

《心理学报》审稿意见与作者回应

题目：发展性阅读障碍儿童的新词习得及其改善

作者：白学军，马杰，李馨，连坤予，谭珂，杨宇，梁菲菲

第一轮

审稿人 1 意见：本研究选题有一定的新意，研究设计合理，研究发现有一定的启发意义。文章存在以下问题：

意见 1：文章的理论意义和理论框架不是十分清晰，导致最后对结果的解释较多处于猜测层面，比如，用正字法、语音意识、语素意识缺陷解释阅读障碍儿童新词学习低效现象，缺乏直接的证据。文章测查了儿童的正字法意识、语素、语音意识，但是一直没有提到这些测验结果的作用，也没有用到。如果作者倾向于用这些技能缺陷解释阅读障碍儿童的新词学习问题，可以尝试进行一下相关分析，看看是否有关联。其他几个测验结果在后面的研究中也并没有用到。

回应：根据审稿专家建议，我们将正字法意识、语音意识、言语工作记忆、快速命名及阅读流畅性等个体测验成绩和儿童在新词上的首字注视时间、凝视时间和总注视时间三个眼动指标分别做了相关。结果发现，正字法意识、语音意识和言语工作记忆与各项眼动指标之间的相关不显著，而数字和字母快速命名成绩和新词的总注视时间呈显著正相关，阅读流畅性成绩与新词的凝视时间呈显著负相关，检验结果已补充至实验 1 结果部分（请见文中第 11 页“2.2.2 各项行为测验与眼动指标之间的相关分析”）。基于该结果，对讨论部分做了相应修改，请见文中第 15 页第 3-4 段。

意见 2：元语言学意识测验具体是如何做的，各个测验的信度如何？语素意识测验写的是音节测验，是什么意识？为什么阅读障碍儿童在这个测验上与控制组没有差异？

回应：元语言学意识测验均来自于前人研究中采用的经典测验任务，这些测验已被广泛运用到发展性阅读障碍儿童的评估与筛选中（林传鼎，张厚粲，1986；凌文铨，滨治世，1988；王晓辰，2010；邹艳春，2003；Denckla & Rudel, 1974；Li, Shu, McBride-Chang, Liu, & Peng, 2012；Li & Wu, 2015；Shu, Meng, Chen, Luan, & Cao, 2005）。各个测验的具体实测过程如下所示：

（1）正字法意识测验。采用 Li 等人（2012）的测试，共 90 个字：包括遵循正字法规则的假字 45 个，如“珥”和违反正字法规则的非字 45 个，如“女”。测验要求儿童判断哪些是真字，哪些不是。每个字判断正确计 1 分，判断错误计 0 分，满分 90 分。

（2）语音意识。包括音节判断任务、音位删除任务、和声母/韵母/声调辨别任务。

音节判断任务采用邹艳春（2003）的实验材料。每组材料由两个双字词构成。其中一半材料中的两个双字词均包含一个相同音节的字，如爬山-衬衫；另一半材料则不包含相同音节的字，如母鸡-猫咪，共 36 组。测验要求被试判断两个词中是否有相同音节的字，答对 1 个计 1 分，答错计 0 分，满分 36 分。需要说明的是，音节判断任务是对整体性语音单元的分析与加工能力的评估，属于语音意识测验的范畴。非常抱歉，在上一稿中出现写作错误，现已纠正，请见文中第 4 页倒数第 2-3 行。

音位删除任务。采用 Shu 等人 (2005) 的材料, 包括首音删除、中音删除和尾音删除。测验要求被试删除某个位置的音, 然后将剩下的音读出来。例如, 首音删除测试: “请你说 shu1 但不用说开头的音”, 答案是“u1” (数字代表声调)。答对 1 题计 1 分, 首音删除和中音删除测验满分为 8 分, 尾音删除测验满分 10 分。

声母/韵母/声调辨别任务。采用 Shu 等人 (2005) 的材料。声母辨别、韵母辨别和声调辨别各包含 12 组测题, 每组测题包含四个字, 其中一个字的声母/韵母/声调和其他三个字不同。测验要求被试写下声母/韵母/声调不同的字的序号。每一正确答案计 1 分, 满分 36 分。

