

创造性评价的影响因素综述

张俊俊

西南大学心理学部, 重庆
Email: 1034649694@qq.com

收稿日期: 2021年1月18日; 录用日期: 2021年2月17日; 发布日期: 2021年2月25日

摘要

众多研究证明, 创造性无论是对于个人发展还是科技发明乃至社会进步都有着至关重要的作用, 而创造性评价作为衡量创造性产物创新水平高低过程中的关键环节, 对于后续创造性的激发具有重要的意义。当前, 尽管创造性评价的方法日趋客观, 但由于影响因素纷繁复杂, 学界依旧缺乏规范统一的评价标准。本文主要探讨影响创造性评价的多种因素, 包括评分者、创作者和文化背景等。明确创造性评价的影响因素, 对于今后合理控制来自测评方面的误差, 规范创造性评价的系统性的标准具有重要意义。

关键词

创造性评价, 影响因素, 评分者, 创作者

A Review of the Influencing Factors of Creativity Evaluation

Junjun Zhang

Department of Psychology, Southwest University, Chongqing
Email: 1034649694@qq.com

Received: Jan. 18th, 2021; accepted: Feb. 17th, 2021; published: Feb. 25th, 2021

Abstract

Numerous studies have proved that creativity plays a vital role in personal development, scientific and technological inventions and social progress. Creativity evaluation, as a key link in the process of measuring the level of innovation of creative products, is of great significance to the subsequent stimulation of creativity. At present, although the methods of creativity evaluation are becoming more and more objective, the academic circle still lacks standardized and unified evaluation standards due to the complex influencing factors. This article mainly discusses a variety of factors

that affect creativity evaluation, including raters, creators, and cultural backgrounds. Clarifying the influencing factors of creativity evaluation is of great significance for the reasonable control of errors from the evaluation and standardization of the systemic standards of creativity evaluation in the future.

Keywords

Creativity Evaluation, Influencing Factors, Raters, Creators

Copyright © 2021 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

创造性是很多心理学家研究的主题，关于该方面的研究也是数不胜数。对于创造性的定义，不同的学者从不同的角度对它进行了解释。著名的古希腊哲学家亚里士多德最早对创造性进行了定义：产生前所未有的事物。而研究创造性的先驱心理学家吉尔福特(Guilford)则认为在产生创造性产品的过程中，创造性思维是创造性的具体表现，它的核心是发散思维，即“从给定的信息中产生信息，从同一来源中产生各种各样、许许多多的输出” [1]。随着时代的发展及心理学家的不断探索，现在被广泛认可的创造性的概念是：创造性是一种产生既新颖(novel)又适用(useful)的产品的能力[2]。由此也发展出了许多测量创造性的工具，比如多用途测验(Alternate Use Task, AUT)、托兰斯创造性思维测验(Torrance Tests of Creative Thinking, TTCT)、远距离联想测验(Remote Association Test, RAT)和创造性人格测验等。

在创造性产物输出之后，对其创新性水平评价自然成为一个重要的把关环节。尽管创造性的测量工具多样，创造性评价的方法日趋完善，但创造性评价仍然没有一个确定可靠的标准。所谓“一千个人眼中有一千个哈姆雷特”，对于同一个创造性产物，由于评价主体和创作者不同，可能会导致不同的评价结果。评价主体即评分者，是对创造性产物进行衡量的执行者，如同裁判和法官的作用一般，非常关键。而创作者是与产物一体的个体，不可分割。除了评价个体方面，文化背景等外部环境对创造性评价结果产生的影响也不容忽视。文化的差异造就了思想的不同，看待事物的眼光自然有所差异。总之，这些内外部变量是创造性评价过程中必然要涉及的因素，它们共同促进评价结果的进行与完成。

