

《心理学报》审稿意见与作者回应

题目：汉语发展性阅读障碍儿童的形音整合能力：字形和字音的相互影响

作者：刘梦连，杨滢晖，赵一帆，毕鸿燕

第一轮

审稿人 1 意见：

研究通过两个实验，考察了阅读障碍儿童形-音整合过程中，形对音和音对形的影响。在阅读障碍的研究中研究单通道的较多，研究跨通道的较少，这是本研究的一个重要新颖之处。但我对论文仍然有以下主要建议，需要作者进一步修改。

意见 1：论文研究问题的提出部分，建议逻辑再明晰一些。目前执行该研究的必要性感觉不充分，创新点感觉没有完全体现出来。

回应：感谢审稿专家的意见，结合审稿专家在文中的批注，我们重新梳理了引言部分关于问题提出的逻辑，以突出本研究的必要性和创新性，并使用蓝色字体标出，具体可见 P.3 line 7—P.4 line 8：

“已有研究结果不一致的原因可能有两种。第一，这可能与 DD 在形-音整合过程中字形和字音的相互影响的表现不同有关。已有研究多使用神经影像的技术，采用被动任务，或要求被试仅对形音一致的刺激对进行判断，这些研究所使用的范式都无法直接说明 DD 字形和语音相互影响是否存在异常。仅有一篇 fMRI 研究以汉语 DD 儿童及 CA 儿童为被试，使用视听跨通道呈现的字形判断任务，设置了形音一致，形音不一致和视觉单通道条件，要求被试对呈现的汉字进行真假字判断(Yang et al., 2020)。该研究在神经层面发现 DD 在涉及形音整合的脑区激活异常，但并未在行为层面发现 DD 与 CA 的组间差异。这在一定程度上说明 DD 字音对字形加工的影响存在异常。然而，这一结论需要行为层面的研究进一步论证。Nash 等人(2016)和 Clayton 等人(2018)研究使用的语音合法性判断任务，行为层面的结果发现 DD 字形对语音加工的影响与 CA 组不存在差异。以上研究结果存在争议的原因可能是探查了 DD 形-音整合的过程中的不同方面，即字形和语音之间的相互影响。DD 可能在形-音整合过程中，表现出语音对字形加工的影响存在异常，但字形对语音加工的影响不存在异常。

第二，DD 在形-音整合过程中，字形和语音的相互影响可能均存在异常，但不同研究的形音刺激对呈现的时间间隔不同导致出现矛盾的结果。在 DD 字形对语音加工影响的研究

中，字形早于语音 500 毫秒呈现(Clayton et al., 2018; Nash et al., 2016)，但语音对字形加工影响的研究则是形音刺激同时呈现(Yang et al., 2020)。有研究表明 DD 具有更大的视听时间整合窗口，即当视听刺激的时间间隔较大时，DD 的整合能力比正常儿童更好(Francisco et al., 2017; Hairston et al., 2005)。已有研究发现 DD 字形对语音加工的影响与 CA 组不存在差异(Clayton et al., 2018; Nash et al., 2016)，可能是形音刺激的时间间隔变大提高了 DD 的形音整合能力，进而抵消了其字形对语音加工影响的异常。那么，DD 在形-音整合过程中，字形和语音的相互影响是否都存在异常，抑或是仅语音对字形加工的影响存在异常，这需要今后的研究进一步控制了形音刺激的时间间隔后，探查 DD 在形-音整合过程中，字形和字音相互影响的表现。

基于以上论述，本研究以 DD 儿童及其对应的同年龄正常儿童对照组（CA 组）为被试，在形音刺激对同时呈现的条件下，设计两个实验分别在行为层面探查 DD 在形-音整合过程，字形和语音的相互影响。实验 1 探查 DD 儿童语音对字形加工的影响是否存在异常；实验 2 探查 DD 儿童字形对语音加工的影响是否存在异常。在行为层面对 DD 的字形和字音相互影响进行探讨，将有利于进一步加深人们对 DD 形-音整合缺陷认知机制的认识，即 DD 的形-音整合缺陷可能源于其整合过程中字形和语音相互影响的异常。同时，本研究也进一步拓展了 Blomert(2011)提出的形-音自动整合假说。该假说仅从形-音整合的角度解释了 DD 产生的原因，本研究则基于该假说，进一步从字形和语音相互影响的角度入手，对 DD 的可能原因进行探讨。”