(3) 言语短时记忆测验。包括数字广度 (顺背) 和 (倒背) 任务、以及汉字广度任务。

数字广度 (顺背+倒背) 任务。选自《韦氏儿童智力量表 (中文版)》(林传鼎等, 1986)。顺背任务共 8 组测题, 每组包括两串同样位数的不同数字序列, 数字位数从 3 到 10 依次递增。主试按序读出每题的数字序列, 要求被试对该数字序列进行复述。一串数字序列复述正确, 计 1 分; 连续两串数字序列均复述错误, 计 0 分并停止测验。满分 16 分。倒背任务共 6 组测题, 数字位数从 2 到 7 依次递增。主试按序读出每题的数字序列, 被试从后往前按倒序对该数字序列进行复述。满分 12 分。数字广度 (顺背+倒背) 任务满分共 28 分。

汉字广度任务。采用王晓辰(2010)的材料。共 9 组测题, 每组包括两串字数相同、无意义的汉字串, 汉字字数从 2 到 10 依次递增。主试播放录音, 读完一串汉字后, 要求被试对该字串进行复述。一个汉字串复述正确计 1 分; 连续 4 串汉字序列均复述错误, 计 0 分并停止测验, 满分 18 分。

(4) 快速自动命名测验。包括数字、字母和颜色快速自动命名。

数字快速自动命名。根据 Denckla 等人 (1974) 的经典范式, 采用 2、6、9、4、7 五个数字随机构成一个 5 行*10 列的矩阵。测验要求被试按照从上到下、从左到右的顺序, 又快又准地逐个读出数字。记录被试在测验中的反应时; 字母快速自动命名采用 a, o, s, p, d 五个字母随机构成一个 5 行*10 列的矩阵, 施测及记录方法同数字快速自动命名; 颜色快速自动命名任务改编自 Stroop 任务中的卡片 B (凌文轻等, 1988), 采用蓝、绿、红、黄、黑五种颜色的圆点随机构成一个 5 行*10 列的矩阵, 施测及记录方法同数字快速自动命名。

(5) 阅读流畅性测验测验, 包括一分钟快速读词和三分钟句子阅读任务。

一分钟快速读词。采用王晓辰(2010)的测验材料。测题包括 90 个词语, 要求被试在一分钟内, 又快又准地逐一读出各词。根据被试读完的词数以及所用的时间计算被试每分钟读的词数。

三分钟句子阅读。根据 Li 等人(2015)的范式, 测题为自编的 80 个简单陈述句。要求被试在三分钟内, 按照顺序又快又准地阅读句子, 并对句子内容作出判断。记录被试三分钟内读过的字数。

发展性阅读障碍儿童在音节判断测验中的成绩与控制组儿童无显著差异, 表明本研究所筛选出的发展性阅读障碍儿童在整体语音单元的分析与加工上无缺陷。事实上, 本研究所筛选出的发展性阅读障碍儿童仅在语音加工的音位删除方面存在缺陷。

意见 3: 前言中罗列的自然阅读中新词学习的四个认知阶段, 与本研究的关系没有得到密切体现, 本研究的发现是否符合这些认知阶段, 阅读障碍个体是哪个认知过程出现了问题? 前

言需要增加理论框架和理论问题的构建，目前的写法侧重于研究发现的文献介绍和描述，显得比较松散。

回应：非常感谢审稿专家在前言写作思路上的建议，这将大幅提升本论文的理论深度。我们在前言部分做了如下修改：

第一，将发展性阅读障碍儿童存在的缺陷和汉语阅读新词学习的四个认知阶段进行了有效结合：前言部分，基于汉语发展性阅读障碍儿童在字词识别上存在、语音加工缺陷、正字法意识缺陷、快速命名缺陷以及阅读流畅性差，推测出他们在新词学习第二个认知阶段，即利用词汇内部线索和语境线索推断新词语义可能更加困难。请见文中第2页第1段第5-9行。

第二，对现有关于新词学习的文献进行了深入论述，结合汉语阅读中新词学习的认知加工过程，解释了正常儿童的新词学习效率低于成人的原因。基于该解释，推测出发展性阅读障碍儿童在新词学习时的可能表现和特点，请见文中第2页第2段。