在创造性评价过程中，无论是评分者、创作者或者文化背景，都可能会影响着评分标准的宽严和评分者一致性的高低，对于评价结果的可靠性有重要的影响。目前，创造性评价结果的偏差主要表现在两个方面：一方面，由于评分者评分标准的宽松或严格，导致对于评价材料的高估或者低估现象，比如在创造性多用途测验中对于创造性观念的新颖性的评分，如果按照相对客观的评价方法即创造性观念的出现频率来评分，即频率越低则视为越新颖，但是在现实测量中，往往会出现许多“独一份”的创造性观念，那么它们真的全都是新颖性水平最高的吗？如果此时仍然刻板遵守“频率为王”的评分标准，结果必然会获得最高分，无疑可能会出现高估现象。但是如果此时综合考虑这些观念的实用性及可操作性等标准，则可以在一定程度上减少误差。另一方面，评分者一致性也是衡量创造性评价结果的重要指标，即两个或两个以上评分者对于同一批材料的评价结果的一致程度，其本质上是相关关系，评分者一致性越高，则评分者信度越高，评价结果也就越可靠。其测量指标主要有 Spearman 相关系数、Kappa 系数、列联系数和肯德尔和谐系数、克隆巴赫 α 系数等。因此，本文基于以上两个方面，对影响创造性评价结果的多种因素进行归纳分析，方便在今后的测评中加以控制，以便提高对创造性评价的标准一致性。

2. 评分者的影响

由于创造性产物的领域及种类多样性,如发散思维测验,这些测验并没有固定的答案,所以目前对于创造性的评价主要是人为评价,但是人为的参加必定会掺杂主观意识,这就会导致评分者效应(rater effects) [3] [4]。

2.1. 个人评价与人际评价

根据被评价材料的所属不同,Runco 和 Smith 将创造性观念评价分为个人评价(intrapersonal evaluation)和人际评价(interpersonal evaluation),前者指个体对自己想法产物的评价,而与此相对,后者指的是个体对他人想法产物的评价[5]。有研究表明,个体对于自己的创造性想法及成果往往高估,而对于他人的观念却倾向于低估,尤其是独特性高的想法[6]。同样的结果也在姜竹卿的研究中得到了验证[7]。分析产生造成这种差异的原因可能是:自我优越感,个体倾向于认为自己的东西要比他人的好,这是一种自我服务偏差,因此对于自己的创造性想法不自觉地进行高估,评分标准相对宽松,而对他人的想法评分标准更为严格。该研究结果揭示了在创造性评价中,选取除自己之外的评分者以提高评分的客观性。

2.2. 评分者的领域匹配与知识经验

领域匹配即评分者的知识领域与被评价材料是高度相通的,比如音乐家评价歌曲,美食家品尝鉴定菜品。很多研究表明,领域匹配的评分者比不匹配的评分者要获得更高的评分一致性。例如,在 Amabile 1982 年的一项研究中,心理学家、艺术教师和艺术家同时对一些儿童创作的拼图进行了评分,结果不同领域的专家得到了不同水平的评分者一致性,尽管三个领域的评分者一致性都达到了 0.70,即被认为是可接受的最低的评分者一致性系数水平,但是发现心理学家的评分者一致性最低,为 0.73,而艺术家们的评分者一致性系数则高达 0.88 [8] [9] [10]。但是另一方面,有研究者发现在不同领域的评分者也可能达到较高的评分者一致性,比如心理学家、老师和作家[11]。对于这种结果,推测可能是由于这三种领域的评分者可能是存在着一些比较类似的认知经验的,因此,才会有比较高的评分者一致性水平。其实,领域匹配性与评分者的知识经验是有密切相关的,知识丰富称为专家,知识匮乏则为新手,对于专家和新手对创造性的评价差异,以前的 CAT 研究已经探讨了用非专家替代专家是否合适,他们发现,与专家评分者相比,新手不可能产生与专家们一样足够高的共识,即新手评价不能代替专家评价。考夫曼和贝尔等人(Kaufman, Baer)发现在评价 200 多首诗歌的整体创造力时,新手(106 名大学生)似乎比 10 位专家诗人表现出稍高的一致性,即 0.94 对 0.83 [12]。但是也有研究并没有发现专家和新手评分一致性水平的差异,如 Freeman 等人比较了三名专家和三名新手评分者在使用同感评估技术(CAT)对时尚插图进行评价时的结果,发现专家小组之间的评分一致性并没有显著高于新手评分者[13]。Amabile 也指出,在她的一些研究中,专业知识似乎没有像预期的那样增加评判的可靠性,并指出“在裁决的可靠性方面,艺术家与非艺术家没有明显的优势”,而非艺术家和艺术家评分者在评定拼贴画时,表现出相当高的一致性[8]。Amabile 承认,这可能是由于正在评估的任务的性质所致,拼贴可能足够简单,以至于大多数人对它们的熟悉程度都很高,而更专业的领域(例如医学)可能需要更多的专业经验。同样,这也是考夫曼、贝尔和科尔(Kaufman, Baer, and Cole)提出的一个观点,他们发现,新手和专家之间的对于短篇小说的评价一致性相关比对诗歌的更高,这也表明了领域或任务的交互作用[14]。

