

治疗效果评定量表在中国大学生群体中的信效度检验

余壮^{1,2}, 孙启武¹, 江光荣¹

(1.华中师范大学心理学院暨湖北省人的发展与心理健康重点实验室, 武汉 430079; 2.上海应用技术大学大学生心理健康教育中心, 上海 201418)

【摘要】 目的:检验治疗效果评定量表(ORS)中文版的信度和效度。**方法:**分别从临床和非临床大学生群体中进行取样并收集数据。非临床样本为武汉市两所高校的202名大学生,该样本连续4次(间隔一周)同时施测OQ-45和ORS中文版问卷,以此检验ORS的信效度;利用参与某高校心理咨询的245名新当事人检验ORS对心理干预的敏感性。**结果:**ORS的内部一致性在0.80以上;连续4次测量结果显示,ORS一周后的重测信度为0.80~0.84;ORS与OQ-45中文版的相关为-0.60~-0.59;ORS的得分能够有效区分临床群体与非临床群体,且对心理干预足够敏感。**结论:**ORS中文版具有良好的心理学测量属性,可用于测量中国大学生当事人的心理治疗效果。

【关键词】 治疗效果评定量表; 信度; 效度

中图分类号: R395.1

DOI: 10.16128/j.cnki.1005-3611.2017.02.017

Reliability and Validity of Chinese Version of Outcome Rating Scale

SHE Zhuang^{1,2}, SUN Qi-wu¹, JIANG Guang-rong¹

¹School of Psychology, Central China Normal University and Key Laboratory of Human Development and Mental Health of Hubei Province, Wuhan 430079, China; ²Shanghai Institute of Technology, Shanghai 201418, China

【Abstract】 Objective: To validate the Chinese version of Outcome rating scale(ORS). **Methods:** Both clinical and non-clinical subjects were recruited in the current study. The nonclinical group consisted of 202 undergraduate and graduate students, who received four measures in a month(once a week) for examining the reliability and validity. Clinical group is made up of 245 clients from a university counseling center, this sample was used to examine the sensitivity to change of ORS in those clients receiving psychotherapy intervention. **Results:** The Cronbach's alpha coefficients of ORS was above 0.80. The one-week test-retest reliability was 0.80-0.84 for the ORS total score. The correlations between OQ-45 total score and ORS total score were in range of -0.60~-0.59. The ORS effectively distinguished dysfunctional and functional clients, and had good responses to psychotherapy intervention. **Conclusion:** The ORS is a reliable and valid measure for Chinese university student and sub-clinical clients.

【Key words】 Outcome rating scale; Reliability; Validity

1996年,Howard等^[1]提出将治疗效果评估应用于临床实践。Howard等认为,在心理治疗过程中,通过对治疗效果进行系统的监控,并将当事人的反馈提供给治疗师,能够帮助治疗师评估当事人是否按照预期进展,以及当前的治疗方案是否恰当。这将有利于当事人更好的从治疗中获益。该观点提出后,引发界内广泛关注。之后,许多研究者投入该领域的研究,最初的工作重点即为开发能够监控当事人治疗效果的工具。

迄今为止,临床学者已发展出数种用来收集当事人反馈的测量工具,这包括心理治疗效果问卷(Outcome Questionnaire-45, OQ-45)^[2]、治疗效果评定量表(Outcome Rating Scale, ORS)^[3]、临床疗效评估量

表(Clinical Outcomes in Routine Evaluation-outcome Measure, CORE-OM)^[4]、咨询中心心理症状评估量表(Counseling Center Assessment of Psychological Symptom, C-CAP)^[5]和治疗效果包(Treatment Outcome Package, TOP)^[6]等。然而,在这些效果监控工具中,仅有OQ-45和ORS被临床试验证实有效^[7]。

OQ-45是Lambert等人响应Howard等人的号召,发展出的首个治疗效果测评工具^[2]。随后,Lambert及其同事利用OQ-45开展了系列临床随机试验,其研究结果验证了Howard的观点,即通过效果监控能够有效提高当事人的治疗效果,特别是那些早期进展不佳(not on track, NOT)当事人的治疗效果^[8]。随后,OQ-45开始应用于临床实践。但由于OQ-45条目偏多,施测与计分较为费时,当事人与治疗师的接受度均不佳^[3,9],因而在临床应用中受到限制。

有鉴于此,Miller和Duncan在前人研究的基础

【基金项目】 教育部人文社科青年基金项目(13YJC190021)资助; 2016年国家社科基金重大投标项目:国民心理卫生素养及其提升机制与对策(16ZDA232)