第三，补充了本研究的另一个实验目的，即通过同时设置生理年龄匹配组和阅读水平匹配组，是为了区分发展性阅读障碍儿童的新词学习效率是滞后还是变异两种发展模式，请见文中第3页第2段。

意见4：首次注视时间、凝视时间、总注视时间上的差异是否可以、以及如何代表加工过程的质的方面。由于句子理解、新词语义归类都无显著区别，注视时间上的差异可能未必代表加工过程的本质不同，应该重点分析注视模式是否有差异，尤其是与阅读水平控制组的不同。

回应：（1）基于阅读领域方面的眼动研究成果（闫国利，熊建萍，臧传丽，余莉莉，崔磊，白学军，2013；Blythe et al., 2012；Liang et al., 2017；Liang, Blythe, Zang, Bai, Yan, & Liversedge, 2015），目前，研究者达成的共识是首次注视时间反映的是词汇加工早期阶段的特征；凝视时间反映的是词汇通达早期阶段的特征；总注视时间反映了词汇加工总体加工的特点，包括后期语义的整合过程。具体描述请见文中第8页第1-5行。

（2）句子理解问题的设置是为了督促读者认真阅读，因此题目非常简单。三组儿童在句子理解问题的正确率上无显著差异，只能表明三组儿童均认真阅读了实验语句；语义类别判断题目的设置，目的是为了考察儿童通过八个语境的学习，是否可以掌握“新词”的语义类别信息。由于语义类别题目是在主观上反映了新词学习的最终结果，不能反映新词学习的内在认知加工过程，且本研究中测试的语义类别属于较为广泛的表层语义类别，很容易被掌握。八个语境的学习之后，语义类别判断成绩可能达到天花板效应，因此，三组儿童在主观测试结果中未表现出显著差异。本研究目的是在保证儿童在参加实验时态度认真、可以学到新词基本的“语义类别信息”的前提下，通过更为客观的眼动记录技术，实时地监测发展性阅读障碍儿童及其匹配组儿童在新词学习中的认知加工过程，通过对比三组儿童对新词进行重复学习时眼动模式的变化，考察发展性阅读障碍儿童在新词学习的内在认知加工过程。

（3）根据审稿专家建议，在讨论部分，增加了发展性阅读障碍儿童和阅读水平匹配组儿童在新词学习眼动模式变化上的讨论，在此基础上区分了发展性阅读障碍儿童新词学习能力低下的原因，即是由于发展相对滞后，还是存在变异的发展模式。请见文中第15页第2段第2-7行。

意见5：语音特征、正字法特征为什么是低水平文字特征，尤其是语音特征，对于书面文字，只有进行了语音提取才能获得语音特征，语音激活为什么被成为低水平文字特征？

回应：基于 Taft, Zhu 和 Peng (1999) 提出的词汇识别模型，词汇识别包括字形、语音和语义加工三个加工单元，语义加工建立在字形和语音加工基础之上。其中字形加工从下到上又包括视觉特征、部件、汉字和词汇四个层次。因此，相对于语义加工，在一定程度上语音和正字法特征属于较低水平的加工。不过，语音特征和正字法特征的加工均建立在汉字视觉特征提取（如视觉特征，部件等）的基础上。我们非常同意审稿专家的观点，即语音特征及正字法特征均不属于低水平的文字特征，已在文中将有关说法进行了修改。

意见6：“该研究结果可能和发展性阅读障碍儿童自身的缺陷有关，即语音意识缺陷 (Ho et al., 2000) 和正字法意识缺陷 (Ho et al., 2002) 等，导致他们不能有效获取汉字的语音、正字法信息等，表现出较低的词汇加工效率”，“自身的缺陷有关”的说法比较奇怪。另外，和前面提到的问题一样，如果认为是语音意识、正字法意识问题，可以利用测到的成绩，验证一下这种推论是否成立。另外，如果是由于单个字形音加工、语音通达的问题（语义是要新学习的），那么可以看看阅读障碍儿童在这些单词/单字阅读上是否存在问题，如果是这个问题，那还是归结到汉字形音转换的问题上。排除这个问题之后，或者在这个问题以外（因为我认为阅读障碍组一定词汇或者单字阅读速度存在一定的落后），可以进一步考察是否存在新语义构建的问题以及这个问题的认知机制是什么？

