

A Review on Evaluation Methods of Creative Thinking

Zhuqing Jiang, Ling Liu, Zhan Xu

Faculty of Psychology, Southwest University, Chongqing
Email: kira72@163.com

Received: Jan. 4th, 2019; accepted: Jan. 17th, 2019; published: Jan. 24th, 2019

Abstract

Creativity is the ability to produce novel, original, unusual and appropriate solutions to complex problems. It plays an important role in promoting scientific and technological progress, artistic creation and commercial development. Its core is creative thinking, which can be divided into divergent thinking and convergent thinking. This study starts from the four factors of creative divergent thinking proposed by Guilford (1967), sorts out the evaluation methods of divergent thinking, and summarizes them into objective and subjective, at the same time, analyzes the methods from the perspective of subjective and objective perspectives. Looking forward to the future research direction with their advantages and disadvantages respectively at the end.

Keywords

Creative Thinking, Divergent Thinking, Creative Evaluation

创造性思维评价方法的研究综述

姜竹卿, 刘玲, 徐展

西南大学心理学部, 重庆
Email: kira72@163.com

收稿日期: 2019年1月4日; 录用日期: 2019年1月17日; 发布日期: 2019年1月24日

摘要

创造性是对复杂问题能够产生新颖的、独创的、不寻常的、适宜的解决方案的能力, 对促进科技进步、艺术创作及商业发展等起着重要作用, 其核心是创造性思维, 分为发散思维和聚合思维。本研究从 Guilford (1967) 提出的创造性发散思维的四个指标出发, 对发散思维的评价方式进行了梳理, 归纳为客

观和主观两种评价方法, 从主客观的角度进行了分析。最后, 结合各自存在的优点和缺陷对未来研究进行了展望。

关键词

创造性思维, 发散思维, 创造性评价

Copyright © 2019 by author(s) and Hans Publishers Inc.

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

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



Open Access

1. 创造性思维

创造性(Creativity), 也称创造力, 源于拉丁文 *creare* 一词。意指创造、创建、生产和造就等。早在古希腊时期, 亚里士多德(Aristotle)就将“创造”定义为“产生前所未有的事物”, 这一定义既包括了精神领域, 也包括了物质世界。虽然这一定义代表了许多研究者的看法, 但是不同的心理学家对创造性的理解和使用存在很大差异, 因此, 迄今为止创造性还没有一个统一、精确的定义。然而, 多数研究者一致认同, 创造性可指个体产生新颖且适用的想法或产品的能力(Runco & Jaeger, 2012)。创造性对促进科技进步、艺术创作及商业发展等起着重要作用, 创造性的核心是创造性思维(张庆林&曹贵康, 2004)。创造性思维包括两种不同的思维方式, 即发散思维(Divergent Thinking, DT)和聚合思维(Convergent Thinking, CT) (Ashton-James & Chartrand, 2009; Chamorro-Premuzic & Reichenbacher, 2008; Colzato, Ozturk, & Hommel, 2012)。Guilford (1950)在智力的三维结构模型中首次区分出发散思维(DT)和聚合思维(CT)。发散思维(DT)是指回答开放式问题或对产生多种想法或解决方案时的认知过程(Guilford, 1950, 1967), 个体沿不同方向进行思考, 重新组织已有信息, 产生新颖的、不寻常的且适用的想法, 强调从不同的角度着手并沿着不同的方向思考, 从而产生多个可行的答案, 而聚合思维(CT)强调综合各方面因素并朝着某个固定的方向进行思考, 从而产生唯一的正确答案。其中发散思维是创造性思维的核心。

