识整体评估技术

方法简介:班组情景意识整体评估技术(Team Situation Awareness Global Assessment Technique, TSAGAT)为研究创伤复苏环境中情景意识可靠性而发展的,其目标是在多学科创伤治疗班组中创造并验证治疗方法的有效性。

实验过程:TSAGAT 使用人体病人模拟(HPS)实验,由 4 个创伤班组完成 2 个单独的创伤情节模拟。学生、初级住院医生、高级住院医生和医护人员这 4 个创伤班组中每组分别分配有外伤组长、护士和气道管理人员。测量以个别 SAGAT 方法作为基础[23]。个别 SAGAT 由各个领域的专家开发,并包含共享和简单的知识问题。使用 SAGAT 实时评估班组,并使用视频审查与传统检查列表进行评估。TSAGAT 计算为个体 SAGAT 评分的总和,并与传统的检查表核心进行比较。研究结果表明随着班组经验的增加,TSAGAT 得分得到提升,班组之间对 TSAGAT 和交互影响因素的差异具有统计意义。

方法总结:TSAGAT 是第一个将 SA 和 HPS 纳入创伤复苏多学科班组绩效的有效的和可靠的评估工具,TSAGAT 可以补充或改进现有的创伤和重症监护评估方法和课程,并为其他外科教育领域的班组评估提供模板。但 TSAGAT 只是简单地将评分计算为个体 SAGAT 评分的总和,可以视为将班组情景意识看作个体情景意识的总和,缺少共享信息、交流协调的过程计算。

二　TSA 测量方法分析和评价标准

(一)TSA 测量方法分析

文章将根据测量内容,从以下几个方面对 TSA 的 12 种测量方法进行分析,包括测量方法的优缺点、测量中班组成员是否有交流协调,测量是以过程还是以结果为导向进行、和该方法是否可应用于实地测验或模拟实验,见表 2。以过程为导向是指参与者在某一时刻对 SA 关键要素的意识;以结果为导向是指参与情景评估的过程,用于生成内存或在结果中的表示。

通过总结分析以上 12 种 TSA 测量方法,可得出如下结论:

(1)测量方法技术多样,各有不同的优缺点,且每个方法都存在争议,没有一个具有普遍适用性的 TSA 测量方法。

(2)针对数字化核电厂的 TSA 测量方法较少。

(3)大部分测量方法中只是简单地将 TSA 归为个体 SA 之和,TSA 的构成有待进一步优化。

表 2 TSA 测量方法分析汇总

方法 \ 分析	优点	缺点	有交流协调的过程	从过程分析	从结果分析	可实地测验	可模拟实验
CAST	(1)提出了设置人为障碍的概念; (2)运用低纬度合成过程来理解班组,有班组协调且不独立于个人。	(1)需要有经验的班组人员进行实验; (2)测量中两组被试专业知识能力、抗压能力等是否一致并未作出解释和调整。	√	√	—	—	√
SACRI	(1)基于 SAGAT 创新了量表; (2)提出了三层次 SA 水平的交互作用。	(1)测试和评估水平有些粗糙; (2)需要有经验的班组人员进行实验。	—	√	—	√	√
CARS	(1)提出了基于心智模型的 4 个层次的 SA; (2)在量表与 4 层次 SA 完成后进行再一次的自我评估,有效提高了回答可靠性。	量表中具体问题的适应性及回答的确信度还需进一步确认。	—	—	√	√	√
TSAGAT	第一个有效地将 SA 与 HPS 结合的评估工具。	(1)简单地将 TSA 计算为个体 SA 之和; (2)缺乏交流协调的过程。	—	—	√	—	√
DUO WOMBAT	(1)具有电子游戏的外观和感觉,可操作性极强; (2)可在实验前评价预估双操作班组中开发良好 SA 水平的潜力。	(1)只能从自身的实验环境中评估 SA 水平; (2)有固定的实验平台,不能与实验设置耦合; (3)可能很大程度上改变实验的真实性。	—	√	—	—	—
CSA	(1)非常注重沟通交流的过程; (2)有专门开发的特色信息传递矩阵; (3)对错误类型和决策提示非常敏感,适合现场测试。	(1)对专家评价的依赖程度过高,影响可信度; (2)过多的语言表达可能影响被试非语言方面的行为。	√	—	√	√	√
IMA	(1)以过程为导向,注重过程参与; (2)有测量个体之间共享交互的数据。	(1)一定程度上依赖注视点技术; (2)很大程度上限制了被试的机动性; (3)过多的语言表达可能影响被试非语言方面的行为。	√	√	—	—	√
TARGETs	(1)预设任务,培训与操作更为便利; (2)可用于实地测量,也可用于模拟环境。	(1)预编脚本有一定的局限性; (2)未给出班组流程的指示; (3)可能错误评估被试 SA 或对 SA 造成不利影响; (4)过多的语言表达可能影响被试非语言方面的行为。	√	—	√	√	√
SAW-IM	(1)提出了同行评分的操作; (2)提出了分段评分,可有效提高 SA 的可靠性; (3)在 CLSA 方法的基础上优化了工作负荷与 SA 比例的结构。	(1)需要有经验的班组人员进行实验; (2)可能改变研究任务的执行,降低实验环境的真实感。	—	—	√	—	√
FASA	(1)更注重“过程”而不是“结果”; (2)可检测 SA 潜在的认知过程影响,可作为 SA 水平的诊断工具。	(1)量表中可能忽略了一些认知因素; (2)缺乏语言沟通与协调。	—	√	—	√	√
IPTE	该测量方法还在进一步研究。 (1)考虑了协调沟通的共享 SA; (2)设置控制变量的实验,可适用于需要控制相关变量的模拟环境。	(1)记录响应时间可能会有偏差; (2)很大程度上依赖于嵌入式探测设备,机动性差,不可实地测量。	√	—	√	—	√