回应：非常感谢审稿专家在讨论部分给予的建设性建议。针对上述建议，文中做了如下修改：第一，将讨论中不合适的说法“自身的缺陷有关”删除；第二，基于个体测验和新词注视的眼动行为的相关分析，摒弃了之前推测性的观点“该研究结果可能和发展性阅读障碍儿童自身的缺陷有关，即语音意识缺陷 (Ho et al., 2000) 和正字法意识缺陷 (Ho et al., 2002) 等，导致他们不能有效获取汉字的语音、正字法信息等，表现出较低的词汇加工效率”，并在总讨论中基于汉语阅读新词加工的认知加工阶段及相关分析结果，对发展性阅读障碍新词学习低下的原因进行了宏观解释，请见文中第 15 页最后 1 段；第三，关于发展性阅读障碍儿童是否同时在新词语义构建上存在缺陷，需要更多的实验证据支持。我们将基于本研究提供的基础数据，展开后续系列研究，针对新词学习的各个认知加工阶段设计相应的实验，以期明确发展性阅读障碍儿童新词学习低下的真正原因。

意见7：首次探讨统计指标时请明确指出是对新词的首次注视时间、凝视时间和总体注视时间。

回应：已修改，请见第 7 页倒数第 1 行。

意见8：研究2和研究1的对比，可以推论阅读障碍儿童在新词学习上的问题之一是词汇切分。是否可以分析一下该词汇切分问题与研究1的发现“早期加工的延迟和晚期加工效率的低下”是否有逻辑联系。

回应：根据审稿专家建议，在实验 1 讨论部分，我们根据新词学习的认知加工阶段模型，结合发展性阅读障碍儿童存在的认知加工缺陷，论述了发展性阅读障碍儿童新词学习效率低下的原因。但由于本研究没有直接测试三组儿童在自然阅读中的词切分能力，因此无法将词切分能力的数据和实验 1 发现的发展性阅读障碍儿童的新词加工特点结合起来做进一步分析。不过，从前期相关研究发现，词间空格的插入主要提高了正常读者（包括三年级儿童和汉语二语学习者，见 Bai, Liang, Blythe, Zang, Yan, & Liversedge, 2013; Blythe et al., 2012）在新词相对晚期的加工效率（具体体现在凝视时间和总注视时间上）。本研究发现，对于发展性阅读障碍儿童，词间空格的促进作用同时表现在新词加工的早期和晚期阶段（体现在首次注视

时间、凝视时间和总注视时间上)。该结果在一定程度上证明了词间空格的插入同时改善了发展性阅读障碍儿童早期加工的延迟和晚期加工效率的低下。

意见 9: 讨论存在超越数据的解释, 需要在理论上深化和增加针对性。

回应: 非常感谢审稿专家在讨论中给予的建议, 我们在讨论部分做了如下几方面修改: 第一, 删除部分超越数据解释的观点, 在数据支持的前提下, 尽可能对结果做出客观解释, 具体为: 依据个体测验成绩和学习新词时眼动行为的相关分析, 并结合阅读中新词学习的认知加工阶段, 对发展性阅读障碍新词学习效率低下做出了客观解释, 请见文中第 15 页第 3-4 段; 第二, 指出由于本研究是关于发展性阅读障碍儿童词汇学习的第一篇研究, 虽然发现发展性阅读障碍儿童在新词学习时的效率低下, 并通过实验证明了可行的干预途径, 然而关于发展性阅读障碍儿童新词学习效率为何低下, 则需要更多的后续实验证据, 请见文中第 16 页第 1 段。

.....

