

林靖凯 涂冬波.生命意义感结构及其自适应测量:基于 Bifactor 模型 [J].江西师范大学学报(自然科学版) 2022 46(3):282-290.

LIN Jingkai ,TU Dongbo.The measurement structure of meaning of life and its computerize adaptive test: based on the Bifactor model [J].Journal of Jiangxi Normal University(Natural Science) 2022 46(3):282-290.

文章编号:1000-5862(2022)03-0282-09

生命意义感结构及其自适应测量:基于 Bifactor 模型

林靖凯^{1 2} 涂冬波^{2*}

(1.北京师范大学教育学部 北京 100875 2.江西师范大学心理学院 江西 南昌 330022)

摘要:该文探讨了生命意义感的测量结构及生命意义感的计算机自适应测量(CAT-MLM)对传统纸笔测验测量信度及测量精度的提升效果,并进行了相应的实验.实验结果表明:CAT-MLM 不仅具有较好的测量信度和效度,还能在不损失测量精度的条件下减少测试的题量,从而提升测试的效率;在相同测验题长下,CAT-MLM 能显著提高传统纸笔测验的信度和精度.

关键词:计算机化自适应测验;项目反应理论;生命意义感

中图分类号:B 841.7 文献标志码:A DOI:10.16357/j.cnki.issn1000-5862.2022.03.11

0 引言

近年来,由于积极心理学兴起,所以对于“生命意义感”的研究热度逐步增加,而且社会竞争越来越激烈,这使得不少人感到压力剧增,在这快节奏的竞争逐渐迷失自我,导致生命意义感缺失,这一现象也使得越来越多研究人员投入生命意义感的研究中.关于生命意义感,V.E. Frankl^[1]认为“对意义的寻求是人类在生活中的首要动机”,人类有在生命寻找意义和价值的内在需求,意义和目的感的缺乏会使人深陷痛苦^[1-2];后来也有许多学者从不同角度对生命意义感进行论述,但一直以来没有得到一个统一的界定.直到2016年对生命意义感进行了总结,学术界才基本认可了生命意义感包含理解、目的和价值^[3].

从积极心理学的意义出发,李金珍等^[4]认为心理学有必要去关注在人性中的积极面,致力于帮助普通人生活得更健康、更美好,促进个人、团体和社会的繁荣.尽管人们的生存环境和内外在条件存在种种困难,但绝大多数普通人能过上一种相对满意的、有尊严的生活,而正是这些普通人构成了社会的基础.生命意义感是积极心理学的一个重要研究内容,对生命意义感测量方式的研究也符合积极心理

学研究的价值导向;大量研究证实,生命意义感的缺乏会对个体或他人造成危及生命等严重后果.因此,为了减轻缺乏生命意义感的个体给自己或他人带来的伤害,生命意义感的测量、识别及诊断就显得十分必要.为此,国内外大量学者开发了相关的测评工具,这些量表所包含的维度、测量角度并不完全一致,相互之间存在些许差异,但目的都是生命意义感(或被称为生活意义感)这一特质.有些项目直接测量了生命意义感,有些项目并非直接测量生命意义感.如有些项目会对生命意义感的来源或对影响生命意义感的因素进行测量,并取得了较好的效果.从这方面可以看出生命意义感的测量是多元的.V.E. Frankl^[1]表示生命可以从3个方面来获得意义:(i)从人们所给予生活的东西中获得,即从人们的创造物中获得;(ii)从人们对世界所求取的东西中获得,即从人们认为有价值的事情中获得;(iii)从人们对命运所采纳的立场中获得,这一立场是人们深信而不可改变的.盛正群^[5]认为:从生存的悲剧性经验中、从痛苦、死亡和愧疚中也可以找到生命的意义.如从“相信生活有终极目的和意义”以及“致力于创造性工作”是从生命意义的来源等角度进行测量.又如“我有自己人生意义的理论体系,这使我真正懂得活着的意义”“我有清晰的人生目标和目的”“我认为自己的人生很重要”这3个项目就分别从生命意义感的理

收稿日期:2022-01-10

基金项目:国家自然科学基金(31660278)资助项目.

通信作者:涂冬波(1978-)男,江西南昌人,教授,博士,博士生导师,主要从事心理统计与测量研究.E-mail: tudongbo@aliyun.com

解、目的和价值^[3] 3 个成分对生命意义感进行直接测量。另外,影响生命意义感的因素众多,如 R.F. Baumeister^[6]指出:生命的意义对个体的自我感和同一性起着重要的作用。缺乏意义和价值就会导致自我的分裂。这种情况通常会在边缘性人格障碍的患者身上发现,在生命意义感量表^[7]中的“从兼职或学生工作等日常生活中,我更了解和肯定自己”“我在付出中获得了对于自我的肯定与喜悦”便是从该角度对生命意义感进行的测量。在以往对生命意义感的研究中,有研究人员认为其是一个 1 维结构,也有研究人员则认为其是多维结构。因此,生命意义感的测量结构有待进一步验证,而且结合以往量表测量角度的多元性和测量结果的相关性来看,有理由相信,生命意义感可能不是简单的 1 维或多维结构,而是含有一般因子和特殊因子的 Bifactor 结构,在这种结构中每一个项目都对应 1 个特殊因子和 1 个一般因子。本文利用计算机自适应测验题库体量大、内容覆盖面广的优势,试图从更全面的多元角度,探索一个更加有效合理的测量结构。

