

《心理学报》审稿意见与作者回应

题目：部分呈现条件下类别分类学习中的信息表征机制

作者：王瑞明，林哲婷，刘志雅

第一轮

审稿人 1 意见：

意见 1：总体而言，论文选题有理论意义，在问题提出和研究思路上的确有一定的创新。论文行文流畅，文献综述较全面，较好地反映了相关研究进展。实验过程规范，变量操纵与控制较严格。数据收集和分析方法等基本正确，可靠，能够围绕研究目标进行系统的检验。但是，论文还存在诸多问题值得商榷及需要改进。论文最大的理论创新在于反驳部分呈现条件下表征非诊断性信息的原因是注意广度的变化所致，而归因为“被试进行了自动化的推理学习”，以此解释上述加工过程中的信息表征机制。但是综合实验 1 与实验 2 的结果看，验证了部分呈现条件下的分类学习能够同时表征诊断性信息和非诊断性信息，其中实验 2 主要证明了表征非诊断性信息不受注意广度的影响，而缺少存在自动化推理过程的直接证据。论文结论若成立，尤其需要回答的是，为什么这一推理过程是自动加工而非控制加工？

回应：感谢审稿人指出了本文在“自动推理加工”表述上的不够严谨，文中改用“推理加工”，并增加了更为慎重的词语，如“可能”来表述结论。实质上实验 2 只证明了表征非诊断性信息不受注意广度的影响，确实缺少存在自动化推理过程的直接证据，起初采用这种表述，是为了较清楚地阐述本研究的思路和假设的来源。为此，按照审稿人的意见在讨论部分分析了该处的不足，补充了“尽管实验二证明了表征非诊断性信息不受注意广度的影响，但仍然缺少证明被试确实进行了推理加工的直接证据，特别是被试如何自动地进行推理加工，这个问题还有待进一步的深入研究。”

意见 2：在缺少进一步深入证据的前提下，论文分析“而部分呈现条件下的分类学习则类似于推理学习的模式，是一种自上而下的演绎加工方式”显得主观。首先，推理本身可以分为归纳与演绎。部分呈现条件下的分类学习不排除有自上而下的加工可能，但是实验并不能肯定只存在演绎而无归纳！部分呈现条件下，完全可能既有自上而下的加工（缺失特征），又有自下而上的加工（已知特征）。因此，基于现有证据对信息表征机制的揭示还不清晰，有待深入。

回应：这部分的讨论分析确实过于主观，还是在表述上不够慎重。考虑到分析的逻辑上还是清楚的，改用了更加慎重的语言来表达我们的推测。具体为“结合前人的研究，我们推测认为，全部呈现条件下的分类学习可能是一种自下而上的归纳加工方式，对个别性的特征信息进行加工，进而学得一般性的类别知识；而部分呈现条件下的分类学习则类似于推理学习的模式，可能是一种自上而下的演绎加工方式，先对一般性信息和若干个别性的特征信息进行加工，进一步推导某些个别性的特征信息。”

意见 3：论文在讨论部分提出“如果类别学习的学习结果是原型，不同典型性类别结构的学习效率应该存在显著的差异”，从而否定了原型理论的假设。而以往研究已经表明，学习效

率不仅表现在次数上，还可以考虑时间。遗憾的是，实验并未报告学习的反应时间在各条件下的差异。多一项观测变量，有助于多角度揭示问题。

回应：审稿人提出了一个好的思路。当前，在探索类别学习效率上，几乎都采用了“达标的学习轮次”这个次数指标，而很少采用反应时。这可能是有一些被试，反应很快，但学很久才学会，这部分被试应该归为效率低的被试，如果采用反应时，这部分被试就归为效率高的被试了。是否可以考虑两个指标结合，次数少且时间短，这确实是一个更精细的效率指标，还可能有利于分析不同的学习结果。鉴于学习效率不是本研究探索的主要问题，因此没有收集这方面的数据。

意见 4：文中还存在其他小问题。例如，每种实验条件下的被试数量，是否有数据剔除及其依据等交代不够清楚。多处英文拼写错误，“trial”应该是“trial”。实验 1 最后一段表述错误，“在部分呈现条件下的分类学习中，被试也能学习到诊断性信息，也就是说，特征的缺失使被试关注到了诊断性特征信息”文字中应该是非诊断性信息。文中统计分析时先 t 检验比较基线，在进行单因素三水平的 F 检验，是否可以考虑 2×3 的 F 检验？等等。

