

## 五因素取向人格测验的发展与比较

陈基越<sup>1</sup> 徐建平<sup>1</sup> 黎红艳<sup>1</sup> 范业鑫<sup>1,2</sup> 路晓兰<sup>1</sup>

(<sup>1</sup>北京师范大学心理学院; 应用实验心理北京市重点实验室, 北京 100875)

(<sup>2</sup>首都医科大学宣武医院, 北京 100053)

**摘要** 五因素取向人格测验是目前评价个体人格特质时被广泛应用的一大类重要人格测验工具。该类测验数量众多, 长度各异, 形式多样。根据其题目呈现的特点, 可将现有的 30 余种五因素取向人格测验分为自陈式、核检表式、非言语式三类。在不同研究和实践情境下, 各个五因素取向人格测验具有各自独特的优劣势, 信效度的变化都不同。在实际应用中, 应根据不同研究需求区别使用, 并注意测量目的、施测次数、测验长度、被试特点、文化差异性对测验结果的影响。当前, 五因素取向人格测验的开发正向着简明、全面、跨文化方向发展。

**关键词** 五因素取向; 大五; 五因素模型; 人格测验  
分类号 B848

五因素取向人格测验是目前评估个体的外倾性、宜人性、责任心、神经质和开放性五种人格特质时被广泛应用的一类重要人格测验工具, 其理论基础是大五人格结构和五因素模型。像 NEO-PI R、NEO-FFI、BFI 等测验已经被翻译成多国语言在不同文化下使用(Schmitt, Allik, McCrae, & Benet-Martinez, 2007)。近些年来, 许多研究者依据大五人格结构和五因素模型, 相继开发出了多种五因素取向人格测验。中国学者也据此理论编制了中国大五人格测验(王孟成, 戴晓阳, 姚树桥, 2010a, 2010b, 2011), 还在检验西方“大五”人格理论的同时, 针对中国人人格特质进行了本土化探索, 开发出了具备中国文化特点的人格测量工具(崔红, 王登峰, 2003; 宋维真, 张建新, 张建平, 张妙清, 梁觉, 1993; 王登峰, 崔红, 2003)。本文旨在回顾并梳理多种五因素取向人格测验工具, 描述它们的编制、特点并进行比较, 让国内的研究者全面地了解当前国内外五因素取向人格测验总体发展状况, 掌握五因素取向人格测验的完整信息, 为研究和实践中选用人格测验提供依据, 为新的人格测验的开发提供借鉴。

### 1 五因素取向人格测验的理论基础

Goldberg (1981)提出的大五人格结构(Big Five)与 Costa 和 McCrae (1985)提出的人格五因素模型(Five-Factor model, FFM), 被认为是五因素人格理论产生的标志和人格研究的共识。它们的出现被称为是近几十年来“人格心理学领域的一场静悄悄的革命”, 一些研究者甚至认为它们适用于全人类(John, Angleitner, & Stendorf, 1988)。

大五人格的提出源于人格词汇学的研究(Allport, 1931; Cattell & Eber, 1950; Digman, & Inouye, 1986; Eysenck, 1968; Fiske, 1949; Norman, 1963, 1967; Tupes & Christal, 1961)。Goldberg (1981)提出用“大五”人格结构这个概念概括人格的五个基本维度, 即外倾性(Extraversion)、宜人性(Agreeableness)、责任心(Conscientiousness)、情绪稳定性(Emotional Stability)和文化(Culture), 并编制了相应的特质描述形容词表(trait descriptive adjectives, TDA) (Goldberg, 1990, 1992)。大五人格是采用人格词汇学研究取向提出的人格结构。

与此同时, Costa 和 McCrae 根据其他人的研究和自己的个人构想, 采用了一种与 Goldberg 截然不同的研究取向, 提出了五因素人格模型。他们(1978)先提取出 Eysenck (1968)人格结构中的神

收稿日期: 2014-02-28

通讯作者: 徐建平, E-mail: xujp@bnu.edu.cn

经质(Neuroticism)和外倾性(Extraversion)两个维度,并接受 Rogers (1958)、Cattell 和 Eber (1950)人格开放性(Openness)理论的影响,构建了 NEO 三因素模型,形成了 NEO 体系的最初框架,并编制了测量这 3 个维度的人格测验 NEO Inventory (McCrae & Costa, 1983)。之后,受大五人格结构的影响,Costa 和 McCrae (1985)又将大五人格结构中的宜人性、责任心两个维度纳入到原有的 NEO 体系中,编制了测量这五个维度的人格测验 NEO-PI (NEO-Personality Inventory)。后来,他们(1991)又根据这五个人格维度的含义,按照自上而下的原则,确定了每个维度下面包含的 6 个层面(facet),编制了 NEO-PI 的修订版 NEO PI-R (Revised NEO Personality Inventory) (Costa & McCrae, 1989, 1992; McCrae & Costa, 1996)。可以看出,五因素模型是在大五人格结构以及其他学者的理论基础上,结合研究者个人构想产生形成的。

实际上,五因素模型测量的维度与大五人格结构中的五维度基本对应。两者之间是否应该加以区分,至今尚有争议。Goldberg (1993)认为,两种模型具有更多的相似之处:都有五个维度;虽然两种模型对“情绪稳定性”和“神经质”命名不同且含义相反,但测量的内容实际相同;责任心维度也十分相似。但同时,Goldberg 也指出,两种模型中至少有两处差异不可忽视:温暖(Warmth)在五因素模型中属于外倾性维度的一个层面,而在大五人格结构中属于宜人性的一个层面;五因素模型中经验开放性(Openness to Experience)这一维度,在大五人格结构中定义为聪慧性(Intellect)和想象力(Imagination)。一些研究者将大五人格结构和五因素模型区别对待,指出两种理论在历史渊源、内容形式、基本性质、研究走向等方面都存在一定差异(尤瑾,郭永玉,2007)。但另一些研究者则认为五因素模型是在吸收大五人格结构结果的基础上,经过不断修订后才形成的,因而在实际应用中大五人格结构和五因素模型不加区分,且 NEO PI-R 相比于 TDA 更易进行,所以应用更加广泛(王登峰,崔红,2008a)。近些年,许多研究者(Block, 1995; Cobb-Clark & Schurer, 2012; Oh, Wang, & Mount, 2011)都将大五人格结构和五因素模型同等对待,本文采用 Block (1995)提议的“五因素取向”(Five-Factor Approach, FFA)统合之。

## 2 五因素取向人格测验的发展及分类

从上世纪 90 年代至今,研究者以五因素人格取向为理论基础相继编制开发出了许多不同类型的人格测验,一些测验经过研究者的不断验证,具有良好的信效度和适用性,应用比较广泛,其中部分测验还开发了简缩版。根据题目呈现的形式可以将这些测验分为语句自陈式、词汇核检表式和图片非言语式。下面将分别介绍这些测验的编制、修订、应用推广等,以概览和梳理五因素取向人格测验的发展脉络。

### 2.1 自陈式:叙述人格表现的语句

在这类五因素取向人格测验中,每个题目均采用了自陈式人格测验中惯常采用的句子或短语来呈现,如“我认为我自己喜欢与他人合作”。其中,以 NEO 和 BFI 测验为典型代表。

NEO-PI 是最早编制的五因素取向人格测验,测试神经质(N)、外倾性(E)、开放性(O)、宜人性(A)和责任心(C)这 5 个人格维度。但因宜人性和责任心两个维度被融合进五因素模型的时间较晚,每个维度中仅包含 2 个层面,而神经质、外倾性、开放性三个维度融合进入 NEO 体系的时间较早,每个维度中包含 6 个层面。由此就形成了包含 5 个维度、22 个层面、181 道题目的 NEO-PI (Costa & McCrae, 1978, 1985)。

由于该测验题目数量多,施测时间过长,因此 Costa 和 McCrae (1989)删减 NEO-PI 编制了它的简版 NEO-FFI (NEO Five-Factor Inventory)。在 NEO-FFI 的 5 个维度中,每个维度 12 道题目,共 60 道题目,这些题目都是在 NEO-PI 中每个维度上载荷最高的题目,并且确保其尽量覆盖各个维度下的层面。它具有良好的信度,分量表内部一致性信度在 0.68~0.86 之间,重测信度在 0.75~0.83 之间,且它与同时编制的 NEO-PI R 的相关很高,因此具有较高的会聚效度(Costa & McCrae, 1992)。目前,NEO-FFI 已被翻译成多种语言并被验证其有效性,是测量人格五因素模型最广泛的工具之一(McCrae & Costa, 2004)。国内研究者在测量五因素人格时,大多采用的也是这个测验(罗杰,戴晓阳,2012),迄今采用这个工具的中文研究文献达到了 232 篇。

由于 NEO-PI 中宜人性、责任心两个维度下的层面没有其他三个维度完整,因此国际上已经

不再使用该测验,更多使用的是 Costa 和 McCrae (1989, 1991)后来编制的 NEO-PI 的修订版 NEO-PI R。该测验包括 5 个维度,每个维度下包含 6 个层面,共 30 个层面,每个层面包含 8 道题目,整个问卷共 240 道题目。验证发现该问卷具有良好的信度,分量表内部一致性信度在 0.86~0.92 之间,重测信度在 0.63~0.81 之间,在配偶评定和同伴评定间,具有良好的会聚效度和区分效度(Costa & McCrae, 1989, 1992; McCrae & Costa, 1996)。问卷编制完成并经过验证后,开始推广到其他国家。在跨文化研究中,相比于词汇学取向的形容词式人格测验,NEO-PI R 这种每道题目是一条语句的自陈式测验更具优势,因为单独的一个形容词,很难在其他语言中找到一个对应的一模一样的替代词汇。自陈式的语句题目相比于形容词则更加具体和情境化,更容易被翻译成其他语言并很好地表达原意(Bond, Nakazato, & Shiraishi, 1975; Eysenck, 1983)。在上世纪 90 年代,除了英语版本之外,它相继被翻译成西、葡、意、荷、德、法、俄、希腊、日、朝、中、菲和克罗地亚语、希伯来语等语种(Hoekstra, Ormel, & De Fruyt, 1996; Martin et al., 1997; McCrae & Costa, 1997; McCrae, Costa, Costa, Bond, & Paunonen, 1996; McCrae, Costa, Del Pilar, Rolland, & Parker, 1998; Piedmont & Chae, 1997; Psychological Assessment Resources, 1994; Tsaousis, 1998; Tsaousis & Kerpelis, 2004)。在国内也有相应的中文版本,但缺乏发表的修订报告文献。目前,有 119 篇发表的中文文献是将其当做工具使用。

由于有研究者对 NEO-FFI 的测量学属性提出质疑,尤其是它的因素结构(Egan, Deary, & Austin, 2000),因此,McCrae 和 Costa (2004)从 NEO-PI R 中选取了 14 道题目,替换了 NEO-FFI 中一些因子载荷低或不易理解的题目,修订形成了 NEO-FFI-R。验证结果显示,相比于 NEO-FFI,NEO-FFI-R 在信度和结构效度方面都有了适当的改善。到目前为止,使用 NEO-FFI-R 测验在国内发表的中文研究文献较少,只有 13 篇,如王萍萍、许燕、陈咏媛和李佳慧(2010)以及朱相华等人(2012)的研究。

在使用 NEO-PI R 的过程中,研究者发现被试群体,尤其是青少年群体对一些题目理解存在着偏差,因此对 NEO-PI R 进行了修订,形成了 NEO-PI-3(NEO Personality Inventory-3)。该测验的

