

《心理学报》审稿意见与作者回应

题目：创造性思维中语义搜索过程：基于答案累积时间函数和语义相似性的量化分析

作者：陈嫣然 梁正 赵庆柏 黄宇 李松清 于全磊 周冶金

第一轮

审稿人 1 意见：

创造性思维的产生过程主要包括想法生成和评价两个阶段，想法生成的关键取决于语义搜索的过程。本文通过对比新颖和寻常条件下发散思维测验在答案累计时间函数和语义相似性等量化参数上的差异，考察了创造性思维中的语义信息搜索过程。论文选题新颖，也有创新之处。在研究范式上，基于 Hass 的研究，结合创造性思维最核心的特点（新颖型+适宜性），设置对照组，改用了 AUT 任务考察创造性思维产生过程的语义搜索策略，更能揭示相关机制。在分析方法上，采用语义相似性的聚类分析对创造性思维中语义搜索的特征进行了详实的挖掘和分析。整体来看，研究范式新颖，方法科学，数据分析深入，研究结果可靠，证据比较充分，较好的支撑了研究结论，丰富和深化了对该主题的理解。以下问题和建议，供参考。

意见 1: 作者提到“其中个体在知识网络中的信息搜索过程尤为重要(Silvia, Beaty, & Nusbaum, 2013)，它为后续的新颖信息选择和整合奠定基础。”，“创造性思维的语义搜索更多的发生在发散性观点生成阶段。”等，建议作者应该先简要介绍创造性思维产生的主要阶段，在此基础上提出语义搜索在创造性思维产生过程中的重要作用，并明确语义搜索与“创造性思维”之间的关系。

回应: 感谢专家提出的宝贵建议。本文在第一段基于创造的生成-选择模型和一些相关实证研究结果，对创造过程进行了简单的描述，即创造是在广泛信息激活和搜索的基础上，选择新颖信息，形成新颖信息联结的过程。同时指出，语义信息的搜索是新颖信息产生的来源，在创造性思维中发挥关键作用。原文第二段直接提出了创造性思维的语义信息搜索更多发生在发散性观点生成阶段，确实少了关于发散思维与语义信息搜索的关系。因此在修改中的第二段起始增加了关于为何语义搜索主要在发散性思维阶段的简要论述。

意见 2: 从文本来看，引言第二段和第四段相关度较高，建议将两段内容整合，进一步完整阐述该理论模型是如何解释“创造性”语义搜索过程，并明确指明创造性思维语义搜索过程在哪些方面具有不同的量化特征。

回应: 感谢专家提出的宝贵建议。本文介绍发散性思维语义搜索过程的思路是，先介绍部分研究者强调的发散性思维的无意识加工特征，从而引出语义网络的激活扩散模型，并进一步介绍语义搜索的负加速和聚类等量化特征，然后指出其他研究者强调的执行功能在发散思维中的作用，从而提出发散性思维中语义搜索过程和自由回忆任务不同，并具体体现在负加速和聚类等特征参数的差异上。在原文的描述中，介绍负加速和聚类特征的第三段在和第二段的衔接上确实不够紧密，因此在修改稿增加了上下衔接句，同时也在第四段尾明确指出不同的量化特征就是负加速和聚类等相关指标。

我们考虑了专家的建议，尝试将第二段与第四段整合。不过那样的话，第三段关于量化特征的介绍就无处安放，毕竟涉及负加速和聚类特征现有的研究主要是关于自由回忆任务

的，与发散性思维的无意识加工特征关联更强。

意见 3: 引言第五段提到的个体语义记忆结构在文中略显突兀，建议将文中所列举的高创造力个体所具有的语义网络特征与研究主题相结合，并做到上下段连贯衔接。另外，随机漫步模型描述过于概括，为了便于读者理解，建议进一步具体解释随机漫步的发生过程和机制。

回应: 感谢专家的宝贵建议。根据专家的建议，我们重写了该段。在修改稿中，特别注意了这部分和上下文的关系，并对随机漫步模型进行了解释。

意见 4: 本文立足于“自由回忆任务”与“发散思维任务”的加工过程的差异，认为时间聚类方法并不适用于发散思维任务，据此提出了本文的改进和创新之处。但是，文中只提到了自由回忆任务的负加速特征，建议进一步通过比较概括说明为什么采用本研究的范式更能揭示创造性思维产生过程的语义搜索机制，从而形成更完整的证据链，并凸显本研究的创新之处。