现有的量表是采用传统纸笔测验的施测手段,而且几乎都是基于经典测量理论(classical test theory, CTT)开发出来的。然而,传统纸笔测验存在一些先天性的局限,如为了使分数具有可比性,所有的人都必须回答相同的项目,这意味着被试常常需要回答与他们心理特质水平不匹配的项目,从而降低了测试的动机、兴趣及测量的精度。对于大学生生命意义感的监测(尤其是对生命意义感缺失个体的监测),需做到定期测量和长期追踪观察,很明显,传统纸笔测验很难做到这一点。因为每一次对传统纸笔测验进行测量,被试都要做相同的题目,在多次重复测量之后被试可能形成作答定势,所以很难排除过往测量的干扰;而且,传统纸笔测验题量较大,测量耗费大量的时间,对大学生群体进行普测也会耗费大量的人力和物力。

然而,计算机自适应测试(computerized adaptive testing, CAT)技术的产生能够很好地解决上述问题。CAT 可以根据每位被试的反应情况和心理特质水平,智能地从题库中挑选出与被试特质水平相匹配的题目,从而做到对被试的自适应测量,即“千人千卷”的个性化测试。CAT 在每次测量中只有少数题目曝光,大大延长了测验的使用寿命;同时 CAT 还能减少题长,减轻被试负担,提高了测试的动机和水平,并在一定程度上降低被试的厌烦情绪。总的来说, CAT 是一种同时兼顾了测量效率和测量精度的新型测量技术,目前已被广泛应用于心理测评领域,

如艾森克个性问卷(成人版)计算机自适应测验^[8]、特质焦虑计算机化自适应测验^[9]等。

因此,为了进一步克服目前国内外关于生命意义感纸笔测量的不足,本文拟在项目反应理论(IRT)框架下,采用 CAT 技术探讨生命意义感的计算机化自适应测验(CAT-MLM),以此来克服传统纸笔测验的一些缺陷;同时,本文还进一步验证了 CAT-MLM 的科学性及有效性,并且探讨了 CAT-MLM 对传统纸笔测验在测量信度及测量精度上的提升效果,这将为生命意义感的测量提供一种新视角和新技术。

1 方法

1.1 使用的量表

计算机化自适应测验的开发离不开大型的题库支持,本文 CAT-MLM 初始题库的项目主要来自通用的 4 个知名生命意义感量表^[10-11],包含生活信念、生活热情、生活成就、生活目的和生活价值等 5 个子维度;生命意义感量表^[7,12]包含意志自由、追求意义的意志和生命的意义等 3 个子维度;生活目的测试^[13-14]包含生活感受、生活目标、生命态度和自主感等 4 个子维度;生命意义问卷^[5,15]包含成就、人际关系、公平、追求、信仰、家庭、亲密关系、自我超越、自我接受和快乐等 10 个子维度。CAT-MLM 初始题库共计 109 题,保留每个项目原始计分方式,同时本研究使用以上量表的中文版量表。

另外,还有一份效标量表,即人生意义问卷(meaning in life questionnaire, MLQ),它是由王鑫强汉化修订^[16]的,该量表共 10 题,采用 Likert 7 点计分,该问卷在美国和日本大学生样本中表现出良好的信效度^[17]。

1.2 被试

将在初始题库中所有项目组成的测验,通过网络问卷和纸笔测试的形式进行发放和回收,并根据该数据构建生命意义感题库。

被试来自中国 30 个省级行政区(澳门、天津、西藏和新疆除外)的大学生,共调查 1 101 人,删除 10 个项目以上未作答的被试 82 人。根据测谎题的作答与最佳答案的差值,3 道测谎题差值总和最大的前 5%删除,上述 2 种策略共删 137 人,剩余有效被试 964 人。被试年龄最小值为 17,年龄最大值为 29,年龄平均数为 19.98($S_D = 1.51$);男生人数为 349,女生人数为 615;大一学生人数为 230,大二学生人数为 294,大三学生人数为 333,大四学生人数为 107;

文科学生人数为 417,理工科学生人数为 474,艺术与体育学生人数为 46,医学学生人数为 27;城镇学生人数为 524,农村学生人数为 440.

1.3 研究方法

1.3.1 实验 1: CAT-MLM 题库开发及 CAT-MLM 算法有效性验证 (i) 基于 Bifactor 模型的因素分析. 将 964 份有效数据随机分成数量相等的 2 份,分别作为探索样本和验证样本.利用探索样本进行探索性因素分析,以各项目的因素负荷作为指标,对题库的项目进行筛选.根据 CFI、NNFI 等拟合指标,选取拟合最好的模型,对验证样本进行验证性因素分析,并将该模型用于后续 CAT 的开发.

(ii) IRT 模型选择.在 IRT 中,选择合适的 IRT 模型进行数据分析是保证数据分析结果准确的前提.本文采用常用的偏差(-2LL、AIC、BIC 统计量)等^[18-19]相对拟合指标进行模型拟合比较,以选择相对更拟合数据的模型.-2LL、AIC、BIC 统计量的值越小表示模型拟合越好.本文考察的 IRT 模型主要有使用较广泛的等级反应模型(GRM)^[20]和拓广分部评分模型(GPCM)^[21],并从这 2 个模型中选择一个与实测数据拟合更佳的模型作为随后数据分析的 IRT 模型.

(iii) 项目分析.对题库进行项目质量分析,以确保最终留在题库中的项目具有较高的质量.项目质量分析主要包括项目区分度、项目拟合度(item-fit)检验以及项目功能差异(DIF)检验.

项目区分度表示具有相似分数的人在潜在特质上的差异程度,为了保证题库的高质量,本文删除区分度过低的项目^[22].

项目拟合(item-fit)检验用于考察项目是否使用的 IRT 模型拟合.本文采用 $S-\chi^2$ 统计量进行项目拟合检验.在 IRT 模型中,当项目的 $S-\chi^2$ 检验对应的 p 值小于临界值 0.01 时,该项目拟合较差,应考虑删除该项目^[23].

采用 Logistic 回归方法进行 DIF 检验,并使用 McFadden's pseudo R^2 的变化量进行评价,当 R^2 的改变量大于 0.02 时,项目存在 DIF^[24].