维度、层面、总题目量都与 NEO-PI R 相同,但为了使被试更易理解,替换了其中 19 个层面的 37 道题目(McCrae, Costa, & Martin, 2005)。在被替换的题目中,有 25 道题目是因为包含了青少年无法理解的单词,如“lackadaisical”和“panhandler”;有 12 道题目是因为不论在青少年群体还是成人团体中,题总相关均小于 0.30。替换后的 37 道题目的平均题总相关由原来的 0.28 提高至 0.37。NEO-PI-3 的内部一致性信度相比于 NEO-PI R 有略微提高,结构效度和会聚效度几乎保持未变。之后,McCrae 和 Costa (2007)缩减 NEO-PI-3 形成了 60 道题目的简版 NEO-FFI-3。在这个简版中,有 59 道题目与 NEO-FFI-R 一样,仅替换了一道题目,它只出现在 NEO-PI-3 中(“I’m hard-headed and tough-minded in my attitudes”替换为“I have no sympathy for beggars”),目前,这一测验国内尚无应用。

1991 年,John, Donahue 和 Kentle 编制了测量五因素人格特质的测验“大五人格问卷”BFI (Big Five Inventory)。该测验虽然依据的是词汇学取向的大五人格理论,但为了详尽描述人格特质以提高作答一致性,每道题目被编制成一个短语而非词汇。例如开放性维度中的特质词汇“original”在 BFI 中被编制为“is original, comes up with new ideas”。它包括 44 道题目,主要是在当大五人格维度下各层面的差异不是研究者主要关注的问题时使用。运用这个简明测验,使大五人格特质的测量更加有效、灵活。虽然它的各个维度仅包含 8~10 道题目,但并没有因其题目的减少而大大削弱其内容覆盖度。例如,在该测验的宜人性维度中,包括了 NEO PI-R (Costa & McCrae, 1992)测验里相同维度中 6 个层面的 5 个。该问卷具有良好的测量学属性,在美国和加拿大样本中,BFI-44 各分测验内部一致性信度的范围在 0.75~0.90 之间,3 个月后重测信度的范围在 0.80~0.90 之间,自评他评一致性在 0.47~0.67 之间,区分效度均低于 0.20 (John, Naumann, & Soto, 2008)。随着 BFI 在国际上的广泛应用,研究者对其进行修订,产生了不同言语版本的 BFI。从 1991 年至 2014 年,检索 Proquest 数据库,BFI 已有德(BFI-German)、法(BFI-Fr)、荷(BFI-Dutch)、韩(BFI-Korean)、意(BFI-Italian)、匈(BFI-Hungarian)、挪(BFI-Norwegian)等 7 种语言的修订版本,并有

相关研究文献发表(Denissen, Geenen, Van Aken, Gosling, & Potter, 2008; Farkas & Orosz, 2013; Fossati, Borroni, Marchione, & Maffei, 2011; Kim et al., 2010; Lang, Lütke, & Asendorpf, 2001; Plaisant, Srivastava, Mendelsohn, Debray, & John, 2005)。为了研究 BFI 的跨文化一致性,也有研究者将 BFI 翻译成 28 种语言在 56 个国家进行施测,其中就包括中文。由于其短小易用,测试全面, BFI 目前已经成为国际上各类研究中测查五因素人格的主要人格测验工具(Schmitt et al., 2007)。

虽然 BFI-44 的编制是以简洁方式全面测量人格结构为目的,但随着研究者对测验简洁需求的进一步提高,一些研究者缩减 BFI-44 修订并形成了多种缩减版。Rammstedt 和 John (2007) 从 BFI-44 每个维度中提取了 2 道题目,组成了一个由 10 道题目组成的 BFI-10(Big Five Inventory-10)。BFI-10 问卷包括英语和德语两种版本。它选取题目时遵循的原则是:(1)每个维度的两道题目是对同一人格特质的相反描述;(2)每道题目尽可能广地覆盖所属维度的各个层面,但叙述内容不能过于冗长;(3)挑选与该维度相关高、在因素分析中因子载荷高且没有交叉载荷的题目。他们(2007)给受测者同时施测了 BFI-10 和其它几种五因素取向人格测验,检验其测量学属性,结果显示, BFI-10 与 BFI-44 对应的各维度的平均相关达到了 0.83, BFI-10 的重测信度为 0.75, 自评他评一致性为 0.40, 各维度平均相关为 0.11, 与 NEO PI-R 的平均会聚效度为 0.67, 各指标与 BFI-44 相比虽然略显不足但仍可接受。BFI-10 作为 BFI-44 的简版,在测验时间有限的情况下具有更大的优势,且其跨文化一致性也得到了验证。除了 BFI-10, BFI-44 还有一些其它的缩减版,譬如 BFI-25(John, Donahue, & Kentle, 1991)、BFI-S(Gerlitz & Schupp, 2005; Hahn, Gottschling, & Spinath, 2012; Lang, John, Lütke, Schupp, & Wagner, 2011)、BFI-K(Kovaleva, Beierlein, Kemper, & Rammstedt, 2013; Rammstedt & John, 2005)和 BFI-20(Engvik & Clausen, 2011),但使用不如 BFI-10 广泛。

除以上两类典型的自陈式五因素取向人格测验, Hendriks, Hofstee 和 De Raad (1999) 编制了五因素人格测验(Five-Factor Personality Inventory, 简称 FFPI)。该测验包含 100 道题目,每道题目是一个简短且具体的陈述句,采用 5 点李克特量表

计分。该量表内部一致性信度在 0.81~0.86 之间,重测信度在 0.74~0.83 之间,自评-他评一致性在 0.54~0.73 之间,与 NEO PI-R 的会聚效度除开放性维度较低外(0.20),其余维度均较好(0.69~0.83)。同时,测验相继被翻译成了巴西语、中文、克罗地亚语、希伯来语、匈牙利语、意大利语、日语、波兰语、斯洛伐克语、西班牙语、瑞典语和罗马尼亚语。

研究者以国际人格题目库(International Personality Item Pool) (IPIP, 2001)为基础,采用五因素取向也编制了一些自陈式人格测验,这些测验包括 IPIP-NEO、IPIP-BFFM 和 Mini-IPIP。IPIP-NEO 同 NEO PI-R 一样测量了人格五因素的 30 个层面,每个层面 10 道题目,共 300 道题,采用 5 点李克特量表计分,其平均内部一致性信度为 0.80,与 NEO PI-R 的平均会聚效度为 0.73 (Goldberg, 1999; Johnson, 2000)。IPIP-BFFM (IPIP-Big Five Factor Markers)测验包括 100 题和 50 题两个版本,平均内部一致性信度分别为 0.90 和 0.84,与 Goldberg 早期编制的 TDA 的平均会聚效度分别为 0.70 和 0.67 (Goldberg, 1999, 2001)。微型国际人格题库(Mini-IPIP)由 Donnellan, Oswald, Baird 和 Lucas (2006) 缩减 IPIP-BFFM-50 测验的题目修订形成,包含 20 道题。该测验具有可接受的内部一致性信度(均在 0.60 以上),与焦虑、抑郁、行为抑制等指标的效标关联效度同其他五因素取向人格测验(BFI、IPIP-BFFM)类似, CFA 结果显示它与大五人格结构模型拟合不好,进行 EFA 发现,多个题目存在交叉载荷,部分分测验的信度指标并不理想,研究者指出 Mini-IPIP 还需要进一步修订(Cooper, Smillie, & Corr, 2010; Baldasaro, Shanahan, & Bauer, 2013)。

国内外研究者编制的其他一些自陈式五因素取向人格测验还有:TIPI、SIMP、YPF 和 CBF-PI。TIPI (Ten-Item Personality Inventory)是一个 10 道题目的自陈式五因素取向人格测验,是 Gosling, Rentfrow 和 Swanm (2003)在前人(Goldberg, 1992; John, Donahue, & Kentle, 1991)编制的人格测验基础上开发的一种简式人格测验。TIPI 这一测验编制的原则是:(1)保证每道题能够覆盖对应维度尽可能多的人格层面;(2)每道题都是对应维度最极端的表述;(3)每道题不分好与坏;(4)不使用否定意义的题目。验证发现 TIPI 具有良好的测量学属

性。也有研究者同时施测 TIPI 与 IPIP-BFFM-50, 发现两者的会聚效度在 0.48~0.82 之间, 并通过检验 TIPI 各维度得分与内外控制倾向、自我监测、生活满意度、特质焦虑的相关, 验证了 TIPI 的效标关联效度(Ehrhart et al., 2009)。之后, 该测验又被翻译成葡萄牙语、西班牙语、土耳其语、韩语、日语等在其他国家进行研究使用(Atak, 2013; de Francisco Carvalho, Nunes, Primi, & da Silva Nunes, 2012; Ha, 2013; Oshio, Abe, Cutrone, & Gosling, 2013; Ruiz, Oberst, Gosling, Rusiñol, & Luser, 2013)。

Woods 和 Hampson (2005)认为, 在研究中使用单题目的测验节省时间, 十分经济, 又可以避免被试长时间测试所引发的疲劳或无聊, 因此, 他们以 IPIP、BFI-44(John et al., 1991)和霍根人格量表(Hogan Personality Inventory, HPI) (Hogan & Hogan, 1997)为基础, 挑选出合适的人格特质形容词或解释, 编制了一个 5 道题的单题目人格测验 SIMP (Sing-Item Measures of Personality)。在这个测验中, 每道题采用意义相反的两极性的人格特质短语来测量五个人格维度。挑选这些短语素材的依据是: (1)这些短语由人格特质描述和行为指标共同组成。前者如“健谈的”(talkative), 后者如“很难拒绝别人”(finds it difficult to say no)。(2)为了使测验能够覆盖更广的层面, 不使用相互对立的描述。例如, “有组织的”(organized)和“无组织的”(disorganized)。(3)同时, 为了使覆盖层面广且描述不要太复杂, 每一极使用的描述词汇不超过五个。(4)避免使用极端性的形容词。Woods 用 SIMP 以及其他 4 种五因素取向人格测验进行施测, 结果显示 SIMP 与其他 4 种测验的平均会聚效度为 0.61, 重测信度为 0.71, 其信度(0.50)、自评-他评一致性(0.38)和区分效度(0.11)都可以接受, 可推广使用。

人格调查表(Your Personality Form, YPF) (Bernard, Walsh, & Mills, 2005)只包含 5 道题, 1 道题目测量一个维度。每道题采用 7 点李克特量表计分, 在 1、4、7 三个数字计分点上有关于该维度对应于低、中、高分的言语描述, 该言语描述来源于 NEO PI-R 操作手册。将该测验与 NEO PI-R 进行多元回归时发现, 两种测验仅能对责任心和神经质两个维度解释同样的方差, 在剩余三个维度上, NEO PI-R 比 YPF 能够解释的方差更多, 因此 YPF 的有效性还需进一步探讨。

随着经济全球化以及文化领域国际交流的日趋密切, 基于相同理论编制测验工具也成为国际学术和跨文化研究的迫切需求, 中国学者(王孟成等, 2010a, 2010b)以五因素人格取向为理论框架, 也编制了一个“中国大五人格问卷”(The Chinese Big Five Personality Inventory, CBF-PI)。测验的完整版包括 134 道题目, 测查了 22 个层面特质。这个人格测验适合中国人语言表达习惯, 内部一致性信度在 0.83~0.91 之间, 间隔 5 周的重测信度均在 0.80 以上, 采用主成分分析法得到了满意的因子结构, 与 NEO PI-R 的效标关联效度在 0.45~0.62 之间, 具有良好的信效度, 并且具有自主版权。之后, 研究者又开发出了 CBF-PI 的简式版 CBF-PI-B (Chinese Big Five Personality Inventory brief version), 它包含 40 道题目, 内部一致性信度在 0.76~0.81 之间, 间隔 10 周后的重测信度在 0.67~0.74 之间, 因子结构较为清晰, 可解释总变异的 43.49%, 与完整版 CBF-PI 的会聚效度均在 0.85 以上, 具有较好的信效度, 可用于相关研究(王孟成等, 2011)。