通讯作者:江光荣, E-mail:grjiang@yeah.net

上,共同开发出一个简式治疗效果测量工具——治疗效果评定量表(outcome rating scale, ORS)^[3],以便在临床实践中监控当事人的治疗进展情况。ORS以OQ-45为基础,其条目主要改编自OQ-45的三个方面:个人(individual)、关系(relational)和社会(social)。通常认为,当事人在这三方面的改变是测量治疗效果的有效指标^[2,10]。此外,Miller和Duncan还创造性地将视觉模拟量表(visual analog scale)引入ORS的计分。与评定式量表相比,视觉模拟量表具有一定的表面效度^[3]。ORS共有4个条目,施测时当事人只需在相应条目下的10cm线段(无刻度)上快速做出标记,即可完成作答。一般情况下,当事人1分钟内即可完成。因此,ORS的施测与计分均变得相当的方便。然而,ORS仅凭4个条目是否能够测量治疗效果呢?有部分学者对此提出质疑^[11]。但多项研究表明,ORS具有良好的信效度指标。根据Miller等报告的结果,ORS的内部一致性系数为0.93,一周后的重测信度(非临床群体)为0.80~0.84。ORS与OQ-45的平均偏相关系数为0.53~0.69,表明其具有中高强度的效标关联效度。此外,ORS能够区分临床与非临床群体,并且在临床应用中表现出良好的敏感性^[3,12]。此后,ORS受到广泛关注,它不仅广泛应用于各类临床试验^[13-15],而且还受到临床工作者的青睐^[16]。

目前,ORS已发展出成人、青少年及儿童三个版本,并被译成25国语言,各版本均可供个人免费下载及使用^[17]。ORS在中国群体中是否具有同等的心理学测量属性,目前尚未获得相关研究的支持。因此,本研究在征得原作者同意后,拟在中国大学生群体中检验ORS的信度和效度。

1 方 法

1.1 被试

本研究分别从临床和非临床两个群体中进行取样。武汉市两所高校的202名本科生和研究生构成本研究的非临床样本。其中男性48名,女性154名,年龄跨度为17~28岁,平均年龄 20.67 ± 1.89 岁。临床样本来自某高校心理咨询中心的来访者。研究对象为研究期间前往该中心接受心理咨询的新个案(本科生或研究生)。最初有245人参与研究。首次ORS测量的基本信息如下:男生57名,女生185名,其中3人性别信息缺失。当事人年龄跨度为16~32岁,平均年龄 21.80 ± 2.96 岁。

1.2 测量工具

1.2.1 治疗效果问卷(OQ-45) OQ-45是一个自陈量表,包括45道李克特5点计分题,分别测量主观不适、人际关系和社会角色表现三个方面,由此获得三个分量表分及总分(180分)。得分越高,表明心理问题越严重。该量表具有较高的内部一致性和重测信度^[18,19],是世界范围内广泛应用的治疗效果测量工具。本研究采用骆宏等人修订的OQ-45中文版。其研究表明,OQ-45在中国当事人群体中信效度指标良好^[18]。

1.2.2 治疗效果评定量表(ORS) ORS是由Miller和Duncan共同开发的可替代OQ-45的极短问卷^[3]。其条目主要来自OQ-45评估当事人机能的三个方面:身心健康、人际关系和生活社交,另增加一个条目评估整体生活状态。因此,ORS共有4个条目。量表使用10cm视觉模拟评分,每一条目要求当事人在10cm线段上做标记,越往左标记代表状况不好,越往右标记代表状况良好。计分使用直尺测量,对应的线段长度即为该条目分数。量表总分处于0~40分,低分表明痛苦程度高,高分表明痛苦程度低。本研究使用的ORS中文版在相关网站下载^[17],并从ORS作者Duncan处获得研究授权。经3位临床方向的教授审定,网站提供的中文版ORS量表内容简洁明了,故未做任何修改。

1.3 研究程序

1.3.1 非临床样本 按照前人惯例,连续进行4次施测^[3],每次施测间隔1周。施测时候,被试同时填写中文版ORS和OQ-45问卷。另为保护被试隐私,每次施测,我们均提前将ORS与OQ-45一起装入信封。所有被试均为班级环境施测,现场作答完毕后由主试当场收回。四次施测程序一致,只是在第三测施测完毕后,统一安排赠予被试一份礼品作为对参与研究的感谢。前后4次测量获得有效问卷的人数分别为202人、175人、160人、167人。被试人数减少是因为被试未能参加该次测量或问卷作答不完整而未纳入数据统计。

1.3.2 临床样本 在某大学心理健康中心施测ORS问卷。主要由中心助理邀请前来中心咨询的新当事人(首次会谈前)参与研究,在征得当事人同意后,将装入信封内的ORS问卷交给当事人。本研究是一项综合研究的一部分,ORS问卷以每次会谈为基础连续施测。研究时间为2015年10月至2016年4月。临床样本的主要目的为检验ORS的敏感性,因此,当事人至少需要参加2次会谈,并完整作答首次及最后一次会谈的ORS问卷。最终,共有186名