(iv) CAT-MLM 算法与效果验证.在建立完成 CAT-MLM 题库的基础上,对 CAT-MLM 的相关算法进行设置并检验其效果.CAT 算法主要涉及初始题选取、选题策略、能力估计方法及终止规则等方面.本文 CAT-MLM 的初始题采用随机法选取,选题策略采用最大信息量法(maximum information, MI),能力参数估计方法采用期望后验方法(EAP),终止规则采用定长终止规则(即被试的用题量达到了事先

设定的要求则终止测试).

实验 1 以 CAT-MLM 真实的题库参数以及真实被试在所有项目上纸笔测验的真实作答为基础,模拟真实被试在 CAT-MLM 上的自适应过程,从而检验 CAT-MLM 的测量效果.真实被试共计 964 人,这些真实被试完成了题库中所有项目的纸笔作答,模拟被试在 CAT-MLM 上的自适应作答.目前在国际上有 2 种运用较广的双因子 CAT 测验设计^[25],本文采用 Bifactor CAT^[26]设计,具体过程如下:从第 1 位真实被试开始,进行其在一般因子上的 CAT 程序,第 1 题从题库中随机选取,并将该被试在纸笔测验中对该题的作答作为其在 G 因子 CAT 上的作答,根据该作答估计该被试特质水平,并根据估计的特质水平(θ)采用最大信息量的方法选取下一试题给被试作答,同样将该被试在纸笔测验中对该题的作答作为其在 G 因子 CAT 上的作答,依此进行直至达到了事先设定的终止规则,结束 G 因子的测试;在 G 因子 CAT 中已经被执行的项目同时测量了 1 个一般因子和 1 个特殊因子,根据被试已作答的项目估计被试在特殊因子 F_1 上的水平值,并作为 F_1 因子 CAT 的水平初始值,根据该水平初始值选择被试在 F_1 因子上的所要作答的项目,然后采用与 G 因子 CAT 相同的方法进行 F_1 因子的 CAT 选题与测试,直至达到 F_1 因子的终止规则,结束 F_1 因子的测试;以同样的方法,测试下一个因子;接下来其余被试依次按此方法模拟,直至完成所有被试在所有因子(包含 1 个一般因子和 3 个特殊因子)上的 CAT-MLM 测试.最后,将 964 位被试在 CAT-MLM 一般因子测试中估计的特质水平(θ_1)与使用在题库中所有题目的纸笔测验所估计的特质水平(θ_2)求相关系数,再计算 964 位被试在 CAT-MLM 各因子测试上的误差均值和边际信度,以考察 CAT-MLM 的测量效果.

同时,为了考察 CAT-MLM 的测量效度,964 位被试还完成了在效标量表上的作答,将这 964 位被试在 CAT-MLM 一般因子上的测量结果与在效标量表 MLQ 上的测量结果进行比较分析,并求取其校标关联效度.

1.3.2 实验 2: CAT-MLM 对传统纸笔测验测量精度的提升.在 CAT-MLM(一般因子)的测验长度与传统纸笔测验(即传统生命意义感量表)相同的情况下,实验 2 探讨 CAT-MLM 是否能提供比传统量表更高的测量精度以及更高的测量信度,即在相同测验长度下, CAT-MLM 是否能显著提升测量的精度及信度,从而探讨 CAT-MLM 的性能及其优势.实验 2 选用了用于构建 CAT-MLM 的 4 个量表,即 PIL、

SLMS、MLM 和 PMP, 它们的题数分别为 20 题、20 题、23 题、46 题。因此在实验 2 中 CAT-MLM 的一般因子测验的终止规则分别设置为 20 题、20 题、23 题和 46 题。

实验 2 在 CAT 算法的选择上与实验 1 一致, 实验 2 将在 4 种终止规则下 CAT-MLM 的结果与具有同等题长的传统量表的结果进行比较, 重点比较它们的测量误差、测验信度和信息量。根据被试特质水平和测验对于被试的测量误差的关系做散点图, 然后依据相同方法做特质水平-信度散点图和特质水平-信息量散点图, 以此比较在相同题长下这 2 种测量形式对于不同特质水平被试的测量误差、测验信度和信息量差异。

1.4 分析软件

探索性因素分析(EFA)和验证性因素分析(CFA)均采用软件 Mplus7, 其余分析利用 R 语言的各种软件包进行。如模型选择和项目拟合检验等采用 mirt 包、DIF 检验采用 lordif 包、CAT 算法采用 catR 包。

2 结果

2.1 实验 1: CAT-MLM 题库开发及 CAT-MLM 算法有效性验证

2.1.1 因素分析 采用双因子模型(Bifactor model)对探索样本进行探索性因素分析, 并在多个竞争模型(分别为 2 个、3 个、4 个和 5 个特殊因子模型)中选择出拟合指标最好的模型(3 个特殊因子的 Bifactor model), 其比较拟合指数(CFI)为 0.91、近似误差均方根(RMSEA)为 0.034、标准化残差均方根(SRMR)为 0.045。在删除因素负荷过低或为负的项目后, 剩余 92 题; 然后利用验证样本对该模型进行验证性因素分析, 其 CFI 为 0.90、RMSEA 为 0.035、SRMR 为 0.050。模型如图 1 所示(该示意图已将项目按照维度顺序重新编号)。这一结构说明: 目前国内外关于生命意义量表测量了 1 个共同因素(即整体生命意义感), 同时包含了 3 个相对独立的特殊因素(即生活感受、生命意志以及生活价值)。