## 2.2 核检表式: 描述人格特质的词汇

依据人格五因素取向编制的这种类型的人格测验, 它们采用的是描述人格的词汇。这种核检表式的人格测验又可细分为两种类型: 一种可称为单极形容词量表, 需要受测者依据自己的情况对每个形容词进行评分, 代表测验有 TDA-100 和 Mini-Markers; 另一种是两极形容词量表, 每道题目由两个意义相反的极端形容词组成, 受测者需根据自己的情况在两个形容词之间选择一个分数, 代表测验有 TDA-35(TBI)和 BFLQ。

人格特质描述性形容词表(trait descriptive adjectives, TDA)是 Goldberg 根据自己的研究结果构建的一种形容词表式的五因素取向人格测验(Goldberg, 1990, 1992)。它包含 100 道题, 每道题是一个形容词, 每个因素下有 20 个形容词。要求受测者按 9 点李克特式量表对每个人格特质形容词进行评定。其分量表具有良好的内部一致性信度, 在 0.82~0.97 之间, 且结构效度良好。Goldberg 认为, TDA 是测量人格五因素的最佳选择(Goldberg, 1992)。他还编制了 TDA 的 35 题版, 它也被称为 TBI (Transparent Bipolar Inventory) (Goldberg, 1992)。该测验将 TDA-100 中的 100 个形容词筛选配对, 组成了 35 组意义相对的两极形容词对,

每个维度下包含7组词对。

为了缩减施测TDA所用的时间,有研究者以TDA-100(Goldberg, 1992)为基础开发了一套简式的五因素取向人格形容词量表 Mini-Markers (Saucier, 1994)。Saucier 使用前人关于TDA的12个数据样本进行数据分析,提取出载荷最高最能代表人格五因素的形容词组成 Mini-Markers,该量表包含40个描述人格的形容词,采用9点李克特量表计分,内部一致性信度(0.78~0.83)略低于TDA (0.84~0.90),题目间一致性(0.31~0.38)高于TDA (0.21~0.31),完成该量表仅需5分钟左右。

还有一种是大五人格定位问卷(Big Five Locator Questionnaire, BFLQ)。它由Howard, Medina和Howard (1996)编制,包含25组两极形容词,每个维度包括5组两极形容词。例如在宜人性维度中,一组形容词中的一极是“友好”(outgoing),另一极则是“冷淡”(cool),受测者需在两极之间的5点量表上打分。量表在编制完成后逐渐流传使用,近些年颇受组织行为学领域研究者喜爱并被广泛使用(Ahmed & Hasnain, 2011; Jeswani & Dave, 2012; Sinha & Srivastava, 2013),在国内也有研究者使用(秦昕,牛丛,黄振雷,徐敏亚,2011)。

除以上3种采用人格词汇应用十分广泛的核检表式的五因素取向人格测验外,研究者还编制了其他一些核检表式的人格形容词量表。Trapnell和Wiggins (1990)修订人际形容词量表(Interpersonal Adjective Scales),形成了同时测量人际领域内特质与五因素人格模型的人际形容词量表大五修订版(Interpersonal Adjective Scales Revised-Big Five, IASR-B5),该量表是一个包含124道题的形容词清单,受测者在8点李克特量表上评定每个形容词描述自己的准确性。各维度内部一致性信度范围在0.87(A)~0.92(N)之间。与NEO-PI和HPI结合分析,具有良好的聚合效度和区分效度,能够同时高效地测量人际领域内特质和人格五因素。Duijsens和Diekstra (1995)编制了包含23道题的两极形容词测验(23 Bipolar Big Five, 23BB5),该测验采用7点李克特量表计分,语言为荷兰语。除开放性维度外(0.49),23BB5其他维度内部一致性信度良好(0.71~0.85),因素分析显示结构效度良好。Perugini和Leone (1996)以之前研究结果中的285个人格特质形容词为基础,筛选出50个形容词编制成采用7点李克特量表计分的简短版大

五形容词核检表 SACBIF (Short Adjectives Checklist to Measure Big Five)。该测验内部一致性信度在0.70~0.86之间,因素分析结果显示具有良好的结构效度。Shafer (1999)依据人格五因素模型编制了包含30道题的两极形容词评定量表(Brief Bipolar Markers-FFM),受测者在7点或者9点李克特量表上进行评定。结果显示该量表具有较高的平均单一因子载荷(0.62),并具有可接受的内部一致性信度(0.79)。五因素模型评定表FFM-RF (Five-Factor Model Rating Form)包括30道题,依据五因素人格模型编制(Mullins-Sweatt, Jamerson, Samuel, Olson, & Widiger, 2006),其测量不仅覆盖人格5维度,也测量了5个维度包含的30个层面,即每道题测量了一个层面。验证结果表明,该测验具有可接受的内部一致性信度(0.51~0.78),会聚效度和区分效度良好。

### 2.3 非言语式: 了解人格的非言语图片

为了消除文化因素在人格测验中的影响,提高其文化公平性,研究者采用非言语材料,也在努力构建其他非言语式的五因素取向人格测验。比如FF-NPQ,这是一种图片式的五因素人格测验,其目的是想尽可能地避免人格测验中不同文化和语言的影响。Paunonen等人为了在跨文化研究中避免翻译误差造成的偏差,以早期编制的非言语人格问卷(Nonverbal Personality Questionnaire, NPQ) (Paunonen, Jackson, & Keinonen, 1990)为基础,从测量Murray (1938)提出的16种心理需要的136道图片式题目中选择了部分题目,结合自编的一些题目编制形成,并命名为非言语五因素人格问卷(Five-Factor Nonverbal Personality Questionnaire, FF-NPQ) (Paunonen & Ashton, 1998; Paunonen, Ashton, & Jackson, 2001; Paunonen, Zeidner, Engvik, Oosterveld, & Maliphant, 2000)。FF-NPQ中的题目除了一小部分是重新编制之外,绝大多数题目是通过项目分析和因素分析从NPQ中挑选出来的,这些题最能代表人格五因素模型。在题目测试时,首先呈现一张如图1所示的黑白图片,图片中那个长满头发的人是图片的中心人物,图片描绘的是该人物在一个具体情境中表现出某种与其人格特质相关的特定行为,通常图片中至少会有一人或多人。要求受测者观看每幅图片,假定自己是这个图片里的中心人物,然后对自己做出该人物行为的可能性按照7点李克特式量表从“非常不

可能”到“非常可能”进行评定。早先的研究发现 (Paunonen & Ashton, 2001), 在 304 名加拿大学生样本中, FF-NPQ 五个分量表的平均内部一致性信度为 0.80; 与 NEO-FFI 相应分量表的平均会聚效度为 0.52; 使用 FF-NPQ 在大学室友间进行自评-他评, 分数也呈显著正相关, 平均值为 0.41; FF-NPQ 中各分量表间的相关均没有超过 0.25, 相关绝对值的平均数为 0.14; 与饮酒、吸烟、约会迟到等 14 个行为指标的效标关联效度的平均值为 0.25; 因此可以认为 FF-NPQ 问卷具有良好的信效度。作为一种非言语式人格量表, FF-NPQ 在跨文化研究领域很受欢迎。因为它无需翻译题目, 只需翻译指导语, 可以很容易地对不同语言群体被试进行施测。研究者在加、英、芬、德、挪、波和俄等不同国家的样本中施测 FF-NPQ, 发现其具有较好水平的内部一致性信度(7 个国家的  $\alpha$  系数平均值分别为 0.75、0.79、0.66、0.69、0.71、0.68 和 0.72)和可接受的聚合效度(7 个国家的平均聚合效度分别为 0.50、0.55、0.40、0.45、0.55、0.49、0.41) (Paunonen et al., 2001)。同时, 它不仅在跨文化研究中具有优势, 也特别适合那些阅读、理解有障碍的被试(Paunonen, 2003)。需要注意的是, 测验中的图片是以西方人的生活方式为基础绘制的, 如果要在东方文化情境中使用这一测验, 需要对其图片的普适性进行深入探究。严格讲, FF-NPQ 介于传统的构造性测验和投射测验之间, 因为它使用的测试刺激材料虽然是图片式的, 需要受测者去想象, 但受测者的反应采用的

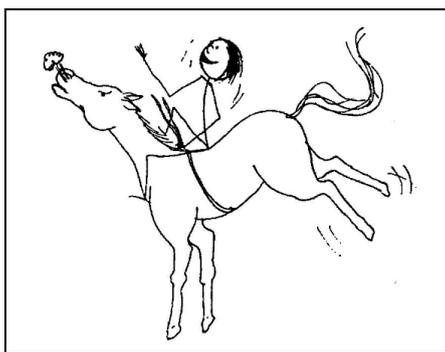


图 1 FF-NPQ 题目示例

资料来源: Paunonen, S. V., Jackson, D. N., & Ashton, M. C. (2004). *NPQ Manual: Nonverbal Personality Questionnaire (NPQ) and Five-Factor Nonverbal Personality Questionnaire (FF-NPQ)*. Sigma Assessment Systems.

是量化的评定式量表, 可以做出比较客观的自评或他评, 其差别主要取决于受测者对图片情境及其人物行为的理解, 而不像其他投射式人格测验那样, 测试刺激材料虽然相同, 但受测者的理解和反应都非常主观, 比较发散, 主试对其反应的评定难度也较大。到目前为止, FF-NPQ 不仅在国内的相关研究较少, 在国内还从未开始使用。若将其引入国内进行修订并使用, 从丰富国内的人格测验工具类型, 促进发展人格测验以及跨文化研究的角度看非常有益。目前, 本研究小组已经完成了该测验在中国的试测。

除了以上按照测验材料呈现形式划分的三大类五因素人格测验之外, 研究者们对前述有些五因素取向人格测验也编制了对应的短式。在相关研究比如组织行为研究中, 当人格因素并非研究者感兴趣的主要变量, 或者受测者作答时间有限, 或研究对测验工具的测量学属性要求不严格时, 就可以用这些短式人格测验测量人格五因素。甚至还出现了仅用一道题目去测量一个人格因素的测验(Gosling et al., 2003; John et al., 1991; Woods & Hampson, 2005)。在前面介绍的自陈式和形容词式的测量工具中, 有些测验就属于这类短式五因素取向人格测验。其中, NEO-FFI-R、NEO-FFI-3、BFI-10、BFI-S、BFI-K、Mini-Markers、Mini-IPIP 就是 NEO-PI、NEO-PI R、NEO-PI-3、BFI-44、TDA 和 IPIP-BFFM 对应的缩减版。也有专门为此目的, 根据研究需要, 依据五因素模型理论编制出的简版人格量表。这些专门编制的简短式人格量表有 BFLQ、TIPI、SIMP、Your Personality Form 和 FFMRM 等。它们使用经济方便、节约时间、又能防止被试在接受测试过程中产生的疲劳感和无聊感。验证性研究也发现, 它们均具有可接受的测量学属性, 但均低于其他人格量表, 并因此受到一些研究者质疑 (Credé, Harms, Niehorster, & Gaye-Valentine, 2012; Miller, Gaughan, Maples & Price, 2011)。

图 2 是对前述各种五因素取向人格测验发展脉络的图示。需要注意的是, IASR-B5 测验除了调查人格的五因素外, 还测查了人际领域内特质。根据图 2, 也可以发现, 随着五因素人格模型的确立, 当人格五因素量表的完整版成熟后, 研究者基于人格五因素开发人格测验工具的研发工作逐渐减缓。随着一些研究对于节约成本的需要,