审稿人 2 意见: 研究运用眼动技术, 探查了发展性阅读障碍儿童在阅读中新词学习效率低下的原因; 同时探讨了词间空格是否可以提高发展性阅读障碍儿童的新词学习效率, 结果发现, 发展性阅读障碍儿童在词间空格条件下的新词学习效率已达到正常儿童水平。但是, 研究存在一些问题:

意见 1: 被试。从数据看, 似乎阅读障碍来自五年级, 年龄匹配组也来自五年级, 而能力匹配组来自三年级。如果是这样, 筛查对象三年级的 991 人是为什么? 完全没有必要。另外, 应该具体说明被试的学校信息及如何计算的年级平均值(合并学校还是不同地区不同学校的分开算, 各个学校的学习分布情况等)等。

回应: 基于前人对于阅读水平匹配组的筛选标准(Meng, Cheng-Lai, Zeng, Stein, & Zhou, 2011; Shu, Mcbride-Chang, Wu, & Liu, 2006), 即阅读水平匹配组的筛选一般在低于发展性阅读障碍儿童 1.5 个年级中进行。本研究是对施测学校的所有三年级小学生进行识字量测试和联合瑞文推理测试, 然后筛选出阅读水平匹配组被试。此外, 根据发展性阅读障碍的定义和筛选方法, 发展性阅读障碍儿童的阅读水平差于同年级正常儿童。由于不同地区不同学校的学生水平不同, 因此我们在进行发展性阅读障碍儿童的筛选时是以学校为单位, 不同地区不同学校分开进行筛选。被试的学校信息以及年级平均值的计算方法等信息, 已补充在文中第 4 页表 1。

意见 2: “句子通顺性的平均值为 4.31 (SD=0.29); 句子难度的平均值为 4.47 (SD=0.35)。表明句子非常通顺且容易理解。”把新词放在句子中, 没学习的小学生有这种反应, 很令人费解。

回应: 为了排除句子框架中新词的存在可能会降低句子通顺性以及增加句子难度, 在具体评定中, 借鉴 Liang 等人(2015, 2017)的做法, 我们将每个句子框架中的“新词”替换为与“新词”同语义类别的熟悉成员, 如“富丽端庄的环米的花瓣不易脱落。”在具体评定中, 变为“富丽端庄的牡丹的花瓣不易脱落。”具体评定过程已补充在文中第 6 页倒数第 2 段第 3-5 行。

意见 3: 似乎能力匹配组的名称有些不合适, 哪些能力匹配了? 阅读? 建议写明。

回应: 根据审稿专家的建议, 已将文中所有的“能力匹配组”改为“阅读能力匹配组”。

意见 4: 实验 1 中的一些眼动差别并没有带来新词学习效果的差异（行为测试结果三组均没差异），怎么就说阅读障碍新词学习效率低？

回应: 语义类别判断题目的设置，目的是为了考察儿童通过八个语境的学习，是否可以掌握“新词”的语义类别信息。由于语义类别题目反映的是新词学习的结果，不能反映新词学习的内在认知加工过程，且本研究中测试的语义类别属于较为广泛的表层语义类别，非常容易被掌握。八个语境的学习之后，语义类别判断成绩可能达到天花板效应，因此，三组儿童在主观测试结果中未表现出显著差异。

本研究目的是在保证儿童在掌握新词基本的“语义类别信息”的前提下，通过更为客观的眼动记录技术，实时监测发展性阅读障碍儿童及其匹配组儿童新词学习的认知加工过程，对比三组儿童对新词进行重复学习时眼动模式的变化，进一步考察发展性阅读障碍儿童新词学习的内在认知加工过程。结果发现，在新词学习的眼动模式变化上，发展性阅读障碍儿童在时间指标上呈不断下降趋势，而生理年龄匹配组儿童和阅读水平匹配组儿童成阶梯式下降模式，即开始下降幅度较大，随后幅度逐渐变小。表明发展性阅读障碍儿童在阅读中学习一个新词时，需要更多的语境。

意见 5: 实验 1 和实验 2 的关系不明确，实验 1 中没有发现阅读障碍学习效率的低下，只是眼动特点和其他两组不同，但不影响其学习效果，为什么要进行实验 2？