除了专家和新手,现实生活中可能更多的是“准专家”,比如相对于本科生来说,研究此领域的研究生可能会具有更多的知识经验,但是又不像专家那样精通。Kaufman、Gentile 和 Baer 发现,准作家在对他人创作诗歌和短篇小说的评价中与专家作家的评价同样匹配,这些准专家作家的评级与专家的评级之间存在很高的相关性[15]。同样, Kaufman 和 Baer 等人对具有不同专门知识水平的评分者进行的两

项研究中,对写作和工程产品设计任务的作品进行评级,发现准专家可替代专家,但仅适用于写作任务,这表明专业知识和领域或任务可能存在相互交叉[16]。这些不同的结果表明,有的时候专门的知识经验尤其重要,但是在某些较为公知的领域,专业知识似乎效应减弱,对此,今后的研究重点则可以去探究导致专家和新手评价差异性的其他因素。

2.3. 评分者的认知与人格特征

众多研究发现,具有高智力水平的个体往往创造性水平也会更高,刘焯琴和刘玲等人研究表明流体智力通过影响大五人格中的开放性间接地正向预测创造性倾向中的好奇心[17]。可见认知水平和人格特征会影响创造性水平,同理,评分者的这些特征也会对创造性的评价结果产生影响。比如,有研究表明,评分者的创造性水平越高,评价得越准确,但是与此同时,他们会对产品的新颖性存在低估的现象[18]。除此之外,也有研究者探讨了认知负荷对评价结果的影响,尽管认知心理学中有关于选择过载和认知负荷的文献[19][20],但是没有已知的创造性测评研究发现评分者在相应一段时间内能够处理多少数量的创造性想法,Amabile 研究结果发现,评分者在一段时间内处理的数据越多,花费时长就要相应延长,疲劳程度就越大,比较的能力就受到损害。另外,Karwowski 等人在 2016 年重复了 Jauk、Benedek、Dunst 和 Neubauer 的研究;这两项研究也都探讨了创造力评价和智力之间的关系[21][22]。在人格特征方面,Tan 等人的研究中先对大学生进行了大五人格测验和创造性测量,之后以他们作为评分者,对儿童的乐高积木作品进行评分。结果发现,人格测验中宜人性维度得分高和创造性水平高的评分者,其评分标准会更为宽松[23]。同样,Benedek 等的研究则发现了高开放性和高智商的评分者能够更好地发现独特性想法[24]。这启示我们,高认知水平和高创造性的个体不仅能够自己生成更具有创造性的产物,还能够善于成为他人独特性产物的“伯乐”。

2.4. 评分者的人数

除了评分者个体的内部特征以外,评分者的人数也可能是影响创造性评价的重要因素。Amabile 指出,很难指定一个理想的数字,只是它应该在很大程度上根据达成可接受的一致性所需的条件来确定。然而,根据测量工具、测量过程等的不同,很难提前计划有多少数字可能实现这一目标。Silvia 等人于 2008 年也曾提出,除了模糊的直觉之外,没有明确的指导方针:“一个显然是不够的;20 个似乎是过分的”[25]。基于此,Kaufman 和 Plucker 等人曾建议,“就大多数目的而言,5~10 名评分者人数足够”[26]。然而,实际研究中使用的评分者人数差别很大。比如,在 Amabile 的多项创造性研究中,评分者人数从 3 到 21 不等。还有一些研究只用了两名评分者。如 Daly 等人的研究中,要求两名评分者对 439 名工科学生设计的创造性作品进行评分,结果评分者一致性系数为 0.70 [27]。当然也有一些研究使用了更多数量的评分者,例如,Valgeirsdottir、Onarheim 和 Gabrielsen 就招募了多达 134 名普通公众评分者对两种手机产品的创造性进行评分[28]。