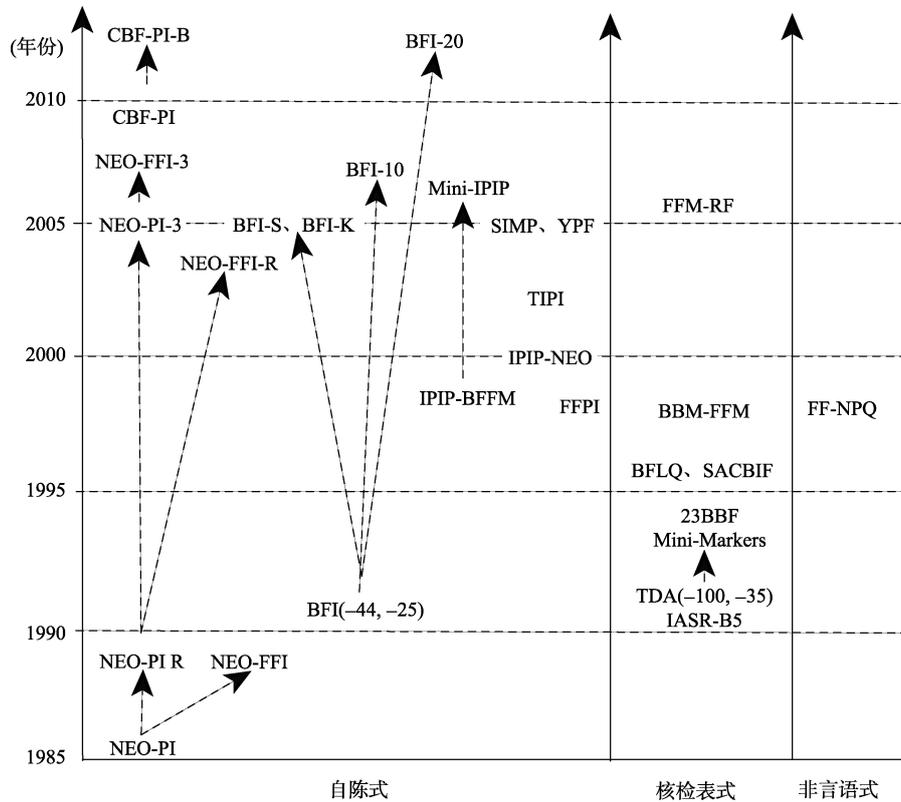


图 2 五因素取向人格测验发展脉络图

以及一些特殊的编制目的, 研究者编制各种不同形式人格问卷时也更倾向于注意控制题目数量, 使测量工具尽量简明又全面。

前述的五因素取向人格测验测试的对象大多只涉及成人, 也有研究者借鉴人格五因素理论针对儿童青少年群体进行研究, 发现了小五人格(little five)结构。研究者通过量表评定法(Digman & Inouye, 1986)、Q 分类法(Robins, John, & Caspi, 1994)、访谈法(Havill, Allen, Halverson, & Kohnstamm, 1994)发现, 在儿童青少年群体中, 他们具有与成人人格相似的五因素结构。不同研究者对其各维度命名虽有差异但意义相同, 为了与成人的大五人格结构相区别, 将其称为小五人格结构(John, Caspi, Robins, Moffitt, & Stouthamer-Loeber, 1994), 并据此模型编制了小五人格测验。John 等人(1994)运用 Q 分类法, 研究了 484 名平均年龄为 10.2 岁的男孩, 依据人格五因素定义从加州儿童 Q 分类程序(California Child Q-Set, CCQ)的 100 道题中选取了 78 道题进行施测, 经过修订形成了包含外倾

性、宜人性、责任心、神经质、开放性五个维度共计 48 道题的 CCQ 版儿童青少年大五人格量表(CCQ Big Five Scales)。该量表除开放性分量表内部一致性信度较低(0.53)外, 其余分量表内部一致性信度良好(0.71~0.83), 并能有效预测儿童不良行为及其学业表现。Mervielde 和 De Fruyt(1999)依据词汇学取向, 编制了包含外倾性、责任心、神经质、仁慈(Benevolence)、想象力(Imagination)五个维度, 18 个层面, 144 道题的儿童分层人格问卷(Hierarchical Personality Inventory for Children, 简称 HiPIC), 该问卷具有良好的内部一致性信度(0.76~0.89), 并且具有稳定的结构效度。Lounsbury 等(2003)针对 6~8 年级儿童编制了包含 85 道题的青少年大五人格问卷 ABFI (Adolescent Big Five Inventory), 该量表具有较好的内部一致性信度(0.72~0.85), 在外倾性、开放性和宜人性三个维度上具有较好的自评-他评一致性(0.31~0.68)。在国内, 也有研究者(周晖, 钮丽丽, 邹泓, 2000)依据大五人格结构, 编制了适合中国儿童青少年人格

测评的工具中学生人格五因素问卷(Five-factor Personality Questionnaire for Middle School Students)。该问卷包含外向性、宜人性、谨慎性、情绪性和开放性五个维度, 90 道题目, 内部一致性信度在 0.68~0.89 之间, 与 NEO PI-R 的会聚效度在 0.59~0.75 之间。从评定形式上看, CCQ 和 HiPIC 这两个测验是由教师和父母通过对孩子或自己孩子的观察进行评定, 题目都是通过一段语句呈现, 因此可以视为是一种“他评”式问卷。

### 3 不同类型五因素取向人格测验的一致性比较

为了清楚地显示一些应用广泛的五因素取向人格测验最初编制时的测量属性, 便于不同类型测验测量属性的比较, 表 1 对前述若干应用较广的测验最初编制时的测量属性进行了小结。根据

表 1 内容, 可以发现, 在测验最初编制时, 在完整版测验中, 自陈式测验(如 NEO PI-R、NEO PI-3、FFPI 等)与核检表式测验(如 TDA)的内部一致性信度差异不大, 均在 0.81 以上, 可以认为完整版的自陈式测验和核检表式测验拥有同样的可靠性; 而非言语式测验 FF-NPQ 相比于前两类测验, 内部一致性信度较低, 在 0.75~0.82 之间, 这也可能与测验项目数量有关, 因为完整版非言语式测验 FF-NPQ 只有 60 个题目, 远小于 NEO 系列测验(240 个题目)和 TDA(100 个题目), 因此当我们想要尽可能完整、有效地测评个体的人格特质时, 应尽量选择完整版的自陈式语句测验或核检表式测验。在缩减版的测验中, 当测验题目数量大致相等时, 自陈式测验的内部一致性信度略微高于核检表式测验的测量学属性, 例如 BFI-44 (0.75~0.90)与 Mini-Markers (0.78~0.83)。虽然缩减版测验的测量学

表 1 五因素取向人格测验最初编制时测量学属性比较

类型	名称(开发年代)	编制者	编制时的测量属性
自陈式	NEO-PI R (1992)	Costa 和 McCrae	$\alpha$ 系数 0.86~0.92; 重测信度 0.63~0.81
	NEO-FFI (1992)	Costa 和 McCrae	$\alpha$ 系数 0.68~0.86; 重测信度 0.75~0.83; 与 NEO-PI R 的会聚效度 0.76~0.91
	NEO-FFI-R (2004)	McCrae 和 Costa	$\alpha$ 系数 0.73~0.84; 与 NEO-PI R 的会聚效度: 0.81~0.89;
	NEO-PI-3 (2005)	McCrae 等	$\alpha$ 系数 0.87~0.92; 与 NEO-PI R 的会聚效度 0.98~0.99
	NEO-FFI-3 (2007)	McCrae 和 Costa	$\alpha$ 系数 0.72~0.86; 与 NEO-PI-3 的会聚效度 0.89~0.93
	BFI-44 (1991)	John 等	$\alpha$ 系数 0.75~0.90; 重测信度 0.80~0.90; 与 NEO-FFI 的会聚效度 0.83~0.97
	BFI-10 (2007)	Rammstedt 和 John	各维度部分-整体相关 0.74~0.89; 重测信度 0.68~0.83; 与 NEO-PI R 的会聚效度 0.51~0.79
	FFPI (1999)	Hendriks 等	$\alpha$ 系数 0.81~0.86, 重测信度 0.74~0.83, 与 NEO PI-R 的会聚效度 0.2~0.83, 自评-他评一致性 0.54~0.73
	IPIP-BFFM (2001)	Goldberg	平均 $\alpha$ 系数 0.84 (50 题版), 0.90 (100 题版); 与 TDA 的平均会聚效度 0.67 (50 题版), 0.70 (100 题版)
	Mini-IPIP (2006)	Donnellan 等	$\alpha$ 系数 0.65~0.77; 与 IPIP-BFFM 的会聚效度 0.85~0.93; CFA 模型拟合不好; EFA 很多题目存在交叉载荷
	TIPI (2003)	Gosling 等	重测信度 0.62~0.77; 与 BFI 的会聚效度 0.65~0.87
	SIMP (2005)	Woods 和 Hampson	重测信度 0.60~0.78; 与 BFI 的平均会聚效度 0.62
	CBF-PI (2010)	王孟成、戴晓阳等	$\alpha$ 系数 0.83~0.91; 重测信度 0.83~0.93; 与 NEO-PI R (中文版)的会聚效度 0.45~0.62
核检表式	TDA (1992)	Goldberg	$\alpha$ 系数 0.82~0.97
	Mini-Markers (1994)	Saucier	$\alpha$ 系数 0.78~0.83
	IASR-B5 (1990)	Trapnell 和 Wiggins	$\alpha$ 系数 0.68~0.86; 与 NEO-PI R 的会聚效度 0.65~0.76
	23BBF (1995)	Duijsens 和 Diekstra	$\alpha$ 系数除开放性维度(0.49)均在 0.70 以上
	SACBIF (1996)	Perugini 和 Leone	$\alpha$ 系数 0.70~0.86; 与 TDA-100 的会聚效度 0.56~0.85
	BBM-FFM (1999)	Shafer	$\alpha$ 系数 0.72~0.89
	FFM-RF (2006)	Mullins-Sweatt	$\alpha$ 系数 0.51~0.78; 与 NEO-PI R 的会聚效度 0.38~0.78
非言语式	FF-NPQ (1998)	Paunonen 等	$\alpha$ 系数 0.75~0.82; 与 NEO-FFI 的会聚效度 0.45~0.59

属性大都良好,但相比于完整版测验都有所降低。

上述内容仅是对测验最初编制时所报告的测量属性进行的初步直观比较,而在实际应用中这些测验的具体表现如何?它们测查的到底是不是相同的五因素人格特质呢?测验工具是否都是可靠的?研究者根据自己的研究目的应该选择哪种测验进行施测最能经济有效地完成研究?对于此类问题,也就是测验间结构和所测量的特质一致性、测验的可靠性以及在不同研究情境下的优劣势,研究者们相继采用多种五因素取向人格测验,对不同被试进行施测,通过对不同测验工具的内部一致性信度、会聚效度、区分效度、效标关联效度的测查,进行深入的比较研究,获得了对这类问题的一些结果和解释。

### 3.1 形容词核检表式与语句自陈式五因素取向人格测验间的比较

Goldberg 和 Kilkowski (1985)研究发现,当使用人格形容词测量人格特质时,受测者回答的一致性要低于把形容词改为详细具体的描述语句时受测者回答的一致性,因此一些学者将形容词式与自陈式两类测查五因素人格的测验进行了比较研究。Mooradian 和 Nezlek (1996)使用 PCA 和 CFA 等统计分析手段比较了 NEO-FFI 和 Mini-markers,发现两种测验测量的结构相似,但并不完全一致,两种测验相同维度只共享了 25%~50% 的方差。Palmer 和 Loveland (2004)对形容词核检表式的 Mini-Markers 和短语自陈式的 IPIP-BFFM 进行比较。结果显示 Mini-Markers (0.74~0.87)的内部一致性信度虽然略低于 IPIP-BFFM (0.77~0.92),但仍可接受。会聚效度和效标效度的结果支持核检表式测验 Mini-Markers 可以像短语自陈式测验一样去测量人格五因素。Woods 和 Hampson (2005)使用 TDA-100、TDA-35、Mini-Markers 等形容词量表与采用语句描述的自陈式五因素人格测验 BFI-44 进行对比,结果显示两种类型的测验内部一致性信度均在 0.80~0.85 之间,无统计差异。其中, TDA-100、Mini-markers 和 BFI-44 各测验的区分效度比较接近,在 0.12~0.17 之间,仅 TDA-100 较大,为 0.22; 4 种测验两两之间的会聚效度也可接受,在 0.63~0.72 之间。可以认为,形容词核检表式人格测验与语句自陈式人格测验两者都可以很好地测量五因素人格特质。