2. 发散思维的测量任务及测量指标

针对发散思维的研究, 已经发展出了各种测量发散思维能力的工具, 常见的发散思维任务有以下几种: 一是多用途任务(Alternate Use Task, AUT), 即在一定时间内尽可能多的列出一个常见物品的用途, 比如“砖头”(Gilhooly, Fioratou, Anthony, & Wynn, 2007; Guilford, 1967); 二是后果任务(Consequences Task), 即尽可能多的想象“如果某件事情发生会造成什么后果”, 比如重力消失会造成什么后果; 三是故事生成任务(Story Generation Task), 即根据给定的提示词生成一个完整故事(Howard-Jones, Blakemore, Samuel, Summers, & Claxton, 2005); 四是问题发现任务(Problem Finding Task), 根据某段文字或某张图片来提出多个可能的问题(Cheng, Hu, Jia, & Runco, 2016); 还有托兰斯创造性思维测验(Torrance Tests of Creative Thinking)、南加利福尼亚大学测验、芝加哥大学创造性测验、威廉姆斯创造性测验(Williams Creativity Assessment Packet, CAP)等等。发散思维测验已经成为测量创造性过程与与创造性有关的技能中最常用的手段(Kaufman & Sternberg, 2007; Plucker, Qian, & Wang, 2011)。本研究中的发散思维任务的评价以 AUT 任务为例。

根据 Guilford (1950)的智力三维模型, 发散思维具有四个主要特征。一是独特性或新颖性(Originality or Novelty or Uniqueness), 指的是对事物的看法或想法超乎寻常或独出心裁的新奇程度。创造性高的个体, 往往能突破常规和经验的束缚, 提出超乎寻常的新观念。二是流畅性(Fluency), 指在一定时间内连续地

产生想法和解决方案的数量。创造性高的个体,心智活动流畅顺利,能够在有限时间内提出较多的想法,即反应迅速而且数目较多。三是变通性或灵活性(Flexibility),指对问题能多方向、多角度思考的灵活程度。具有高变通性的个体往往不易受到思维定势的影响,思维灵活多变、举一反三。例如“砖头”一词,变通性较低的个体一般会想到“铺路、砌墙”等,拘泥于“建筑材料”这一范畴,但是变通性高的个体则可能做出不同类别的反应,例如“打外星人、当椅子、刻石雕”等。四是精细性(Elaboration),指想法的详细或细节明确具体的程度,例如对“砖头”的反应为“把砖头磨成粉末之后画沙画”这种详细的描述。这四个特点也经常作为测量发散思维能力的操作指标(Guilford, 1967; Radel, Davranche, Fournier, & Dietrich, 2015; White & Shah, 2016),也有一些研究只用到前面三个指标。由此可以看出,发散思维的测验不仅仅看反应的数量而且看重反应的创造性质量(Silvia, 2008)。

3. 发散思维任务的评价方法

长久以来,研究者们发展了很多种发散思维任务的评价方式,其中争议最多的来自于对独特性评价。通过整理文献,对独特性的评价方法按照评分的依据标准大致可以归纳为两种,即客观的评价和主观的评价。客观的评价是指评分过程中不依赖于评分者和被试的主观想法,仅仅是在所有被试的答案库中,运用一些数学的方法进行独特性评价。主观的评价往往需要请除主试和被试以外的其他人来评价,依靠评价者的主观感受或经验来对被试的答案进行独特性评分。

3.1. 客观的评价方法

根据 Guilford 对创造性的早期探索,将独特性定义为“不寻常的”(Uncommon)、“远距离的”(Remote)和“聪明的”(Clever) (Wilson, Guilford, & Christensen, 1953)。其中,对“不寻常的”(Uncommon)的操作性定义为个体产生的想法在统计上对于该被试所属的群体来说是不常见的。具体而言就是,将一组被试的所有反应汇集在一起,统计每一个反应在组内出现的频率,按照出现频率的高低赋予相应的权重,将出现频率更低的答案给予更高的权重。因此,那些得分高的个体,往往就是那些给出了不常被他人提及的答案的个体。例如,在该研究中,被试要求对 6 个常见物品写出 6 个非常规的用途。410 名被试的所有反应答案经过分类、统计和加权,采用 5 重制,对于最不常提到的回答权重为 5,其次是最不常提到的回答权重为 4,以此类推,对于最常提到的回答权重为 1,将每个被试的权重分数相加即可得到该被试的独特性分数。