2.1.2 题库建设 在尝试 GRM 和 GPCM 模型之后, GRM 模型的-2LL 统计量值为 21 9687.6、AIC 统计量值为 221 095.5、BIC 统计量值为 224 524.8(这里-2LL 是在 logistic 回归中的最大似然估计, AIC 是赤池信息准则, BIC 是贝叶斯信息准则), GPCM 模型在经过 500 次迭代后, 仍然无法收敛, 这说明 GPCM 模型不适用于本研究。因此, 选择 GRM 模型有助于后续的研究。

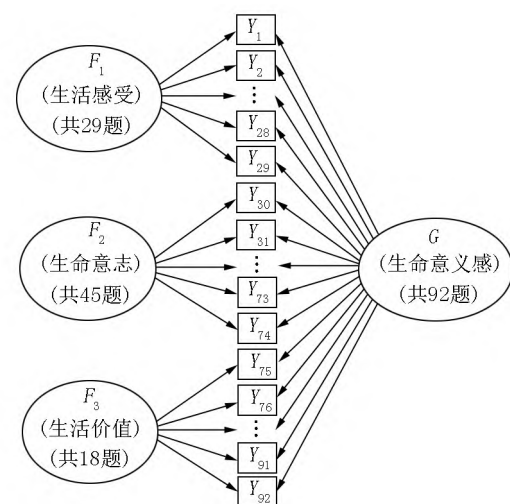


图 1 CAT-MLM 模型示意图

运用 GRM 模型进行参数估计, 在 92 题中有 5 题区分度过低, 为了提高题库的质量, 将这 5 题删除, 剩下 87 题。对剩余 87 个项目进行项目拟合检验, 发现所有项目的显著性 p 值均小于临界值 0.01。将男性作为参照组, 女性作为目标组, 考察项目是否存在性别 DIF。所有项目的 McFadden's R^2 变化量均高于临界值 0.02, 即项目在性别水平上均不存在 DIF。最终的 CAT-MLM 题库构建完成, 共包含 87 个项目。整个题库的平均区分度参数为 1.46($S_D = 0.82$)。这说明整个题库具有较高的质量; 同时项目的位置参数(难度参数)的取值范围为 $[-6.24, 6.80]$, 覆盖范围远大于能力值标准正态分布的“ 3σ 原则”的区间 $(-3, 3)$, 适合测量不同特质水平的被试。结果如表 1 所示。

表 1 在题库中各项目的来源

量表	维度	题号
PMP	成就	18、21、26、27
	人际关系	7、14、19、20、22、25、35
	公平	11、23、30、31
	追求	2、5、6、9、10
	信仰	3、4、15、16
	家庭	1、32、33、34
	亲密关系	8
	自我超越	12、28、29
	自我接受	24
	快乐	13、36、37
SLMS	意志自由	38、39、40、41、42
	求意义的意志	43、44、45、46、47、48、49
	生命的意义	50、51、52、53
MLM	生活信念	57、61、64、67
	生活热情	54、55、56、68
	生活成就	58、59、63、69
	生活目的	62、66
	生活价值	60、65、70
PIL	生活感受	71、72、74、75、78、79、80、82、86
	生活目标	73、76、77、81、85、87
	生命态度	84
	自主感	83

表 1 呈现了在最终题库中 87 个项目所测量的子维度,从表 1 可知 4 个原量表的所有子维度均有项目被纳入题库,即在题库中的 87 个项目测量了 4 个原量表的所有子维度。

2.1.3 CAT-MLM 算法与效果验证 表 2 和表 3 分别为 G 因子和特殊因子(包括 F_1 因子、 F_2 因子、 F_3 因子)基于真实被试模拟的 CAT 在定长终止规则下各项指标的结果,其中 G 因子的定长终止规则为做完 20 个项目即停止测试, F_1 、 F_2 、 F_3 因子的定长终止规则为做完属于本因子题库 60% 的项目即停止测试,表 2 和表 3 所呈现的指标结果可以反映 CAT-MLM 题库质量及其 CAT 算法的效果。表 2 包括了一般因子在定长终止规则下的平均测量误差、测量的边际信度,以及使用题库所有题目所估计的特质值与使用定长终止策略估计的特质值的相关系数和一般因子在定长终止规则下与效标量表 MLQ 的效标关联度;表 3 包含 3 个特殊因子在定长终止规则下的误差均值和边际信度。

从表 2 可知:使用题库中所有项目对一般因子进行估计,其误差均值和边际信度分别为 0.10 和 0.99,效标关联度高达 0.753,这也能够说明本文所开发的 CAT 题库有较高的质量。若使用 20 题定长终止为一般因子 CAT 的终止规则,则测验整体的误差均值和边际信度与使用整个题库基本相当,分别为 0.15 和 0.98,在此终止规则下,估计出的特质值与使用整个题库估计出的特质值相关系数高达 0.978,且效标关联度为 0.749,这说明使用此终

止规则能够使测量结果稳定、精确、有效。

表 3 为该 CAT 的各特殊因子分测验的精度指标 F_1 、 F_2 、 F_3 因子的误差均值分别为 0.34、0.37、0.45,均不大于 0.45;3 个因子的边际信度分别为 0.89、0.86、0.80,均不小于 0.80。

总之,实验 1 的结果说明 CAT-MLM 有较好的测量信度和效度,同时也说明 CAT-MLM 所采用的算法可行、科学、效果好。

表 2 在定长 20 题终止规则下一般因子 CAT 的各项效果指标

终止规则	误差均值	IRT 边际信度	与整个题库估计结果的相关	效标关联度
做完整个题库	0.10	0.99	1.000	0.753**
定长终止	0.15	0.98	0.978	0.749**