意见 2：实验描述部分建议增加更多的对于被试和材料的细节的描述。

回应：感谢审稿专家的建议。我们已经补充了对被试和材料的细节描述，并在文中用蓝色字体标出，具体如下：

P.4 line 14—line 16

“本实验包括 59 名三到六年级汉语儿童，符合统计要求。其中 30 名（男 25 名，女 5 名）DD 儿童，29 名（男 18 名，女 8 名）同年龄正常对照组儿童。卡方检验的结果显示两组儿童在各个年级人数的分布不存在显著差异($\chi^2 = 0.47, p = 0.93$)。……”

P.4 line 25—P.5 line 5

“视觉刺激材料包括 30 个真字和 45 个假字。真字均选自小学三年级至五年级语文课本。假字是在真字部件的基础上，打乱顺序重新组合成符合正字法规则的刺激（例如：扌同）。所

有真字和假字均为左右结构，字体为“黑体”，大小为3cm×3cm。真假字在笔画数、部件频率上匹配。在实验过程中，每个真字在三种条件下各呈现1次，每个条件各30试次，共90试次；每个假字在跨通道条件(假字和字音一同呈现，例如：才者-jǐǎ)和视觉单通道条件下各呈现1次，每个条件各45试次，共90试次。正式实验共180试次，真字和假字出现的概率均等。……”

P.7 line 20—P.7 line 23

“64名四到六年级汉语儿童参与了本实验，符合统计要求。其中包括32名(男26名，女6名)DD儿童，32名(男19名，女13名)同年龄正常对照组儿童。卡方检验的结果显示两组儿童在各个年级人数的分布不存在显著差异($\chi^2 = 0, p = 1$)。……”

P.8 line 7—line 12

“……正式实验共包括197个试次，真实字音和非真实字音出现的概率均等。字音刺激由一位普通话标准的女性主试以DELL笔记本自带录音软件录制，采用GoldWave声音软件处理声音文件，声音刺激参数为取样率44100Hz，采样为16位。视觉刺激材料为105个真字，均选自三年级至五年级语文课本，为左右结构的汉字。其中，30个汉字与听觉刺激的真字发音一致、30个汉字与听觉刺激的真字发音不一致、45个伴随非真实字音出现。形音一致条件和形音不一致条件呈现的汉字在字频和笔画数上匹配。”

意见3：建议根据汉语的特点对结果进行讨论。

回应：感谢审稿专家的建议，结合您在文中的批注，我们根据本研究的结果以及汉语阅读的特点，在讨论部分增加了相关的讨论，并使用蓝色字体标出，具体如下：

P.12 line 13—line 22

“第三，本研究的结果也可能与汉语的语言特异性有关。不同于拼音文字，汉字不存在明确的形-音对应规则。汉语约有1200个基本音节(卢偃, 2001)，但常用字却多达3500个。这使得在汉语阅读过程中，形到音的对应往往是一对一的(多音字除外)(如，“熟”对应“shú”)，但音到形的对应则存在一个音对多个汉字的情况(如，“shú”则可对应“熟、塾、赎”等字)，从而造成读者在阅读过程中字形对语音的约束更大(字对应明确的音)，但字音对字形的约束则较小(字音很大程度上难以对应具体的字形)。因此，字形对语音的加工的影响对汉语读者来说有更加重要的作用。并且，已有的众多研究也发现视觉加工对汉语阅读

