

医学新生网络测评SCL-90量表的参考值和信效度研究

强 景¹, 陈井婷^{2*}, 李文福^{1*}

¹济宁医学院精神卫生学院, 山东 济宁

²济宁医学院, 山东 济宁

收稿日期: 2022年4月10日; 录用日期: 2022年5月12日; 发布日期: 2022年5月19日

摘 要

目的: 揭示医学专业新生网络施测症状自评量表(SCL-90)的参考值和信效度, 探讨时代变迁对问卷信度的影响。方法: 以某医学院校2010~2019年间10个年级的35446名新生的SCL-90问卷的测验数据(有效数据33673份, 有效率95.0%)为样本, 对SCL-90问卷进行内部一致性信度、内容效度和结构效度分析。结果: ① 男生和女生的心理问题发生率分别为23.89%和23.87%, 城市和农村学生心理问题发生率分别为23.54%和24.11%, 心理问题发生率在不同性别和不同生源地上的差异均无统计学意义(P 值均 > 0.05); ② 以题总相关为指标的项目分析发现80道题目与总分相关在0.40~0.69之间(均 $P < 0.01$); ③ 10个因子的内部一致性信度系数在0.70~0.89之间; ④ 躯体化、强迫、人际关系、抑郁、偏执、精神病性、睡眠及饮食7个因子的信度系数与测评年份的相关系数在0.70~0.99之间(均 $P < 0.01$); ⑤ 以因子分与总分相关为指标的内容效度分析发现10个因子与总分之间的相关系数在0.76~0.92之间(均 $P < 0.01$); ⑥ 以各因子分之间的相关系数为指标的结构效度分析发现因子分间的相关系数为中等或中等偏上(0.55~0.82); ⑦ 来自10个因子的12道题目之间的相关均具有统计学意义(均 $P < 0.01$), 来自同一因子的两道题目间的相关甚至小于与其他因子题目间的相关。结论: ① SCL-90问卷的内部一致性信度、内容信度较高, 但结构效度存在问题; ② 网络施测没有导致问卷信度的下降。

关键词

症状自评量表, 信度, 效度, 网络施测

Network-Based Survey Exploring the Reference Value, Reliability and Validity of SCL-90 among Medical Freshman

Jing Qiang¹, Jingting Chen^{2*}, Wenfu Li^{1*}

¹School of Mental Health, Jining Medical University, Jining Shandong

*通讯作者。

²Jining Medical University, Jining ShandongReceived: Apr. 10th, 2022; accepted: May 12th, 2022; published: May 19th, 2022

Abstract

Objective: To explore the marginal value, reliability and validity of Symptom Check List 90 (SCL-90) and assess the effect of network-based survey on reliability. **Method:** The item analysis, internal consistency reliability, content validity and construct validity of SCL-90 were analyzed based on 35446 freshmen's data (including 33673 valid data) of 10 grades in the past 10 years (2010~2019) from a medical university. **Results:** ① The rate of mental disturbances of male and female students was 23.89% and 23.87%, while the rate of mental disturbances of urban and rural students was 23.54% and 24.11%. The rate of psychological problems between male and female or urban and rural students had no significant difference; ② With regard to internal consistency, 80 item-total correlations were more than 0.4 ($P_s < 0.01$); ③ The Cronbach's α of all ten factors reached 0.70~0.89; ④ The correlation coefficients between α of somatization, obsessive-compulsive, interpersonal sensitivity, depression, paranoid ideation, psychoticism and sleep-diet and the data collected year were 0.70~0.99 ($P_s < 0.01$); ⑤ With regard to content validity, the correlation coefficients between the scores of ten factors and total score were 0.76~0.92 ($P_s < 0.01$); ⑥ With regard to construct validity, the correlation coefficients between the score of each factor and another were moderate or even higher (0.55~0.82); ⑦ The correlations between the 12 items from 10 factors were statistically significant ($P_s < 0.01$), and the correlation coefficients between the 2 items from the same factor were lower than that between the 2 items from the different factors. **Conclusion:** ① The internal consistency reliability and content validity were satisfied, while the construct validity was somewhat of a question. ② There were no effects of network-based survey on the reliability of SCL-90.

Keywords

Symptom Check List 90, Reliability, Validity, Network-Based Survey

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>

Open Access

1. 引言

近年来频发的大学生焦虑、抑郁甚至自杀等事件,使大学生的心理健康问题成为学校、家庭和社会关注的焦点。医学生作为在院校特点、学制期限和学习繁重程度等方面异于其他专业的学生群体,承受了更大的心理压力,易于诱发其心理健康问题(李茵等, 2016; 辛素飞等, 2019)。另一方面,日益紧张的医患关系也增加了医学生的心理负担,影响其心理健康(邵一鸣等, 2016; 王雨桐等, 2018)。症状自评量表(Symptom Check List 90, SCL-90)虽已广泛用于测量中国人各类群体的心理健康水平,但尚未有研究探讨医学专业新生心理健康的敏感指标,以揭示医学生心理健康水平的参考值。研究医学新生心理健康的参考值有利于早期发现其心理健康问题,进行早期心理健康教育,进而预防由心理健康问题导致的悲剧。

信度和效度是心理测验的核心问题(易晓明, 2020),是评价心理测量优劣的常用指标。作为最常用的心理健康测评工具之一, SCL-90 问卷已广泛用于测量当代中国人(刘媛媛等, 2018)、医学专业大学生(辛素飞等, 2019)、老年人(辛素飞等, 2020)、独生与非独生子女(廖友国, 连榕, 2020)、新冠疫情下民众的