续表

方法 \ 分析	优点	缺点	有交流协调的过程	从过程分析	从结果分析	可实地测验	可模拟实验
SALIANT	(1)可变化场景,适用性强; (2)可用于实地测量,也可用于模拟环境; (3)对被试的要求不高,可操作性强。	(1)必须多次开发不同的场景等关键要素; (2)依赖于参与被试的理解与意愿。	—	—	√	√	√

(4)大部分实验测量方法在测量的过程中忽略了班组成员的交流协调过程。交流协调是测量共享因素的重要内容。

(5)大部分测量方法是以结果为导向进行分析,只注重了班组成员 SA 在某一时刻的最后状态,而未注重情景评估的程序与内容,即以过程为导向的分析。从过程分析能够更明确地测量 TSA。

(6)12 种 TSA 测量方法几乎都可以作为模拟实验对 TSA 进行测量,而大部分方法在实地测量方面都有局限。

(二)TSA 测量方法评价标准

TSA 测量方法是进行 TSA 测量实验操作的规则、步骤和途径,因此,评估 TSA 测量方法的好坏对 TSA 测量实验有着重要意义。本文将采用等级评估法作为评价指标,即确定每一个评价指标的的评分等级标准,将每个评价等级标准分为三等(高、中、低分别对应于 3 分、2 分、1 分),通过整合大量测量实验方法文献与专家评估相结合的方法,对以上 12 种 TSA 测量方法分别从方法整体、实验流程、测量结果三个方面的多层次内容进行评价,为开发出有效性、可靠性更高的 TSA 测量方法奠定基础。

1. 从方法整体层次出发

在 TSA 测量中,能够测量 TSA 的实验方法将指导人员如何进行测量。一个好的 TSA 测量方法,应该能够准确地测量所需实验数据,并拥有详细易懂的步骤说明,使实验人员能够更简单地完成实验。具体标准如下:

(1)有效性:它是指测量工具或手段能够准确测出所需测量的事物的程度。

(2)指导性:方法必须给予分析人员进行分析的指导,使其明白工作的先后顺序,避免重复和资源浪费[38]。

(3)具体性:方法的指导步骤应该尽可能具体化,并附有合适的说明,不能过于概括,如通过决策流程图形式说明[30]。

(4)可理解性:方法应该通俗易懂,减少职员的培训。同时易使不同的人员使用该方法进行分析能得到一致的结果[30]。

2. 从实验流程出发

TSA 测量实验流程是指导实验操作的重要部分。在实验过程中可以设置 TSA 测量操作和获取各种有关 TSA 的数据,因此从前期培训到后期数据结果的获取都需要一套正确可用的流程来保障实验的流程的执行。具体标准如下:

(1)前期培训复杂程度:实验操作熟悉需要时长是否过长,培训次数是否过多,培训效果是否良好,通过率是否过低等。

(2)可操作性:测量实验在具体实施前及过程中的培训,测量方法在运用起来是否好用,是否操作简便、易上手,在实地或模拟实验的运用可行度的高低。

(3)实验数据获取的难易程度:实验数据获取等待时间是否较短或适中不影响操作,数据的偶然性差错发生概率是否较低。

3. 从测量结果出发

测量结果是由测量所得的赋予班组成员 TSA 被测量的值,即 TSA 的获取数据。要分析 TSA 的高低与可靠性,所获得的 TSA 数据就必须具备可靠性与真实性的条件,这样才能有效反映出 TSA 的不足与可提高之处。