回应: 感谢专家提出的宝贵建议。我们对论文的前言部分进行了较大的修改，以更好地展现本研究的思路和创新之处。论文增加了创造性联想理论和创造性执行理论。根据创造性联想理论，创造性思维的语义搜索过程是一种从近处到远处的激活扩散联想加工，该过程是自发的、不受控制的。这一过程符合语义网络的激活扩散模型，会表现出与自由回忆相似的负加速特征以及聚类特征。而根据创造性执行理论，创造性思维是一个自上而下的过程，个体需要认知控制以更为有效地搜索语义网络。正是由于抑制控制的参与，创造性思维的语义过程会表现出与自由联想不同的语义搜索过程，具体表现为答案生成的速度更慢，答案与题目的语义关联度更低，答案生成的聚类程度更低。在论文的上一版中，前言部分对语义相似性相关指标介绍较少，在修改稿中，我们增加相关的描述，并在对实验任务和范式介绍的基础上，明确提出了本研究的假设，以更好地阐释创造性思维产生的语义搜索机制。

意见 5: 关于“长时记忆的提取依然是依赖于语义网络中概念连接的远近和强度”，在创造力背景下这一观点的提出需要更多的理论支撑或实验证据。

回应: 感谢专家的宝贵意见。我们对此句话的阐述进行了修改以更准确的达意，同时加入了相关理论支撑。我们将此句修改为：“发散思维过程中的长时记忆提取可能依然受到语义网络中概念连接的远近和强度的影响”，并加入了两篇参考文献(Beaty et al., 2017; Volle, 2018)。在 Beaty 等人(2017)的研究中提到，记忆为创造性思维提供了基础，创造性思维在语义记忆网络中进行激活扩散，以形成新颖和有用的联想。在 Volle(2018)的文中提到，产生原创的观点依靠自发的从近处到远处的激活扩散的联想加工，可以通过激活不寻常的联想而产生新颖想法。除此之外，Olsson(2011)的顿悟重配理论也指出，创造性信息激活依赖于信息间的联结强度，反馈可以改变信息间的激活强度，以抑制寻常信息、激活新颖信息。

意见 6: 创造性语义搜索过程是一个较为复杂的认知过程，涉及到不同方面的量化特征，以“探索创造性思维的语义搜索过程”做为研究主题是有学术价值的。但在整体问题提出的思路上建议更加凸显创造性思维语义搜索过程的策略和特征。建议围绕研究目的，基于相关研究进展进行论证，尽可能提出研究假设。

回应: 感谢专家的宝贵建议。现已修改了论文的整体问题提出的思路，并提出了研究假设具体如下。

本研究的目的在于通过描绘发散性思维中答案生成在累积函数和语义相似性等一系列参数上的量化指标，揭示创造性思维语义搜索过程中答案生成的时间累积特征和搜寻语义场的空间特征。根据创造性联想理论，发散性思维以语义网络为基础，其语义搜索过程可能具有语义激活扩散模式的相关特点，累积答案个数的时间函数可能符合双曲线函数。但由于创

创造性观点的产生往往需要远距离语义的激活,创造性任务的语义搜索速度可能比寻常任务更慢。据此提出假设 1: 新颖 AUT 的语义信息搜索过程会出现负加速的时间特征,且搜索速度较寻常 AUT 更慢。同时,创造性执行理论认为,创造性思维的语义搜索过程需要认知控制等功能来抑制近距离的优势反应,而进行远距离搜索以激活新颖想法。据此提出假设 2: 新颖 AUT 中个体更多地搜索了距离题目更远的语义场,其答案与题目之间的语义相似性显著低于寻常 AUT。除了避免就近搜索,由于答案之间较高的相似性会降低个体对答案新颖性的评价(Acar & Runco, 2019),因此创造性思维中个体可能同样避免在某一语义场中寻找类似答案的局部搜索行为。据此提出假设 3: 新颖 AUT 条件中答案聚类程度显著低于寻常 AUT 条件。

意见 7: 在实验设计上,设置设计报动物名任务作为练习实验的目的是什么呢?其与正式实验的有何关联。另外,答案评价时两两配对过程和原则是怎样的呢?两种任务条件是否进行了平衡以排除顺序效应?