存在关键作用, 汉语DD 更容易表现出视觉字形加工缺陷(Chung et al., 2008; Meng et al., 2011; Qian & Bi, 2014)。本研究发现汉语 DD 儿童仅在字形对语音加工的影响上存在异常, 可能是因为 DD 字形加工存在异常, 从而导致其对语音加工的影响减弱。”

意见 4: 建议讨论增加对研究结果的理论意义和应用价值的讨论。

回应: 感谢审稿专家的建议, 我们已经在讨论部分增加了对研究结果的理论意义和应用价值的探讨, 并用蓝色字体标出, 具体如下:

P.12 line 23— line 29

“本研究的结果存在一定的理论和实践意义。在已有研究结论的基础上, 本研究的结果创新性地揭示了在形-音整合的过程中, 汉语 DD 儿童表现出字形对语音加工影响的异常, 但语音对字形加工的影响不存在异常。从理论意义上来讲, 本研究的结果进一步对汉语 DD 的形-音整合缺陷的内在机制进行了阐释, 即 DD 字形加工能力较差, 进而字形对语音加工影响存在异常, 导致阅读困难。从实践意义上讲, 本研究的结果为今后 DD 阅读能力的矫治提供了理论基础, 即可以通过训练 DD 字形识别和其利用字形加工语音的能力, 进而提高其阅读能力。”

意见 5: 论文的设计只有 CA, DD 两组, 没有阅读水平控制组, 因而不能对形-音整合是否是 DD 的缺陷做结论, 作者需要再讨论中写明这一 limitation

回应: 感谢审稿专家的建议, 我们在正文中讨论部分增加了局限性, 并用蓝色字体标出, 具体如下:

P.12 line 30—P.13 line 8

“本研究也存在一定的局限性。第一, 本研究以汉语 DD 儿童为被试, 并设置了同年龄对照组, 但没有招募同阅读水平的正常儿童作为对照组。这就导致在形-音整合过程中, DD 儿童字形对语音加工的影响存在困难这一现象可能是源于其缺乏足够的阅读经验, 而不是 DD 儿童固有的缺陷。将来的研究需要进一步与同阅读水平对照组进行比较, 以明确 DD 字形对语音加工的影响的困难是否是其固有缺陷。第二, 本研究没有考虑时间间隔的可能影响。已有研究发现 DD 当视听刺激的时间间隔变大时, 其仍然具有一定的视听整合能力, 正常读者的视听整合能力则消失(Francisco et al., 2017; Hairston et al., 2005)。那么, 如果形音刺激有一定时间间隔时, 其字形和语音的相互影响又会有怎样的表现呢? 后续研究需要更系统的操作这些影响因素, 全面探查 DD 在形-音整合过程中字形和语音的相互影响。”

意见 6: 行文有一些错误, 以及需要规范的地方, 见批注。

回应: 感谢审稿专家的辛苦工作和宝贵意见, 我们已经根据您的标注修改了文章, 具体见文中对批注的回复。同时, 我们仔细修改了文章, 认真校对, 杜绝低级错误的发生。

意见 7: 英文摘要需要进一步修改。

回应: 感谢审稿专家的建议, 我们已对英文摘要进行了认真修改, 并请英语专业人士进行修改。一些细节问题是我们的疏忽, 抱歉! 我们已经根据审稿人的意见进行了修改, 并用蓝色字体标出, 具体可见英文摘要部分。

意见 8: 其它细节问题见对论文的批注。

回应: 感谢审稿专家的细心修改, 我们已经修改了您在文中标注出来的错误, 具体可见文中批注部分以及对批注的回复。

.....

审稿人 2 意见:

本文探讨形音整合过程中阅读障碍儿童和控制组儿童形音之间的作用差异。研究有意义, 聚焦性好。论文存在以下问题:

意见 1: 论文写作表述需要进一步精确, 比如开头三句, “一部分人群”, “即”、“即”, “存在问题”等表述都不是很精确。第 5 页倒数第 7 行, “对应的” CA 具体什么是对应的?