心理健康(苏斌原等, 2020)等方面, 其信度和效度问题也引起了比较多的关注(Yu et al., 2019; 陈树林, 李凌江, 2003; 谢华, 戴海崎, 2006)。杨雪花和戴梅竞(2000)使用纸笔施测的方式探讨了 SCL-90 在大学生中的信度和效度, 结果发现其分半信度、内部一致性信度和内容效度较好, 但其结构效度不佳。施玲燕等(2013)使用纸笔测试方式也发现 SCL-90 在大学生群体中的信度和效度较好。最近一项大样本研究也采用纸笔测验形式探讨了 SCL-90 问卷的信度和效度, 结果发现其组内相关系数(intraclass correlation coefficients)、题总相关系数和因子拟合度较高。以上研究表明, 纸笔测验形式的 SCL-90 问卷在中国大学生中的信度和效度较好。

自网络普及以来, 使用手机、电脑等进行心理测验已非常普遍(Musch & Reips, 2000)。由于网络测验在数据收集经济性、数据录入和数据取样便利性、答题形式简便性等方面比纸笔测验具有更大优势(王道阳, 刘争光, 2018), 纸笔测验方式有逐渐被网络测验取代的趋势(Raffaelli et al., 2016)。心理测验在网络施测环境下的信度和效度问题也成为国内外学者关注的热点(Buchanan et al., 2005; 蔡华俭等, 2008)。如有研究探讨了生活满意度量表在网络测验环境下的可靠性问题, 发现测验完成时间、测试环境噪音、被试情绪等影响测试质量。最近的一项研究(易晓明, 2020)探讨了纸笔测试和网络测试对 16PF 问卷中文版信度的影响, 研究采用上海某大学跨度 16 年的 53750 名大学生的测验数据, 对 16PF 问卷中文版的信度和效度进行分析, 结果发现网络测试条件下的信度优于纸笔测试条件, 问卷信度系数的波动在 16 年间变化较小。但到目前为止尚未有研究探讨网络施测对 SCL-90 问卷信度和效度的影响。

因此, 有必要对网络测验条件下 SCL-90 问卷的信度和效度进行系统检验。首先, 作为 1984 年被引进国内的心理健康测评工具, SCL-90 问卷在诸多心理健康测评工具中的使用频次最高(李坚, 2007), 虽然该问卷在纸笔测试条件下的信度和效度已得到大量研究的证实(施玲燕等, 2013; 杨雪花, 戴梅竞, 2000), 但尚没有研究探讨网络施测条件下 SCL-90 问卷的信度和效度。其次, 越来越多的研究开始关注时代变迁对大学生心理健康的影响(李文福等, 2020; 辛自强等, 2012), 以比较不同年代大学生心理健康水平是否发生系统变化, 而时代变迁会导致问卷信、效度的改变(易晓明, 2020), 但尚没有研究关注网络施测环境下时代变迁对 SCL-90 问卷信度和效度的影响。综上所述, 本研究拟对某医学院校 2010~2019 年间共 10 个年级学生的 SCL-90 问卷的施测数据进行分析, 以揭示网络施测条件下 SCL-90 问卷的信度和效度及时代变迁的影响, 并确定 SCL-90 问卷的医学新生常模, 以便提供该人群心理健康筛查的标准, 为有效心理健康教育提供参考。

2. 对象与方法

2.1. 对象

采用北京心海导航教育科技股份有限公司开发的学生心理档案系统, 对某医学院校 2010~2019 级的本科新生进行调查。测试在每年新生参加军训后的 10~12 月份, 以班级为单位进行网络施测, 测试时每个页面呈现 1 个题目。10 年来共测试 35,446 人, 其中有效问卷 33,673 份, 有效回收率为 95.0%。其中男生 14301 人, 女生 19,372 人, 分别占 42.5%和 57.5%; 农村生源 20,036 人, 城市生源 13,637 人, 分别占 60.6%和 39.4%。各年度下研究对象的分布情况见表 1。

Table 1. The distribution of gender and place of birth [n(%)]

表 1. 性别和生源地分布[n(%)]

| 测试年份 | 总样本 | 男 | 女 | 城市 | 农村 |
|------|------|--------------|--------------|--------------|--------------|
| 2010 | 2672 | 1110 (41.54) | 1562 (58.46) | 877 (32.82) | 1794 (67.14) |
| 2011 | 3300 | 1226 (37.15) | 2074 (62.85) | 1021 (30.94) | 2279 (69.06) |
| 2012 | 3181 | 1182 (37.16) | 1999 (62.84) | 1166 (36.66) | 2015 (63.34) |

Continued

| | | | | | |
|------|------|--------------|--------------|--------------|--------------|
| 2013 | 3080 | 1195 (38.80) | 1885 (61.20) | 1243 (40.36) | 1837 (59.64) |
| 2014 | 3388 | 1409 (41.59) | 1979 (58.41) | 1396 (41.20) | 1992 (58.80) |
| 2015 | 3058 | 1339 (43.79) | 1719 (56.21) | 1324 (43.30) | 1734 (56.70) |
| 2016 | 3747 | 1485 (39.63) | 2262 (60.37) | 1481 (39.52) | 2266 (60.48) |
| 2017 | 3817 | 1723 (45.14) | 2094 (54.86) | 1654 (43.33) | 2163 (56.67) |
| 2018 | 3619 | 1791 (49.49) | 1828 (50.51) | 1659 (45.84) | 1960 (54.16) |
| 2019 | 3811 | 1841 (48.31) | 1970 (51.69) | 1815 (47.63) | 1996 (52.37) |