(1)可靠性:指的是采取同样的方法对同一对象重复进行测量时,其所得结果相一致的程度。

(2)真实性:测量实验的分析结果符合实际,经得起检验。

(三)比较结果分析

从方法整体来看,以上 12 种测量方法都可以对 TSA 进行测量,但是 CAST、SACRI 和 CARS 方法更为成熟,尤其是 CARS 方法,它将每个测量步骤中的详细分级说明都已表述出来。而 TSAGA 和 SALIANT 方法可能依赖于被试的意愿,在测量准确度上有待改进,属于中等偏上的测量方法。DUO WOMBAT 方法可用来在实验前预估 SA 水平的高低[24],FASA 方法可作为诊断工具来确定 SA 水平高低的原因,这两个方法都对 TSA 的测量有重要意义。另外,CSA[25] 和 SAW-IM 方法在测量评分的过

程中任务执行可能会受影响,因此可能改变测量的有效性,这两种测量在数据有效性水平上还需提高。IMA[26-27]和 TARGETs[28]方法受环境和方法限制,且过于依赖语言沟通,可能会影响到被试 SA 的有效性。IPTE 方法还处于开发中,它可适用于研究变量控制的环境,对研究 TSA 很有意义。

从实验流程来看,以上 12 个测量方法在流程指导方面有较大差异,其中 CARS,TSAGAT 和 TARGETs 方法的流程更为具体,且对被试水平要求不高,可操作性强。CAST,SACRI 和 SAW-IM 需要有经验的操作员参与实验,对被试要求很高,实验成本较高。DUO WOMBAT,CSA,IMA 和 FASA 没有详细的实验步骤方法,但操作简单且对被试水平要求不高,因此完成实验难度较低。而 IPTE,SALIANT 在数据获取上较为困难,对实验设置与操作的要求较高。

从测量结果来看,SACRI、CAST、CARS 和 IMA[29]、FASA 方法在可靠性与真实性上评分较高,特别是 SACRI 方法的置信度最高,实验被试、被试行为、实验流程等内容都对数据的影响较小。TSAGAT、CSA、TARGETs 和 SALIANT 方法数据结果可靠性较低,数据易受多种因素影响,特别是在专家参与度过高的实验中,数据的主观性较强。另外,DUO WOMBAT 和 SAW-IM 方法所需要的活动可能会在很大程度上改变实验的真实性。IPTE 方法还处于开发中,实验流程还未确定,无法获取测量结果信息。

评价结果表明,各种 TSA 测量方法都有自己的优劣之处,还需进一步完善。为了发展一种更为有效的数字化核电厂 TSA 测量方法,增强 TSA 测量方法的可用性,建议如下。

在方法整体上:(1)不能只简单地将 TSA 归为个体 SA 之和,应考虑 TSA 的构成,包括个体 SA、共享 SA 与班组 SA 等,对这三个部分分别进行实验测量,然后综合考察 TSA;(2)在实验测量的过程中需要更加注重交流协调的过程,明确交流协调在测量共享 SA 与班组 SA 中的作用;(3)应注重以过程为导向的测量,从过程内容更加详细明确地分析情景评估的程序,更加注重被试的时段状态和参与感,而不是只注重班组成员 SA 在某一时刻的最后状态;(4)大部分方法在实地测量方面都有所局限,应开发既可应用于模拟实验又可应用于实地实验的测量方法,为实地测量数字化核电厂的 TSA 提供更大的可能性。

在实验流程上:(1)设置实验时,需要提供有效降低被试的经验、知识水平的可能,以减小实验成本和实验前期培训的复杂程度,缩短实验周期;(2)应设置通俗易懂的实验流程,提高实验的可操作性;(3)需注意实验过程中被试的行为是否会影响到实验数据的获取(真实有效性),应解决或排除对实验有影响的流程步骤;(4)实验设置需要更加清晰地明确数据的获取来源,以降低实验数据获取的难度。(5)根据 TSA 不同情景(如小破口事件)不同阶段的特点设置不同的问卷,并体现团队合作的交互性,如个体 SA、班组 SA 与共享 SA 的差异。

在测量结果上:(1)在实验设置过程中应尽量降低各种因素(如:实验步骤、语言交流、专业参与度等)对实验结果的影响;(2)实验活动应参照数字化核电厂现实环境来安排,并无外在因素改变实验进程,以保证实验的真实性。(3)实验结果不能单纯计为个体情景意识之和,也不能单纯记为个体 SA、共享 SA 与班组 SA,可根据三者之间的交互关系进行计算。