回应: 感谢专家提出的问题。实验设置报动物名是为了让被试熟悉实验程序的操作界面和流程,在论文的实验流程部分有做说明,在修改稿中的实验任务部分我们也增加了说明。

答案相似度评价过程是:在同一 AUT 任务中将所有答案进行两两配对,并按随机顺序呈现给被试进行相似度评价。相关表述已在修改稿的相应位置进行了添加。

实验流程中每个被试均是先完成寻常 AUT 任务,后完成新颖 AUT 任务,并未进行顺序的平衡。之所以这样操作,是由于本实验的自变量是通过任务要求来进行操纵,在寻常 AUT 中要求被试尽可能多的想出有用的用途,随着语义网络的扩散,以及被试自发的新颖性寻求动机的作用,被试依然会产生出新颖答案,如果在完成寻常 AUT 之前要求被试先完成新颖 AUT,那么新颖的任务要求可能会对后续的寻常 AUT 任务产生启动或者暗示,从而影响寻常 AUT 的语义信息搜索过程。当然,未进行顺序匹配,可能会产生一定的练习效应,不过从新颖 AUT 条件的平均解答时间显著长于寻常 AUT 条件这一结果可以推测,可能存在的练习效应并未对新颖 AUT 任务产生能影响最终结果的助益。相关的讨论已添加至修改稿讨论的不足部分。

意见 8: 讨论部分主要参考了 Hass (2017) 和 Kenett (2019) 的两篇文章对研究结果进行了解释和说明,在部分研究结论的推论上增加理论依据和文献支撑。如:“在新颖 AUT 中答案新颖性与答案个数并.....简单的由近及远的激活扩散过程。”“可以推测,在新颖和寻常 AUT 任务.....可能是新颖条件下答案个数少于寻常条件的原因。”“这说明,寻常 AUT 任务中就近搜索.....在远离题目的语义场内可能会进行局部就近搜索。”等。建议进一步在前言或讨论中指明答案生成时间、答案个数、答案间聚类、语义相似度等量化指标之间的关系能够反映出何种语义搜索的加工方式,并对前人相关研究进展做出系统的梳理和补充,为实验结论的提出提供更有力的证据支持。

回应: 感谢专家的宝贵建议。根据专家的建议,我们对论文的前言和讨论进行了较大的修改。之前的版本更多的是在描述和解释数据结果,以数据驱动为主,缺乏对前人研究系统的梳理。修改之后的版本在创造性思维语义搜索过程的理论推测和阐释方面增加了大量内容。前言部分在对先前研究进行梳理的基础上,明确提出了研究假设,更好地明确了数据指标与语义搜索方式的关系。讨论部分则是呼应前言的理论假设,以先理论阐释、后数据支持的方式展开讨论。

意见 9: 建议在讨论中围绕本研究的实验设计和研究过程进一步补充研究局限性及未来展望。

回应：感谢专家的宝贵建议。现已添加研究局限性和展望，具体如下：

本文存在以下不足和可以继续探究的问题：首先，实验流程中每个被试均是先完成寻常 AUT 任务，后完成新颖 AUT 任务，并未进行顺序的平衡。之所以这样操作，是由于本实验的自变量是通过任务要求来进行操纵，在寻常 AUT 中要求被试尽可能多的想出有用的用途，不过随着语义网络的扩散以及被试自发的新颖性寻求动机的作用，被试依然会产生出新颖答案，如果在完成寻常 AUT 之前要求被试先完成新颖 AUT，那么新颖的任务要求可能会对后续的寻常 AUT 任务产生启动或者暗示，从而影响寻常 AUT 的语义信息搜索过程。当然，未进行顺序匹配，可能会产生一定的练习效应，不过从新颖 AUT 条件的平均解答时间显著长于寻常 AUT 条件这一结果可以推测，可能存在的练习效应并未对新颖 AUT 任务产生能影响最终结果的实质性助益。其次，创造力依赖于联想(即自发激活语义网络)和控制(即策略性搜索语义网络)过程的之间的协调平衡(Taylor & Barbot, 2021)。在创造性思维中执行控制发挥了重要作用，抑制寻常的凸显信息，激活新颖的远距离信息。未来的研究可进一步探究执行功能在创造性思维的动态语义搜索过程中的具体作用，为相关理论提供量化实证支持。此外，Gilhooly 等人(2007)发现，被试在 AUT 任务中的除了会基于长期记忆中检索已知的用途，还可能使用了一些策略，如利用物体的属性(如形状、质地、重量)来构建用途。因此，未来研究可以进一步探究思维策略在创造性思维语义搜索过程中的作用。