2.2. 研究工具

SCL-90 (Derogatis et al., 1973; 汪向东等, 1999)包括 90 个关于个体感觉体验、情感、思维、意识、行为以及生活习惯、人际关系等方面的问题, 并可以分成躯体化、强迫、人际关系、抑郁、焦虑、敌对性、恐怖、骗纸、精神病性和睡眠及饮食共 10 个因子。每个题目均为 1~5 的 5 点评分, 分别对应无、轻度、中度、偏重和严重。因子均分是该因子所包含的所有题目总分除以题目个数, 得分越高表明心理问题越严重。参照孟健男等(2019)的做法, 以因子均分 ≥ 2 分或量表总分 ≥ 160 分为心理问题检出标准, 量表总分 ≥ 250 分作为严重心理问题检出标准。

2.3. 统计学分析

使用 SPSS 22.0 统计软件包进行各个题目得分与总分之间的 Spearman 相关分析, 各因子和总分的内部一致性信度(Cronbach's alpha)分析、各因子的内部一致性信度与数据收集年份的 Spearman 相关分析、各因子均分的 Pearson 相关分析、各题目得分的 Spearman 相关分析以及源于 10 个因子的 12 道题目得分的 Spearman 相关分析。

各个题目得分与总分之间的相关分析用于项目分析, 以检验各题目的区分度。内部一致性信度分析的结果用于评估 SCL-90 问卷的信度情况以及时代变迁与信度的关系。如果各因子的信度系数与数据收集年份显著相关, 可以推测时代变迁对问卷施测有影响, 反之没有影响。因子均分之间的相关分析可以用于检验问卷的结构效度, 该方法被广泛用于因素独立性的验证(易晓明, 2020)。各因子均分之间的相关系数绝对值越高, 各因子之间的相互独立性越差。因子分与总分之间的相关用于检验问卷的效度, 因子分与总分之间的相关越高, 表明问卷的内容效度越好, 各因子分与总分之间的相关系数超过各因子分之间的相关系数是结构效度的指标。各题目得分之间的相关分析可以用于因素独立性假设的检验, 虽然以往研究指出该方法可能会高估测题与因子之间的关联, 但为了显示题目之间的相关趋势, 本研究以图的形式显示题目之间的不同相关程度, 并进一步以 10 个因子中的 12 道题目(因子 2 和因子 9 各 2 道题目, 其余因子各 1 道题目)为代表进行检验。如果各因子测题符合因素独立性假设, 同一因子下的题目间相关应较强, 而不同因子下的题目间相关应较弱(易晓明, 2020)。即不同因子下题目之间的相关系数绝对值大于同一因子下题目之间的相关系数绝对值, 就可以认为问卷的结构效度较差。

3. 结果

3.1. SCL-90 问卷在医学新生上的参考值

3.1.1. 心理问题检出情况

从收集到的有效问卷中筛选有心理问题和严重心理问题的学生, 发现 8041 名学生存在心理问题, 占比 23.88%, 633 名学生存在严重心理问题, 占比 1.09%。在 8041 名心理问题学生中, 男生和女生心理问

题发生率分别为 23.89% 和 23.87%，差异无统计学意义($\chi^2 = 0.0006, P = 0.980$)；城市和农村学生心理问题发生率分别为 23.54% 和 24.11%，差异无统计学意义($\chi^2 = 1.46, P = 0.226$)。2010~2019 年间，不同性别和生源地的发生率检验结果如表 2 所示。

Table 2. Incidence of gender and place of birth of each year

表 2. 各年份不同性别和生源地的发生率检验

| 年份 | 性别 | | | 生源地 | | |
|------|--------------|--------------|--------|--------------|--------------|--------|
| | 男性(%) | 女性(%) | | 城市(%) | 农村(%) | |
| 2010 | 234 (21.08) | 365 (23.37) | 1.95 | 181 (20.64) | 418 (23.30) | 2.44 |
| 2011 | 204 (16.64) | 317 (15.28) | 1.06 | 135 (13.22) | 386 (16.94) | 7.32** |
| 2012 | 208 (17.60) | 431 (21.56) | 7.27** | 209 (17.92) | 430 (21.34) | 5.37* |
| 2013 | 248 (20.75) | 413 (21.91) | 0.58 | 268 (21.56) | 393 (21.39) | 0.01 |
| 2014 | 330 (23.42) | 408 (20.62) | 3.80 | 283 (20.27) | 455 (22.84) | 3.18 |
| 2015 | 309 (23.08) | 420 (24.43) | 0.76 | 308 (23.26) | 421 (24.28) | 0.43 |
| 2016 | 442 (29.76) | 659 (29.13) | 0.17 | 459 (30.99) | 642 (28.33) | 3.06 |
| 2017 | 529 (30.70) | 651 (31.09) | 0.07 | 497 (30.05) | 683 (31.58) | 1.03 |
| 2018 | 392 (21.89) | 398 (21.77) | 0.01 | 354 (21.34) | 436 (22.24) | 0.43 |
| 2019 | 520 (28.25) | 563 (28.58) | 0.05 | 516 (28.43) | 567 (28.41) | 0.00 |
| 总体 | 3416 (23.89) | 4625 (23.87) | 0.00 | 3210 (23.54) | 4831 (24.11) | 1.46 |