由此看来,对于创造性测评中评分者人数的采用标准需要进一步探索,虽然根据研究任务和变量的不同,所需要的评分者人数会存在差异,但是从管理学上来讲,合理利用人力资源,既要达到较高的评分者一致性,同时也要控制人力、金钱等资源的合理投入,未来研究仍然需要细致研究。

3. 创作者的影响

行为研究表明,作者信息在艺术评价过程中起着重要的作用。特别是,当作者信息反映出高声望时,艺术品的美学和创造力评级就会膨胀。这种现象被称为威望偏见。

这种偏见广泛存在于多种艺术形式中。例如,一项关于音乐偏好的研究,要求学生对比贝多芬作品的质量进行评分。这首音乐被称为是由“专业”或“学生”作曲家创作的。结果发现,被表现出更高的偏

好是“专业”作曲[29]。在评价对绘画的偏好方面也发现了类似的偏见[30] [31] [32]。此外, Kirk、Skov、Hulme、Christensen 和 Zeki 的一项研究的被试行为表明, 他们更喜欢从著名的美术馆获得的艺术品, 而不是由实验者使用照片编辑软件生成的艺术品[33]。总之, 这些研究表明, 作者信息体现的威望, 对于创造性的评价起着重要的作用。此外, 创造者年龄的信息也会显著影响评分者的主观评价, 具体来说, 评分者对于儿童作品的评价分数显著高于对成人的评价分数[34]。这说明评分者得知创作者的年龄之后, 会将此与作品联合起来, 考虑多种外部因素, 从而采用各自的评分方法。与此类似的结果也出现在了韩建涛等人的新近研究中, 该研究中将创作者的年龄分别虚构为儿童、大学生、成年人和老年人四种情况, 结果发现, 对于年龄是儿童组的创造性想法, 评分者倾向于给出高的独特性和实用性得分, 而对于虚构为老年组的想法, 则给出更低的分, 验证了人们固有的“小而灵通, 老而顽固”的刻板印象[35]。除了对年龄的刻板印象, 创作者的性别信息也会影响创造性评价的结果, 譬如, Lebeda 和 Karwowski 发现, 创作者为男性的作品会获得更高的评分[36], 这也符合性别刻板印象的结果[37], 即认为男性在很多认知方面的表现优于女性。此外, 为了探究种族对评价的影响, Kaufman、Baer、Agars 和 Loomis 等人的研究则发现了评分者对于白种人创作者的更高评分[38]。总之, 在创造性的测评过程中, 大多数情况下虽然评分者是来自外部的, 但是如果给予创作者的个人信息, 仍然会对评价结果产生影响, 因此, 在今后研究中探索评分者与创作者的交互作用显得尤为重要。

4. 文化背景的影响

来自同一文化背景的评分者往往产生类似的创造性评价排名。但是, 如果从更大的历史和地理范围来考虑, 创造性评估的情况就会变得非常不同, 以至于有些文化甚至没有一个关于创造力的词如阿拉伯语[39], 而另一些文化则有两个或两个以上如波兰语[40]。这引人思考, 这意味着什么? 是否意味着一个人不能根据一个普遍的、科学的定义来评价创造力, 而不能从这些特定的环境之外, 作为一个独立的、公正的观察者来评价创造力? 为了研究创造性的跨文化评价, Lan 和 Kaufman 的研究中选取了美国和中国的评分者对同一批材料进行评价, 结果表明, 中国评分者更加偏好在一定社会条件下的发挥, 而来自美国的评分者则不管社会条件的限制, 更加喜欢大胆不落窠臼的想法[41], 即关于创造性的两个关键要素, 美国评分者更加重视新颖性, 而中国评分者更加在意实用性[42]。该结果的差异与东西方的文化背景有关, 与西方相比, 东方人更加传统谨慎。这种结果也在 Hong 和 Lee 的研究中得到了进一步验证, 在对建筑的创造性进行评价时, 同时作为评分者, 在相同情况下, 东亚人比美国人对新颖建筑的评分更低[43]。但是, 由于世界广阔, 文化形式多种多样, 创造性评价的跨文化研究显得难以达到全面, 因此目前关于创造性评价的跨文化研究并不是很多。总之, 跨文化评价是创造力研究中的一个难题, 就像心理学和相关学科的交叉一样。毫无疑问, 在该方向的研究沿途有许多陷阱, 但它们可以转化为机会, 以加深我们对文化和创造性的交互影响的理解。