意见 10：文章还存在一些写作规范和参考文献引用格式的问题，如补充被试的基本人口学信息（如性别、年龄等）；表 6 中需标明指明括弧里的值代表的含义；“2 个 AUT 任务”的表述易于与实验条件混淆。此外还有个别存在错别字的地方，请作者自查。如引言最后一句“本研究选择语义相似度作为答案生成聚类算法的基础，已解决时间聚类算法可能不适用于研究创造性思维语义搜索过程的问题。”应该为“以解决”。

回应：感谢专家的宝贵意见。现已添加被试基本人口学信息、修改表格，为防止“2 个 AUT 任务”的表述与实验条件混淆，将该句改为“被试在两种实验条件下需要各完成 2 个相应的 AUT 任务”。另外我们也修改了错别字和一些细节。

.....

审稿人 2 意见：

该论文使用发散性思维测验，对比被试在新颖性的多用途任务（AUT）和常规用途上的任务表现，记录了想法生成的数量和时间，想法之间的语义相似度，并通过非线性拟合和聚类分析来研究新颖性 AUT 的语义搜索特征，发现新颖性 AUT 的语义检索发生在较远距离的语义场内，而非由近及远的激活扩散过程，上述发现从新的更加细化的角度揭示了新颖想法产生的搜索过程，揭示了具有理论价值的实验现象。

意见 1：即使要求被试尽量生成高新颖的答案，在答案中也一定存在低新颖的部分，或者虽然高新颖，但是不怎么有效的部分，在这种情况下是怎么处理的？另外新颖 AUT 和常规 AUT 的对比是否真的分离出了新颖性特征？

回应：感谢专家提出的问题。正如专家所言即使要求被试生成新颖且实用的答案，在实验中被试也会生成低新颖的答案。这也体现了创造性思维语义激活过程的真实特点，即发散思维依赖于联想和执行控制过程之间的协同合作，自上而下的执行功能与自下而上的联想激活在创造性思维中均发挥着关键作用(Beaty, Silvia, et al., 2014; Benedek et al., 2017; Kenett, Beaty, et al., 2016)，低新颖的答案主要反映了自下而上的联想过程。因此，我们将所有答案都纳入到了分析中。

在寻常 AUT 中，随着语义网络的由近及远的扩散以及被试自发的新颖性寻求动机的作

用, 被试也会产生出新颖答案, 而在新颖 AUT 中, 被试也会产生出部分低新颖性答案, 因此新颖 AUT 与寻常 AUT 的比较并非 1 和 0 的质性比较, 二者是在新颖性上的量化差异。结果显示, 新颖 AUT 和寻常 AUT 两个条件间平均答案新颖性具有显著差异($t(38) = 9.08, p < 0.001, \text{Cohen's } d = 1.45$)。而且, 如前所述, 创造性思维的语义搜索也并不是单一的新颖答案产生过程, 还包含了低新颖信息的激活。因此, 我们认为, 采用新颖 AUT 和常规 AUT 的对比的方式, 可以较好地探究创造性思维中语义搜索过程。

意见 2: 在讨论部分, 基本上只聚焦于对目前研究结果的解读, 缺乏与本领域相关前人研究的关联性探讨以及可能的对未来研究的影响。

回应: 感谢专家提出的宝贵意见。本问题与第一位审稿专家的讨论部分问题 1 相似。之前的版本更多的是在描述和解释数据结果, 以数据驱动为主, 缺乏对前人研究系统的梳理。修改之后的版本在创造性思维语义搜索过程的理论推测和阐释方面增加了大量内容。前言部分在对先前研究进行梳理的基础上, 明确提出了研究假设。讨论部分则是呼应前言的理论假设, 以先理论阐释、后数据支持的方式展开讨论。

意见 3: 正如作者所指出的, 对 AUT 语义距离的衡量, 主流方法一般是使用潜在语义分析, 不知是否可以以此指标重新进行聚类, 看一看结果是否稳定?