相关研究证明,形容词核检表式与语句自陈

式五因素取向人格测验均具有良好的内部一致性信度,并具有较好的会聚效度和效标关联效度,但仍有一点需要研究者注意。采用形容词核检表式的人格测验中的一些题目,譬如,“热切”、“精力集中”等可能会让受测者错误地理解为近期的个人状态而非个人长期持久性的特质,在采用语句的自陈式测验中,因为每道题都有详尽的叙述,如“有时我感到自己毫无用处”,给出了比较详细的语境,让被试产生误解的可能性降低。因此,核检表式的人格形容词评定量表,可能会因被试近期的一些情绪变化而造成测量误差,这在使用中需要引起注意,应在指导语中加以说明。而从测验的简洁性来看,每道题目为一个形容词或一组两极形容词的核检表式测验,显然要比语句式自陈测验更加简洁清晰。

### 3.2 非言语式与言语式五因素取向人格测验间的比较

作为一种测量五因素人格特质的非言语测验,FF-NPQ 具有其独特优势,在跨文化研究中施测十分方便。在 FF-NPQ 逐渐被研究者熟悉并使用的同时,研究者也对其适用性以及与其他言语式五因素人格测验的一致性进行了比较研究。为了检验人格五因素对行为的预测作用以及不同形式量表的一致性,Paunonen (2003)对 418 名被试施测了 FF-NPQ、NEO-FFI、NEO PI-R 三种不同形式的测验,并记录了受测者的一些行为指标,如饮酒吸烟行为、社交活动等。结果显示 3 种五因素人格测验均具有良好的内部一致性信度(0.72~0.87),对行为指标的预测均良好,且具有一致性。例如,3 种测验中的外倾性维度与饮酒行为呈显著正相关(0.14~0.33),神经质维度与不恰当的性行为呈显著正相关(0.20~0.53); FF-NPQ 与 NEO-FFI、NEO PI-R 的平均会聚效度均为 0.55。然而,FF-NPQ 与 NEO-FFI、NEO PI-R 在测验内各维度间的相关具有一定差异。例如,FF-NPQ 中宜人性维度与开放性维度的相关系数高达 0.42,而在后两个测验中两个维度的相关系数分别为 0.09 和 0.26。Moore, Schermer, Paunonen 和 Vernon (2010)对 232 对同卵双生子和 64 对异卵双生子施测 NEO PI-R 和 FF-NPQ,探讨了基因和环境因素对言语式和非言语式五因素人格测验测量结果的影响。该研究结果显示,FF-NPQ 平均内部一致性信度为 0.76,低于 NEO-FFI 的 0.90,但仍可接受;两个测验的

会聚效度在 0.40~0.56 之间, 测验内各维度间的相关系数(区分效度)均呈中低程度的相关, 表明言语式与非言语式测验间具有较好的一致性。

相关研究证明, 虽然非言语式五因素取向人格测验具有可接受的内部一致性信度, 但低于言语式五因素取向人格测验, 使其测验结果的可靠性不如言语式测验工具。但从效标关联效度指标来看, 非言语测验与言语测验能同等程度地预测行为指标。

### 3.3 简短版与完整版五因素取向人格测验间的比较

一些学者将一部分简短版五因素人格测验与 NEO PI-R 等一些未简化的完整版的五因素取向人格测验进行了比较研究, 以检验简短式量表的测量学属性。Dwight, Cummings 和 Glenar (1998) 对 Goldberg (1992) 的 TDA-100 及其简短版 Mini-Markers (Saucier, 1994) 进行比较, 结果发现 Mini-Markers 的内部一致性信度(0.75~0.90)略低于 TDA (0.84~0.90), 但仍可接受。两种测验的效标关联效度基本没有差异, 责任心维度均可以预测学生的学业成就。Dwight 认为当施测时间有限或者问卷不宜过长的研究条件下可以使用 Mini-Markers 替代 TDA-100。Gosling 等人(2003)对同一批被试施测 TIPI、BFI-44、NEO PI-R 进行比较研究, 结果发现 TIPI 与 BFI-44 的平均会聚效度为 0.77, 平均区分效度为 0.20, 十分良好, 但 TIPI 与 BFI-44 的会聚效度要低于 NEO PI-R 与 BFI-44 的会聚效度。重测结果显示, 虽然 TIPI 有可接受的重测信度(0.72), 但低于题目数量更多的 BFI-44(0.80)。所以 TIPI 虽具有可接受的测量学属性, 但仍低于题目数量更多的量表。Miller 等(2011)的研究发现, BFI-44 与 NEO PI-R 所测量的人格特质并不一致, 例如宜人性维度中, NEO PI-R 包括了“诚实”和“谦虚”这两个层面, 而 BFI-44 却没有包含。这既有可能是因为 BFI-44 起源于大五人格结构理论, NEO PI-R 起源于五因素模型, 其理论基础不同, 也有可能是因为 BFI-44 最初编制的初衷是缩减测验题目数量, 因此造成每一维度覆盖层面不广所造成的。因此, 研究者认为在今后的研究中若使用简短版五因素取向人格测验需谨慎对待。Credé 等人(2012)检验了简短版五因素取向测验的测量结果, 测验包括 5 题的 SIMP 和 YPF, 10 题的 TIPI 和 BFI-10, 20 题的 Mini-IPIP, 30 题的 BBM-FFM 和

40 题的 Mini-Markers。结果随着测验题目数量的增加, 其内部一致性信度也在不断提高; 与任务绩效、组织公民行为及反工作行为等效标的回归分析显示, 简短版测验容易使研究者低估人格特质对于影响人们行为的重要性, 因为使用简式测验会使测验结果增加一类错误和二类错误的概率, 并认为在研究中应该尽量使用长一些的测量工具。

相关研究证明, 简短版测验虽然具有可接受的测量学属性, 但是在内部一致性信度、效标关联效度等指标上低于完整版测验, 且测验所覆盖的测量层面也不如完整版测验广泛全面。但有一点不可忽视, 施测时间简短是简短版测验的优势。

## 4 西方五因素取向人格测验与中国本土人格测验的关系

早在西方人格五因素模型开始被人们广泛认识之前, 我国一些研究者在 1990 年就开始探索建构符合中国本土文化和国情的人格测验工具《中国人个性测量表》(宋维真等, 1993)。在研究文献中, 也有研究者将《中国人个性测量表》(Chinese Personality Assessment Inventory, CPAI)称为《中国人个性量表》(甘怡群, 张妙清, 宛小昂, 孙增霞, 2005)、《中国人人格测量表》(张建新, 张妙清, 梁觉, 2003)、《中国人格评定量表》(曹慧, 张建新, 2007)或《跨文化(中国人)个性测量表》(Cross-cultural (Chinese) Personality Assessment Inventory) (范为桥, 张妙清, 张建新, 张树辉, 2011)。正当 CPAI 的研究进入到探索人格特质结构的关键时刻时, 西方的五因素取向人格模型开始被人们广泛认识并影响到整个人格特质研究领域。这样, 两个人格特质结构碰撞在了一起, 就出现了一个有趣的问题: 两者测量的结果是相互包容、还是相互区分?

研究者施测了 CPAI 和 NEO-PI, 联合因素分析显示, 抽取出的六因素能够作有意义的解释, 其中 4 个因素分别与 N、E、A、C 因素相包容, 第五个因素仅容纳 O 因素, 而 CPAI 中的“人际关系”因素则未与任何 NEO-PI 中的因素相包容, 独自形成了第六个因素(宋维真等, 1993; Cheung et al., 2001)。后来, 研究者对其归纳总结提出了中国人人格结构的六因素假说(张建新, 周明洁, 2006)。在西方人群中使用 CPAI 的英文版也发现了类似的结构(张树辉, 2002)。这些研究表明西方的五因素人格特质在某种程度上能够描述中国人的人格特

质,但并不能完全解释中国人的行为倾向,CPAI中的人际关系因素对五因素模型做了补充,能够完整地描述和解释中国人的人际关系特质。在实际使用时,为了测查中国人的人格六因素模型,相关研究都同时施测CPAI和NEO PI-R两种工具(张建新等,2003;Zhang & Bond,1998)。

随着不断的应用(甘怡群等,2005;孔伟,曹佳,胡文祥,2002),结合中国本土以及跨文化研究与应用的实践经验,考虑到CPAI未能反映西方五因素人格模型所强调的“开放性”人格特质,研究团队于2001年在中国内地和香港对CPAI进行了修订与重新标准化研究,CPAI-2增加了6个本土化的与开放性相关的分量表(张妙清,张树辉,张建新,2004;Cheung et al.,2008)。为了解决测验在使用过程中题目数量过大的问题,研究者也修订形成了仅包含120题的CPAI-2简式量表(曹慧,张建新,2007)。研究者还结合青少年人格特点,基于CPAI-2理论框架,编写了CPAI青少年版(Cross-cultural (Chinese) Personality Assessment Inventory – Adolescent version,CPAI-A),并在两岸三地青少年被试中进行了施测(范为桥,张妙清,张树辉,梁觉,2008;张妙清,范为桥,张树辉,梁觉,2008)。作为第一个被译成韩、日、英、荷、罗等多个不同语言版本的中文人格测验,CPAI的跨文化一致性在亚洲、美洲和欧洲各文化背景中也得以检验(Born & Jooren,2009;范为桥等,2011)。

西方一些五因素取向人格测量工具,在中国也得到了修订并广泛使用。例如,NEO PI-R、NEO-FFI在中国分别由戴晓阳、姚树桥、蔡太生和杨坚(2004)、戴晓阳和吴依泉(2005)以及张建新修订,TDA和BFI-44在中国也得到了验证和广泛使用(刘邦惠,2009;罗杰,戴晓阳,2012;曾垂凯,时勘,2007;周洁,2010)。引进的西方五因素取向人格测验在中国人群中使用时,研究发现存在着诸如文化偏差等问题(王登峰,崔红,2005a;2005b;2006),因此一些研究者开始探索构建符合中国文化的人格结构并编制相应测验(崔红,王登峰,2003;王登峰,崔红,2003)。

王登峰、方林和左衍涛(1995)依据词汇学方法,根据中文人格特质形容词进行因素分析获得了中国人人格七因素结构,7个因素分别为外向性、行事风格、善良、才干、人际关系、情绪性和处世态度,并包含15个次级因素,其中除了善良和处

世态度之外,其他五因素对应于西方五因素人格特质。根据七因素模型,编制了包括123个形容词,6点计分的中国人人格形容词评定量表(QZPAS)。经检验,QZPAS具有良好的内部一致性信度、重测信度以及效标效度(王登峰,崔红,2003;崔红,王登峰,2003,2004)。之后,依据QZPAS又编制出能够快速测查七因素模型的简式中国人人格形容词评定量表(QZPAS-SF)。该量表包括60个形容词,内部一致性信度(0.51~0.65)和重测信度(0.52~0.83)虽然低于QZPAS,但均可接受,且效标关联效度达到了显著水平(崔红,王登峰,2006)。为了更准确地测量人格结构,研究者除了编制形容词评定量表外,还根据形容词的含义编写句子,并编制了中国人人格量表(QZPS),因素分析结果同样得到了7个因素,与QZPAS结构一致。该量表包含180道题,经验证具有良好的信效度(王登峰,崔红,2003,2004)。研究者还比较了QZPS与NEO PI-R的效标关联效度,发现在中国党政领导干部人群中,NEO PI-R所解释的工作绩效的变异只有QZPS的三分之一(王登峰,崔红,2008b)。