5. 其他因素的影响

除了以上谈到的影响创造性评价的几个因素之外, 研究者们不懈探索, 从不同角度分析了其他影响因素。比如, 关于创造性评价时情境因素的影响, Birney、Beckmann 和 Seah 于 2016 年提出了个体 - 任务 - 情境三维创造性评价模型, 该模型认为创造性评价并不是一个完全单维独立的过程, 在此过程中, 由作为个体的评分者甚至创作者、评价任务和情境因素共同促成[44]。之后, Zhou 等人的研究则由此进一步揭示, 创造性测评受到评分者(调节定向)、观点(创造性水平)和情境(损失或收益)三者之间的交互影响[45]。不仅如此, 张红和任靖远的研究则发现了创造性产品评价中的从众效应, 采用“初次评价 - 从众诱导 - 延时再评”的实验程序, 结果揭示, 对于新颖性和实用性的评价都存在从众效应[46]。而姜竹卿

则从评价时所采用的参考系角度分析,即他人作为评分者,分别参考单个被试的答案集和所有被试的答案集进行评价时,结果发现,两者并不存在显著差异[7]。由于人为评价创造性的主观性导致评价结果不够客观,所以有研究者从评价工具方面进行了研究,徐建平等人利用 Python 语言设计图形类发散思维测验机器评分程序,对发散思维测验的答案进行机器评分,并与人工评分进行比较,发现了机器评分的可行性及良好的效度[47],这为发散思维测验实现自动评分提供了可靠的技术支持。这也为将来规范性评价的统一标准的智能化提供了可能,给研究者带来了信心。

6. 总结与展望

综上所述,创造性评价的准确性和一致性受到多种因素的影响,评分主体涉及到的评分者和创作者很容易给评分过程造成系统误差,外部的文化背景因素产生的影响也值得探索。具体可总结为以下几点:1) 评分者的内外部方面,个人评价要高于人际评价,即自己对自己的创造性观念存在高估现象。2) 如果为外部评分者,评分者的领域匹配和知识经验影响着评分一致性,领域匹配的评分者一致性水平要高于领域不匹配的评分者一致性。知识经验方面,专家在某些领域的一致性显著高于新手的评分者一致性,而在另一些领域却无显著差异,表明知识经验和领域任务存在交互作用。3) 评分者认知特征方面,智力和创造性水平高的个体创造性评价更加准确;人格特征方面,大五人格中的尽责性和开放性也会影响创造性评价。4) 目前研究中对于创造性的人数并没有统一的标准,这需要研究者们重点关注并继续探索。5) 关于创作者方面,存在年龄刻板印象和性别刻板印象,即对具有儿童或男性创作者信息的创造性产物进行高估。6) 此外,文化背景和评分情境等的不同也会造成评分结果的差异。具体说来,西方评分者侧重新颖性的评价,而东方评分者更加在意实用性。对于这些已知的影响因素,我们要致力于探索背后的心理机制并加以调控,比如可以采用深化评分者认知研究、整合不同层面评分者效应的研究、拓展创造力测评方法和技术等方法减少误差[48],目前,对于创造性的评价方法仍然是以评分者的主观评价为主,而这个过程必然是评分者的认知、评价及决策过程,因此深入探索评分者这一系列认知过程背后的心理机制甚至是脑机制都有必要意义,未来研究可以采用脑电及核磁技术更加细致地监测这一过程的动态变化,细化到动用的每个脑区,以此不仅可以加深对评分者认知的深入研究,还可借助生物脑科学方面的技术,进一步促进交叉学科的发展。值得我们注意的一点是,关于评分者的人数标准,学界至今并无一个固定的说法,并且根据任务的不同,所需的评分者人数也会有所差异,并且要同时考虑到评价的准确性和人力资源的合理使用,关于此方面的研究也需要进一步拓展,可作为将来创造性评价研究的一个重点。然而,要实现创造性评价的绝对客观甚至是自动化并不是绝无可能,我们可以发挥现代计算机智能化的技术,如前文所述,利用电脑软件 Python 等的语句编程,采用机器统一评分,现有的研究已经取得了初步成果,虽然还不够完善,但这已经为以后的研究者提供了信心。此外,创造性评价的影响因素之间往往存在着复杂的交互作用,关于此方面的研究方兴未艾,但是尚未有一个固定的答案,它们正在等待研究者揭开神秘的面纱。