注：*表示在 0.05 水平(双侧)有统计学意义；**表示在 0.01 水平(双侧)有统计学意义，下同。

3.1.2. 医学院校新生 SCL-90 问卷的参考值

由于在性别上，除躯体化、强迫、人际关系、精神病性和睡眠及饮食因子外，其他因子的差异均达到统计学显著(P 值均 < 0.05)；在生源地上，除躯体化、强迫因子外，其他因子的差异均达到统计学显著(P 值均 < 0.05)，因此，分别从不同性别和生源地上计算各因子的参考值。参照已有研究(施玲燕等, 2013)，计算各因子 95% 的百分位数作为参考值，结果如表 3。

Table 3. Reference value of SCL-90 of medical students

表 3. 不同性别不同生源地医学生 SCL-90 问卷的参考值

| 性别 | 生源地 | 人数 | 躯体化 | 强迫 | 人际关系敏感 | 抑郁 | 焦虑 | 敌对性 | 恐怖 | 偏执 | 精神病性 | 睡眠及饮食 | 总分 | 阳性项目个数 | 阳性项目均分 |
|----|-----|-------|------|------|--------|------|------|------|------|------|------|-------|--------|--------|--------|
| 男 | 城市 | 6231 | 2.00 | 2.80 | 2.67 | 2.31 | 2.30 | 2.33 | 2.00 | 2.50 | 2.20 | 2.29 | 200.00 | 72 | 2.83 |
| | 农村 | 7886 | 2.00 | 2.80 | 2.78 | 2.31 | 2.30 | 2.33 | 2.14 | 2.33 | 2.30 | 2.29 | 201.00 | 74 | 2.78 |
| 女 | 城市 | 7244 | 2.08 | 2.70 | 2.67 | 2.38 | 2.30 | 2.33 | 2.14 | 2.33 | 2.20 | 2.14 | 199.75 | 74 | 2.73 |
| | 农村 | 11929 | 2.00 | 2.70 | 2.67 | 2.38 | 2.30 | 2.33 | 2.29 | 2.17 | 2.20 | 2.29 | 198.00 | 75 | 2.68 |

3.2. 信度分析

3.2.1. 项目分析

参照以往研究(Yu et al., 2019)做法，对各个题目得分与总分进行 Spearman 相关分析。结果表明，除了第 4、5、12、13、15、16、35、47、64、82 题在 0.34~0.39 之间外，其余 80 道题目均在 0.40~0.69 之

间, 所有相关系数均在 0.01 (双侧)水平有统计学意义。根据 Nunnally 和 Bernstein (1994)的建议, 题总相关系数超过临界值 0.3 时, 表明项目信度较好。

3.2.2. 内部一致性信度

利用各年度测量数据, 分别计算各因子的内部一致性系数(克伦巴赫 α 系数), 结果如表 4 所示。结果表明, SCL-90 问卷的整体信度较高, 各因子每年份的信度系数均大于 0.7 (Kline, 2020), 只有睡眠及饮食因子在 2010~2015 年间的信度系数略小于 0.7, 表明各因子的信度较好。

Table 4. Reliability coefficient of SCL-90 factors among each year (Cronbach alpha)
表 4. 各年度下 SCL-90 各因子的信度系数(克伦巴赫 α 系数)

| 因子 | 2010 | 2011 | 2012 | 2013 | 2014 | 2015 | 2016 | 2017 | 2018 | 2019 | 总体 |
|-------|------|------|------|------|------|------|------|------|------|------|------|
| 躯体化 | 0.70 | 0.82 | 0.85 | 0.85 | 0.85 | 0.85 | 0.86 | 0.87 | 0.87 | 0.88 | 0.77 |
| 强迫 | 0.83 | 0.81 | 0.82 | 0.82 | 0.83 | 0.83 | 0.82 | 0.84 | 0.84 | 0.85 | 0.86 |
| 人际关系 | 0.81 | 0.82 | 0.82 | 0.83 | 0.84 | 0.84 | 0.85 | 0.85 | 0.85 | 0.86 | 0.83 |
| 抑郁 | 0.82 | 0.86 | 0.88 | 0.88 | 0.88 | 0.88 | 0.89 | 0.90 | 0.90 | 0.91 | 0.84 |
| 焦虑 | 0.88 | 0.80 | 0.84 | 0.84 | 0.84 | 0.84 | 0.85 | 0.86 | 0.86 | 0.86 | 0.89 |
| 敌对性 | 0.82 | 0.77 | 0.78 | 0.78 | 0.77 | 0.80 | 0.81 | 0.80 | 0.81 | 0.82 | 0.85 |
| 恐怖 | 0.76 | 0.70 | 0.74 | 0.74 | 0.74 | 0.74 | 0.74 | 0.75 | 0.78 | 0.77 | 0.80 |
| 偏执 | 0.71 | 0.72 | 0.72 | 0.73 | 0.75 | 0.74 | 0.76 | 0.77 | 0.78 | 0.78 | 0.74 |
| 精神病性 | 0.78 | 0.75 | 0.78 | 0.79 | 0.79 | 0.77 | 0.79 | 0.80 | 0.82 | 0.82 | 0.79 |
| 睡眠及饮食 | 0.68 | 0.67 | 0.68 | 0.70 | 0.68 | 0.68 | 0.71 | 0.71 | 0.72 | 0.73 | 0.70 |