注: ** 表示在 0.01 水平上相关性显著。

表 3 定长 CAT 的各特殊因子的测量误差和信度

特殊因子	误差均值	IRT 边际信度
F_1	0.34	0.89
F_2	0.37	0.86
F_3	0.45	0.80

2.2 实验 2: CAT-MLM 与传统纸笔测验的比较

实验 2 的结果如图 2~图 4 所示。图 2 呈现了在相同测验长度下 CAT-MLM(一般因子)和传统量表的测量误差散点图。图 2 结果显示:在题数相同的情况下,绝大多数被试在 CAT 上的测量误差比在传统量表上的测量误差更低,即 CAT-MLM 有效降低了测验误差。其中,与 SLMS、PIL、MLM、PMP 相比, CAT-MLM 的平均测量误差分别降低了 41.13%、39.30%、27.50%、29.20%。

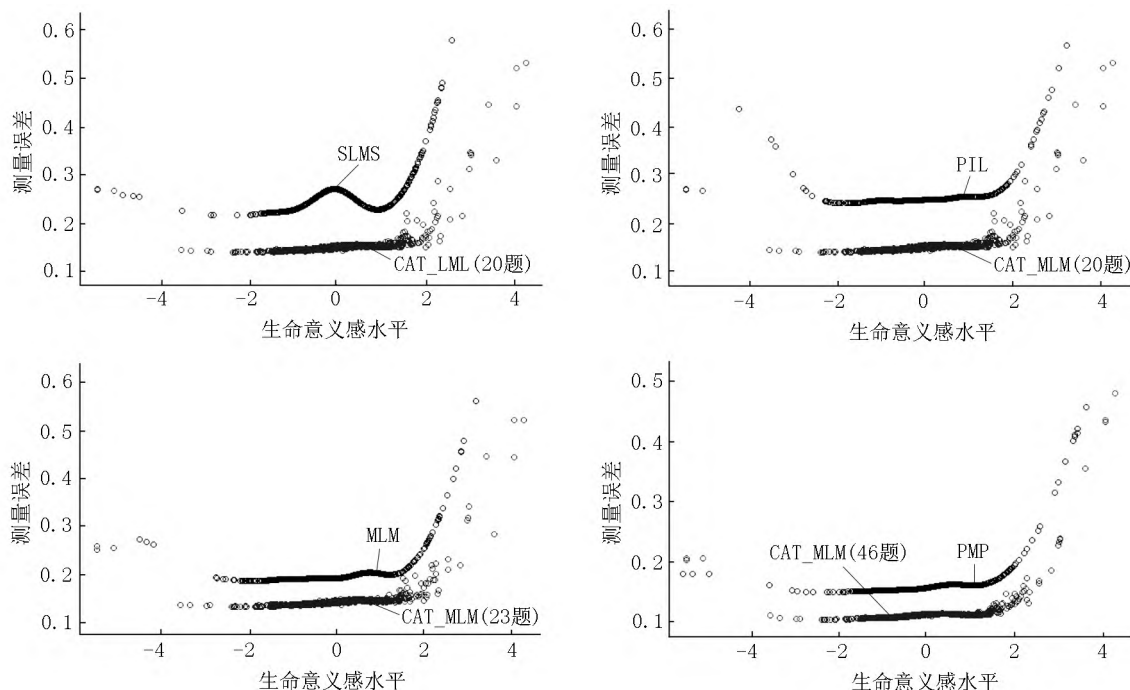


图 2 在各定长下 CAT-MLM 与其相同题数的传统纸笔测验的测量误差散点图

图3呈现了在相同测验长度下CAT-MLM(一般因子)和传统量表的测量信度散点图.图3结果显示:在测量题数相同的情况下,绝大多数被试在CAT上的测量信度比在传统量表上的测量信度更

高,即CAT-MLM有效提升了测量信度.其中,与SLMS、PIL、MLM、PMP相比,CAT-MLM的边际信度分别提升了5.10%、4.41%、2.09%、1.43%.