本文通过对创造性评价过程影响因素研究的梳理,有利于今后的研究者规避影响创造性测评的各种不利因素,为创造性评价发展出更为客观、准确的评价标准,特别是为跨文化背景下的创造性测评搭建一座桥梁,然而,这还需要研究者们持之以恒的努力。

参考文献

- [1] Guilford, J.P. (1950) Creativity. *American Psychologist*, 5, 444-454. <https://doi.org/10.1037/h0063487>
- [2] Sternberg, R.J., Amabile, T.M., Lubart, T.I., et al. (1999) *Handbook of Creativity*. Cambridge University Press, New York.
- [3] Wolfe, E.W. (2004) Identifying Rater Effects Using Latent Trait Models. *Psychology Science*, 46, 35-51.

- [4] Wolfe, E.W. and McVay, A. (2012) Application of Latent Trait Models to Identifying Substantively Interesting Raters. *Educational Measurement: Issues and Practice*, **31**, 31-37. <https://doi.org/10.1111/j.1745-3992.2012.00241.x>
- [5] Runco, M.A. and Smith, W.R. (1992) Interpersonal and Intrapersonal Evaluations of Creative Ideas. *Personality and Individual Differences*, **13**, 295-302. [https://doi.org/10.1016/0191-8869\(92\)90105-X](https://doi.org/10.1016/0191-8869(92)90105-X)
- [6] Blair, C.S. and Mumford, M.D. (2007) Errors in Idea Evaluation: Preference for the Unoriginal? *The Journal of Creative Behavior*, **41**, 197-222. <https://doi.org/10.1002/j.2162-6057.2007.tb01288.x>
- [7] 姜竹卿. 创造性评价: 参考系、评价主体和评价指标的作用[D]: [硕士学位论文]. 重庆: 西南大学, 2019.
- [8] Amabile, T.M. (1982) Social Psychology of Creativity: A Consensual Assessment Technique. *Journal of Personality and Social Psychology*, **43**, 997-1013. <https://doi.org/10.1037/0022-3514.43.5.997>
- [9] Amabile, T.M. (1996) Creativity in Context: Update to the Social Psychology of Creativity. Westview, Boulder.
- [10] Kline, P. (2000) The Handbook of Psychological Testing, 2nd Edition, Vol. 25, Routledge, London.
- [11] Baer, J., Kaufman, J.C. and Riggs, M. (2009) Rater-Domain Interactions in the Consensual Assessment Technique. *International Journal of Creativity & Problem Solving*, **19**, 87-92. <https://doi.org/10.1037/e492272008-001>
- [12] Kaufman, J.C., Baer, J., Cole, J.C. and Sexton, J.D. (2008) A Comparison of Expert and Nonexpert Raters Using the Consensual Assessment Technique. *Creativity Research Journal*, **20**, 171-178. <https://doi.org/10.1080/10400410802059929>
- [13] Freeman, C., Son, J. and McRoberts, L.B. (2015) Comparison of Novice and Expert Evaluations of Apparel Design Illustrations Using the Consensual Assessment Technique. *International Journal of Fashion Design, Technology and Education*, **8**, 122-130. <https://doi.org/10.1080/17543266.2015.1018960>
- [14] Kaufman, J.C., Baer, J. and Cole, J.C. (2009) Expertise, Domains, and the Consensual Assessment Technique. *Journal of Creative Behavior*, **43**, 223-233. <https://doi.org/10.1002/j.2162-6057.2009.tb01316.x>
- [15] Kaufman, J.C., Gentile, C.A. and Baer, J. (2005) Do Gifted Student Writers and Creative Writing Experts Rate Creativity the Same Way? *Gifted Child Quarterly*, **49**, 260-265. <https://doi.org/10.1177/001698620504900307>