3.2.3. 各因子信度系数与数据收集年份的相关分析

对各因子的克伦巴赫 α 信度系数与数据测量年份进行相关分析, 由表 5 可知, 躯体化、强迫、人际关系、抑郁、偏执、精神病性、睡眠及饮食因子的信度随测评年份逐渐提高, 焦虑、敌对性和恐怖因子的信度系数随年份没有变化, 表明测量年份不会导致问卷信度降低。

Table 5. The relationship between reliability coefficient and year (spearman correlation coefficient)
表 5. SCL-90 各因子信度系数与年份的相关系数(Spearman 相关)

| | 躯体化 | 强迫 | 人际关系 | 抑郁 | 焦虑 | 敌对性 | 恐怖 | 偏执 | 精神病性 | 睡眠及饮食 |
|--------|--------|--------|--------|--------|------|------|------|--------|--------|--------|
| 数据收集年份 | 0.90** | 0.77** | 0.99** | 0.92** | 0.39 | 0.51 | 0.58 | 0.99** | 0.85** | 0.87** |

3.3. 效度分析

3.3.1. 内容效度

各因子分与总分之间的相关可以用于考察量表的内容效度(王极盛等, 1999), 由表 6 可知, 各因子分与总分之间的相关系数在 0.76~0.92 之间, 表明该量表的内容效度较好。

3.3.2. 结构效度

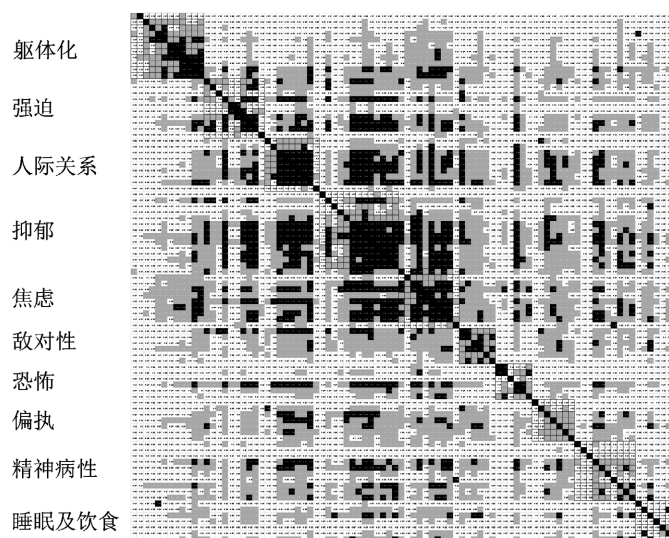
各因子分与总分的相关是否超过各因子分之间的相关是衡量量表结构效度的检验方法之一(王极盛等, 1999), 由表 6 可知, 各因子分之间的相关系数在 0.55~0.82 之间, 接近各因子分与总分之间的相关系数(在 0.76~0.92 之间), 表明该量表的结构效度欠佳, 各因子间的独立性不好。

Table 6. The relationship among each factor of SCL-90
表 6. SCL-90 各因子均分之间的 Pearson 相关分析结果

| 因素 | 相关系数(<i>r</i>) | | | | | | | | | | | |
|----------|------------------|------|------|------|------|------|------|------|------|------|------|--|
| | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 | |
| 1 总分 | 1.00 | | | | | | | | | | | |
| 2 躯体化 | 0.80 | 1.00 | | | | | | | | | | |
| 3 强迫 | 0.88 | 0.64 | 1.00 | | | | | | | | | |
| 4 人际关系 | 0.89 | 0.60 | 0.78 | 1.00 | | | | | | | | |
| 5 抑郁 | 0.92 | 0.68 | 0.79 | 0.81 | 1.00 | | | | | | | |
| 6 焦虑 | 0.91 | 0.74 | 0.77 | 0.78 | 0.82 | 1.00 | | | | | | |
| 7 敌对性 | 0.80 | 0.62 | 0.64 | 0.68 | 0.70 | 0.70 | 1.00 | | | | | |
| 8 恐怖 | 0.76 | 0.57 | 0.64 | 0.68 | 0.67 | 0.70 | 0.55 | 1.00 | | | | |
| 9 偏执 | 0.84 | 0.62 | 0.71 | 0.77 | 0.74 | 0.74 | 0.71 | 0.59 | 1.00 | | | |
| 10 精神病性 | 0.89 | 0.66 | 0.74 | 0.77 | 0.79 | 0.78 | 0.67 | 0.61 | 0.75 | 1.00 | | |
| 11 睡眠及饮食 | 0.82 | 0.69 | 0.67 | 0.66 | 0.72 | 0.73 | 0.64 | 0.57 | 0.65 | 0.71 | 1.00 | |

注：所有相关系数均在 0.01 (双侧)水平有统计学意义。

研究继续对所有题目得分进行相关分析，作为量表结构效度的另一指标。由图 1 可发现，同一因子下的题目相关较高，但也有较多不同因子下的题目相关较高，如人际关系与抑郁因子以及抑郁与焦虑因子。为了量化不同因子间题目的相关情况，表 7 显示分属于不同因子的 12 道题目得分之间的相关系数(均 P 小于 0.01)，属于精神病性因子的第 62 题和第 77 题之间的相关系数几乎比这两道题目与其他 9 个因子的相关都低，属于强迫因子的第 51 题和第 28 题之间的相关系数也表现出同样的趋势，表明因素间的独立性较差。