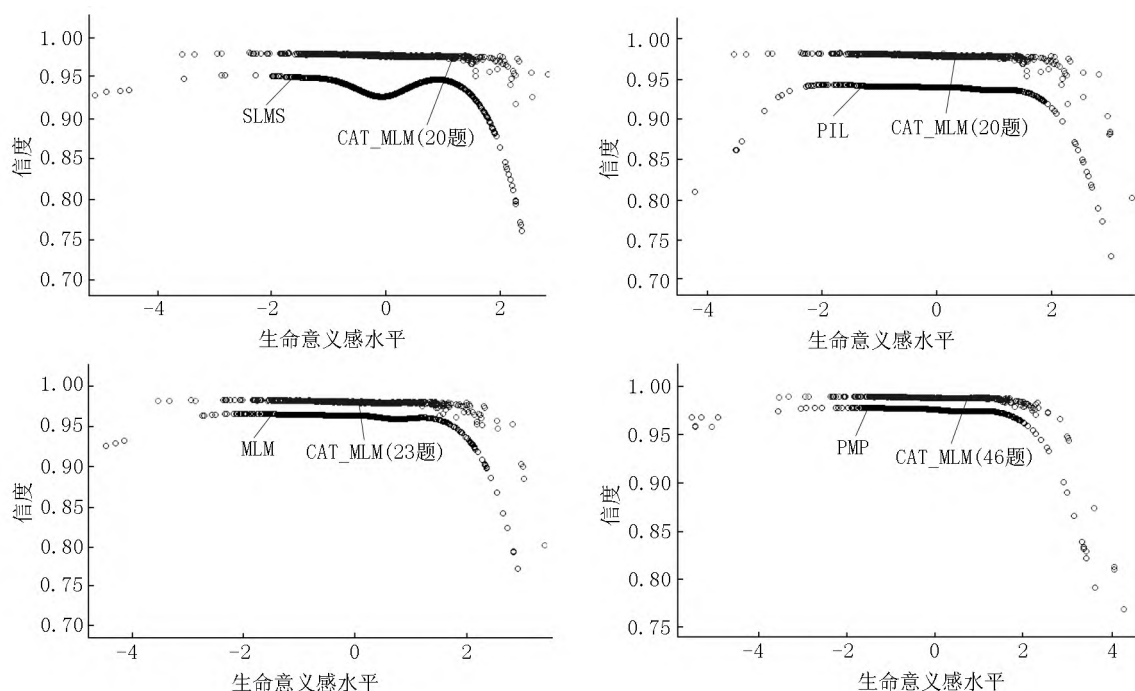


图3 在各定长下CAT-MLM与其相同题数的传统纸笔测验的测量信度散点图

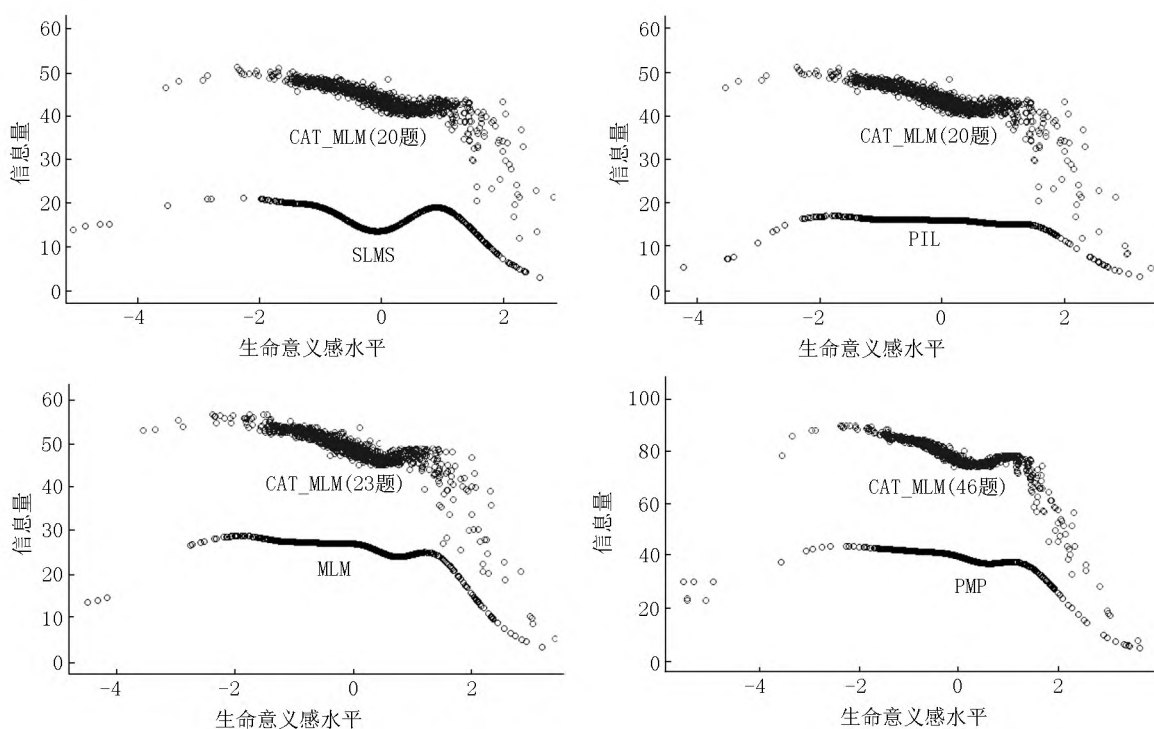


图4 在各定长下CAT-MLM与其相同题数的传统纸笔测验的信息量散点图

图4为CAT-MLM使用定长终止规则在各种题长下对于所有真实被试的信息量与其题长相等的传统纸笔测验量表对于所有真实被试的信息量的散点图.由图4可以看出:无论是传统纸笔测验量表还是

计算机化自适应测验,对于生命意义感水平在 $[-2, 2]$ 上的被试都有较高的信息量.除了题数较多的PMP外,其余独立量表对于生命意义感很低的极端被试的信息量都较低(信息量低于25意味着测验标准

误差大于 0.2),而与其题长相同的 CAT-MLM 能够在生命意义感水平较低的被试群体中有较好的表现。相比传统纸笔测验量表, CAT-MLM 对于生命意义感低的极端被试在测量信息量上有相当大的提升。从图 4 不难看出:在相同题长的情况下,对于在 [-4, 4] 上的所有被试, CAT-MLM 的信息量几乎大于任何一个传统纸笔测验量表的信息量。由此可见, CAT-MLM 对于所有生命意义感水平的被试信息量都有所提升,尤其是对于生命意义感水平低于 -2 的极端被试, CAT-MLM 更是有飞跃性的提升。

3 结论与讨论

本研究的目的是:探索出一个更有利于生命意义感测量,且拟合度更佳的测量模型,并依据此模型,建立一个基于 IRT 的生命意义感的计算机自适应测验(CAT-MLM),再验证 CAT-MLM 对传统纸笔测验在测验精度及测量信度上的提升。实验 1 基于 964 位真实被试的作答数据,运用 Bifactor 模型对所有项目进行探索性因素分析和验证性因素分析,最终选取在所有竞争模型中拟合最佳的模型,用于后续 CAT 的开发,最终选取的模型包含 1 个一般因子和 3 个特殊因子。在 IRT 框架下,通过项目拟合检验、项目功能差异分析等方法,最终选取了包括符合测量学要求且项目质量较高的 87 题来构建 CAT-MLM 题库。研究发现:在本实验开发的题库中拥有 87 个项目的 CAT-MLM 效果良好;使用定长终止规则和最大信息量选题法对真实被试进行模拟测量,一般因子的平均测量误差为 0.15、边际信度高达 0.98,一般因子的平均误差均不大于 0.45、边际信度均不小于 0.80,并且该 CAT 有良好的效标关联度,能高效地对生命意义感进行测量。以上多方面结果说明本研究开发的 CAT-MLM 算法可行、科学、效果好。另外,从表 1 可以看出:所有原量表的子维度均有题目被保留下来,并纳入最终题库,各子维度的测量角度存在些许差异,但都测量了生命意义感这一共同因子。本文利用了 CAT 题库题量大、覆盖面广的优点,对生命意义感的测量结构进行了重新探索。本文构建的 Bifactor 模型在一定程度上说明生命意义感的测量是多角度、多元的,但最终都能有效测量生命意义感这个一般因子。从理论上分析,由于每个个体所处的外部环境和自己心理的内部环境都十分复杂,导致生命意义感的影响因素众多,所以生命意义感的测量内容十分多元。由于以往生命意义