- [16] Kaufman, J.C., Baer, J., Cropley, D.H., Reiter-Palmon, R. and Sinnett, S. (2013) Furious Activity vs. Understanding: How Much Expertise Is Needed to Evaluate Creative Work? *Psychology of Aesthetics, Creativity, and the Arts*, **7**, 332-340. <https://doi.org/10.1037/a0034809>
- [17] 刘烯琴, 刘玲, 宋宜颖, 刘嘉. 流体智力对创造性倾向的影响: 人格的中介作用[C]//第十七届全国心理学学术会议论文摘要集, 2014: 1145-1147.
- [18] 葛樱. 个体的创造力水平对创造性产品评价的影响及其潜在机制[D]: [硕士学位论文]. 上海: 华东师范大学, 2018.
- [19] Miller, G.A. (1956) The Magical Number Seven plus or minus Two: Some Limits on Our Capacity for Processing Information. *Psychological Review*, **63**, 81-97. <https://doi.org/10.1037/h0043158>
- [20] Scheibehenne, B., Greifeneder, R. and Todd, P.M. (2010) Can There Ever Be Too Many Options? A Meta-Analytic Review of Choice Overload. *Journal of Consumer Research*, **37**, 409-425. <https://doi.org/10.1086/651235>
- [21] Karwowski, M., Dul, J., Gralewski, J., Jauk, E., Jankowska, D.M., Gajda, A., Benedek, M., et al. (2016) Is Creativity without Intelligence Possible? A Necessary Condition Analysis. *Intelligence*, **57**, 105-117. <https://doi.org/10.1016/j.intell.2016.04.006>
- [22] Jauk, E., Benedek, M., Dunst, B. and Neubauer, A.C. (2013) The Relationship between Intelligence and Creativity: New Support for the Threshold Hypothesis by Means of Empirical Breakpoint Detection. *Intelligence*, **41**, 212-221. <https://doi.org/10.1016/j.intell.2013.03.003>
- [23] Tan, M., Mourgues, C., Hein, S., MacCormick, J., Barbot, B. and Grigorenko, E. (2015) Differences in Judgments of Creativity: How Do Academic Domain, Personality, and Self-Reported Creativity Influence Novice Judges' Evaluations of Creative Productions? *Journal of Intelligence*, **3**, 73-90. <https://doi.org/10.3390/jintelligence3030073>
- [24] Benedek, M., Nordtvedt, N., Jauk, E., Koschmieder, C., Pretsch, J., Krammer, G. and Neubauer, A.C. (2016) Assessment of Creativity Evaluation Skills: A Psychometric Investigation in Prospective Teachers. *Thinking Skills and Creativity*, **21**, 75-84. <https://doi.org/10.1016/j.tsc.2016.05.007>
- [25] Silvia, P.J., Winterstein, B.P., Willse, J.T., Barona, C.M., Cram, J.T., Hess, K.I., Richard, C.A., et al. (2008) Assessing Creativity with Divergent Thinking Tasks: Exploring the Reliability and Validity of New Subjective Scoring Methods. *Psychology of Aesthetics, Creativity, and the Arts*, **2**, 68-85. <https://doi.org/10.1037/1931-3896.2.2.68>
- [26] Kaufman, J.C., Plucker, J.A. and Baer, J. (2008) Essentials of Creativity Assessment.
- [27] Daly, S.R., Seifert, C.M., Yilmaz, S. and Gonzalez, R. (2016) Comparing Ideation Techniques for Beginning Designers. *Journal of Mechanical Design*, **138**, Article ID: 101108. <https://doi.org/10.1115/1.4034087>
- [28] Valgeirsdottir, D., Onarheim, B. and Gabrielsen, G. (2015) Product Creativity Assessment of Innovations: Considering

