

《心理学报》审稿意见与作者回应

题目：反向对抗逻辑范式的创立与证实——人工语法 PDP 对抗逻辑的改进

作者：张剑心 汤旦 李莹丽 刘电芝

第一轮

审稿人 1 意见：研究指出“人工语法对抗逻辑范式不能排除相似性和辨别力，导致自动反应不纯粹或无法被检测”。研究者在以往研究的基础上，对原有人工语法对抗逻辑范式进行改进，尝试创立“反向对抗逻辑范式”，并设计实验对这一范式进行验证。纵观整篇研究，以下几点尚待改进。

意见 1：从研究本身来讲，一个实验分量相对来讲比较单薄，是否有与文章中的实验成系列的实验可考虑补充进来。

回应：感谢专家的意见！本文尚无法增加新的实验，有三个原因。（1）笔者致力于建立和验证反向对抗逻辑范式，还没有做新的实验。非常感谢专家的指导性意见！使笔者获得了应该设计系列实验的科研意识，认识到应该对一个问题递进深入或扩展探索下去。这促使笔者深入思考，发现进一步研究一方面可以将反向对抗逻辑范式改造成为被试内设计，就能测量到同一个被试人工语法学习的受控反应（意识成分）和自动反应（无意识成分）；另一方面可采用反向对抗逻辑范式研究人工语法学习和意识机制。已将这些展望添加在了 3.3 的讨论中，敬请专家进一步指导！（2）人工语法学习的对抗逻辑范式存在诸多缺陷，如无法排除形式相似性和辨别力影响，无法检测到真实的自动反应，获得的受控反应不纯粹等。本研究在实验设计和统计分析时逐一考虑并弥补了这些缺陷，因此实验内容和工作量都较大，最终本文用一个实验已能有效地解决这些问题，达到了研究目的。（3）为了清晰地呈现和解决这些问题，本文引言需要详细论述，实验和数据分析部分内容都较多，已达到了临界篇幅，无法增加新的实验了。

意见 2：所有的统计检验均缺少效应量。

回应：感谢专家的详细审阅，笔者疏忽了，已添加了效应量。

意见 3：第 5 页第 17-18 行“实验材料使用两种不同的限定状态语法 AB 和非法字符串 U”表达似乎稍有歧义。

回应：已改成了“实验材料使用两种不同的限定状态语法 A 和 B 以及非法字符串 U”，不知是否清楚了？

意见 4：文章中有的地方用“内隐序列学习的 PDP 对抗逻辑范式”，有的地方又用“人工语法对抗逻辑范式”，均是指代“PDP 对抗逻辑范式在人工语法中的引入”，希望作者能把这些概念统一起来，以免给读者带来困惑。

回应：原版确实没有解释清楚，已改为“内隐序列学习中的对抗逻辑范式”和“人工语法学习中的对抗逻辑范式”来区分表述。内隐序列学习和人工语法学习是内隐学习的两个不同的经典范式，对抗逻辑范式在内隐序列学习中应用得很成功，产生了包含和排除任务，有效地分离出了受控反应和自动反应；但是在人工语法中却没有获得成功。两者的区别见 1.1 最后一段蓝色字体。

意见 5：作者在自检报告第一条的回答中，最后一句指出：“后来 Wan, Dienes 和 Fu（2008）以及 Norman（2011）都采用了 Dienes 等（1995）的反向对抗逻辑范式，结果都没有发现自动反应。”这里提到“反向对抗逻辑范式”，文章中又说“反向对抗逻辑范式”是自己首创的，请作者思考并作出解释。

回应：原版产生了笔误，已改正了。Dienes 等（1995）建立的是对抗逻辑范式，Wan, Dienes 和 Fu（2008）以及 Norman（2011）采用的都是其对抗逻辑范式，遗憾的是都没有成功获得自动反应。

意见 6：按照作者的思路，对“人工语法对抗逻辑范式的缺陷”的分析的第一点的分析似乎还可进一步深入。请作者思考。

回应：非常感谢专家的启发性和指导性意见！笔者进一步深入思考，发现 Dienes 等（1995）的实验对语法 A 和 B 分别连贯学习，加深了对两者深层语法结构的掌握与辨别，很可能是规则学习；而测量阶段非法字符串 U 又与语法 A 和 B 都相似，进一步激发了内隐辨别系统，导致其提取必须排除熟悉性和片段相似性的影响，因此很可能是对已获得了的语法 A 和 B 的深层语法结构的提取。