表 3 TSA 测量方法评价

标准	方法	CAST	SACRI	CARS	TSAGAT	DUO WOMBAT	CSA	IMA	TARGETs	SAW-IM	FASA	IPTE	SALIANT
方法整体	有效性	3	2	2	1	2	2	2	1	2	3	2	2
	指导性	3	3	3	3	3	2	2	2	3	2	—	3
	具体性	2	2	3	2	2	3	2	1	2	2	—	2
	可理解性	3	3	3	3	3	2	1	2	3	1	3	3
实验流程	前期培训复杂程度	1	1	2	3	3	2	2	3	1	2	—	2
	可操作性	2	3	3	2	1	3	2	3	2	3	2	3
	数据获取难易程度	2	3	3	3	3	2	2	2	2	2	1	1
测量结果	可靠性	2	3	2	1	2	1	2	1	3	2	—	1
	真实性	3	3	3	2	1	2	2	2	1	3	—	3

三　结论与讨论

文章对 12 种测量班组情景意识的方法进行描述汇总,首先简单描述了测量方法的优缺点与是否有交流过程等方面的特点,让读者简单明了地了解复杂的实验方法,并进行比较分析。另外文章从方法本身、方法测量过程与结果等方面设置多个标准,来对 12 种班组情景意识测量方法进行评价,评价指标分类较为具体,使评价更具有特色与针对性,易于发现测量技术中可取之处和不足之处,可为未来开发新的班组情景意识提供借鉴。尽管 12 种方法具有多样性和复杂性等特点,但通过建立的评价标准可对方法进行优劣识别,便于实验设置人员根据测量需求、实验环境等实际情况选择合适的测量方法或组合实验(选取不同实验方法中适合的部分进行融合使用),也可作为 TSA 测量技术创新的理论基础。但文章中评价只是依据建立的评价指标来进行评价,带有一定的主观性,评估结果正确与否后期需要进一步检验。

[参考文献]

[1] SEUNG WOO LEE,JINKYUN PARK, AR RYUM KIM, POONG HYUN SEONG. Measuring situation awareness of operation teams in NPPs using a verbal protocol analysis [J]. Annals of Nuclear Energy,2012,43.

[2] 何旭洪, 黄祥瑞. 工业系统中人的可靠性分析:原理、方法与应用[M]. 北京:清华大学出版社, 2007.

[3] WOODHOUSE R,WOODHOUSE R A. Navigation errors in relation to controlled flight into terrain (CFIT) accidents[C]//Proceedings of the 8th International Symposium on Aviation Psychology. Columbus:[s. n.],1995.

[4] ENDSLEY M R. A taxonomy of situation awareness errors [C]//Western European Association of Aviation Psychology 21st Conference. Dublin,Ireland:[s. n.],1994.

[5] KABER D B, ENDSLEY M R. Team situation awareness for process control safety and performance[J]. Process Safety Progress,1998,17(1):43-48.

[6] 李鹏程,张力,戴立操,蒋建军. 复杂工业系统中班组情景意识的研究进展与发展趋势[J]. 原子能科学技术,2017,51(5):879-889.

[7] BANBURY S, TREMBLAY S. A cognitive approachto situation awareness: Theory and application [M]. London: Routledge,2004.

[8] KIM S K, PARK J Y, BYUN S N. Crew resource management training for improving team performance of operators in Korean advanced nuclear power plant [C]//IEEE International Conference on Industrial Engineering and Engineering Management. Hongkong, China:[s. n.], 2009: 2055-2059.

[9] SCHWARTZ D. Training for situational awareness [C]// The Fifth International Symposium on Aviation Psychology. Columbia, Ohio:[s. n.],1991.

[10] JAMIE C, GORMAN, NANCY J, COOKE, JENNIFER L. WINNER. Measuring team situation awareness in decentralized command and control environments[J]. Ergonomics,2006,49(12-13).

[11] GORMAN J C, COOKE N J, PEDERSON H K, et al. Coordinated Awareness of Situation by Teams (CAST): Measuring Team Situation Awareness of a Communication Glitch[J]. Proceedings of the Human Factors and Ergonomics Society Annual Meeting, 2005, 49(3):274-277.

[12] 石玉生,黄伟芬,田志强. 班组情景意识的概念、模型及测量方法[J]. 航天医学与医学工程,2017,30(6):463-468.

[13] 李鹏程,戴立操. 数字化核电厂操纵员的情景意识理论、测量与可靠性研究[M]. 北京:中国原子能出版社,2018:39-42.