回应：按照审稿人的意见介绍了每种实验条件下的被试数量，改正了审稿人指出的写作错误，并再次对全文进行了错别字的检查和修正。我们认为也可以考虑进行 2×3 的 F 检验，两种检验的结果是一致的，“呈现方式”这个变量差异均不显著。但考虑到如果选择报告 2×3 的 F 检验，另外一个变量的主效是显著的，可能还需要报告交互作用和简单效应检验，而这些检验数据和我们主要想揭示的问题没有紧密的关系。此外，Taylor 和 Ross（2009）研究也采用这样的统计方法，故而选择报告单因素三水平的 F 检验，与设计单因素设计保持一致。

审稿人 2 意见：

意见 1：本研究采用全部与部分呈现物体特征的方法考察了类别学习过程中对不同类型的信息的加工。研究采用先分类后典型性（相似性）评定的实验步骤，发现全部呈现条件下的分类学习只能表征诊断性信息，而部分呈现条件下的分类学习能够表征两种信息。另外，作者在实验二中操纵了实验任务的记忆广度，发现对类别学习没有显著影响。据此，作者推论部分呈现条件下表征非诊断性信息的原因在于被试进行了自动化的推理学习，而非注意广度的变化。总体来看，研究有一定的理论与技术创新，研究方法较恰当，结论可靠，但存在如下问题需要改进。整个文章易读性需要大大提高，我看了几遍也不能完全读懂作者分析实验数据的过程，也不清楚作者如何通过数据来得出自己的推论。下面的一些意见会进一步提到这些问题。

回应：按照审稿人的意见，增加了数据的逻辑分析，并对部分文字进行了修改。具体措施包括：第一，突出了数据分析的目的和作用；第二，调整了部分数据次序，突出了重点的数据；第三，增加了两个实验间的逻辑分析；第四，通读全文，修葺了部分不当的用词。

意见 2：作者的研究是在 Taylor 和 Ross（2009）基础上的推进，建议作者在前言中对这一研究的介绍与分析再详细一些。另外，在第 4 页最后一段中作者提到：Taylor 和 Ross（2009）根据第二种假设，改变了实验条件的研究范式，采用呈现缺失特征的样例进行分类学习任务，结果表明...，审稿人不明白什么是“改变了实验条件的研究范式”。作者在第 5 页第一段最

后一句提到: ...而全部呈现条件下的被试要注意 6 个特征, 注意的广度变大, 而被试能注意到的却有限, 只能将注意资源放在了诊断性特征上, 导致了实验结果的不一致。建议作者说明具体是什么结果不一致。

回应: 本研究是在 Taylor 和 Ross (2009) 基础上的推进, 实验一基本属于重复实验, 唯一不同的是学习后的迁移材料增加了 14 个低典型的评定材料 (见表 1)。按照审稿人意见介绍了 Taylor 和 Ross (2009) 的研究, 并且修整了部分不清楚的表述, 具体内容如下: 而 Taylor 和 Ross (2009) 认为需要对类别学习做一个更全面的理解, 为此, 他们针对分类学习不能学到非诊断性信息提出两种假设。一是非诊断性特征不能促进分类学习; 二是分类学习中也能够学习到非诊断性特征, 只是实验室的不自然的条件难以观察到这个结果。Taylor 和 Ross (2009) 根据第二种假设, 设计了新的类别学习任务, 用缺失特征的样例进行分类学习, 并指出这种任务更接近自然的条件, 绝大部分的事物都存在一些被遮挡了的特征。结果表明, 部分呈现条件下的分类学习也能表征典型性信息, 即分类学习任务中仍然可以学到非诊断性信息。作者用注意广度来解释这种现象, 认为由于注意的资源有限, 部分呈现条件下分类学习的被试需要注意的特征较少, 此时就有更多的注意资源分配到非诊断性特征上, 而在全部呈现条件下, 被试要注意更多的特征, 注意的广度变大, 而被试的注意资源有限, 只能将注意资源放在了作用性更大的诊断性特征上。