回应: 感谢审稿专家指出问题, 抱歉, 因为我们的表述不清让审稿专家疑惑了。我们已经修改了相关的内容, 并蓝色字体标出, 如下:

P.1 line15—line 17

“然而, 有一类特殊人群并不能获得良好的阅读能力, 即发展性阅读障碍(Developmental Dyslexia, DD), 指智力正常并接受正常的教育, 无器质性损伤, 但在阅读上存在显著困难的个体(Lyon et al., 2003)。”

P.3 line 30—P.4 line 2

“基于以上论述, 本研究以 DD 儿童及其对应的同年龄正常儿童对照组 (CA 组) 为被试, 在形音刺激对同时呈现的条件下, 设计两个实验分别在行为层面探查 DD 在形-音整合过程, 字形和语音的相互影响。”

意见 2: 论文研究形音整合中形音的相互影响, 其目的是揭示什么理论问题, 有什么理论意义? 论文没有给出相应的介绍。

回应: 感谢审稿专家的意见, 结合另外一位审稿人的意见, 我们在讨论部分加入本研究理论意义的探讨, 同时, 我们也在引言部分的最后一段加入了揭示的理论问题以及理论意义, 并用蓝色字体标出, 具体如下:

P.4 line 3—line 8

“在行为层面对 DD 的字形和字音相互影响进行探讨, 将有利于进一步加深人们对 DD 形-音整合缺陷认知机制的认识, 即 DD 的形-音整合缺陷可能源于其整合过程中字形和语音相互影响的异常。同时, 本研究也进一步拓展了 Blomert(2011)提出的形-音自动整合假说。该假说仅从形-音整合的角度解释了 DD 产生的原因, 本研究则基于该假说, 进一步从字形和语音相互影响的角度入手, 对 DD 的可能原因进行探讨。”

意见 3: 文献综述的第三段说是通过 ERP 研究形音进程, 但是内容中只字未提时间进程。似乎并未与第二段有什么区别。

回应: 感谢审稿专家的意见, 这一段是我们写作上的问题, 抱歉。我们已经在文中加入了脑电研究中关于时间进程的内容, 并用蓝色字体标出, 具体如下:

P.2 line15—line 21

“结果发现, 跨通道条件下, CA 儿童的 MMN 波幅比听觉单通道条件显著增强, DD 组则未发现这种变化, 但表现出晚期辨别负波的增强(Late Discrimination Negative, LDN)。研究者认为, 早期的 MMN 和晚期的 LDN 都可以反应被试的形-音整合能力。DD 只表现出晚期 LDN 波幅的增强说明其形-音整合的时间进程更慢, 自动化程度较差。之后, 有研究额外增加了图形和语音刺激对作为偏差刺激(Mittag et al., 2013), 以成人 DD 和 CA 组为被试, 结果与 Froyen 等人(2011)一致, 即无论是不一致的形音刺激还是图形和语音刺激, DD 的 MMN 波幅增强都弱于 CA 组, MMN 的潜伏期更长……”

意见 4: 实验设计了假字条件, 数据分析和结果部分并未使用, 设计假字条件的意义和目的是什么? 仅仅是控制“是否”反应吗?

回应: 感谢审稿专家提出的意见。正如审稿专家所说, 我们设置假字的条件就是为了控制被试的“是否”反应。在视听整合的研究中, 已有多项研究采用这种实验设计的方法 (Clayton

et al., 2018; Nash et al., 2016; Yang et al., 2020), 并且假字的条件也没有纳入到最后的数据分析。

Clayton, F. J., & Hulme, C. (2018). Automatic activation of sounds by letters occurs early in development but is not impaired in children with dyslexia. *Scientific Studies of Reading*, 22(2), 137–151.

Nash, H. M., et al. (2016). Are the literacy difficulties that characterize developmental dyslexia associated with a failure to integrate letters and speech sounds? *Developmental Science*, 20, e12423.

Yang, Y., et al. (2020). An audiovisual integration deficit underlies reading failure in nontransparent writing systems: an fMRI study of Chinese children with dyslexia. *Journal of Neurolinguistics*, 54, 1000884.