当事人符合此要求,并进入数据分析。

1.4 统计方法

利用SPSS17.0对数据进行统计分析。

2 结 果

2.1 基础数据

首先分析了非临床样本和临床样本中的平均数和标准差(见表1)。

表1 非临床与临床样本的描述统计

	n	测量工具	\bar{X}	SD
非临床样本	202	ORS	30.25	6.41
		OQ-45	54.62	15.01
临床样本	245	ORS	22.03	6.97

2.2 信度

2.2.1 内部一致性信度 非临床样本的Cronbach's α 为0.90($n=202$),临床样本的Cronbach's α 为0.84($n=245$)。

2.2.2 重测信度 计算连续两次测量分数间的相关系数。统计结果显示,在非临床群体中,ORS的重测信度为0.79~0.82(见表2);在临床群体中,ORS的重测信度为0.55~0.63(见表3)。

表2 非临床样本中ORS与OQ-45的重测信度

	1-2次施测 ($n=175$)	2-3次施测 ($n=160$)	3-4次施测 ($n=167$)
ORS非临床样本	0.79**	0.82**	0.82**
OQ-45非临床样本	0.79**	0.81**	0.80**

注:** $P<0.01$,下同。

表3 临床样本中ORS的重测信度

	1-2次施测 ($n=186$)	2-3次施测 ($n=151$)	3-4次施测 ($n=127$)
ORS临床样本	0.63**	0.60**	0.55**

2.3 效度

2.3.1 效标关联效度 在四次同时测量过程中,ORS与OQ-45总分之间相关系数保持相对稳定,为-0.60~-0.59。见表4。

2.3.2 区分效度 作为治疗效果的测量工具,ORS需要能够区分临床群体与非临床群体。利用独立样本 t 检验对临床与非临床样本的初测分数进行差异性检验,结果显示,两个群体中的ORS初测平均分差异显著($t=-12.86, P<0.00$),临床群体的得分要显著低于非临床群体。

2.3.3 对改变的敏感性 根据Lambert和Bergin^[20]的观点,如果效果量表是有效的,那么它应该对心理干预敏感,即它应该能够反映心理干预所带来的改

变;而在非治疗群体中则保持稳定。在临床样本中,当事人至少需要参与两次心理咨询才可比较治疗效果。本研究中,总计有186个样本符合此要求。分别取当事人第一次会谈及最后一次会谈的ORS分数为治疗效果的前后测,采用相关样本 t 检验分析前后测差异,结果发现前后测分数差异显著,ORS后测分数显著高于前测分数。在非临床样本中,将被试的第1次和第4次ORS分数作为前后测,进行相关样本 t 检验。结果显示,非临床群体中前后测分数差异不显著。

表4 ORS与OQ-45间相关

第1次施测 ($n=202$)	第2次施测 ($n=175$)	第3次施测 ($n=160$)	第4次施测 ($n=167$)
-0.59**	-0.60**	-0.60**	-0.60**

注:OQ-45得分越高,表明心理问题越严重;ORS得分越高,表明心理问题越轻,所以二者相关为负;** $P<0.01$ 。

表5 ORS的敏感性检验

	n	前测 \bar{X} (SD)	后测 \bar{X} (SD)	t	P
非临床样本	166	30.51(6.26)	31.03(5.96)	1.735	>0.05
临床样本	186	21.27(6.90)	28.30(7.09)	12.89	<0.00

3 讨 论

一般说来,4个条目的ORS其信度应低于OQ-45。但国外研究发现,ORS与OQ-45有着相近的内部一致性系数^[3]。本研究亦得到类似结果。ORS具有很高的内部一致性,反映出量表各条目间是高度相关的,这表明ORS更应被看作是对心理痛苦的整体测量而非包含分量表的多维测量工具^[3]。研究结果还显示,与OQ-45相比,ORS亦有可比的信度。在四次测量中,ORS与OQ-45的重测信度都较为接近。但临床群体的重测信度要低于非临床群体,这可能是由于ORS对心理干预较为敏感所致。

鉴于ORS是基于OQ-45发展出来,二者理应存在较高的相关。但研究结果却低于预期:ORS与OQ-45间仅为中等程度的相关。原因可能有二:一为ORS的条目较为简洁(仅4个条目),这导致它遗漏了部分信息;一为二者使用的计分方式不同所致。此外,临床与非临床大学生样本数据的 t 检验分析表明,ORS具有较好的区分效度,能够有效区分这两个群体。最后,我们检验了ORS的敏感性:临床样本中ORS前后测分数差异显著,而非临床样本中ORS前后测分数无明显改变。这表明ORS对心理干预是敏感的。