意见 3: 在前言最后一段, 作者描述了本研究的两个实验的目的及假设。实验一似乎是一个重复实验, 要验证 Taylor 和 Ross (2009) 的研究结果。实验二是作者的创新点, 作者试图用实验证明 Taylor 和 Ross (2009) 的解释是不正确的, 但作者并没有在前言中十分清楚地说明如何来用实验证明自己的假设。

回应: 按照审稿人意见说明研究的意图, 在实验一结束段和实验二的开始段进行了修改, 内容如下: 为何缺失条件下分类学习能够加工非诊断信息, 而且学习效率还更高, Taylor 和 Ross (2009) 研究的解释是因为注意广度的影响, 相对于全部呈现条件下, 在缺失条件下有更充分的注意资源去加工诊断性信息外的其它信息。然而, 如果被试在缺失条件下对缺失特征进行了推理, 也能够解释为何缺失条件下分类学习能够加工非诊断性信息, 而且缺失的特征越多, 加工的程度越深, 这样还进一步解释了为何缺失特征越多, 学习效率反而越好。实验二进一步探索缺失条件下分类学习者对非诊断性信息的加工, 究竟是注意广地还是推理学习造成的。

实验一采用了全部呈现、缺失 1 个特征和缺失 2 个特征这三个水平设计, 这三个条件的注意广度不同, 实验二改变为全部呈现 7 个特征, 全部呈现 5 个特征和部分呈现 5 个特征的三水平设计, 前两个水平间的注意广度不同但推理情况相同, 而后两个水平间的注意广度相同但推理情况不同。如果前者对非诊断性信息加工出现差异显著, 则说明是注意广度的原因; 如果后者对非诊断性信息加工差异显著, 则说明是推理加工的原因。

意见 4: 实验一, 第一段第二行与第三行描述的内容似乎是矛盾的。缺失特征数越多, 学习效果是更好还是更差?

回应: 改正了这部分不太准确的表述, 并把这部分内容放在前言的末尾, 改为: Taylor 和 Ross (2009) 的研究还表明, 在典型性一致的情况下, 部分呈现条件和全部呈现条件达到学习目标所需要的 trials 数目在统计上差异不显著 (74 和 77), 但部分呈现条件还是低于全部呈现条件, 这可能说明, 分类学习任务中缺失特征数越多, 学习效率越高。两个实验继续探讨这个

问题。

意见 5: 实验一是被试间设计，每种条件分配多少被试？实验条件的安排顺序又是如何的？

回应: 按照审稿人意见完善了这部分内容。

意见 6: 在进行典型性评定时，具体是如何评定的，即每次呈现给被试的刺激是什么？要求被试将刺激与哪一个原型进行比较？

回应: 按照审稿人意见补充了典型性评定的说明，内容为：接下来是典型性评定阶段，在屏幕上逐个呈现表 1 中的完整样例图片和类别的标签（Deeger 或 Koozle），要求被试对该图片在这个类别中的典型性进行评定。按 1—7 数字键反应，数字越大代表典型性越高。

意见 7: 对典型性评定的分析很难理解，表二中的 4 3 2 表示什么意思。

回应: 修改了典型性评定的分析这部分的表述，并在各表中增加了脚注，说明 4, 3, 2 表示诊断性特征的数目。

意见 8: 第 8 页，对实验一结果讨论时应该与前人的研究比较。

回应: 按照审稿人的意见进行了比较。

意见 9: 第 9 页第一段第一行，作者指出：...即在部分呈现 4 个特征时，对另外 2 个特征的推理加工程度比呈现 5 个特征条件下推理 1 个特征的加工程度更深，因而分类学习的效果掌握得更加牢固...。作者提到了推理加工，但审稿人很难看出有何推理的认知过程。