相比于西方的五因素取向人格测验,中国学者编制的人格测验除了具有良好的测量学属性之外,更具有中国本土文化特点,能够更好地测查中国人所特有的人格特质,而一些测验,如CPAI、QZPAS,与西方的五因素取向人格测验也具有某些相似之处。在CPAI的本土化与跨文化应用研究中,从理论与实践的双重视角,强调与支持了兼顾文化共通性与特殊性的人格研究方法在人格评估领域的重要性,并促使国内外相关研究者开始思考使用该种人格研究方法以进一步推动人格测验的研究(Cheung, Van de Vijver, & Leong, 2011)。

## 5 选用五因素取向人格测验工具的建议

从词汇学取向的TDA和人格理论取向的NEO PI-R发展至今,衍生出了多种具有较大影响的测查五因素取向人格特质的测验。在整个五因素取向人格测验家族中,根据各个测验题目的特点、形式、作答方式,其中,自陈式这类测验结构最为稳定,信效度指标良好;形容词核检表式测验在题目呈现方式上更为简明清晰,但测验结果有可能被受测者测试时的情绪所干扰;非言语式测验通过图片形式避免了翻译误差,可用于跨文化研究,但因中西方文化差异,其在中国的使用

仍需谨慎处理。研究者编制的那些缩减版五因素取向人格测验,节省了施测时间,虽然其信效度低于完整版测验,但仍可接受使用。下面根据在应用中不同的研究需求给出工具选用建议。

根据不同的研究目的选用适当的测验。当研究目的是为了做出准确地临床诊断,或是研究者需要良好的测量学属性作为实验严谨性的支撑时,完整版测验的高信效度和结构稳定性的优势更为突出,此时建议研究者选用如 NEO-PI-3、IPIP-NEO、CBF-PI、TDA-100 等工具。当研究人格变量只是作为被试群体分类描述,或人格变量只作为一个研究变量,研究者不想占用过多资源对人格变量进行施测,但又不能过分牺牲测量学属性时,建议研究者选用一些中等长度的测验,如 NEO-FFI-3、CBF-PI-B、BFI-44、Mini-Markers、TDA-35 等工具。当研究者进行跨文化研究时,相比于言语式测验,选用非言语式测验 FF-NPQ 可以排除翻译偏差带来的测量误差,更加适用,同时对于那些具有阅读障碍的被试施测也十分便利。但 FF-NPQ 因为其编制时使用的样本为教育水平较高的人群,因此是否适用于教育水平较低的人群还有待验证;该测验在西方国家之间进行跨文化施测,可能比较适合,考虑到东西方国家之间文化差异较大,FF-NPQ 中的一些题目在中国测试,因为有些图片描绘的场景和行为带有浓厚的典型的西方特色,中国被试并不能很好地准确理解图片含义,即使对图片做出了评定,但评定时各自心目中理解或想像不一定相同。反之,如果修改了原始图片,提供给被试的刺激材料发生变化,其跨文化研究的功用会受到削减,国际间比较研究的难度就会增加。因此,在中国推广和使用该测验,必须结合中国文化特点,做进一步修订。

根据研究条件和情境选用适当的测验。在测验施测时间有限、为了防止被试过度疲劳、研究经费紧张的情形下,或是人格变量并不是研究重点关注的变量时,选用缩减版五因素取向人格测验确实能够满足研究需求,如选用 BFI-10、TIPI、SIMP、BFLQ 等工具。但此时需要提醒研究者注意,虽然缩减版与完整版测验具有良好的一致性,在信效度方面可以接受,但均有一定程度的降低,研究中选用缩减版测验,在满足研究关于时间、费用等条件的需求时,可能会以牺牲部分测量学属性为代价。

需要注意的是,以上一些人格测验的测量学属性以及适用性在国内还未进行测试验证,尚未有中文修订版,例如 IPIP-NEO、TDA-100、Mini-Markers、BFI-10、TIPI、SIMP、BFLQ 等,FF-NPQ 虽然不存在言语翻译误差问题,但因东西方文化差异问题,在中国直接使用可能会有问题。这种情形致使国内研究者在研究中选择五因素取向人格测验时,仅局限于从自己知道或已经在国内使用过的那些测验中进行选择,而忽视了那些最新发展,或尚未介绍到国内但已经在外国得到充分应用的那些测验。例如,中国目前已修订的最简短版本的五因素取向人格测验应是 60 道题目的 NEO-FFI,但在外国,比它更短的 10 道题目,甚至 5 道题目的五因素取向人格测验也已经开始广泛使用。而这些短式五因素取向人格测验,在那些对施测时间有限制,或人格因素并非研究主要关注的变量的研究中则会发挥重要作用,但这些工具还没有被国内研究选择使用,也未引起研究者的重视。本文不仅期望让研究者全面了解五因素人格测验及技术,还希望对研究者根据自己的研究需求恰当正确地选择五因素人格测评工具提供帮助,也期望对心理测量学领域研究者提供一定启发,将一些新颖、实用的五因素人格测量工具引入中国,进行适用性检验和修订。在将这些新的测量工具引入中国的同时,也不能忽视工具版权的问题,有的工具如 BFI、IPIP-BFFM,编制者已将其放在网站上供人免费使用,其余大部分工具题目和计分键则是在编制者研究论文中出现,国内研究者在使用时应注意相关工具的版权要求。

五因素取向人格测验在全世界范围广泛应用中,也遇到一些问题:如结合国家与地区文化环境特点的特异性人格特质的探讨。对于五因素取向人格测验面临的“测什么”、“怎样测”、“测得如何”提出了挑战和质疑。中国本土人格测验虽然是根据人格特质理论独自开发的具有中国本土文化特点的人格测验,但在发展过程中,随着西方五因素取向人格测验被广泛接受,两者之间不可避免地碰撞在了一起,对于本土测验后期的发展产生了影响,因此西方五因素取向人格测验对于中国本土人格测验的发展所产生的影响也是不可忽略的。如何超越五因素人格,对人格做出全面认识、深刻剖析、精准测量,这将是五因素取向人格特质理论和测验作为起点,进行人格测验研

究和应用实践时, 需要面临和思考的重要问题。

## 参考文献

- 曹慧, 张建新. (2007). 中国人格评定量表(CPAI-2)简式量表修订[摘要]. 第十一届全国心理学学术会议论文摘要集, 11, 501.
- 崔红, 王登峰. (2003). 中国人人格结构的确认与形容词评定结果. *心理与行为研究*, 1(2), 89-95.
- 崔红, 王登峰. (2004). 中国人人格形容词评定量表(QZPAS)的信度, 效度与常模. *心理科学*, 27(1), 185-188.
- 崔红, 王登峰. (2006). 中国人人格形容词评定量表(简式)的信度与效度. *第四军医大学学报*, 27(4), 294-296.
- 王登峰, 崔红. (2008). 领导干部的人格特点与工作绩效的关系: QZPS与NEOPI-R的比较. *心理学报*, 40(7), 828-838.
- 戴晓阳, 姚树桥, 蔡太生, 杨坚. (2004). NEO个性问卷修订本在中国的应用研究. *中国心理卫生杂志*, 18(3), 171-174.
- 戴晓阳, 吴依泉. (2005). NEO-PI-R在16-20岁人群中的应用研究. *中国临床心理学杂志*, 13(1), 14-18.
- 范为桥, 张妙清, 张树辉, 梁觉. (2008). 香港中学生人格特质的性别差异及其发展性分析. *心理学报*, 40(9), 1002-1012.
- 范为桥, 张妙清, 张建新, 张树辉. (2011). 兼顾文化共通性与特殊性的人格研究: CPAI及其跨文化应用. *心理学报*, 43(12), 1418-1429.
- 甘怡群, 张妙清, 宛小昂, 孙增霞. (2005). 用中国人个性量表(CPAI)预测国有企业中高层管理者的绩效. *应用心理学*, 8(3), 35-39.
- 孔伟, 曹佳, 胡文祥. (2002). 火箭推进剂作业人员CPAI测评. *健康心理学杂志*, 10(5), 391-397.
- 刘邦惠. (2009). 罪犯人格障碍症状与人格特征的相关研究. *中国临床心理学杂志*, 17(2), 169-170.
- 罗杰, 戴晓阳. (2012). “大五”人格测验在我国使用情况的元分析. *中国临床心理学杂志*, 19(6), 740-742.
- 秦昕, 牛丛, 黄振雷, 徐敏亚. (2011). 甲流了解程度, 疫苗安全感知, 接种行为及其影响机制. *心理学报*, 43(6), 684-695.
- 宋维真, 张建新, 张建平, 张妙清, 梁觉. (1993). 编制中国人个性量表(CPAI)的意义与程序. *心理学报*, 25(4), 400-407.
- 王登峰, 崔红. (2003). 中国人人格量表(QZPS)的编制过程与初步结果. *心理学报*, 35(1), 127-136.
- 王登峰, 崔红. (2004). 中国人人格量表的信度与效度. *心理学报*, 36(3), 347-358.
- 王登峰, 崔红. (2005a). 中国基层党政领导干部个性与工作绩效关系研究——兼论中国人人格量表(QZPS)与西方大五人格量表(NEO PI-R)的预测效度. *第十届全国心理学学术大会论文摘要集*.
- 王登峰, 崔红. (2005b). 中国人有没有独立的“神经质”人格维度?. *西南师范大学学报(人文社会科学版)*, 31(3), 25-31.
- 王登峰, 崔红. (2006). 中国人的“开放性”——西方“开放性”人格维度与中国人的性格. *西南师范大学学报: 人文社会科学版*, 32(6), 1-10.
- 王登峰, 崔红. (2008a). 中西方人格结构差异的理论与实证分析——以中国人人格量表(QZPS)和西方五因素人格量表(NEOPI-R)为例. *心理学报*, 40(3), 327-338.
- 王登峰, 崔红. (2008b). 领导干部的人格特点与工作绩效的关系: QZPS与NEOPI-R的比较. *心理学报*, 40(7), 828-838.
- 王登峰, 方林, 左衍涛. (1995). 中国人人格的词汇研究. *心理学报*, 27(4), 400-406.
- 王孟成, 戴晓阳, 姚树桥. (2010a). 中国大五人格问卷的初步编制I: 理论框架与信度分析. *中国临床心理学杂志*, 18(5), 545-548.
- 王孟成, 戴晓阳, 姚树桥. (2010b). 中国大五人格问卷的初步编制II: 效度分析. *中国临床心理学杂志*, 18(6), 687-690.
- 王孟成, 戴晓阳, 姚树桥. (2011). 中国大五人格问卷的初步编制III: 简式版的制定及信效度检验. *中国临床心理学杂志*, 19(4), 454-457.
- 王萍萍, 许燕, 陈咏媛, 李佳慧. (2010). 地震灾区高中生外向性对主观幸福感的影响: 社会期望的调节作用. *中国特殊教育*, (6), 65-68.
- 尤瑾, 郭永玉. (2007). “大五”与五因素模型: 两种不同的人格结构. *心理科学进展*, 15(1), 122-128.
- 曾垂凯, 时勤. (2007). 大五人格因素与企业职工工作倦怠的关系. *中国临床心理学杂志*, 15(6), 614-616.
- 张建新, 张妙清, 梁觉. (2003). 大六人格因素的临床价值——中国人人格量表(CPAI), 大五人格问卷(NEOPI)与MMPI-2临床量表的关系模式. *中国心理卫生协会第四届学术大会论文汇编*, 4, 17-22.
- 张建新, 周明洁. (2006). 中国人人格结构探索——人格特质六因素假说. *心理科学进展*, 14(4), 574-585.
- 张妙清, 张树辉, 张建新. (2004). 什么是“中国人”的个性? ——《中国人个性测量表 CPAI-2》的分组差异(英文). *心理学报*, 36(4), 491-499.
- 张妙清, 范为桥, 张树辉, 梁觉. (2008). 跨文化(中国人)个性测量表—青少年版(CPAI-A)的香港标准化研究-兼顾文化共通性与特殊性的人格测量. *心理学报*, 40(7), 839-852.
- 张树辉. (2002, October). 中国人个性测量表在西方样本的初步研究. 第四届华人心理学家研讨会论文, 北京.
- 周晖, 钮丽丽, 邹泓. (2000). 中学生人格五因素问卷的编制. *心理发展与教育*, 16(1), 48-54.
- 周洁. (2010). 大五人格问卷的结构效度分析. *社会心理学*, 25(1), 22-29.
- 朱相华, 乔娟, 李娇, 赵后锋, 周勤, 王成东, 耿德勤. (2012). 儿童期虐待, 大五人格对医学生自尊水平影响的路径分析. *中国健康心理学杂志*, 20(1), 84-86.
- Ahmed, A., & Hasnain, N. (2011). Correlates of personality, diversity, decision-making and affective well being of management students [Special Issue]. *Journal of the Indian Academy of Applied Psychology*, 37, 71-77.
- Allport, G. W. (1931). What is a trait of personality?. *The*