Wallach 和 Kogan (1965)提出了一个聪明而简单的方法,即将所有的样本量汇集起来,给每一个样本赋予 0 或 1 分,如果某个样本只被一个被试提出,这个唯一的答案将得 1 分,而其余的样本都得 0 分。这个方法有很多的优点,第一,他只需要一个评分者就可完成。第二,他比 Guilford 所提出的给每一个样本按照出现频率赋值的方法更简单。第三,它有一个直截了当的解释——创造性的答案就是一个独特的唯一的答案。无独有偶,TTCT 也运用过类似的方法。他们将常见的标准的一些答案汇集在一起,成为一个答案库,将这个答案库之外的回答赋予一个分数(Torrance, 2008),然后将这些分数加起来作为独特性得分。其他研究中,研究者对样本中少于 5%的答案赋值 1 分,对所有其他答案赋值 0 分(Milgram & Milgram, 1976),然后把这些答案的值相加。尽管表面上存在差异,Wallach 和 Kogan 的独特性得分和托兰斯的独特性得分都采用了相同的心理测量模型:人们会因为统计上不常见的反应而得分,而这些分数是可以相加的。

然而这种评价方式存在三个严重的缺陷。第一,独特性(独特答案的个数)会与创造性思维的流畅性(所有答案的总个数)混淆(Clark & Mirels, 1970, Hocevar, 1979a, Hocevar, 1979b; Hocevar & Michael, 1979)。对 Wallach 和 Kogan 数据的潜变量再分析发现,潜在流畅性和独特性变量之间的相关性 $r = .89$ (Silvia, 2008)。在最新标准的 TTCT 测试中,独特性和流畅性同样高度相关 $r = 0.88$ (Torrance, 2008)。并且,随着样本量

的增加,流畅性和独特性之间的相关性也会增加,因此这种混淆并不能通过大样本的方法来缓解。第二,对独特性得分的界定和解释并不清楚(Silvia, 2008)。创造性的反应不仅仅是看他是不是唯一的,它们也必须适合于当前的任务(Sawyer, 2006)。那些古怪的、荒诞的、不合适的答案很难从答案库里面过滤出来,导致研究者在给 0 分或者 1 分的时候就容易将“创造性”和“随意性”相混淆。同时,那些平常的反应会从这种评分方法中溜走,从而降低创造性评分的可靠性。例如,“铺一条砖路”显然是砖的一种常规用途,但在小样本的创造性研究中,它可能就是独一无二的答案。第三,独特性的得分受样本量大小的影响(Silvia, 2008)。一个答案在小样本中比在大样本中更有可能被认为是独特的。样本库中的答案个数是被试数量的函数,随着样本容量的增加,两个人给出相同答案的概率也会增加。

3.2. 主观的评价方法

既然客观的评价方式存在这些缺陷,那么有什么方法能够克服这些缺陷呢?研究者们开始重新思考对创造性发散思维任务主观评价的价值。在最早的创造力研究中,Guilford 的研究团队就曾经采用评分员为他们的发散性思维任务打分。将训练有素的评分人员作为评价者来对被试的反应做评价,从而对创造性打分,这是被广泛使用的办法。为了评估创造力成分中的“聪明度”(Cleverness),Guilford 让被试完成情节标题生成任务(Plot Titles test),给被试呈现两个简短的故事,被试需要在规定时间之内写出尽可能多的合适的标题。然后由两名评分员在 1~5 点的范围内对标题进行打分(Christensen, Guilford, & Wilson, 1957),或由 3 名评分员在 0~6 点的范围内打分(Wilson et al., 1953)。对于创造力成分中“远距离程度”(Remoteness)的评价,评价者对被试完成的后果任务(Consequences Task)进行 1-3 点的远近程度评价。但是这类方法直到二十世纪 80 年代才被偶尔使用。