意见 5: 作者根据参考文献计算了 CBS, 虽然有一定道理。但是由于 DD 儿童语音通达本身就与 CA 不同, 其干扰作用也会比 CA 大或者小 (或许是小, 因为其自动化通达能力弱, 但是需要实验检验), 如果两方面都不同, 再把二者相减除以不一致条件, 很可能并不能得到很理想的增益作用。

回应: 感谢审稿专家的意见, 这个问题是本研究数据分析部分值得探讨的问题。一致性获益分数(CBS)是一个衡量被试的字形和语音整合的指标。目前, 已有多项研究使用 CBS 值衡量被试的跨通道的加工能力(Evens, 2020; Stürmer et al., 2014), 主要是反映跨通道刺激所引起的促进作用。本研究结果中, CBS 值的结果与方差分析的结果相辅相成, 也体现了 CBS 可以真实反映被试的跨通道刺激的加工能力。但是, 审稿专家的看法确实有道理, 在后续研究中还需要进一步考察这一指标的合理性问题。

意见 6: 讨论和结论中需要对“在字形对语音加工影响的任务下表现得更差。以上结果说明 DD 儿童在形音整合的过程中, 仅在字形对语音加工的影响中表现出缺陷”进行理论意义的讨论, 不仅仅是比较文献上的一致或者不一致。另外如果说有缺陷, 最好是与 RA 组进行对照。

回应: 感谢审稿专家提出的意见。结合您和另一位审稿专家的意见, 我们在讨论部分加上了本研究结果的理论和实践意义的探讨。此外, 我们的确是没有设置 RA 组, 无法得出 DD 儿童存在字形对语音加工的影响存在缺陷的结论, 因此我们将“缺陷”改为了“异常”, 并且在研究局限性中指出了这一问题, 并用蓝色字体标出, 具体如下:

P.12 line 23—P.13 line 8

“本研究的结果存在一定的理论和实践意义。在已有研究结论的基础上，本研究的结果创新性的揭示了在形-音整合的过程中，汉语 DD 儿童表现出字形对语音加工影响的异常，但语音对字形加工的影响不存在异常。从理论意义上来讲，本研究的结果进一步对汉语 DD 的形-音整合缺陷的内在机制进行了阐释，即 DD 字形加工能力较差，进而字形对语音加工影响存在异常，导致阅读困难。从实践意义上讲，本研究的结果为今后 DD 阅读能力的矫治提供了理论基础，即可以通过训练 DD 字形识别和其利用字形加工语音的能力，进而提高其阅读能力。

本研究也存在一定的局限性。第一，本研究以汉语 DD 儿童为被试，并设置了同年龄对照组，但没有招募同阅读水平的正常儿童作为对照组。这就导致在形-音整合过程中，DD 儿童字形对语音加工的影响存在困难这一现象可能是源于其缺乏足够的阅读经验，而不是 DD 儿童固有的缺陷。将来的研究需要进一步与同阅读水平对照组进行比较，以明确 DD 字形对语音加工的影响的困难是否是其固有缺陷。第二，本研究没有考虑时间间隔的可能影响。已有研究发现 DD 当视听刺激的时间间隔变大时，其仍然具有一定的视听整合能力，正常读者的视听整合能力则消失(Francisco et al., 2017; Hairston et al., 2005)。那么，如果形音刺激有一定时间间隔时，其字形和语音的相互影响又会有怎样的表现呢？后续需要更系统的操作这些影响因素，全面探查 DD 在形-音整合过程中字形和语音的相互影响。”

第二轮

审稿人 1 意见：

作者对评审人提出的意见进行了细致地一一回答，并对论文进行了认真修改。论文相对初稿有了很大的提升。我只剩一个小的建议：建议作者把对部分问题的回答整合进对文稿的修改中，而不仅仅是回答问题。比如年龄段选择的理由、实验材料选择 3-5 年级课本的字的理由等等。