感量表的开发者有些从生命意义感本身进行直接测量,有些从生命意义感的来源或影响因素进行间接测量,所以以往量表的开发者开发出了测量角度差异较大,但测量目标高度重合,结果高度相关;从统计分析的角度看,对被试在原量表各子维度的得分进行相关性分析,结果发现这些子维度两两之间都具有显著的中度或高度相关。从理论和统计测量的角度都能有力地支持本研究所开发的 Bifactor 测量模型,也能解释 Bifactor 模型对生命意义感的测量为何有更好的拟合度。由于传统纸笔量表需要被试完成所有的项目,要考虑到被试的答题负担,所以要控制量表的题量,便难以同时涵盖如此多的测量角度,而将 CAT 运用于类似于生命意义感这种测量角度多元的特质的测量,就能显示出 CAT 题库内容覆盖面广的优点。由于生命意义涵盖广泛的内容与诸多因素相关,所以更广的覆盖面无疑能使生命意义感的测量有更好的效果,要有如此多元的测量角度和内容覆盖面,这是传统量表难以做到的。因此,尽管 CAT-MLM 从不同的测量视度出发,但是它们均是从直接测量和间接测量 2 个角度来测量人的生命意义感这一相同特质。综合这些结果也可以看出生命意义感不是简单的多维结构,而是一种被较少用于生命意义感研究的双因子结构,同时,也能够较好地解释长期以来生命意义感的结构和成分为何一直难以统一。实验 2 在与传统量表相同题长的情况下考察 CAT-MLM 是否能有效提高测量的精度及测量的信度。实验结果显示:与传统纸笔测验的量表相比,无论是测量误差、边际信度还是信息量, CAT-MLM 都有明显差异, CAT-MLM 对传统纸笔测验量表的测量误差降低幅度最多达到 41.13%,对传统纸笔测验量表的边际信度提升幅度最高达到 5.10%。比较不同测量方式对于所有被试的信息量后发现:相比传统纸笔测量, CAT-MLM 对绝大多数被试具有较高的测量精度,尤其是对生命意义感水平低于 -2 的极端被试的信息量有极大提升。

本文开发的 CAT-MLM 对于传统纸笔测量的测量精度有极大提升。CAT-MLM 作为生命意义感的一种新的测量工具,利用其高精度、高效率和对生命意义感极低的被试十分敏感的特点,可以被用来较好地解决前文提到的对在大学生群体中缺失生命意义的个体的筛查问题,以识别出生命意义感水平过低的人,及早预防和干预,并对所有在校大学生进行定期普查。利用 CAT-MLM 进行普查的优势是速度快、

效率高,同时 CAT 可以减少题目的曝光率,增长题库使用寿命,防止被试前后多次重复回答同样的题目,形成回答定势,导致测验准确度降低。而且,CAT-MLM 能较好地弥补传统纸笔测量对于极端被试测量精度极低的缺陷,对于在普查中期望发现严重缺乏生命意义感的个体具有较高的敏感度,一旦发现生命意义感过低的学生,尽早采取必要措施,减少自残、自杀等悲剧事件的发生。