回应: 按照审稿人意见进行了修改，见文中实验一末尾两段，还增加了部分数据支持。

意见 10: 作者在第 12 页讨论这推理过程时，似乎并不能从本研究中找到数据支持。

回应: 实质上实验 2 只是证明了表征非诊断性信息不受注意广度的影响，究竟是什么影响，我们推测是推理加工的影响，因为在部分呈现特征下，即使不要求被试去推理缺失了的特征，被试也可能会整体性地去把握该事物。诚然，原文指出的是一种自动化推理过程，但还缺乏自动化推理过程的直接证据，起初采用这种表述，是为了较清楚地阐述本研究的思路和假设的来源。为此，按照审稿人的意见在讨论部分分析了该处的不足，补充了“尽管实验二证明了表征非诊断性信息不受注意广度的影响，但仍然缺少证明被试确实进行了推理加工的直接证据，特别是被试如何自动地进行推理加工，这个问题还有待进一步的深入研究。”

第二轮

审稿人 1 意见:

意见 1: 没有详细地描述实验任务，如什么叫部分呈现？

回应: 对这个实验任务进行了补充描述，具体见红字部分。“采用了三维立体的刺激作为材

料，由于在两维的显示屏上呈现，必然使到刺激的背面的特征被遮挡，并通过这种方法来操纵缺失特征来进行分类学习”。

意见 2：不能有“一定程度”这种模糊性的描述。

回应：删除了这个不严谨的用词。

意见 3：需要说具体一些，即为何前两水平间是注意广度不同但推理相同？

回应：按照审稿人的意见进行了修改，具体为：“学习阶段的学习材料是线性分离的家族相似性类别结构（linearly separable category, Yamauchi & Markman, 2000），这种类别结构每个样例都和所属的原型有一个特征不同。实验一采用了全部呈现、缺失 1 个特征和缺失 2 个特征这三个水平设计，因此这三个学习条件的被试分别面对了 6 个特征，5 个特征和 4 个特征的刺激进行分类学习，所以三个条件的注意广度是不同的；实验二改变为全部呈现 7 个特征，全部呈现 5 个特征和部分呈现 5 个特征的三水平设计，前两个水平间的注意广度不同（7 个特征和 5 个特征）但推理情况相同（都是全部呈现），而后两个水平间的注意广度相同（都是 5 个特征）但推理情况不同（全部呈现和部分呈现）。如果前者对非诊断性信息加工出现差异显著，则说明是注意广度的原因；如果后者对非诊断性信息加工差异显著，则说明是推理加工的原因。”

意见 4：、讨论 4.1 的三个批注意见。第一，这一点并不是本研究的创新？第二、具体是什么条件？第三、两个实验都只是发现了数据上的细微差异，没有达到统计上的显著水平？

回应：按照审稿人意见逐一进行了修改，讨论 4.1 的内容如下：

4.1 部分呈现条件下的分类学习者能够学习非诊断性信息。以往研究认为，在类别学习任务中，分类学习者只对诊断性信息敏感，而推理学习者对于典型性信息（包括诊断性和非诊断性信息）更加敏感。在这些研究中，实验材料往往是完整呈现的，然而，在日常生活中，我们对物体进行分类和预测时常常不能了解到该物体的全部特征，如看到一只狗，我们知道它有四条腿，两只眼睛，会跑会跳，但是我们并不清楚它是公的还是母的，它是什么品种等，不过，这并不影响我们对它进行分类。

我们基本沿用了 Taylor 和 Ross (2009) 所用的研究范式，采用缺失特征的材料进行实验，并对材料进行了改进，验证了部分呈现条件下的类别分类学习的学习者也是能够学习到非诊断性信息的。所以，不仅在推理学习中可以学习到非诊断性信息，在部分呈现条件下的分类学习中也可以学习到非诊断性信息。

意见 5：讨论 4.2 的批注意见，补充文献？

回应：补充了疏漏的文献，具体如下：结合前人研究结果，特别是分类和推理两种学习模式比较研究（刘志雅和莫雷，2006；2009），我们推测认为，全部呈现条件下的分类学习可能是一种自下而上的归纳加工方式，对个别性的特征信息进行加工，进而习得一般性的类别知识；而部分呈现条件下的分类学习则类似于推理学习的模式，可能是一种自上而下的演绎加工方式，先对一般性信息和若干个别性的特征信息进行加工，进一步推导某些个别性的特征信息。

意见 6：这段讨论是否与本研究的主要目的有关联？

回应: 这部分内容虽然和主要的研究目的关联度不大,但我们觉得还是要把研究的结果和当前主要的理论模型(原型理论)进行比较,所以还是保留了这部分内容。

意见 7: 为何不回顾前言中的问题与假设来讨论?