回应: 如第 4 条回答所述，本研究的实验目的是从新词学习的内在认知加工过程角度，考察发展性阅读障碍儿童的新词学习特点。实验 1 结果发现，虽然在最为表层的语义类别判断正确率上，发展性阅读障碍儿童并未和正常儿童表现出显著差异。但在新词学习的眼动模式变化上，发展性阅读障碍儿童在时间指标上呈不断下降趋势，而生理年龄匹配组儿童和阅读水平匹配组儿童成阶梯式下降模式，即开始下降幅度较大，随后幅度逐渐变小。表明发展性阅读障碍儿童在阅读中学习一个新词时，需要更多的语境。

在实验 1 讨论部分（请见文中第 12 页第 1 段），我们根据新词学习的认知加工阶段和中文词汇切分和识别模型，论述了发展性阅读障碍儿童新词学习效率低下的原因。结果发现，发展性阅读障碍儿童的新词学习效率低下，其首要可能性便是和汉语文本呈现方式的特殊性有关，即没有视觉词切分线索帮助读者成功进行新词切分。

以往研究发现，词间空格的插入能够促进对词汇的切分和识别，主要提高了正常读者（包括三年级儿童和汉语二语学习者，见 Bai et al., 2013; Blythe et al., 2012）在新词相对晚期的加工效率（具体体现在凝视时间和总注视时间上）。因此，实验 2 在实验 1 的基础上，探讨词间空格是否能有效提高发展性阅读障碍儿童的新词学习效率。结果发现，对于发展性阅读障碍儿童，词间空格的促进作用同时表现在新词加工的早期和晚期阶段（同时体现在首次注视时间、凝视时间和总注视时间上），表明词间空格的插入改善了发展性阅读障碍儿童“早期加工的延迟和晚期加工效率的低下”。

意见 6: 被试问题。实验 2 中有部分是重复利用实验 1 的被试，有些是新被试。研究者应明确报告新旧被试的人数。这样，对旧被试而言，有一般的实验材料是已经学习过的，导致了材料对不同被试的难度不同。

回应: 实验 2 的被试信息已补充至文中第 12 页最后 1 段。具体内容为：由于实验 1 中的 18 名被试（包括 8 名发展性阅读障碍儿童、5 名阅读水平匹配儿童、5 名生理年龄匹配儿童）因家长及学校原因，不愿参加随后的实验。因此，在实验 2 中，根据相同筛选标准又增加了 12 名被试（包括 6 名发展性阅读障碍儿童、3 名阅读水平匹配儿童、3 名生理年龄匹配儿童）。实验 2 共包括发展性阅读障碍、生理年龄匹配和阅读水平匹配组各 20 人，共计 60 人。

对被试分配进行详细说明：实验材料是由 30 个新词构成的 240 个实验语句。在实验 1 中，

所有被试在正常无空格条件下阅读一半实验材料（由 15 个新词构成的 120 个实验语句）；在实验 2 中，同时参与实验 1 的旧被试在词间空格条件下阅读了另一半实验语句，而新筛选出的 12 名被试分别在正常无空格条件和词间空格条件下各阅读一半实验语句。因此，每个被试在正常无空格和词间空格条件下分别阅读了相同数量的实验语句，且每个实验语句都只阅读了一次，不存在对于每个被试的难度不同这一问题。非常抱歉，在上一稿中未将被试分配进行详细说明。

意见 7：实验 2 的结果存在与实验 1 同样的问题，行为结果没有任何差异，眼动结果能说明什么问题？

回应：实验 2 设置阅读理解判断题和语义类别判断题目的目的和实验 1 相同，是为了督促儿童要认真阅读实验语句；语义类别判断题目的设置旨在考察儿童对新词最表层语义信息的掌握程度。八个语境的学习之后，三组儿童不论在词间空格文本，还是在正常无空格文本中，语义类别判断题目的正确率均达到较高水平，且无显著差异。

实验 2 的目的是在保证儿童能学到新词“语义类别信息”以及阅读态度认真的前提下，通过眼动记录技术，实时监测发展性阅读障碍儿童及其匹配组儿童在不同文本呈现方式下对新词的认知加工，对比三组儿童在词间空格和无空格条件下对新词注视的眼动行为，考察词间空格在发展性阅读障碍儿童新词学习中的作用。通过分析三组儿童在不同呈现方式下对新词的注视时间，发现词间空格显著缩短了发展性阅读障碍儿童在新词上的首次注视时间、凝视时间以及总注视时间，表明词间空格作为一种视觉词切分线索，显著提高了发展性阅读障碍儿童在新词各个认知加工阶段的效率。