4 参考文献

- [1] FRANKL V E. Man's search for meaning: an introduction to logotherapy [M]. New York: Washington Square Press, 1963.
- [2] 刘思斯,甘怡群. 生命意义感量表中文版在大学生群体中的信效度 [J]. 中国心理卫生杂志, 2010, 24(6): 478-482.
- [3] GEORGE L S, PARK C L. Meaning in life as comprehension, purpose, and mattering: toward integration and new research questions [J]. Review of General Psychology, 2016, 20(3): 205-220.
- [4] 李金珍,王文忠,施建农. 积极心理学: 一种新的研究方向 [J]. 心理科学进展, 2003, 11(3): 321-327.
- [5] 盛正群. 大学生生命意义问卷修订 [D]. 广州: 华南师范大学, 2007.
- [6] BAUMEISTER R F. Meanings of life [M]. New York: Guilford Press, 1991.
- [7] WU Hotang, CHOU Meiju, LEI Mengshan, et al. Development of Taiwan college students' sense of life meaning scale [J]. Universal Journal of Educational Research, 2015, 3(8): 536-545.
- [8] 雷辉,戴晓阳,朱熊兆,艾森克个性问卷(成人版)计算机自适应测验的信效度研究 [J]. 中国临床心理学杂志, 2012, 20(3): 325-327.
- [9] 邓远平,戴海琦,罗照盛. 计算机自适应测验在特质焦虑量表中的运用 [J]. 心理学探新, 2014, 34(3): 272-275, 283.
- [10] MORGAN J, FARSIDES T. Measuring meaning in life [J]. Journal of Happiness Studies, 2009, 10(2): 197-214.
- [11] 诸晓. 中学生生命意义感的特点及其与生活事件、社会支持的关系研究 [D]. 南京: 南京师范大学, 2012.
- [12] 梁进龙,崔新玲,吴和堂,等. 大学生生命意义感量表编制及信效度分析 [J]. 中国卫生事业管理, 2018, 35(4): 303-307.
- [13] CRUMBAUGH J C, MAHOLICK L T. An experimental study in existentialism: the psychometric approach to Frankl's concept of noogenic neurosis [J]. Journal of Clinical Psychology, 1964, 20(2): 200-207.
- [14] SHEK D T L, HONG E W, CHEUNG M Y P. The purpose in life questionnaire in a Chinese context [J]. The Journal of Psychology: Interdisciplinary and Applied, 1987, 121(1): 77-83.
- [15] WONG P T P. Implicit theories of meaningful life and the development of the personal meaning profile [M] // WONG P T P, FRY P S. The human quest for meaning: a handbook of psychological research and clinical applications. New Jersey: Lawrence Erlbaum Associates Publishers, 1998: 111-140.
- [16] 王鑫强,张大均. 心理素质与心理健康关系模型构建: 对 PTH 和 DFM 的超越 [J]. 西南大学学报(社会科学版), 2012, 38(6): 67-74, 174.
- [17] 戴晓阳. 常用心理评估量表手册 [M]. 修订版. 北京: 人民军医出版社, 2015.
- [18] AKAIKE H. Stochastic theory of minimal realization [J]. IEEE Transactions on Automatic Control, 1974, 19(6): 667-674.
- [19] SCHWARZ G. Estimating the dimension of a model [J]. The Annals of Statistics, 1978, 6(2): 461-464.
- [20] SAMEJIMA F. Estimation of latent ability using a response pattern of graded scores [J]. Psychometrika, 1970, 35(1): 139.
- [21] MURAKI E. A generalized partial credit model: application of an EM algorithm [J]. Applied Psychological Measurement, 1992, 16(2): 159-176.
- [22] FLIEGE H, BECKER J, WALTER O B, et al. Development of a computer-adaptive test for depression (D-CAT) [J]. Quality of Life Research, 2005, 14(10): 2277-2291.
- [23] REEVE B B, HAYS R D, BJORNER J B, et al. Psychometric evaluation and calibration of health-related quality of life item banks: plans for the patient-reported outcomes measurement information system (PROMIS) [J]. Medical Care, 2007, 45(S1): 22-31.
- [24] CHOI S W, GIBBONS L E, CRANE P K. lordif: an R package for detecting differential item functioning using iterative hybrid ordinal logistic regression/item response theory and Monte Carlo simulations [J]. Journal of Statistical Software, 2011, 39(8): 1-30.
- [25] 刘馨婷,彭思韦,涂冬波. 双因子模型下 CAT 测验优化设计及其效果验证 [J]. 江西师范大学学报(自然科学版), 2019, 43(2): 128-134.
- [26] WEISS D J, GIBBONS R D. Computerized adaptive testing with the bifactor model [EB/OL]. [2021-10-19]. <http://iacat.org/sites/default/files/biblio/cat07weiss%20gibbons.pdf>.

The Measurement Structure of Meaning of Life and Its Computerize Adaptive Test: Based on the Bifactor Model

LIN Jingkai^{1,2}, TU Dongbo^{2*}

(1. Faculty of Education, Beijing Normal University, Beijing 100875, China;

2. School of Psychology, Jiangxi Normal University, Nanchang Jiangxi 330022, China)

Abstract: The main purpose of this study is to explore the measurement structure of meaning of life, then develop a Computerize Adaptive Test of Meaning of Life Measure (CAT-MLM) based on this structure, and discuss its improvement on the reliability and measurement accuracy of traditional paper-pencil test. The experimental results show that CAT-MLM not only has good measurement reliability and validity, but also reduces the number of questions and improves the efficiency of the test without losing the measurement accuracy. Under the same test item length, CAT-MLM can significantly improve the reliability and accuracy of traditional paper and pencil test.

Key words: computerized adaptive test; item response theory; meaning of life

(责任编辑: 冉小晓)

(上接第 276 页)

[19] HE Xuming, SHAO Qiman. A general Bahadur representation of M -estimators and its application to linear regression [J]. The Annals of Statistics, 1996, 24(6): 2608-2630.

[20] HENDRICKS W, KOEBKER R. Hierarchical spline models for conditional quantiles and the demand for electricity [J]. Journal of the American Statistical Association, 1992, 87(417): 58-68.

[21] HALL P, SHEATHER S J. On the distribution of a stu-

dentized quantile [J]. Journal of the Royal Statistical Society: Series B (Methodological), 1988, 50(3): 381-391.

[22] HOSMER DW Jr, LEMESHOW S. Applied survival analysis: regression modeling of time to event data [M]. New York: John Wiley & Sons, 1999.

[23] KOENKER R. Censored quantile regression redux [J]. Journal of Statistical Software, 2008, 27(6): 1-25.

The Piecewise Linear Censored Quantile Regression Model Estimation Based on Smoothing Technique

WANG Xiaogang, CHEN Jiangmeng

(School of Mathematics and Information Science, North Minzu University, Yinchuan Ningxia 750021, China)

Abstract: The smoothing technique is proposed in the piecewise linear censored quantile regression model to solve the problem of change point, the estimator of change point and coefficients are obtained, and the large sample properties of the estimator is derived. The smoothing technique solves the cumbersome calculation and unreal meaning of the grid search method, and remedies the difficulty that linearization technology cannot prove the asymptotic properties. The validity and robustness of the estimation are verified by Monte Carlo simulation with homoscedasticity and heteroscedasticity, fixed and random censoring at different quantiles. The empirical analysis of drug abuse data shows that the recurrence interval and treatment time have a positive effect, and the recurrence time has a change point at 0.498 (0.5 quantile). The treatment time before 0.498 is longer than after 0.498, that is, the treatment in the first half of the time is more effective.

Key words: smoothing technique; piecewise linear; censored quantile regression model; change point

(责任编辑: 曾剑锋)