二十世纪 80 年代,Amabile (1982)提出了一种共识性评价技术(Consensual Assessment Technique, CAT)。这项技术不是建立在任何创造性理论之上的,它的主要思想是,对创造性的最佳评价者应该是相关领域公认的专家(Experts) (Amabile, 1982; Hennessey, Amabile, & Mueller, 2011)。专家们不需要经过指导或小组讨论,而是凭借他们自己的经验对什么是创造性做出独立的评判(Baer, 1994c; Baer & McKool, 2009; Kaufman et al., 2007)。评价结束之后,专家之间的内部一致性程度就会作为评价者之间的可靠性得分,达到要求的信度值之后,取每个专家对每个答案给出的平均分作为该被试的最后得分(Amabile, 1982; Baer, 2015; Barbot et al., 2011; Hennessey et al., 2011)。专家们往往表现出较高的内部一致性,一般在 0.72 到 0.96 的范围(Amabile, 1982; Baer, Kaufman, & Gentile, 2004; Brinkman, 1999; Kaufman, Gentile, & Baer, 2005; Kaufman, Lee, Baer, & Lee, 2007)。专家们的专业知识有助于达成共识,因而,随着评价者专业水平的提高,CAT 被证明越来越具有足够的有效性(Kaufman, Baer, & Cole, 2009; Kaufman, Baer, Cole, & Sexton, 2008)。

主观评分的方法具有一定优势,可以克服客观的独特性(Uniqueness)评分所面临的三个问题。首先,评分原则上不与流畅性混淆:因为评分者对每个答案的判断是独立的,产生大量的答案并不一定会提高一个人的创造性得分。其次,从客观的独特性评分方式中溜走的奇特、怪异以及常见的答案都能够被主观的评价者捕获。例如,“铺一条砖路”这种砖头的常见用法,评价者总是会给出较低的分数的。第三,主观评分的方法不受样本量大小的影响。创造性是由评分者设定的标准来评判的,而不是由反应的频率来评判的。因此,无论样本大小,评分者的标准是相同的(Silvia, 2008)。

4. 小结与展望

本研究通过对创造性发散思维的评价方法进行梳理,整理了对创造性发散思维评价方法主要的两种类型,客观评价与主观评价。这两种方法存在各自的优缺点,对创造性发散思维任务的评价方法仍然有

待于探索和优化。

第一，客观评价的方法虽然受到了批判，但是至今仍然被研究者广泛使用，原因在于，这种方法仅仅依赖被试自己的答案在统计上出现的频率这个客观的指标，不受评分者主观想法和评判标准的影响，并且评价起来更省时省力。

第二，主观评价的方法，尤其是 CAT 技术存在诸多优点，被研究者们广泛接受并沿用至今，但是其中也不乏一些问题。例如，如何挑选评价创造性的最佳人选？其次，评判者本身的人格差异和文化差异会不会对创造性评价产生影响？这些都是值得深入探究的问题。

第三，不管是客观评价还是主观评价，对发散思维任务的评价仅仅是对被试的结果给予一个分数，5 点或 7 点评分，焦点都只是被试在规定时间内所写出的结果，而没有考虑被试在完成创造性任务中的认知过程。在后面的研究当中，可以从发散思维任务的过程出发，去探究过程中的创造性表现，或许会是一种全新的评价方式。