[14] DAVID N H, KNUT F, FRODE S, et al. Development of a situation awareness measure to evaluate advanced alarm systems in nuclear power plant control rooms[J]. Ergonomics,1995,38(11):2398-2413.

[15] ENDSLEY M R. Situation awareness in dynamic human decision making: measurement [J]. Human Factors, 1995, 37: 65-84.

[16] ENDSLEY M R. Toward a theory of situation awareness in Dynamic systems[J]. Human Factors,1995, 37(1): 32-64.

[17] MCGUINNESS B, FOY J L. CARS(Crew Awareness Rating Scale): A subjective tool for situational awareness assessment. A Subjective measure of SA: the Crew Awareness Rating Scale (CARS)[R]//Human performance, situation awareness and automation conference proceedings, October 16-19,1999, Savannah, GA.

[18] MCGUINNESS B, EBBAGE L. Assessing human factors in command and control: Workload and situational awareness metrics[C]. Monterey: Paper presented at the Command and Control Research and Technology Symposium, 2001, 39-55.

[19] BRETON R, TREMBLAY S, BANBURY S. Measurement of individual and team situation awareness: A critical evaluation of the available metrics and tools and their applicability to command and control environments [R]. Defence R&D Canada-Valcartier, Technical Report DRDC Valcartier TR 2003-283.

[20] BRETON R, ROUSSEAU R. Situation Awareness Areview of the concept and its measurement [R]. DRDC Valcartier TR 2001-220.

[21] MCGUINNESS B, SYSTEMS B. Quantitative Analysis of Situational Awareness (QUASA): Applying Signal Detection Theory to True/False Probes and Self-Ratings [M]. Command and Control Research and Technology Symposium, 2004.

[22] GAWRON V J. Human performance, workload, and situational awareness measures handbook[M]. CRC Press, 2008.

[23] MICHAEL S CROZIER, HEATHER Y TING, DARRELL C BOONE, et al. O'Regan, Nathalie Bandrauk, Andrew Furey, Cynthia Squires, Joanne Hapgood, Michael P. Hogan. Use of Human Patient Simulation and Validation of the Team Situation Awareness Global Assessment Technique (TSAGAT): A Multidisciplinary Team Assessment Tool in Trauma Education[J]. Journal of Surgical Education,2015,72(1).

[24] MOSIER K L, CHIDESTSER T R. Situation assessment and Situation Awareness in a team setting [C]//Y. Queinnec and F. Daniellou. Designing for Everyone,1991: 798-800.

[25] ANDERSEN H H K., HAULAND G. Mesuring team situation awareness of reactor operators during normal operation: A technical pilot study[R]. Paper presented at the Proceeding of the Fourth International Symposium on Aviation Psychology, Savannah, Georgia,2000.

[26] HELMREICH R L, FOUSHEE H C. Why crew resource management? Empirical and theoretical bases of human factors training in aviation[C]//E. L. Weiner B G Kanki, and R C Helmreich. Cockpit resource management,1993: 3-45.

[27] ADAMS S. Practical considerations for measuring Situational Awareness[C]. Proceedings for the Third Annual Symposium and Exhibition on Situational Awareness in the Tactical Air Environment,1998:157-164.

[28] BANBURY S, HOWES A. Development of generic methodologies for the evaluation of collaborative technologies [R]. DERA Technical Report,2001,005-2927.

[29] FINK A A, MAJOR D A. Mesuring team situation awareness: A comparison of three techniques[C]. Paper presented at the Proceeding of the Fourth International Symposium on Aviation Psychology,Savannah,Georgia,2000.

[30] 李鹏程,张力,王以群. 框架式人误原因分析技术评价研究[J]. 人类工效学,2009,15(1):44-47.

Comparative Analysis of Experimental Methods for Team Situational Awareness Measurement

LI Peng-cheng, JIN Xiao[1]

(*University of South China*, *Hengyang* 421001, *China*)

Abstract: In order to comprehensively analyze and evaluate the advantages and disadvantages of the team situational awareness measurement methods, this article focuses on the four main methods of the team situational awareness measurement method, including CAST, SACRI, CARS and TSAGAT assessment technology for situational awareness. Further, multiple evaluation criteria were set up from the method itself, method measurement process and results, so as to evaluate the 12 situational awareness measurement methods. The results show that the 12 measurement methods have their own characteristics, but the established evaluation criteria can identify the advantages and disadvantages of the method, and it is convenient for the experiment setup personnel to select the appropriate measurement method or combination experiment according to the actual conditions such as measurement requirements and experimental environment, provide ideas and suggestions for the development of the new nuclear power plant team situational awareness measurement method.

Key words: digital nuclear power plant; team situational awareness; measurement method; evaluation criteria