综上可知,ORS在中国大学生群体具有较好的

信效度指标。但ORS仍然存在一些局限。首先,作为一个简短治疗效果测量工具,ORS固然无法像较长的量表(如OQ-45)那样提供丰富的信息。如它遗漏了自杀、酗酒及药物使用等重要信息。其次,ORS由当事人主观评定,这可能影响评分的准确性。但ORS设计之初是为了追踪当事人的治疗进展,其较高的可行性或许可以在某种程度上弥补以上不足。如一项关于ORS和OQ-45的使用依从率的调查发现^[3],一年后ORS的使用率仍保持在89%,而OQ-45的使用率仅为25%。这表明,利用测量工具对治疗效果进行追踪,除需保证量表的信效度外,还应考虑现实层面的可行性。ORS则较好地实现了可用性与可行性之间的平衡。

参 考 文 献

- Howard KI, Moras K, Brill PL, et al. Evaluation of psychotherapy: Efficacy, effectiveness, and patient progress. *American Psychologist*, 1996, 51(10): 1059-1064
- Lambert MJ, Burlingame GM, Umphress V, et al. The reliability and validity of the Outcome Questionnaire. *Clinical Psychology & Psychotherapy*, 1996, 3(4): 249-258
- Miller SD, Duncan BL, Brown J, et al. The outcome rating scale: A preliminary study of the reliability, validity, and feasibility of a brief visual analog measure. *Journal of Brief Therapy*, 2003, 2(2): 91-100
- Evans C, Connell J, Barkham M, et al. Towards a standardised brief outcome measure: psychometric properties and utility of the CORE-OM. *The British Journal of Psychiatry*, 2002, 180(1): 51-60
- McAleavey AA, Nordberg SS, Hayes JA, et al. Clinical validity of the Counseling Center Assessment of Psychological Symptoms-62 (CCAPS-62): Further evaluation and clinical applications. *Journal of Counseling Psychology*, 2012, 59(4): 575-592
- Kraus DR, Seligman DA, Jordan JR. Validation of a behavioral health treatment outcome and assessment tool designed for naturalistic settings: The Treatment Outcome Package. *Journal of Clinical Psychology*, 2005, 61(3): 285-314
- Lambert MJ, Shimokawa K. Collecting client feedback. *Psychotherapy*, 2011, 48(1): 72-79
- Lambert MJ, Whipple JL, Hawkins EJ, et al. Is It Time for Clinicians to Routinely Track Patient Outcome? A Meta-Analysis. *Clinical Psychology Science & Practice*, 2003, 10(10): 288-301
- Brown J, Dreis S, Nace DK. What really makes a difference in psychotherapy outcome? Why does managed care want to know? 1999
- Kazdin AE. Methodology, design, and evaluation in psychotherapy research. 1994
- Duncan BL, Reese RJ. Clinical and scientific considerations in progress monitoring: When is a measure too long? *Canadian Psychology/Psychologie Canadienne*, 2013, 54(2): 135-137
- Bringhurst DL, Watson CW, Miller SD, et al. The Reliability and Validity of the Outcome Rating Scale: A Replication Study of a Brief Clinical Measure. *Journal of brief Therapy*, 2003, 5(1): 23-30
- Reese RJ, Norsworthy LA, Rowlands SR. Does a continuous feedback system improve psychotherapy outcome? *Psychotherapy: Theory, Research, Practice, Training*, 2009, 46(4): 418-431
- Anker MG, Duncan BL, Sparks JA. Using client feedback to improve couple therapy outcomes: A randomized clinical trial in a naturalistic setting. *Journal of Consulting and Clinical Psychology*, 2009, 77(4): 693-704
- Reese RJ, Toland MD, Slone NC, et al. Effect of client feedback on couple psychotherapy outcomes. *Psychotherapy: Theory, Research, Practice, Training*, 2010, 47(4): 616-630
- Boswell JF, Kraus DR, Miller SD, et al. Implementing routine outcome monitoring in clinical practice: Benefits, challenges, and solutions. *Psychotherapy Research*, 2015, 25(1): 6-19
- <https://heartandsoulofchange.com/content/measures/download.php>
- 李钰静, 骆宏. 心理咨询效果评估量表中文版的信度和效度. *中国心理卫生杂志*, 2009, 23(2): 105-107
- 秦佑凤, 胡妹婧. 心理咨询效果问卷(OQ-45)在中国部分人群的试用报告. *中国临床心理学杂志*, 2008, 16(2): 138-140
- Lambert MJ, Bergin AE. The effectiveness of psychotherapy. W: Bergin AE, Garfield SL, (Eds.). *Handbook of psychotherapy and behaviour change*. New York: John Wiley, 1994. 143-189

(收稿日期:2016-07-12)