注：图中题目顺序按照因子进行划分排列，对角线上为 10 个因子内部相关情况，灰色方格表示相关系数大于 0.30，黑色方格表示相关系数大于 0.40。

Figure 1. Correlation coefficient matrix among each item of SCL-90

图 1. SCL-90 各题目的相关系数矩阵

Table 7. The Spearman correlation coefficient among twelve items of ten factors
表 7. 来自 10 个因子的 12 道题目得分的 Spearman 相关结果

| 因子 | 相关系数(<i>r</i>) | | | | | | | | | | | | |
|-------|------------------|------|------|------|------|------|------|------|------|------|------|------|------|
| | 题 56 | 题 51 | 题 69 | 题 30 | 题 33 | 题 67 | 题 50 | 题 43 | 题 62 | 题 59 | 题 77 | 题 28 | |
| 躯体化 | 题 56 | 1.00 | | | | | | | | | | | |
| 强迫 | 题 51 | 0.44 | 1.00 | | | | | | | | | | |
| 人际关系 | 题 69 | 0.39 | 0.39 | 1.00 | | | | | | | | | |
| 抑郁 | 题 30 | 0.43 | 0.43 | 0.44 | 1.00 | | | | | | | | |
| 焦虑 | 题 33 | 0.41 | 0.40 | 0.41 | 0.51 | 1.00 | | | | | | | |
| 敌对性 | 题 67 | 0.35 | 0.34 | 0.39 | 0.36 | 0.36 | 1.00 | | | | | | |
| 恐怖 | 题 50 | 0.39 | 0.44 | 0.40 | 0.44 | 0.44 | 0.32 | 1.00 | | | | | |
| 偏执 | 题 43 | 0.36 | 0.33 | 0.41 | 0.37 | 0.38 | 0.34 | 0.36 | 1.00 | | | | |
| 精神病性 | 题 62 | 0.42 | 0.40 | 0.42 | 0.41 | 0.40 | 0.35 | 0.39 | 0.37 | 1.00 | | | |
| 睡眠及饮食 | 题 59 | 0.33 | 0.34 | 0.36 | 0.37 | 0.36 | 0.38 | 0.33 | 0.33 | 0.35 | 1.00 | | |
| 精神病性 | 题 77 | 0.36 | 0.39 | 0.41 | 0.47 | 0.39 | 0.34 | 0.38 | 0.32 | 0.36 | 0.35 | 1.00 | |
| 强迫 | 题 28 | 0.38 | 0.41 | 0.36 | 0.48 | 0.44 | 0.30 | 0.44 | 0.33 | 0.38 | 0.31 | 0.35 | 1.00 |

注：所有相关系数均在 0.01(双侧)水平有统计学意义。

4. 讨论

4.1. 医学专业新生的 SCL-90 参考值

本研究采用近 10 年的连续测量数据，探讨医学院校新生的心理健康水平参考值。结果发现在 2010~2019 年间的 33673 份有效问卷中有 8041 名学生存在心理问题，占比 23.88%，633 名学生存在严重心理问题，占比 1.09%。该检出率结果与以往研究发现一致(张静, 2016)，表明大学生心理问题的发生率较高，需要引起相关心理健康教育部分的重视。已有研究发现医学新生在强迫症状、人际关系等心理问题检出率上分别为 26.17% 和 20.31% (靳蔚云等, 2016)，其他研究也发现医学专业本科生的阳性检出率为 23.20% (李雪, 2020)，医学专业毕业生的 SCL-90 阳性检出率为 18.70%。本研究的结果进一步表明，医学专业新生的心理问题突出，需要加强心理健康教育和心理卫生干预，以提高其心理健康水平。

研究结果发现，男生和女生心理问题发生率分别为 23.89% 和 23.87%，城市和农村学生心理问题发生率分别为 23.54% 和 24.11%，不同性别和不同生源地学生的心理问题发生率差异均无统计学意义。该结果与以往研究关于大学生心理问题检出率和性别差异的发现一致(施玲燕等, 2013)，表明性别和生源地对医学专业新生的心理问题检出率没有影响。但研究也发现，不同性别医学专业新生在抑郁、焦虑、敌对性、恐怖和偏执因子上的差异具有统计学意义，不同生源地医学专业新生在人际关系、抑郁、焦虑、敌对性、恐怖、偏执、精神病性和睡眠及饮食因子上的差异具有统计学意义。该结果可能是由于不同性别大学生在就业压力等方面存在差异，而导致其心理问题严重程度不同(杜燕红, 2012; 辛自强等, 2012)。其他研究也表明农村生源地和城市生源地大学生持有的教育观念、享有的教育资源以及感受到的生活压力等的差异(罗伏生等, 2009; 周晓琴等, 2013)，使不同生源地大学生的心理健康水平产生差异。结果提示，需要对不同性别和生源地医学专业新生进行有针对性的心理健康教育和干预，以有效促进医学生的心理健康。

4.2. SCL-90 问卷的信度

研究结果发现，SCL-90 问卷的各因子的内部一致性信度系数在 0.70~0.89 之间，仅睡眠及饮食因子

在 2010~2015 年间的信度系数在 0.67~0.70 之间,表明 SCL-90 问卷的信度较好,题总相关分析发现除了第 4、5、12、13、15、16、35、47、64、82 题与总分相关在 0.34~0.39 之间外,其余 80 道题目与总分相关系数均在 0.40~0.69 之间,研究结果表明 SCL-90 问卷的内部一致性信度较高,该结果与以往纸笔测验的发现一致。杨雪花和戴梅亮(2000)使用纸笔施测的方式研究了 SCL-90 问卷在大学生中的信度,结果发现其分半信度和内部一致性信度较好。施玲燕等(2013)也发现 SCL-90 问卷总量表的内部一致性系数为 0.954,各因子的内部一致性系数平均值为 0.765,表明 SCL-90 问卷的信度较好。最近一项大样本研究(Yu et al., 2019)也发现 SCL-90 问卷的组内相关系数(intraclass correlation coefficients)和题总相关系数较高。本研究进一步表明 SCL-90 问卷在网络测验条件下的信度较高,印证了 SCL-90 问卷信度较高的结论。