回应: 感谢专家的宝贵意见。我们在数据分析时考虑过潜在语义分析(LSA)的方法。但是考虑到以下两方面并未选择该方法:

首先, 正如文中提到“在 LSA 中, 词汇与词汇之间的语义相似性计算是基于两个词汇在同一语境(如段落)下同时出现的概率, 这依赖于一个已经客观存在的海量语料库”。因此, LSA 模型的性能在很大程度上取决于所使用的文本语料库的选择和范围(Volle, 2018)。目前, 语义空间的更高级计算模型正在开发中(Mandera et al., 2017), 基于文本语料库分析估计语义距离的有效性尚未确定(Deyne et al., 2016)。而 AUT 任务中答案的精细化水平(即, 描述的细节程度或字数)也会影响基于 LSA 的语义距离度量(Forthmann et al., 2019)。因此该方法有一定的缺陷(Kenett, 2019)。此外, 特别需要指出的是, 据本研究课题组了解的文献, LSA 方法在创造性领域的应用基本都是在英语中。中文的 LSA 方法多用于信息检索和智能化推荐, 此处的语义相似性与心理学研究的语义相似性在概念上未必一致。在先前的另一个研究中, 我们曾同时采用 LSA 和被试评价两个方法评估两个词汇之间的语义相似性, 结果发现二者的相关性仅有 0.3~0.4。因此, 本研究并未尝试使用 LSA 方法。

其次, 对于语义距离的判断可能因个体而异(Yee, 2016)。研究发现, 与缺乏创造力的个体相比, 富有创造力的个体对不相关词对的语义距离估计更近, 低估倾向更明显(Guo et al., 2019; Volle, 2018)。由于 LSA 是一个客观评价语义距离的方法, 其并不适用于本研究探究个体一般的语义搜索过程, 因此使用主观语义距离判断可能更为合适。

Yee, E., Thompson-Schill, S.L. (2016). Putting concepts into context. *Psychon Bull Rev*, 23, 1015–1027.

Guo, J., Ge, Y., Pang, W. (2019). The underlying cognitive mechanisms of the rater effect in creativity assessment: The mediating role of perceived semantic distance. *Thinking Skills and Creativity*, 33, 100572.

Kenett, Y.N. (2019). What can quantitative measures of semantic distance tell us about creativity? *Current Opinion in Behavioral Sciences*, 27, 11–16.

Mandera, P., Keuleers, E., Brysbaert, M., 2017. Explaining human performance in psycholinguistic tasks with models of semantic similarity based on prediction and counting: A review and empirical validation. *Journal of Memory and Language*, 92, 57–78.

Forthmann, B., Oyebade, O., Ojo, A., Günther, F., Holling, H., 2019. Application of Latent Semantic Analysis to Divergent Thinking is Biased by Elaboration. *J Creat Behav*, 53, 559–575.

De Deyne, S., Verheyen, S., Storms, G., 2016. Structure and Organization of the Mental Lexicon: A Network Approach Derived from Syntactic Dependency Relations and Word Associations, in: Mehler, A., Lücking, A., Banisch, S., Blanchard, P., Job, B. (Eds.), Towards a Theoretical Framework for Analyzing Complex Linguistic Networks, Understanding Complex Systems. Springer, Berlin, Heidelberg, pp. 47–79.

Volle, E., 2018. Associative and Controlled Cognition in Divergent Thinking: Theoretical, Experimental, Neuroimaging Evidence, and New Directions, in: Jung, R.E., Vartanian, O. (Eds.), The Cambridge Handbook of the Neuroscience of Creativity. Cambridge University Press, pp. 333–360.