参考文献

- 张庆林, 曹贵康(2004). *创造性心理学*. 北京: 高等教育出版社.
- 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>
- Ashton-James, C. E., & Chartrand, T. L. (2009). Social Cues for Creativity: The Impact of Behavioral Mimicry on Convergent and Divergent Thinking. *Journal of Experimental Social Psychology*, 45, 1036-1040. <https://doi.org/10.1016/j.jesp.2009.04.030>
- Baer, J. (1994). Performance Assessments of Creativity: Do They Have Long-Term Stability? *Roeper Review*, 7, 7-11. <https://doi.org/10.1080/02783199409553609>
- Baer, J. (2015). The Importance of Domain-Specific Expertise in Creativity. *Roeper Review*, 37, 165-178. <https://doi.org/10.1080/02783193.2015.1047480>
- Baer, J., & McKool, S. S. (2009). Assessing Creativity Using the Consensual Assessment Technique. In C. S. Schreiner (Ed.), *Handbook of Research on Assessment Technologies, Methods, and Applications in Higher Education* (pp. 65-77). Hershey, PA: IGI Global. <https://doi.org/10.4018/978-1-60566-667-9.ch004>
- Baer, J., Kaufman, J. C., & Gentile, C. A. (2004). Extension of the Consensual Assessment Technique to Nonparallel Creative Products. *Creativity Research Journal*, 16, 113-117. https://doi.org/10.1207/s15326934crj1601_11
- Barbot, B., Besançon, M., & Lubart, T. I. (2011). Assessing Creativity in the Classroom. *The Open Education Journal*, 4, 124-132. <https://doi.org/10.2174/1874920801104010058>
- Brinkman, D. (1999). Problem Finding, Creativity Style and the Musical Compositions of High School Students. *The Journal of Creative Behavior*, 33, 62-68. <https://doi.org/10.1002/j.2162-6057.1999.tb01038.x>
- Chamorro-Premuzic, T., & Reichenbacher, L. (2008). Effects of Personality and Threat of Evaluation on Divergent and Convergent Thinking. *Journal of Research in Personality*, 42, 1095-1101. <https://doi.org/10.1016/j.jrp.2007.12.007>
- Cheng, L., Hu, W., Jia, X., & Runco, M. A. (2016). The Different Role of Cognitive Inhibition in Early versus Late Creative Problem Finding. *Psychology of Aesthetics Creativity and the Arts*, 10, 32-41. <https://doi.org/10.1037/aca0000036>
- Christensen, P. R., Guilford Press, J. P., & Wilson, R. C. (1957). Relations of Creative Responses to Working Time and Instructions. *Journal of Experimental Psychology*, 53, 82-88.
- Clark, P. M., & Mirels, H. L. (1970). Fluency as a Pervasive Element in the Measurement of Creativity. *Journal of Educational Measurement*, 7, 83-86. <https://doi.org/10.1111/j.1745-3984.1970.tb00699.x>
- Colzato, L. S., Ozturk, A., & Hommel, B. (2012). Meditate to Create: The Impact of Focused-Attention and Open-Monitoring Training on Convergent and Divergent Thinking. *Frontiers in Psychology*, 3, 116. <https://doi.org/10.3389/fpsyg.2012.00116>
- Gilhooly, K. J., Fioratou, E., Anthony, S. H., & Wynn, V. (2007). Divergent Thinking: Strategies and Executive Involvement in Generating Novel Uses for Familiar Objects. *British Journal of Psychology*, 98, 611-625. <https://doi.org/10.1111/j.2044-8295.2007.tb00467.x>
- Guilford, J. P. (1950). Creativity. *American Psychologist*, 5, 444-454. <https://doi.org/10.1037/h0063487>
- Guilford, J. P. (1967). Creativity: Yesterday, Today and Tomorrow. *Journal of Creative Behavior*, 1, 3-14.