回应：感谢审稿专家的建议，我们已经结合审稿人上次和本次的意见，在相应的位置补充了相关的信息，并使用紫色字体标注，具体可见：

P.4 line 4-line 5

“施威茨(2019)指出一年级和二年级儿童阅读障碍的诊断存在一定的不确定性。因此，本实验包括 59 名三到六年级汉语儿童，符合统计要求。”

P.4 line 26-P.5 line 1

“考虑到要保证儿童被试对实验材料具有一定的熟悉性，研究中的真字均选自小学三年级至五年级语文课本。”

P.8 line 12-line13

“形音一致条件和形音不一致条件呈现的汉字在字频（根据《筮骏现代汉语单字频率列表》）和笔画数匹配。”

审稿人 2 意见：“综合以上的结果，本研究发现 DD 儿童在形-音整合过程中，字形对语音加工的影响存在异常，但语音对字形加工的影响与正常儿童相同。”这个结论不像个结论，需要把这个发现背后的意义提炼出来作为研究结论，上面这个只是对研究结果的描述。

回应：感谢审稿专家指出问题，我们已经修改完善了结论部分的内容，并使用紫色字体标注，具体可见：

P. 13 line 10-line 14

“综合本研究的结果，主要得出如下结论：汉语 DD 儿童形-音整合缺陷主要来源于更差的视觉加工技能，表现为 DD 儿童字形对语音加工的影响更弱。”

第三轮

编委复审意见：两位评审专家都认为作者做了认真修改，文章质量有了较大的提高，但仍有一些地方需要修改，比如结论的表述、被试年龄的选择以及实验材料等方面，建议修改后接受发表。

回应：感谢编委会专家的建议，我们已经结合审稿人和编委会的意见，对结论的表述、被试年龄的选择以及实验材料部分的内容进行了修改，具体如下：

P.4 line 12-16

“实验 1 使用 G*power 3.1 软件(Faul et al., 2007)，计算研究所需的样本量，将效应量设置为 0.25， α 设置为 0.05。计算结果表明，为了达到 0.95 的统计检验力，实验 1 至少共需要 54 名被试。有研究者指出一年级和二年级 DD 儿童的诊断存在一定的不确定性（施威茨，2019）。因此，实验 1 包括 59 名三至六年级的儿童，其中 DD 儿童 30 名（男 25 名，女 5 名），同年龄正常对照组（CA 组）儿童 29 名（男 18 名，女 8 名）。”

P.5 line 1-9

“研究使用的真字均选自小学三年级至五年级语文课本。假字是在真字基础上，打乱部件顺序重新组合成符合正字法规则，但实际并非真实存在的汉字刺激（例如：扌同）。所有真字和假字均为左右结构，字体为“黑体”，大小为3cm×3cm。真假字在笔画数、部件频率上匹配。在实验过程中，每个真字在三种条件（形音一致、形音不一致和视觉单通道）下各呈现1次，每个条件30个试次，共90个试次；每个假字在跨通道条件（假字和字音一同呈现，例如：扌者-jǐǎ）和视觉单通道条件下各呈现1次，共90个试次。正式实验共180个试次，包括真字和假字各90试次。听觉刺激材料为105个真实存在的字音，包括30个与视觉刺激材料里的真字发音一致的字音、30个与视觉刺激材料里的真字发音不一致的字音和45个与假字同时呈现的字音。”

P.8 line 6

“正式实验共包括197个试次（包括17个捕捉试次）”

P.13 line 10-12

“综合本研究的结果，本研究主要得出如下结论：汉语DD儿童形-音整合缺陷主要来源于更差的视觉加工技能，具体表现为DD儿童字形对语音加工的影响与正常儿童相比更弱，但语音对字形加工的影响与CA组不存在差异。”

主编意见：借助视听跨通道的字形判断任务，本论文对阅读障碍儿童在形音整合过程中字形和字音的相互影响进行了考察。本论文的选题具有一定新颖性，获得的研究结论具有较强的理论和现实意义。