意见 4: 有一些不符合格式的地方, 例如 p, 小于号, 数字之间需要空格等。表 1 的格式有些问题, 似乎需要重新调整单元格。

回应: 感谢专家提出的问题与建议。现已对全文进行自查, 修改了格式问题。

意见 5: 聚类标准的选取, 即 $d \leq 0.5$ 是否影响实验结论的稳定性? 希望在尾注部分进一步说明不同标准对实验结果的影响。或者直接以不同标准再做分析。

回应: 感谢专家提出的问题与建议。文中选取 $d \leq 0.5$ 作为聚类标准的主要考虑两个方面: 首先, 我们计算了聚类标准为 $d \leq 0.25$ 和 $d \leq 0.75$ 时的结果, 两者结果与 $d \leq 0.5$ 基本一致, 因此我们认为实验结论是稳定的。考虑到 $d \leq 0.25$ 时的聚类标准过高, 去除无法聚类答案的 19 名被试的数据后仅剩 20 人数据, 而当采用宽松的聚类标准 $d \leq 0.75$ 时, 所得结果与 $d \leq 0.5$ 时一致, 因此我们在文中仅呈现了较为适中的 $d \leq 0.5$ 时的计算结果。针对专家提出的建议, 我们在论文中加以说明。

以下是对 $d \leq 0.25$ 、 $d \leq 0.75$ 时的计算结果的呈现:

$d \leq 0.25$:

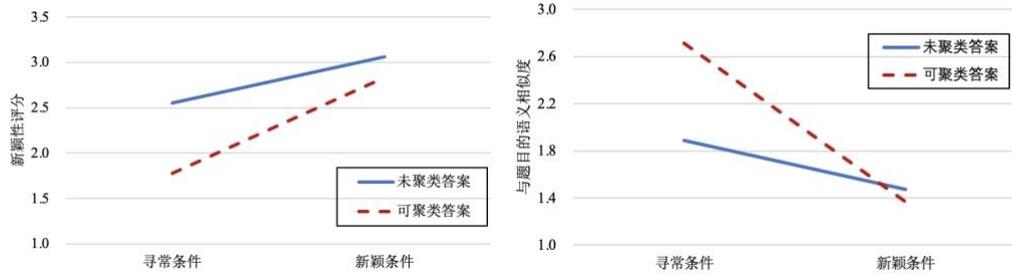
当选择聚类标准为 $d \leq 0.25$ 时, 寻常条件下有 4 名无聚类答案数据的被试, 新颖条件下有 17 名无聚类答案数据的被试, 需删除 19 名被试数据, 仅剩 20 人数据。

以实验条件(寻常条件 vs. 新颖条件)和聚类情况(未聚类 vs. 可聚类)为自变量, 被试平均答案新颖性以及答案与题目的平均语义相似度为因变量, 分别进行 2×2 的重复测量方差分析。方差分析结果显示, 实验条件和聚类情况对新颖性的主效应均显著(实验条件: $F(1, 19) = 19.00, p < 0.001, \text{partial } \eta^2 = 0.50$; 聚类情况: $F(1, 19) = 11.84, p = 0.003, \text{partial } \eta^2 = 0.38$), 两者交互作用边缘显著($F(1, 19) = 4.12, p = 0.057, \text{partial } \eta^2 = 0.18$)。简单效应分析结果显示, 在寻常条件下, 可聚类答案的新颖性评分显著低于未聚类答案($F(1, 19) = 39.10, p < 0.001, \text{partial } \eta^2 = 0.67$); 而在新颖条件下, 可聚类答案与未聚类答案的新颖性评分不存在显著差异($p > 0.05$)。

实验条件和聚类情况对答案与题目的语义相似度主效应均显著(实验条件: $F(1, 19) = 53.71, p < 0.001, \text{partial } \eta^2 = 0.74$; 聚类情况: $F(1, 19) = 12.45, p = 0.002, \text{partial } \eta^2 = 0.40$), 两者交互作用显著($F(1, 19) = 11.92, p = 0.003, \text{partial } \eta^2 = 0.39$)。简单效应分析结果显示, 在寻常条件下, 可聚类答案与题目的语义相似度显著高于未聚类答案($F(1, 19) = 17.02, p < 0.001, \text{partial } \eta^2 = 0.47$); 而在新颖条件下, 可聚类答案和未聚类答案在与题目的语义相似程度上不存在显著差异($p > 0.05$)。

表 1 不同实验条件下可否聚类答案平均新颖性评分

	聚类情况	被试个数	平均新颖性评分 <i>M</i> (<i>SD</i>)	平均各答案与题目的语义相似 度 <i>M</i> (<i>SD</i>)
寻常条件	未聚类	20	2.55 (0.58)	1.89 (0.63)
	可聚类	20	1.78 (0.58)	2.71 (0.54)
新颖条件	未聚类	20	3.06 (0.53)	1.47 (0.31)
	可聚类	20	2.83 (1.10)	1.37 (0.56)