- <https://doi.org/10.1002/j.2162-6057.1967.tb00002.x>
- Hennessey, B. A., Amabile, T. M., & Mueller, J. S. (2011). Consensual Assessment. In M. A. Runco, & S. R. Pritzker (Eds.), *Encyclopedia of Creativity* (Vol. 1, 2nd ed., pp. 253-260). San Diego: Academic Press.
<https://doi.org/10.1016/B978-0-12-375038-9.00046-7>
- Hocevar, D. (1979a). A Comparison of Statistical Infrequency and Subjective Judgment as Criteria in the Measurement of Originality. *Journal of Personality Assessment*, 43, 297-299.
- Hocevar, D. (1979b). Ideational Fluency as a Confounding Factor in the Measurement of Originality. *Journal of Educational Psychology*, 71, 191-196. <https://doi.org/10.1037/0022-0663.71.2.191>
- Hocevar, D., & Michael, W. B. (1979). The Effects of Scoring Formulas on the Discriminant Validity of Tests of Divergent Thinking. *Educational and Psychological Measurement*, 39, 917-921. <https://doi.org/10.1177/001316447903900427>
- Howard-Jones, P. A., Blakemore, S. J., Samuel, E. A., Summers, I. R., & Claxton, G. (2005). Semantic Divergence and Creative Story Generation: An fMRI Investigation. *Brain Research Cognitive Brain Research*, 25, 240-250.
<https://doi.org/10.1016/j.cogbrainres.2005.05.013>
- Kaufman, A. J., Johnston, D. T., Farquhar, J., Masterson, A. L., Lyons, T. W., Bates, S. et al. (2007). Late Archean Biospheric Oxygenation and Atmospheric Evolution. *Science*, 317, 1900-1903. <https://doi.org/10.1126/science.1138700>
- Kaufman, J. C., & Sternberg, R. J. (2007). Resource Review: Creativity. *Change*, 39, 55-58.
<https://doi.org/10.3200/CHNG.39.4.55-C4>
- Kaufman, J. C., Baer, J., & 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>
- Kaufman, J. C., Baer, J., Cole, J. C., & 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>
- Kaufman, J. C., Gentile, C. A., & 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>
- Kaufman, J. C., Lee, J., Baer, J., & Lee, S. (2007). Captions, Consistency, Creativity, and the Consensual Assessment Technique: New Evidence of Reliability. *Thinking Skills and Creativity*, 2, 96-106. <https://doi.org/10.1016/j.tsc.2007.04.002>
- Milgram, R. M., & Milgram, N. A. (1976). Creative Thinking and Creative Performance in Israeli Students. *Journal of Educational Psychology*, 68, 255-259. <https://doi.org/10.1037/0022-0663.68.3.255>
- Plucker, J. A., Qian, M., & Wang, S. (2011). Is Originality in the Eye of the Beholder? Comparison of Scoring Techniques in the Assessment of Divergent Thinking. *Journal of Creative Behavior*, 45, 1-22.
<https://doi.org/10.1002/j.2162-6057.2011.tb01081.x>
- Radel, R., Davranche, K., Fournier, M., & Dietrich, A. (2015). The Role of (Dis)inhibition in Creativity: Decreased Inhibition Improves Idea Generation. *Cognition*, 134, 110-120. <https://doi.org/10.1016/j.cognition.2014.09.001>
- Runco, M. A., & Jaeger, G. J. (2012). The Standard Definition of Creativity. *Creativity Research Journal*, 24, 92-96.
<https://doi.org/10.1080/10400419.2012.650092>
- Sawyer, R. K. (2006). *Explaining Creativity: The Science of Human Innovation*. New York: Oxford University Press.
- Silvia, P. J. (2008). Creativity and Intelligence Revisited: A Latent Variable Analysis of Wallach and Kogan (1965). *Creativity Research Journal*, 20, 34-39. <https://doi.org/10.1080/10400410701841807>
- Torrance, E. P. (2008). *Torrance Tests of Creative Thinking: Norms-Technical Manual, Verbal Forms A and B*. Bensenville, IL: Scholastic Testing Service.
- Wallach, M. A., & Kogan, N. (1965). Modes of Thinking in Young Children: A Study of the Creativity-Intelligence Distinction. *Archives of General Psychiatry*, 3, 329-331.
- White, H. A., & Shah, P. (2016). Scope of Semantic Activation and Innovative Thinking in College Students with Adhd. *Creativity Research Journal*, 28, 275-282. <https://doi.org/10.1080/10400419.2016.1195655>
- Wilson, R. C., Guilford, J. P., & Christensen, P. R. (1953). The Measurement of Individual Differences in Originality. *Psychological Bulletin*, 50, 362-370. <https://doi.org/10.1037/h0060857>

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2160-7273，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：ap@hanspub.org