第二轮

审稿人 1 意见：为了强化和明确本研究的理论意义和价值，请进行以下两项修改：

意见 1：摘要中需要简要写出研究目的和发现的理论意义。

回应：根据审稿专家的建议，已在文中第 1 页摘要部分进行补充。

意见 2：第一条结论中需要增加该发现反映了阅读障碍儿童新词学习能力是发展落后，还是发展变异。如果是变异，这种变异具体是哪种机制。第二条结论需要进一步说明视觉词切分起促进作用的认知机制是什么。

回应：实验 1 结果发现，发展性阅读障碍儿童在学习新词的眼动模式变化既不同于生理年龄匹配组，也不同于阅读能力匹配组，表明发展性阅读障碍儿童学习新词时的阅读模式是变异发展的模式。这种变异发展的模式可能与发展性阅读障碍儿童自身的病理性缺陷（如正字法、语音、快速命名等缺陷）有关。

实验 2 结果发现，空格作为一种视觉词切分信息，促进了发展性阅读障碍儿童的新词习得。这可能是由于空格的插入，帮助了发展性阅读障碍儿童成功进行词切分，且缓解了文本的视觉拥挤效应，因此促进了发展性阅读障碍儿童的新词习得。

根据审稿专家的建议，针对第一条和第二条结论的内容，已在文中进行补充，请见文中第 17 页结论部分。

.....

审稿人 2 意见：研究者对审稿意见做了较好的回复，并对文章进行了修改。但是，还是有一点问题很关键，该研究只能说明不同被试群体的不同眼动模式或特点，而在这种不同模式下，是否有不同的学习效率，似乎不能明确说明。建议在文中注意这些表述。

回应：根据审稿专家的建议，已在文中对有关学习效率的表述进行修改，请见文中第 12 页第 1-2 行，第 15 页总讨论第 1 段第 1-3 行以及第 16 页第 1-2 行。

第三轮

审稿人意见 1：作者对发展变异的解释有一定的推理和猜测成分，因此不要说的那么确定。另外“自身缺陷“的说法不够准确，是否可以进一步明确，比如“自身认知缺陷”或者更加具体化。

回应：根据审稿专家意见，已将第一条结论修改为：“与正常儿童相比，发展性阅读障碍儿童的新词习得较为缓慢，主要表现为新词习得早期加工的延迟和晚期加工效率的低下，这可能是由他们在形-音转换的自动化加工及反映多种低水平认知加工的阅读流畅性上存在缺陷，导致新词习得的阅读模式呈变异发展。”修改请见 p17。

审稿人意见 2：第二条结论,视觉拥挤的解释没有直接的数据支持，可以说成“可能。。。”。

回应：根据审稿专家意见，已将第二条结论修改为：“词间空格作为视觉词切分信息，帮助发展性阅读障碍儿童成功进行词切分，可能在一定程度上缓解了文本的拥挤效应，促进了新词习得。”修改请见 p17。

第四轮

编委意见：作者对审稿人的意见进行了认真回复和修改。仍有以下问题请作者注意：自检报告要求作者写出计划的样本量，实际的样本量。如果二者有差别，请写出理由。作者回答：计划样本量为发展性阅读障碍组、生理年龄匹配组和阅读能力匹配组各 22 人。实际样本量：实验 1，每组 22 人，共计 66 人，达到计划样本量。而正文中，作者写道：筛选出 41 名发展性阅读障碍儿童，筛出率为 4.65%，见表 1。由于学校及学生个人原因，共 22 名发展性阅读障碍儿童参与实验，其中男生 16 人，女生 6 人。既然是经筛查和各种原因，最后得到 22 名被试，作者是如何做到一开始的计划就是 22 人？这个计划的 22 人的样本量又是依据什么确定的？

回应：根据编委意见，已在“自检报告”中将样本量的相关问题进行进一步说明，请见第 10、11 页黄色标记部分。