而 Higham 等（2000）的实验再学习阶段是语法 A 和 B 同时学习，可能无法获得两者深层语法结构，只是熟悉性学习，并且无法辨别开来；甚至把语法 A 和 B 作为一套联合语法 A*B 来学习和存储。而测量阶段非法字符串 U 又与语法 A 和 B 相似性不高，只需要激发熟悉性知识就能完成任务，因此没必要也无法分别提取两套语法规则——因为要么并没有获得语法规则，要么获得的是一套无法分离的联合规则 A*B。

所以学习阶段的差异可能导致两者的人工语法的学习机制和提取机制都具有本质性差异（不同学习机制参考了 Pothos, 2007 的综述）。非常感谢专家的宝贵意见，不知专家对第一点的深入思考是什么呢？期待能获得专家的进一步指导和交流！

意见 7：请作者在行文方面多下功夫（例如“第二，非法字符串 U 的设计有差异”这一标题下的内容表达有堆砌之嫌），让文章表达更流畅，另外，请作者注意文章各部分的衔接，比如可以使理论研究部分与实证研究部分衔接得更好一些。

回应：感谢专家意见！笔者通过仔细阅读，发现第二点分析中实质上包含了两个不同的内容，一是非法字符串 U 的差异导致对其拒绝的难易不同，二是导致对语法 A 和 B 的内隐辨别加工程度不同。因此拆开成了第二和第三点。已加强了各部分的衔接。

审稿人 2 意见：该研究针对以往人工语法学习范式中存在的问题，加以改进，创造了反向对抗逻辑范式，分离了人工语法学习中的自动成分。该范式具有一定的推广价值。

意见 1: 我并非人工语法学习方面的专家，因此无法判断该范式对于人工语法学习研究的真正价值。我更关注的是，研究本身对于初始理论问题的意义；据我所知，人工语法是用于研究内隐学习的工具，那么，比较重要的问题是，该研究对于内隐学习的理论（例如 Pothos (2007) 所总结的那些理论）有何促进，让我们知道了哪些之前不知道的东西。作者于研究范式的细节变化历史上花了大量笔墨，但对于理论背景和研究问题的介绍甚少，让人难以了解其理论价值所在。

回应: 非常感谢专家的指导性意见！专家强调“研究本身对于初始理论问题的意义”，让笔者意识到科学研究不止要说清楚做了什么研究，还应特别重视理论发展、意义和价值。这促使去阅读人工语法学习和意识机制的重要文献，特别是专家推荐的 Pothos (2007) 的综述，结果大大加深了对人工语法学习的认识！专家的意见指导笔者真正初步掌握了这个领域，非常感谢！理论意义已补充在 3.3 的讨论中，摘录如下。

“笔者发现内隐序列学习中对抗逻辑范式应用得很成功(Fu, Dienes, & Fu, 2010)，能够客观和定量地分离出内隐序列学习获得的受控反应（即受控意识）和自动反应（即内隐成分）；但是对于同样作为内隐学习经典范式之一的人工语法学习，对抗逻辑范式却无法成功获得真实的自动反应，导致难以客观定量地探索其内隐和外显成分。这一方面缺乏客观定量的证据来证明人工语法范式确实属于内隐学习，或者至少存在内隐成分，另一方面缺乏客观定量地测量意识的工具，这可能使得人工语法学习的意识研究走向了主观意识领域(Dienes & Scott, 2005)，而在客观意识领域发展缓慢。因此本研究反向对抗逻辑范式的成功，可以为人工语法范式提供一种良好的客观定量的意识测量范式；而作为一种普适性的工具，可以进一步探索片段学习、组块学习、熟悉性、规则学习等不同的人工语法学习机制以及对应的研究范式 (Pothos, 2007)各自产生的受控反应和自动反应有何特征和差异？受控反应和自动反应的组合与互动模式能否作为这些学习机制的良好校标？在此基础上甚至可以将该范式改造成被试内设计，采用受控反应和自动反应的关系来确定每个人工语法字符串属于哪种学习机制和哪种知识。这可能将促进人工语法学习和意识机制研究的发展。当然还可以扩展到其它内隐学习和内隐认知领域，探索抗干扰内隐学习和内隐辨别，以及否定性内隐规则的学习机制。”

意见 2: 此外，作者在背景介绍上存在一些令人费解的地方。例如：作者提到 Higham, Vokey and Prithard (2000) 将 PDP 引入人工语法学习，但下一段却说 Dienes et al. (1995) 使用了类似的范式，那么究竟是谁先引入该范式？

回应: 感谢专家的详细审阅！确实疏忽了没讲清楚，Dienes et al. (1995)是首次将对抗逻辑范式引入到人工语法学习，但是未能成功检测到自动反应；而 Higham, Vokey and Prithard (2000) 首次成功将对抗逻辑范式引入到人工语法学习，发现了受控反应和自动反应，得到了一定程度的认同，同时也受到很多质疑，因为其自动反应很可能是虚假的。详见 1.1。