- Journal of Abnormal and Social Psychology*, 25(4), 368–372.
- Atak, H. (2013). The Turkish adaptation of the Ten-Item personality inventory. *Archives of Neuropsychiatry*, 50(4), 312–319.
- Baldasaro, R. E., Shanahan, M. J., & Bauer, D. J. (2013). Psychometric properties of the Mini-IPIP in a large, nationally representative sample of young adults. *Journal of Personality Assessment*, 95(1), 74–84.
- Bernard, L. C., Walsh, R. P., & Mills, M. (2005). Ask once, may tell: Comparative validity of single and multiple item measurement of the Big-Five personality factors. *Counseling and Clinical Psychology Journal*, 2(1), 40–57.
- Block, J. (1995). A contrarian view of the five-factor approach to personality description. *Psychological Bulletin*, 117(2), 187–215.
- Bond, M. H., Nakazato, H., & Shiraishi, D. (1975). Universality and distinctiveness in dimensions of Japanese person perception. *Journal of Cross-Cultural Psychology*, 6(3), 346–357.
- Born, M., & Jooren, J. (2009). *How does the Cross-cultural Personality Assessment Inventory (CPAI-2) fare in The Netherlands?* Paper presented at the 11th European Congress of Psychology, Oslo, Norway, July.
- Cattell, R. B., & Eber, H. W. (1950). *The 16 personality factor questionnaire*. Champaign, Illinois: Institute for Personality and Ability Testing.
- Cheung, F. M., Cheung, S. F., Zhang, J., Leung, K., Leong, F., & Huiyeh, K. (2008). Relevance of openness as a personality dimension in Chinese culture aspects of its cultural relevance. *Journal of Cross-Cultural Psychology*, 39(1), 81–108.
- Cheung, F. M., Leung, K., Zhang, J. X., Sun, H. F., Gan, Y. Q., Song, W. Z., & Xie, D. (2001). Indigenous Chinese personality constructs is the Five-Factor model complete?. *Journal of Cross-Cultural Psychology*, 32(4), 407–433.
- Cheung, F. M., Van de Vijver, F. J., & Leong, F. T. (2011). Toward a new approach to the study of personality in culture. *American Psychologist*, 66(7), 593–603.
- Cobb-Clark, D. A., & Schurer, S. (2012). The stability of big-five personality traits. *Economics Letters*, 115(1), 11–15.
- Cooper, A. J., Smillie, L. D., & Corr, P. J. (2010). A confirmatory factor analysis of the Mini-IPIP five-factor model personality scale. *Personality and Individual Differences*, 48(5), 688–691.
- Costa, P. T., & McCrea, R. R. (1978). Objective personality assessment. In M. Storandt, I. C. Siegler, & M. F. Elias (Eds.), *The clinical psychology of aging* (pp. 119–143). New York: Plenum Press.
- Costa, P. T., & McCrea, R. R. (1985). *The NEO personality inventory manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrea, R. R. (1989). *Revised NEO Personality Inventory (NEO PI-R) and NEO Five-Factor Inventory (NEO-FFI)*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., & McCrae, R. R. (1992). *NEO PI-R professional manual*. Odessa, FL: Psychological Assessment Resources.
- Costa, P. T., McCrae, R. R., & Dye, D. A. (1991). Facet scales for agreeableness and conscientiousness: A revision of the NEO personality inventory. *Personality and Individual Differences*, 12(9), 887–898.
- Credé, M., Harms, P., Niehorster, S., & Gaye-Valentine, A. (2012). An evaluation of the consequences of using short measures of the Big Five personality traits. *Journal of Personality and Social Psychology*, 102(4), 874–888.
- Denissen, J. J., Geenen, R., Van Aken, M. A., Gosling, S. D., & Potter, J. (2008). Development and validation of a Dutch translation of the Big Five Inventory (BFI). *Journal of Personality Assessment*, 90(2), 152–157.
- de Francisco Carvalho, L., Nunes, M. F. O., Primi, R., & da Silva Nunes, C. H. S. (2012). Evidências desfavoráveis para avaliação da personalidade com um instrumento de 10 itens. *Paidéia*, 22(51), 63–71.
- Digman, J. M., & Inouye, J. (1986). Further specification of the five robust factors of personality. *Journal of Personality and Social Psychology*, 50(1), 116–123.
- Donnellan, M. B., Oswald, F. L., Baird, B. M., & Lucas, R. E. (2006). The mini-IPIP scales: Tiny-yet-effective measures of the Big Five factors of personality. *Psychological Assessment*, 18(2), 192–203.
- Duijsens, I. J., & Diekstra, R. W. (1995). The 23bb5: A new bipolar big five questionnaire. *Personality and Individual Differences*, 19(5), 753–755.
- Dwight, S. A., Cummings, K. M., & Glenar, J. L. (1998). Comparison of criterion-related validity coefficients for the mini-markers and Goldberg's markers of the big five personality factors. *Journal of Personality Assessment*, 70(3), 541–550.
- Egan, V., Deary, I., & Austin, E. (2000). The NEO-FFI: Emerging British norms and an item-level analysis suggest N, A and C are more reliable than O and E. *Personality and Individual Differences*, 29(5), 907–920.
- Ehrhart, M. G., Ehrhart, K. H., Roesch, S. C., Chung-Herrera, B. G., Nadler, K., & Bradshaw, K. (2009). Testing the latent factor structure and construct validity of the Ten-Item Personality Inventory. *Personality and Individual Differences*, 47(8), 900–905.
- Engvik, H., & Clausen, S. (2011). Norsk kortversjon av big five inventory (BFI-20). *Tidsskrift for Norsk Psykologforening*, 48(9), 869–872.
- Eysenck, H. J. (1968). *Eysenck personality inventory*. San Diego, Calif.: Educational and Industrial Testing Service.

- Eysenck, S. B. (1983). One approach to cross-cultural studies of personality. *Australian Journal of Psychology*, 35(3), 381–391.
- Farkas, D., & Orosz, G. (2013). The link between ego-resiliency and changes in big five traits after decision making: The case of Extraversion. *Personality and Individual Differences*, 55(4), 440–445.
- Fiske, D. W. (1949). Consistency of the factorial structures of personality ratings from different sources. *The Journal of Abnormal and Social Psychology*, 44(3), 329–344.
- Fossati, A., Borroni, S., Marchione, D., & Maffei, C. (2011). The Big Five Inventory (BFI). *European Journal of Psychological Assessment*, 27(1), 50–58.
- Gerlitz, J. Y., & Schupp, J. (2005). *Assessment of Big Five personality characteristics in the SOEP* (Research Notes 4). Berlin, German: German Institute of Economic Research.
- Goldberg, L. R. (1981). Language and individual differences: The search for universals in personality lexicons. In L. Wheeler (Ed), *Reviews of personality and social psychology* (Vol. 2, pp. 141–165). Beverly Hills, CA: Sage.
- Goldberg, L. R. (1990). An alternative "description of personality": The big-five factor structure. *Journal of Personality and Social Psychology*, 59(6), 1216–1229.
- Goldberg, L. R. (1992). The development of markers for the Big-Five factor structure. *Psychological Assessment*, 4(1), 26–42.
- Goldberg, L. R. (1993). The structure of phenotypic personality traits. *American Psychologist*, 48(1), 26–34.
- Goldberg, L. R. (1999). A broad-bandwidth, public domain, personality inventory measuring the lower-level facets of several five-factor models. In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe* (Vol. 7, pp. 7–28). Tilburg, The Netherlands: Tilburg University Press.
- Goldberg, L. R. (2001). International Personality Item Pool. <http://ipip.ori.org>.
- Goldberg, L. R., & Kilkowski, J. M. (1985). The prediction of semantic consistency in self-descriptions: Characteristics of persons and of terms that affect the consistency of responses to synonym and antonym pairs. *Journal of Personality and Social Psychology*, 48(1), 82–98.
- Gosling, S. D., Rentfrow, P. J., & Swanm, W. B., Jr. (2003). A very brief measure of the Big-Five personality domains. *Journal of Research in Personality*, 37, 504–528.
- Ha, S. E. (2013). Personality traits and South Koreans' attitudes toward North Korea. *Journal of Pacific Rim Psychology*, 7(1), 12–23.
- Hahn, E., Gottschling, J., & Spinath, F. M. (2012). Short measurements of personality—Validity and reliability of the GSOEP Big Five Inventory (BFI-S). *Journal of Research in Personality*, 46(3), 355–359.
- Havill, V. L., Allen, K., Halverson, C. F., & Kohnstamm, G. A. (1994). Parents use of Big Five categories in their natural language descriptions of children. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 371–386). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Hendriks, A. J., Hofstee, W. K., & Raad, B. D. (1999). The five-factor personality inventory (FFPI). *Personality and Individual Differences*, 27(2), 307–325.
- Hoekstra, H. A., Ormel, J., & De Fruyt, F. (1996). Handleiding NEO Persoonlijkheidsvragenlijsten [Manual NEO Personality Questionnaires]. Lisse, the Netherlands: Swets Test Services.
- Hogan, R., & Hogan, J. (1997). *Hogan Personality Inventory: UK edition manual*. Tunbridge Wells: PCL.
- Howard, P. J., Medina, P. L., & Howard, J. M. (1996). The big-five locator: A quick assessment tool for consultants and trainers. In J. W. Pfeiffer (Ed.), *The 1996 Annual: Volume 1* (pp. 107–122). San Diego, CA.
- International Personality Item Pool (2001). A Scientific Collaboratory for the Development of Advanced Measures of Personality Traits and Other Individual Differences (<http://ipip.ori.org/>). Internet Web Site
- Jeswani, S., & Dave, S. (2012). Impact of individual personality on turnover intention a study on faculty members. *Management and Labour Studies*, 37(3), 253–265.
- John, O. P., Angleitner, A., & Ostendorf, F. (1988). The lexical approach to personality: A historical review of trait taxonomic research. *European Journal of Personality*, 2(3), 171–203.
- John, O. P., Caspi, A., Robins, R. W., Moffitt, T. E., & Stouthamer-Loeber, M. (1994). The "Little Five": Exploring the nomological network of the five-factor model of personality in adolescent boys. *Child Development*, 65(1), 160–178.
- John, O. P., Donahue, E. M., & Kentle, R. L. (1991). *The big five inventory—versions 4a and 54*. Berkeley: University of California, Berkeley, Institute of Personality and Social Research.
- John, O. P., Naumann, L. P., & Soto, C. J. (2008). Paradigm shift to the integrative Big Five trait taxonomy: History, measurement, and conceptual issues. In O. P. John, R. W. Robins, & L. A. Pervin (Eds.), *Handbook of personality: Theory and research* (3rd ed., pp. 114–158). New York, NY: Guilford.
- Johnson, J. A. (2000). Web-based personality assessment. In *71st Annual Meeting of the Eastern Psychological Association*. Baltimore, MD.
- Kim, S. Y., Kim, J. M., Yoo, J. A., Bae, K. Y., Kim, S. W.,