回应: 按照审稿人意见结合了问题与假设来讨论,具体的内容:

4.4 部分呈现条件下的分类学习者进行了推理加工

本研究结果表明,全部呈现条件和部分呈现条件下的分类学习者对于非诊断性信息的敏感程度不同,部分呈现条件下的分类学习者对非诊断信息的变化敏感,这种信息表征上的差异究竟是来自心理机制本身,还是决定于外在的学习形式,或者两者同时起作用,这是我们比较关注的问题。

实验二采用了全部呈现7个特征,全部呈现5个特征和部分呈现5个特征的三水平设计,前两个水平间的注意广度不同(7个特征和5个特征)但推理情况相同(都是全部呈现),而后两个水平间的注意广度相同(都是5个特征)但推理情况不同(全部呈现和部分呈现)。如果前者对非诊断性信息加工出现差异显著,则说明是注意广度的原因;如果后者对非诊断性信息加工差异显著,则说明是推理加工的原因。结果支持了后者的假设。因此可以认为,部分呈现条件下的分类学习者进行了推理加工,由于推理学习能够表征非诊断性信息,因此,在部分呈现条件下的分类学习者能够学习到非诊断性信息。部分呈现条件下,学习者不可能只关注一个或两个维度上的特征就能够顺利进行分类,因为学习过程中出现的样例是不完整的,学习者在关注其他特征信息的同时可能对缺失的特征进行了推理加工。

本研究通过2个实验证明了部分呈现条件下的分类学习中可以表征非诊断性信息,并且这种非诊断性信息的表征不是注意广度造成的,而是引发被试进行了推理学习造成的。当然,部分呈现条件下分类学习中被试能够表征非诊断性信息是被试进行了推理学习的观点还需要进一步的实验证据,特别是这种推理学习是自动化进行的还是策略性发生的,还需要在以后的研究中进一步探讨。

审稿人 2 意见:

意见 1: 作者根据所提的意见进行了针对性的修改,论文的可读性、逻辑性有显著的提升。从目前所完成的两个实验看,达到了验证前人研究,批判其注意广度起作用观点的目的,遗憾的是,作者否定了原有的观点,但在树立自身有关部分学习过程中发生了(自动化)推理的观点方面做得还不到位,没有证据支持该推断。这样就无法“有破有立”,也大大降低了论文的理论创新与贡献。要让被试“进行了推理学习”的观点成立,建议在现有基础上补充实验。例如,考察归纳推理强或弱的被试在典型性评定任务中正确率和反应时有无差异,由此推断推理能力是否在此过程起作用,间接证明发生了推理学习等等。当然,如果作者能够直接证明推理的发生更好,甚至还可以考虑这种推理是渐进的还是顿悟的。这样有利于更深入地揭示机制。

回应: 分类学习和推理学习下诊断性信息和非诊断性信息的加工一直是本课题组关注的重要领域,也在这一领域取得了一定的研究成果。审稿专家提出的补充实验也是本课题组目前在考虑实施的内容。我们也相信,在类别学习领域,还是有很多问题需要去进一步探讨的。就本研究而言,我们主要探讨的是部分呈现条件下类别分类学习中的信息表征机制,特别是部

分呈现条件下的分类学习能够表征非诊断性信息的原因。本研究的实验逻辑和主要发现如下：

类别学习能力是人类的基本能力，目前类别学习的研究主要有两种形式：一是分类学习，二是推理学习。先前研究者一般认为，在推理学习任务中，可以同时学到典型性和诊断性特征信息，而在分类学习任务中只能学到诊断性特征信息，不能学习非诊断性特征信息 (Markman & Ross, 2003)。而 Taylor 和 Ross (2009) 的研究中只使用了全部呈现和部分呈现（缺失 2 个特征）两种条件，发现部分呈现条件下的分类学习仍然可以学会非诊断性信息，并用注意广度来解释这种现象，认为由于注意的资源有限，部分呈现条件下分类学习的被试需要注意的特征较少，此时就有更多的注意资源分配到非诊断性特征上，而在全部呈现条件下，被试要注意更多的特征，注意的广度变大，而被试的注意资源有限，只能将注意资源放在了作用性更大的诊断性特征上。