意见 3: 作者将 Redington (2000) 对 Higham 等人研究的批评作为范式改进的动机之一，实际上 Higham 本人对此也有回应，作者应加以引述。此外，Higham 等人在之后十多年内也有许多新的发展，也应该加以介绍并说明作者所发明的范式相对于他们进步之处。

回应：感谢专家提醒！确实如专家所说，Higham 和 Vokey（2000）对 Redington (2000)的质疑做了回应，但是笔者认为其回应无法排除其自动反应具有虚假成分的可能性——甚至可能是受控反应，详见 1.2 新增的蓝色字体（摘录在下段，以方便专家审阅）。Vokey 和 Higham (2004)还做了神经网络建模，但仍然无法解决这个问题。搜索了 Higham 等人在之后十多年的研究，已没有对抗逻辑范式的进一步论证和扩展，可能就是由于无法解决这个问题。其他研究者也试图建立新的对抗逻辑范式来弥补其缺陷，但都没有获得成功（详见 1.2）。因此本文有必要建立反向对抗逻辑范式。

“Higham 和 Vokey（2000）回应 Redington 说其自动反应并不虚假，至少是相似性导致的熟悉性产生的真实自动反应，与内隐记忆中的基于熟悉性产生的自动反应类似。但笔者并不认同其解释，笔者认为：Higham 等（2000）检测到的自动反应仅仅是因为没有在学习阶段把 A 和 B 辨别开来而已，被试从自动和受控层面都把语法 B 和 A 看做是同一套联合语法 A*B 来学习和存储。因此在对抗条件组，将语法 B 自动判定为合法正是遵从外显任务要求的受控反应！其认知过程并没有产生任何对抗，谈何自动反应呢？”

第二轮

审稿人 1 意见：本文针对人工语法范式的不足，创立了反向对抗逻辑范式，并通过实验加以验证，研究成果具有较大的理论及应用价值，对内隐学习领域的研究具有一定推动作用。经过一审后，作者参照审稿意见已对文章进行了较好的修改。通篇行文流畅，逻辑清晰，理论切入点也非常明确，且实验设计合理有效，数据分析严谨，为所提出的反向对抗逻辑范式提供了较好的理论支持，故同意发表。

需稍作修改的细节：2.5.1, 2.5.3 两部分的结果是否可以通过图表的形式简洁地呈现，这样可省去诸多较为繁琐的语言描述。

回应：非常感谢专家的肯定！专家在一审中给予了详细而极具启发性的指导意见，促使作者进一步思考已有研究缺陷的本质，获得了更深刻的认识——提出联合语法 A*B 的可能性，从而真正揭示了已有研究的不足，使得本研究的推进和价值得以显现。在此过程中作者的思维得到了很大的锻炼提高，非常感谢专家的宝贵意见！已按照专家二审意见，将 2.5.1, 2.5.3 两部分的结果用表加以呈现，这样确实能让读者阅读时清晰直观。为了保持统计分析的规范性和完整性，还是保留了文字叙述。

审稿人 2 意见：经过修改，论文基本解决了审稿人提出的问题；但文字上仍存在晦涩难懂、部分句子不通顺等问题，需要进一步润色。

回应：非常感谢专家的肯定！由于专家在一审中要求作者阐述本研究的理论意义和价值，并推荐了人工语法理论的重要文献，这些非常有建设性的意见指导作者进一步阅读和思考人工语法理论和机制，结果获得了对人工语法理论的初步掌握，从学习和提取的认知机制上分析已有研究不足和本研究的进展，从而较大程度地提升了本文的理论深度和意义。非常感谢专

家的关键指导！已按照专家二审意见，对全文进行了精修，对某些内容进行了解释，改进了某些句子的表述。

第三轮

编委专家意见：作者根据外审专家意见做了认真修改，论文质量有了较大提升，达到了心理学报的质量要求，同意发表。建议：作者应对文字做在进一步修改润色，以增强其可读性。

回应：非常感谢编委专家的肯定和指导！已按照专家意见对文字做了进一步修改润色，在某些不清楚的地方增加了解释说明，对不通顺的句子变换了阐述方式，详见正文紫色字体。另外正文蓝色字体是根据一审专家意见做的修改，红色字体是根据二审专家意见做的修改。敬请主编审阅。

第四轮

主编意见：我仔细阅读了稿件《反向对抗逻辑范式的创立与证实—人工语法 PDP 对抗逻辑的改进》(xb15-199) 及其评审过程，同意责编委和评审专家的意见，同意发表。