- Yang, S. J., & Yoon, J. S. (2010). Standardization and validation of Big Five Inventory-Korean Version (BFI-K) in Elders. *Korean Journal of Biological Psychiatry*, 17(1), 15–25.
- Kovaleva, A., Beierlein, C., Kemper, C. J., & Rammstedt, B. (2013). Psychometric properties of the BFI-K: A cross-validation study. *The International Journal of Educational and Psychological Assessment*, 13(1), 34–50.
- Lang, F. R., John, D., Lütke, O., Schupp, J., & Wagner, G. G. (2011). Short assessment of the Big Five: Robust across survey methods except telephone interviewing. *Behavior Research Methods*, 43(2), 548–567.
- Lang, F. R., Lütke, O., & Asendorpf, J. B. (2001). Testgüte und psychometrische Äquivalenz der deutschen Version des Big Five Inventory (BFI) bei jungen, mittelalten und alten Erwachsenen. *Diagnostica*, 47(3), 111–121.
- Lounsbury, J. W., Tatum, H., Gibson, L. W., Park, S. H., Sundstrom, E. D., Hamrick, F. L., & Wilburn, D. (2003). The development of a Big Five adolescent personality inventory. *Journal of Psychoeducational Assessment*, 21(2), 111–133.
- Martin, T. A., Draguns, J. G., Oryol, V. E., Senin, I. G., Rukavishnikov, A. A., & Klotz, M. L. (1997, August). *Development of a Russian-language NEO-PI R*. Paper presented at the Annual Convention of the American Psychological Convention, Chicago.
- McCrae, R. R., & Costa, P. T., Jr. (1983). Joint factors in self-reports and ratings: Neuroticism, extraversion and openness to experience. *Personality and Individual Differences*, 4(3), 245–255.
- McCrae, R. R., Costa, P. T. (1996). Toward a new generation of personality theories: Theoretical contexts for the five-factor model. In J. S. Wiggins (Eds.), *The Five-Factor model of personality: Theoretical perspectives* (pp. 51–87). New York: Guilford.
- McCrae, R. R., & Costa, P. T., Jr. (1997). Personality trait structure as a human universal. *American Psychologist*, 52(5), 509–516.
- McCrae, R. R., & Costa, P. T., Jr. (2004). A contemplated revision of the NEO Five-Factor Inventory. *Personality and Individual Differences*, 36(3), 587–596.
- McCrae, R. R., & Costa, P. T., Jr. (2007). Brief versions of the NEO-PI-3. *Journal of Individual Differences*, 28(3), 116–128.
- McCrae, R. R., Costa, P. T., Del Pilar, G. H., Rolland, J. P., & Parker, W. D. (1998). Cross-Cultural assessment of the Five-Factor model the revised NEO personality inventory. *Journal of Cross-Cultural Psychology*, 29(1), 171–188.
- McCrae, R. R., Costa, P. T., Jr., & Martin, T. A. (2005). The NEO-PI-3: A more readable revised NEO personality inventory. *Journal of Personality Assessment*, 84(3), 261–270.
- McCrae, R. R., Zonderman, A. B., Costa, P. T., Jr., Bond, M. H., & Paunonen, S. V. (1996). Evaluating replicability of factors in the Revised NEO Personality Inventory: Confirmatory factor analysis versus Procrustes rotation. *Journal of Personality and Social Psychology*, 70(3), 552–566.
- Mervielde, I., & De Fruyt, F. (1999). Construction of the Hierarchical Personality Inventory for Children (HiPIC). In I. Mervielde, I. Deary, F. De Fruyt, & F. Ostendorf (Eds.), *Personality psychology in Europe. Proceedings of the Eight European Conference on personality psychology* (pp. 107–127). Tilburg, the Netherlands: Tilburg University Press.
- Miller, J. D., Gaughan, E. T., Maples, J., & Price, J. (2011). A comparison of Agreeableness scores from the Big Five Inventory and the NEO PI-R: Consequences for the study of narcissism and psychopathy. *Assessment*, 18(3), 335–339.
- Mooradian, T. A., & Nezlek, J. B. (1996). Comparing the NEO-FFI and Saucier's Mini-Markers as measures of the Big Five. *Personality and Individual Differences*, 21(2), 213–215.
- Moore, M., Schermer, J. A., Paunonen, S. V., & Vernon, P. A. (2010). Genetic and environmental influences on verbal and nonverbal measures of the Big Five. *Personality and Individual Differences*, 48(8), 884–888.
- Mullins-Sweatt, S. N., Jamerson, J. E., Samuel, D. B., Olson, D. R., & Widiger, T. A. (2006). Psychometric properties of an abbreviated instrument of the five-factor model. *Assessment*, 13(2), 119–137.
- Murray, H. A. (1938). *Explorations in personality*. New York: Oxford University Press.
- Norman, W. T. (1963). Toward an adequate taxonomy of personality attributes: Replicated factor structure in peer nomination personality ratings. *The Journal of Abnormal and Social Psychology*, 66(6), 574–583.
- Norman, W. T. (1967). *2800 personality trait descriptors: Normative operating characteristics for a university population*. Ann Arbor: University of Michigan, Department of Psychology.
- Oh, I. S., Wang, G., & Mount, M. K. (2011). Validity of observer ratings of the five-factor model of personality traits: A meta-analysis. *Journal of Applied Psychology*, 96(4), 762–773.
- Oshio, A., Abe, S., Cutrone, P., & Gosling, S. D. (2013). Big five content representation of the Japanese version of the Ten-Item personality inventory. *Psychology*, 4(12), 924–929.
- Palmer, J. K., & Loveland, J. M. (2004). Further investigation of the psychometric properties of Saucier's Big Five" Mini-Markers:" Evidence for criterion and construct validity.

- Individual Differences Research*, 2(3), 231.
- Paunonen, S. V. (2003). Big five factors of personality and replicated predictions of behavior. *Journal of Personality and Social Psychology*, 84(2), 411–424.
- Paunonen, S. V., & Ashton, M. C. (1998). The structured assessment of personality across cultures. *Journal of Cross-Cultural Psychology*, 29, 150–170.
- Paunonen, S. V., & Ashton, M. C. (2001). Big Five predictors of academic achievement. *Journal of Research in Personality*, 35(1), 78–90.
- Paunonen, S. V., Ashton, M. C., & Jackson, D. N. (2001). Nonverbal assessment of the Big Five personality factors. *European Journal of Personality*, 15(1), 3–18.
- Paunonen, S. V., Jackson, D. N., & Ashton, M. C. (2004). *NPQ Manual: Nonverbal Personality Questionnaire (NPQ) and Five-Factor Nonverbal Personality Questionnaire (FF-NPQ)*. Sigma Assessment Systems.
- Paunonen, S. V., Jackson, D. N., & Keinonen, M. (1990). The structured nonverbal assessment of personality. *Journal of Personality*, 58(3), 481–502.
- Paunonen, S. V., Zeidner, M., Engvik, H., Oosterveld, P., & Maliphant, R. (2000). The nonverbal assessment of personality in five cultures. *Journal of Cross-Cultural Psychology*, 31, 220–239.
- Perugini, M., & Leone, L. (1996). Construction and validation of a short adjectives checklist to measure Big five (SACBIF). *European Journal of Psychological Assessment*, 12(1), 33–42.
- Piedmont, R. L., & Chae, J. H. (1997). Cross-cultural generalizability of the five-factor model of personality development and validation of the NEO PI-R for Koreans. *Journal of Cross-Cultural Psychology*, 28(2), 131–155.
- Plaisant, O., Srivastava, S., Mendelsohn, G. A., Debray, Q., & John, O. P. (2005). Relations between the French version of the Big Five Inventory and the DSM classification in a French clinical sample of psychiatric disorders. *Annales Médico-Psychologiques*, 163(2), 161–167.
- Psychological Assessment Resources. (1994). *The Revised NEO Personality Inventory: Manual Supplement for the Spanish Edition*. Odessa, FL: Author.
- Rammstedt, B., & John, O. P. (2005). Kurzversion des Big Five Inventory (BFI-K). *Diagnostica*, 51(4), 195–206.
- Rammstedt, B., & John, O. P. (2007). Measuring personality in one minute or less: A 10-item short version of the Big Five Inventory in English and German. *Journal of Research in Personality*, 41(1), 203–212.
- Robins, R. W., John, O. P., & Caspi, A. (1994). Major dimensions of personality in early adolescence: The Big Five and beyond. In C. F. Halverson, G. A. Kohnstamm, & R. P. Martin (Eds.), *The developing structure of temperament and personality from infancy to adulthood* (pp. 267–291). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Rogers, C. R. (1958). A process conception of psychotherapy. *American Psychologist*, 13(4), 142–149.
- Ruiz, V. R., Oberst, U., Gosling, S. D., Rusiñol, J., & Lusaar, A. C. (2013). Translation and validation of the Ten-Item Personality Inventory into Spanish and Catalan. *Aloma: revista de psicologia, ciències de l'educació i de l'esport*, 31, 85–97.
- Saucier, G. (1994). Mini-markers: A brief version of Goldberg's unipolar Big-Five markers. *Journal of Personality Assessment*, 63(3), 506–516.
- Schmitt, D. P., Allik, J., McCrae, R. R., & Benet-Martínez, V. (2007). The geographic distribution of Big Five personality traits patterns and profiles of human self-description across 56 nations. *Journal of Cross-Cultural Psychology*, 38(2), 173–212.
- Shafer, A. B. (1999). Brief bipolar markers for the five factor model of personality. *Psychological Reports*, 84(3), 1173–1179.
- Sinha, N., & Srivastava, K. B. (2013). Association of personality, work values and socio-cultural factors with intrapreneurial orientation. *Journal of Entrepreneurship*, 22(1), 97–113.
- Trapnell, P. D., & Wiggins, J. S. (1990). Extension of the interpersonal adjective scales to include the big five dimensions of personality. *Journal of Personality and Social Psychology*, 59(4), 781–790.
- Tsaousis, I. (1998). The traits personality questionnaire (TPQue): A Greek measure for the five factor model. *Personality and Individual Differences*, 26(2), 271–283.
- Tsaousis, I., & Kerpelis, P. (2004). The Traits Personality Questionnaire 5 (TPQue5): Psychometric properties of a shortened version of a big five measure. *European Journal of Psychological Assessment*, 20(3), 180–191.
- Tupes, E. C., & Christal, R. E. (1961). *Recurrent personality factors based on trait ratings* (USAF ASD Tech. Rep. No. 61–97). Lackland Air Force Base, TX: U.S. Air Force.
- Woods, S. A., & Hampson, S. E. (2005). Measuring the Big Five with single items using a bipolar response scale. *European Journal of Personality*, 19(5), 373–390.
- Zhang, J., & Bond, M. H. (1998). Personality and filial piety among college students in two Chinese societies the added value of indigenous constructs. *Journal of Cross-Cultural Psychology*, 29(3), 402–417.

## The Evolution and Comparison of Personality Tests Based on the Five-Factor Approach

CHEN Jiyue<sup>1</sup>; XU Jianping<sup>1</sup>; LI Hongyan<sup>1</sup>; FAN Yexin<sup>1,2</sup>; LU Xiaolan<sup>1</sup>

*(<sup>1</sup> School of Psychology, Beijing Normal University;*

*Beijing Key Lab for Applied Experimental Psychology, Beijing 100875, China)*

*(<sup>2</sup> Xuanwu Hospital Capital Medical University, Beijing 100053, China)*

**Abstract:** This review aims to examine different aspects of existing personality tests that have been developed to fit the five factor approach. The analyses indicated that there are more than 30 types of five-factor approach personality tests of various lengths and forms. These tests can be classified into three board categories: self-report inventory, adjective checklist and nonverbal inventory. Each test appears to have unique advantages and disadvantages when applied to different research and practical situations. Their reliability and validity also varies across situations. Taken together, these results suggest that researchers should be careful when choosing personality tests and take into account factors such as the test length, the test purpose, and cultural differences.

**Key words:** five-factor approach; big five; five-factor model; personality test