而本研究因为，部分呈现条件下的分类学习能学会非诊断性信息，并不是注意广度的影响，实际上是因为被试在这种分类学习任务时也同时进行了推理学习。这样就跟先前大多数研究者认为的“推理学习任务中可以同时学到典型性和诊断性特征信息，分类学习任务中只能学到诊断性特征信息而不能学习非诊断性特征信息”的观点相一致。

为了证明这一基本假设，本研究实验 1 在重复 Taylor 和 Ross (2009) 的实验的基础上增加了“缺失 1 个特征”这一实验条件，如果部分呈现条件下的分类学习能学会非诊断性信息是注意广度的影响，那么“缺失 1 个特征”和“缺失 2 个特征”条件下非诊断性信息的变化就会有显著差异。实验结果发现，只有部分呈现条件下非诊断性信息有变化，全部呈现条件下非诊断性信息没有变化，这跟 Taylor 等人的研究结果一致，但是“缺失 1 个特征”和“缺失 2 个特征”条件下非诊断性信息的变化没有显著差异，这说明注意广度没有产生影响。

实验 1 虽然证明了部分呈现条件下被试非诊断性信息的学习不是注意广度造成的，但还不能充分说明就是推理学习造成的，因此本研究又设计了实验 2，对这一问题进行进一步的探讨。

实验 2 中自变量有全部呈现 7 个特征、全部呈现 5 个特征和部分呈现 5 个特征三种水平。前两个水平间的注意广度不同（一个水平要注意 7 个特征、一个水平要注意 5 个特征）但推理情况相同（都不需要进行推理学习）。而后两个水平间的注意广度相同（都只需要注意 5 个特征）但推理情况不同（一个水平需要推理学习、一个水平不需要推理学习）。如果前两个水平在非诊断性信息加工上出现差异，则说明是注意广度的影响；如果后两个水平在非诊断性信息加工上出现差异，则说明是推理加工的影响。实验结果是在全部呈现 7 个特征和全部呈现 5 个特征的条件，被试都没有学会非诊断性信息，而只有在部分呈现 5 个特征的条件，被试才学会了非诊断性信息。也就是说，注意广度不同但推理情况相同前两个条件都没有加工非诊断性信息，而注意广度相同但推理情况不同的后两个条件则不同，其中全部呈现 5 个条件没有加工非诊断性信息，而部分呈现 5 个条件加工了非诊断性信息。这个结果说明缺失条件下分类学习者对非诊断性信息的加工，不是注意广度造成的，应该是推理学习造成的。

由此可以看出，本研究虽然在“立”的方面还有需要进一步探讨的问题，比如部分呈现条件下分类学习中发生的推理到底是自动化进行的还是策略性发生的（这是本课题组进一步研究的方向），但是就目前本研究的 2 个实验来看，在一定程度上已经达到了“有破有立”，既证明了部分呈现条件下分类学习中非诊断性信息的变化不是注意广度造成的（见实验 1 中“缺失 2 个特征”和“缺失 1 个特征”两种条件的比较结果，实验 2 中“全部呈现 7 个特征”和“全部呈现 5 个特征”的比较结果），也说明了部分呈现条件下分类学习中非诊断性信息的变化应该是推理学习造成的（见实验 2 中“全部呈现 5 个特征”和“部分呈现 5 个特征”的比较结果）。

根据审稿专家的建议，我们在修改稿中对实验逻辑进行了进一步的说明，对实验结果也进行了更深入的讨论。对有待于进一步研究的问题也明确的进行了说明。请见 P15、16、20、24、25。

总之，本研究虽然没有证明部分呈现条件下的分类学习中被试进行了自动化的推理学习，但是已经证明了部分呈现条件下的分类学习中被试进行了推理学习。表述上进行了如下的修改：

“本研究通过 2 个实验证明了部分呈现条件下的分类学习中可以表征非诊断性信息，并且这种非诊断性信息的表征不是注意广度造成的，而是引发被试进行了推理学习造成的。当然，部分呈现条件下分类学习中被试能够表征非诊断性信息是被试进行了推理学习的观点还需要进一步的实验证据，特别是这种推理学习是自动化进行的还是策略性发生的，还需要在以后的研究中进一步探讨”。