4.3. SCL-90 问卷的效度

各因子的内部一致性信度较高、因子之间彼此独立是问卷结构效度的重要条件(易晓明, 2020)。根据表 4 可知问卷的信度系数在 0.70~0.89 之间,表明问卷不存在信度问题。根据表 6,各因子分与总分之间的相关系数在 0.76~0.92 之间,达到中等以上的相关,表明该量表的内容效度较好。但各因子分之间的相关系数在 0.55~0.82 之间,同样达到中等以上的相关,表明各因子之间的独立性较差。图 1 和表 7 的结果揭示了各因子分之间相关较高的原因。依据图 1,对角线上为各因子所属题目内部的相关模式,对角线之外的为不同因子所属题目的相关模式,可以发现各因子下的题目不仅与同因子下的其他题目相关较高,与其他因子下的题目相关也较高。依据表 7,来自 10 个因子的 12 道题目之间的相关都较高,属于精神病性因子的第 62 题和第 77 题之间的相关系数几乎比这两道题目与其他 9 个因子的相关都低,属于强迫因子的第 51 题和第 28 题之间的相关系数也表现出同样的趋势,说明原本用于测量不同因子的题目在心理意义上的关系过于紧密(易晓明, 2020)。这些关联紧密的题目分别被用于测量不同的心理症状(例如,第 51 题和第 56 题:51、脑子变空了;56、感到身体的某一部分较弱无力),由此可知,不同因子下的题目有一部分关联过于紧密,语义相似性过高,是导致各因子独立性较差的原因,降低了 SCL-90 问卷的结构效度。

4.4. 时代变迁对 SCL-90 信度的影响

揭示 SCL-90 问卷的信度是否随时代变迁而变化是本研究的第三个目的。根据表 5,躯体化、强迫、人际关系、抑郁、偏执、精神病性、睡眠及饮食 7 个因子的内部一致性信度与测评年份的相关系数在 0.77~0.99 之间,焦虑、敌对性和恐怖 3 个因子的信度系数与测评年份的相关没有达到统计学显著,相关系数分别为 0.39、0.51、0.58。相关结果表明 SCL-90 问卷各因子的信度系数没有因为年代变迁而降低,与最近研究发现的 SCL-90 问卷的信度较高的结论(Yu et al., 2019)一致。同时,该结果进一步印证了网络测试与纸笔测试形式不影响问卷信度的结论(蔡华俭等, 2008; 易晓明, 2020)。易晓明(2020)也指出设计合理的网络测验能够提高测验信度,网络测验对受测者的认知资源需求较小、降低了答题出错概率,进而提高了内部一致性信度。另外,本研究使用数据为学校心理健康教育中心通过班级辅导员完成测试任务的下发,学生在测验服从度上较高,也可能使问卷的一致性信度较高。

4.5. 反思和研究展望

本研究通过对 2010~2019 年间某医学院校 33,673 名新生 SCL-90 问卷的测量数据进行分析,发现不同性别和不同生源地学生的心理问题检出率没有差异,并构建了医学院校新生 SCL-90 问卷各因子 95% 的参考值,为医学院校新生的心理健康评估提供参考。研究也发现 SCL-90 问卷在网络施测环境下的项目质量、内部一致性信度和内容效度较好,能作为大规模网络施测大学生心理健康的测评工具。但本研究

也有部分局限,比如研究数据来源于医学本科院校,其他类型院校或非高等教育受测者是否也有这种结果尚需进一步验证。本研究发现 SCL-90 问卷的结构效度欠佳,在以后的研究中可能需要针对不同的数据进行不同的因子划分,而不是使用统一的因子结构划分标准,使研究结果更适合特定的研究对象,为心理健康测评提供更科学有效的工具。

基金项目

本项目受山东省自然科学基金(JY2015BS03)、山东省高等学校“青创科技计划”项目(2019RWF003)、贺林院士新医学科研基金项目(JYHL2019MS03)、济宁医学院教师科研扶持基金(JYFC2019KJ011)的资助。