不同实验条件和聚类情况下答案的新颖性评分(左图)以及题目与答案的语义相似度(右图)

$d \leq 0.75$:

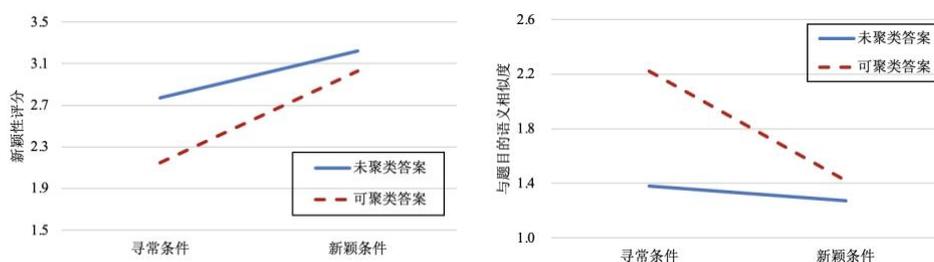
当选择聚类标准为 $d \leq 0.75$ 时，寻常条件下有 2 名无未聚类答案数据的被试，新颖条件下有 1 名无未聚类答案数据的被试，有 4 名无可聚类答案数据的被试，需删除 7 名被试数据，剩余 32 人数据参与后续分析。

以实验条件(寻常条件 vs.新颖条件)和聚类情况(未聚类 vs.可聚类)为自变量，被试平均答案新颖性以及答案与题目的平均语义相似度为因变量，分别进行 $2 * 2$ 的重复测量方差分析。方差分析结果显示，实验条件和聚类情况对新颖性的主效应均显著(实验条件: $F(1, 31) = 37.85, p < 0.001, \text{partial } \eta^2 = 0.55$; 聚类情况: $F(1, 31) = 12.01, p < 0.001, \text{partial } \eta^2 = 0.28$)，两者交互作用显著($F(1, 31) = 4.54, p = 0.041, \text{partial } \eta^2 = 0.13$)。简单效应分析结果显示，在寻常条件下，可聚类答案的新颖性评分显著低于未聚类答案($F(1, 31) = 16.89, p < 0.001, \text{partial } \eta^2 = 0.35$)；而在新颖条件下，可聚类答案与未聚类答案的新颖性评分不存在显著差异($p > 0.05$)。

实验条件和聚类情况对答案与题目的语义相似度主效应均显著(实验条件: $F(1, 31) = 46.36, p < 0.001, \text{partial } \eta^2 = 0.60$; 聚类情况: $F(1, 31) = 50.79, p < 0.001, \text{partial } \eta^2 = 0.62$)，两者交互作用显著($F(1, 31) = 32.48, p < 0.001, \text{partial } \eta^2 = 0.51$)。简单效应分析结果显示，在寻常条件下，可聚类答案与题目的语义相似度显著高于未聚类答案($F(1, 31) = 66.28, p < 0.001, \text{partial } \eta^2 = 0.68$)；而在新颖条件下，可聚类答案和未聚类答案在与题目的语义相似程度上不存在显著差异($p > 0.05$)。

表 2 不同实验条件下可否聚类答案平均新颖性评分

	聚类情况	被试个数	平均新颖性评分	平均各答案与题目的语义相似
			M (SD)	度 M (SD)
寻常条件	未聚类	31	2.77 (1.05)	1.38 (0.46)
	可聚类	31	2.15 (0.50)	2.22 (0.43)
新颖条件	未聚类	31	3.22 (0.69)	1.27 (0.30)
	可聚类	31	3.04 (0.65)	1.42 (0.44)



不同实验条件和聚类情况下答案的新颖性评分(左图)以及题目与答案的语义相似度(右图)

意见 6: 语义关系计算结果这一部分, 所谓各位置上的答案究竟是什么意思需要在先前就明确一下, 以便理解。

回应: 感谢专家的宝贵建议。现已在文中方法部分添加说明“根据各被试答案生成的先后顺序进行排序, 将序号作为每条答案的位置信息。”

第二轮

审稿人 1 意见:

作者详细地回复了我所关心的问题。修改后的论文在理论高度、结果解释和创新性的论证等方面具有大的提升。建议在文字的表达方面再进一步凝练, 个别语句做到简洁精炼。

回应: 感谢专家提出的宝贵建议。现已通读全文, 更改了部分文字表达, 使其更为简洁凝练。请见文中蓝色字。

审稿人 2 意见: 建议修后发表。

意见 1: 建议修改一下摘要, 使之更加简介清晰。在目前摘要的表述中, 在“实验观察”与“该观察带来的理论启示”(说明... ..)之间逻辑上跨越度大, 仅看摘要较难理解: “(1)新颖 AUT 条件中, 语义信息搜索呈现负加速特征, 且搜索速度较寻常 AUT 条件更慢。但答案的新颖性与答案个数不存在显著相关”——“说明创造性思维的语义搜索过程并非简单的由近及远的激活扩散过程”; “(2)新颖 AUT 的序列答案与题目均具有较低的语义相似性, 且显著小于寻常 AUT 条件”——“说明新颖性要求使个体在最初搜索时便开始摆脱题目的语义限制, 避免了就近搜索”; “(3)新颖条件比寻常条件表现出显著更低的答案间聚类, 其中可聚类答案与未聚类答案与题目的语义相似度均较低, 且不存在显著差异, 二者在新颖性上也不存在显著差异”——“这说明, 创造性思维的语义搜索过程会避免在题目附近搜索, 但是在远离题目的语义场内可能会进行局部搜索”。

回应: 感谢专家提出的宝贵意见。在原摘要写作中, 由于缺乏对研究范式和指标的了解, 在结果后直接阐述结论可能的确使得阅读者感到逻辑上跨度较大且难以理解。鉴于此, 我们调整了摘要结构, 将研究结果进行了汇总介绍, 使得阅读者先全面的了解实验结果。后再将整个研究所得结论引出, 使之更为清晰。同时, 我们修改了部分措辞, 使阅读者能够更好地理解研究结论。具体如下:

通过描绘发散性思维测验(物品多用途, AUT)中答案生成在累积函数和语义相似性等一系列参数上的量化特征, 揭示创造性思维的语义搜索过程。结果发现: (1)新颖 AUT 条件中, 语义搜索呈现与自由联想类似的负加速特点, 但搜索速度较寻常 AUT 条件更慢。(2)新颖 AUT 条件中所生成的答案与题目(即物品)均具有较低的语义相似性, 且显著小于寻常 AUT 条件。(3)新颖 AUT 条件中生成的答案比寻常 AUT 条件表现出显著更低的聚类程度, 其中可聚类答案与未聚类答案与题目的语义相似度均较低, 且不存在显著差异, 二者在新颖性上也不存在显著差异。以上结果说明了创造性思维的语义信息搜索过程具有与自由联想类似的激活扩散特征, 但总体搜索速度较慢。新颖性要求使得个体在最初搜索时便开始摆脱题目的语义限制而进行远距离搜索(避免就近搜索), 并倾向于在每个语义场中只生成一个答案(避免局部搜索), 但也可能会在远离题目的语义场中生成多个同类别答案。

意见 2: 在引言中似乎也存在表达上的问题, 不能开门见山地提出问题。建议考虑采用言简

意赅地引出关键问题的叙述风格，而不是“从头娓娓道来”，这使读者可能有难以把握研究关键点的困惑。

回应：感谢专家提出的宝贵建议。为了开门见山地引出关键问题，我们在首段末尾对研究问题的进行了简要阐述，点明研究的关键点是根据以往相关理论，通过量化分析的方法探究创造性思维中语义信息搜索的动态特征。具体如下：

创造是人们产生新颖、适用且有社会价值的产品或想法的过程。在创造性思维过程中，个体会根据问题情境广泛地激活语义概念，组合新颖信息，形成新颖联结，从而创造性地解决问题(Kenett, Gold, et al., 2018; 赵庆柏 等, 2015)。创造性思维的关键在于生成并选择新颖信息(Bink & Marsh, 2000)，其中个体在知识网络中的信息搜索过程尤为重要(Silvia et al., 2013)，它为后续的新颖信息选择和整合奠定基础。因此，本研究在对现有创造性思维中语义信息搜索的相关理论进行梳理分析的基础上，借鉴语义网络激活的量化分析方法，探究创造性思维中语义信息搜索的动态特征。