- the Creative Process. *International Journal of Design Creativity and Innovation*, **3**, 95-106. <https://doi.org/10.1080/21650349.2014.954626>
- [29] Duerksen, G.L. (1972) Some Effects of Expectation on Evaluation of Recorded Musical Performance. *Journal of Research in Music Education*, **20**, 268-272. <https://doi.org/10.2307/3344093>
- [30] Bernberg, R.E. (1953) Prestige Suggestion in Art as Communication. *The Journal of Social Psychology*, **38**, 23-30. <https://doi.org/10.1080/00224545.1953.9711433>
- [31] Gergen, K.J. and Breger, I. (1965) Two Forms of Inference and Problems in the Assessment of Creativity. *Proceedings of the Annual Convention of the American Psychological Association*, **20**, 215-216. <https://doi.org/10.1037/e469432008-107>
- [32] Marcus, S., Opera, P. and Sandulescu, M. (1973) Painting Evaluation as Affected by Properties of the Stimulus. *Revista de Psihologie*, **19**, 225-237.
- [33] Kirk, U., Skov, M., Hulme, O., Christensen, M.S. and Zeki, S. (2009) Modulation of Aesthetic Value by Semantic Context: An fMRI Study. *NeuroImage*, **44**, 1125-1132. <https://doi.org/10.1016/j.neuroimage.2008.10.009>
- [34] Hennessey, B.A. (1994) The Consensual Assessment Technique: An Examination of the Relationship between Ratings of Product and Process Creativity. *Creativity Research Journal*, **7**, 193-208. <https://doi.org/10.1080/10400419409534524>
- [35] 韩建涛, 庞维国. 小而灵通, 老而顽固? 年龄信息对创造性观念评价的影响及其机制[J]. 心理科学, 2020(3): 615-621.
- [36] Lebeda, I. and Karwowski, M. (2013) Tell Me Your Name and I'll Tell You How Creative Your Work Is: Author's Name and Gender as Factors Influencing Assessment of Products' Creativity in Four Different Domains. *Creativity Research Journal*, **25**, 137-142. <https://doi.org/10.1080/10400419.2013.752297>
- [37] Leslie, S.J., Cimpian, A., Meyer, M. and Freeland, E. (2015) Expectations of Brilliance Underlie Gender Distributions across Academic Disciplines. *Science*, **347**, 262-265. <https://doi.org/10.1126/science.1261375>
- [38] Kaufman, J.C., Baer, J., Agars, M.D. and Loomis, D. (2010) Creativity Stereotypes and the Consensual Assessment Technique. *Creativity Research Journal*, **22**, 200-205. <https://doi.org/10.1080/10400419.2010.481529>
- [39] Hennessey, B.A. (2003) Is the Social Psychology of Creativity Really Social? Moving beyond a Focus on the Individual. In: Paulus, P. and Nijstad, B., Eds., *Group Creativity: Innovation through Collaboration*, Oxford University Press, New York, 181-201. <https://doi.org/10.1093/acprof:oso/9780195147308.003.0009>
- [40] Karwowski, M. (2009) I'm Creative, But Am I Creative? Similarities and Differences between Self-Evaluated Small and Big-C Creativity in Poland. *The International Journal of Creativity & Problem Solving*, **19**, 7-26.
- [41] Lan, L. and Kaufman, J.C. (2012) American and Chinese Similarities and Differences in Defining and Valuing Creative Products. *The Journal of Creative Behavior*, **46**, 285-306. <https://doi.org/10.1002/jocb.19>
- [42] Goncalo, J.A. and Staw, B.M. (2006) Individualism-Collectivism and Group Creativity. *Organizational Behavior and Human Decision Processes*, **100**, 96-109. <https://doi.org/10.1016/j.obhdp.2005.11.003>
- [43] Hong, S.W. and Lee, J.S. (2015) Nonexpert Evaluations on Architectural Design Creativity across Cultures. *Creativity Research Journal*, **27**, 314-321. <https://doi.org/10.1080/10400419.2015.1087245>
- [44] Birney, D.P., Beckmann, J.F. and Seah, Y.Z. (2016) More than the Eye of the Beholder: The Interplay of Person, Task, and Situation Factors in Evaluative Judgments of Creativity. *Learning and Individual Differences*, **51**, 400-408. <https://doi.org/10.1016/j.lindif.2015.07.007>
- [45] Zhou, J., Wang, X.M., Song, L.J. and Wu, J. (2017) Is It New? Personal and Contextual Influences on Perceptions of Novelty and Creativity. *Journal of Applied Psychology*, **102**, 180-202. <https://doi.org/10.1037/apl0000166>
- [46] 张红, 任靖远, 刘晨阳, 罗劲. 创造性产品评价中的从众效应[J]. 心理学报, 2019, 51(6): 688-698.
- [47] 徐建平, 段海燕, 李露, 朱黎明. 发散思维测验人工评分与机器评分比较[C]//第二十二届全国心理学学术会议摘要集, 2019: 2048-2049.
- [48] 韩建涛, 刘文令, 庞维国. 创造力测评中的评分者效应[J]. 心理科学进展, 2019, 27(1): 171-180.