参考文献

- 蔡华俭,林永佳,伍秋萍,严乐,黄玄凤(2008).网络测验和纸笔测验的测量不变性研究——以生活满意度量表为例. *心理学报*, 40(2), 228-239.
- 陈树林,李凌江(2003).SCL-90 信度效度检验和常模的再比较. *中国神经精神疾病杂志*, 29(5), 323-327.
- 杜燕红(2012).河南省高职院校大学生心理健康状况的调查研究. *河南科技学院学报*, (8), 93-97.
- 靳蔚云,刘玉龙,辛红梅,郭秀琴(2016).医学生心理健康状况调查分析. *内蒙古医学杂志*, 48(2), 205-207.
- 李坚(2007).心理健康测评工具使用频度分析与思考. *华南师范大学学报(社会科学版)*, (3), 119-122+160.
- 李文福,强景,徐芳芳,李功迎(2020).医学院校大学新生心理症状变迁的横断历史研究. *中国心理卫生杂志*, 34(9), 784-789.
- 李雪(2020).大数据时代下高校大学生心理危机识别与干预机制的构建——基于贵州某医科院校学生自杀态度与心理现状的调查. *法制博览*, (23), 29-31.
- 李茵,苏国范,冀慧玲,牟冬梅,于洗河(2016).医学生生活事件、社会支持与心理健康的调查分析. *中国高等医学教育*, (11), 25-26.
- 廖友国,连榕(2020).独生与非独生子女心理健康变迁的差异——一项横断历史研究. *西南大学学报(社会科学版)*, 46(3), 117-126+203.
- 刘媛媛,武圣君,李永奇,邵峰,苏景宽,刘旭峰(2018).基于 SCL-90 的中国人心理症状现状调查. *中国心理卫生杂志*, 32(5), 437-441.
- 罗伏生,沈丹,张珊明,王小凤,袁红梅,李志强(2009).贫困大学生心理健康状况及其影响因素研究. *中国临床心理学杂志*, 17(3), 272-274.
- 孟健男,司维,邵杰(2019).大学新生 SCL-90 心理测评研究——以某中医药大学药学院 2017 级新生为例. *中国社会医学杂志*, 36(4), 388-391.
- 邵一鸣,孙玉盼,宋静,杨树林(2016).医患关系的媒体舆情对临床医学生的影响. *中国医学伦理学*, 29(5), 864-867.
- 施玲燕,黄水平,卓朗,吴秀娟,赵华硕,耿德勤,张训保(2013).大学生应用 SCL-90 量表的参考值及信效度评价. *中国学校卫生*, 34(2), 223-224.
- 苏斌原,叶苑秀,张卫,林玛(2020).新冠肺炎疫情不同时间进程下民众的心理应激反应特征. *华南师范大学学报(社会科学版)*, (3), 79-94.
- 汪向东,王希林,马弘(1999). *心理卫生评定量表手册(增订版)*. 中国心理卫生杂志社.
- 王道阳,刘争光(2018).SWLS 网络测验中可靠性变量的测量不变性检验. *心理学探新*, 38(1), 80-85.
- 王极盛,李焰,赫尔实(1999).中学生 SCL-90 信度、效度检验与常模建立. *中国心理卫生杂志*, 13(1), 9-11.
- 王雨桐,闫菊娥,黄燕萍,周睿(2018).医患关系的媒体舆情对临床医学生专业学习的影响. *科教导刊*, (12), 139-140.
- 谢华,戴海崎(2006).SCL-90 量表评价. *神经疾病与精神卫生*, 6(2), 156-159.
- 辛素飞,姜文源,辛自强(2019).1993 至 2016 年医学生心理健康变迁的横断历史研究. *心理科学进展*, 27(7), 1183-1193.
- 辛素飞,岳阳明,辛自强(2020).1996 至 2016 年中国老年人心理健康变迁的横断历史研究. *心理发展与教育*, 36(6), 753-761.
- 辛自强,张梅,何琳(2012).大学生心理健康变迁的横断历史研究. *心理学报*, 44(5), 664-679.

- 杨雪花, 戴梅竞(2000). 大学生 SCL-90 信度及效度检验. *中国校医*, (4), 243-245.
- 易晓明(2020). 16PF 中文版的信度、结构效度问题以及测试方式对信度的影响. *中国临床心理学杂志*, 28(2), 326-330.
- 张静(2016). 内蒙古某高校医学生 SCL-90 量表测评结果及综合干预研究. *卫生职业教育*, 34(16), 126-127.
- 周晓琴, 奚晓岚, 程灶火, 季庆, 李萍(2013). 在校大学生心理健康状况调查. *中国临床心理学杂志*, 21(6), 1026-1028.
- Buchanan, T., Ali, T., Heffernan, T. M., Ling, J., Parrott, A. C., Rodgers, J., & Scholey, A. B. (2005). Nonequivalence of On-Line and Paper-and-Pencil Psychological Tests: The Case of the Prospective Memory Questionnaire. *Behavior Research Methods*, 37, 148-154. <https://doi.org/10.3758/BF03206409>
- Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An Outpatient Psychiatric Rating Scale—Preliminary Report. *Psychopharmacology Bulletin*, 9, 13-28.
- Kline, P. (2020). *A Psychometrics Primer*. Free Association Books.
- Musch, J., & Reips, U.-D. (2000). A Brief History of Web Experimenting. In M. H. Birnbaum (Ed.), *Psychological Experiments on the Internet* (pp. 61-87). Academic Press. <https://doi.org/10.1016/B978-012099980-4/50004-6>
- Nunnally, J. C., & Bernstein, I. H. (1994). *Psychometric Theory* (3rd ed.). McGraw-Hill.
- Raffaelli, M., Armstrong, J., Tran, S. P., Griffith, A. N., Walker, K., & Gutierrez, V. (2016). Focus on Methodology: Beyond Paper and Pencil: Conducting Computer-Assisted Data Collection with Adolescents in Group Settings. *Journal of Adolescence*, 49, 1-9. <https://doi.org/10.1016/j.adolescence.2016.02.002>
- Yu, Y., Wan, C., Huebner, E. S., Zhao, X., Zeng, W., & Shang, L. (2019). Psychometric Properties of the Symptom Check List 90 (SCL-90) for Chinese Undergraduate Students. *Journal of Mental Health*, 28, 213-219. <https://doi.org/10.1080/09638237.2018.1521939>