

《心理学报》审稿意见与作者回应

题目：小学低段汉字识别和听写的发展轨迹：语素意识的预测作用

作者：李利平 伍新春 程亚华

第一轮

审稿人 1 意见：

汉字识别和听写是字词水平上两项重要的读写技能。本研究对 127 名一年级儿童进行了两年四个时间点的追踪调查，通过测查儿童的一般认知能力、语音意识、正字法意识和语素意识（包括同音语素意识、复合语素意识），探讨影响儿童汉字识别和听写发展的因素。作者尝试利用潜变量增长模型，考察了低年级儿童汉字识别和听写的发展轨迹，发现汉字识别的线性发展轨迹和听写的曲线发展轨迹，语素意识对儿童汉字识别和听写发展的作用不同。研究的问题明确，方法比较的当新颖，对结果的解释清晰、学术意义较高。以下问题请思考、补充、修改、说明：

意见 1：在文献综述部分（无论拼音文字……）提到了无论拼音文字还是汉语研究都存在“语素意识”对识字能力的影响，但是本文只列举了汉语研究中“语素意识”对识字能力的影响，并没有介绍拼音文字中二者之间的关系研究。

回应：非常感谢专家的建议，使文章逻辑表述更加完整。作者已在文献综述部分第四段添加了关于拼音文字的相关研究：“例如，Apel 和 Lawrence (2011)调查了一年级英语儿童，发现语素意识预测了阅读正常儿童的字词识别能力；Shu 等人(2006)研究也表明语素意识是汉语儿童字词识别的强有力的预测因子，预测的路径系数达到 0.51”，见综述第四段红色字体部分。

意见 2：同样在文献综述部分（语素意识不仅影响汉字识别……）也提到了无论拼音文字还是汉语研究都存在“语素意识”对字词听写的影响，但是本文只列举了拼音文字研究中“语素意识”对听写能力的影响，并没有列举在汉语文字中二者之间的关系研究。

回应：非常感谢专家的建议。根据建议，作者已在文献综述部分第五段添加了关于语素意识和汉语听写的相关研究：“此外，汉语听写的研究(李利平, 伍新春, 熊翠燕, 程亚华, 阮氏芳, 2016)发现在整个小学阶段语素意识均对儿童听写有显著影响，控制了智力、口语词汇、语音意识、正字法意识和快速命名后，语素意识分别解释了低、中、高年级听写的 15%、7%、3%的变异”，见综述第五段红色字体部分。

意见 3：尽管已有的研究表明在阅读中存在马太效应(Bast & Reitsma,1998; Stanovich, 1986)，也有研究表明存在补偿效应(Walczyk & Griffith - Ross, 2007)，那么在识字和听写方面请作者在已有研究的基础上大胆做出明确的实验假设。

回应：非常感谢专家的建议。在文中问题提出部分的最后一段，作者补充了研究假设：“鉴于实践中人们的感知经验认为儿童的起始点对后期发展起着重要作用，因此本研究提出假设 1：汉字识别和听写的发展呈现马太效应；同时，基于语素意识对读写水平的已有研究结果，本研究提出假设 2：语素意识影响汉字识别和听写的发展轨迹”。

意见 4: 正字法意识测验的描述不清楚,从后面的描述统计部分不难看出在本测验中作者只是采用了“假字”一个维度,请在这部分清楚阐明这一点。

回应: 非常感谢审稿专家的意见。根据专家意见,作者在 2.2.3 部分对正字法意识的计分方式进行了补充说明:“由于假字符合正字法规则,因此本研究中假字作为填充材料,其得分不计在正字法意识得分之内,正字法意识测验的最后得分仅包括位置错误、笔画乱写、部件错误三项”。

文中描述统计部分得知,一年级初儿童的正字法意识得分的确不高,平均分仅为 25.8 (总分为 45 分)。位置错误、部件错误、笔画乱写三项平均分分别为 7.69、5.53、12.50,可见刚入小学的儿童能很好地分辨笔画乱写的不是汉字,但是在位置错误、部件错误上得分较低,很容易把其辨识为真字,这和已有研究结果(李虹, 彭虹, 舒华, 2006)一致。一年级儿童正字法意识水平较低,但儿童的正字法意识发展迅速,到三年级时已达到较高水平。

参考文献:

李虹, 彭虹, 舒华. (2006). 汉语儿童正字法意识的萌芽与发展. *心理发展与教育*, (01), 35-38.

意见 5: 本研究关心的问题是儿童汉字识别和听写能力的发展趋势及其影响因素(语素意识),而从统计分析部分可以看出作者列举了其他影响因素(一般认知能力、语音意识、正字法意识、同音语素意识和复合语素意识),并且将一般认知能力、语音意识、正字法意识三者作为控制变量,同音语素意识和复合语素意识作为预测变量,来考察预测变量对汉字识别或听写的起始水平和发展速度的影响。请在综述部分列出关于这些变量的相关研究。

回应: 非常感谢审稿专家的建设性建议。作者已在综述部分列举了这些变量的相关研究:“鉴于已有研究(Apel & Lawrence, 2011; 李虹, 饶夏澍, 董琼, 朱瑾, 伍新春, 2011; Shu et al., 2006; Tong et al., 2009)认为一般认知能力、语音意识、正字法意识影响儿童的读写”。由于三者是控制变量,因此作者仅列举了相关的已有研究的文献,并没有详细陈述研究内容和结果。

意见 6: 还有一些小细节需要注意,例如 2.2.2 语素意识部分参考文献应用处多了一个括号。本研究利用潜变量增长模型探讨汉字识别与听写的发展轨迹以及语素意识对发展轨迹的预测作用,研究极具创新性,且相较于静态考察,儿童语言能力的动态考察更具实践意义。

回应: 非常感谢审稿专家的严谨和认真。作者已对多余的括号进行了删除,同时对文中的其它细节也进行了再次审核。

意见 7: 问题提出部分的文献综述中提到 Apel 等人(Apel, Wilson-Fowler, Brimo, & Perrin, 2012)的横断研究:考察了美国 88 名一年级儿童,结果发现语素意识显著预测了语音障碍儿童和正常儿童的听写测验成绩;但是文稿中的结果显示第一时间点的语素意识没有预测听写的初始水平。这里似乎有冲突,论文讨论部分详细分析了结果中听写成绩不受儿童初始语素意识影响的原因,但并未对与引用文章中出现的不同结果进行讨论。除此之外,综述中还提到 Pittas 和 Nunes(2014)研究结果显示一年级的语素意识预测了八个月后的听写,文中的结果是低年级(一年级初到二年级末的四个时间点的追踪)语素意识对听写发展没有显著影响,此处似乎也有冲突,但并没有讨论结果不同的原因。即使这几篇文章的研究对象是外语,也可从文字差异的角度做适当讨论。

回应: 非常感谢专家的建议,使文章前后呼应,结构上更加完整、严谨。作者已在讨论部分增加了:“此外,本研究结果与已有的研究结果(Apel et al., 2012; Deacon et al., 2009; Pittas & Nunes, 2014)并不一致。可能的原因是这些研究均是对某个时间点上静态的听写水平的预测,

语素意识预测了听写水平但是并不预测听写的发展轨迹。听写发展速度的快慢可能还受其他因素（如记忆、动机、阅读量等）的影响”的内容，来呼应引言中的文献研究。

意见 8：文章的亮点是首次运用潜变量增长模型，考察低年级儿童汉字识别和听写的发展轨迹以及发现二者的补偿效应。但是问题提出部分大多数篇幅是在总结前人关于语素意识的影响的研究，而前人关于汉字识别和听写的发展的研究是从语素意识的影响的研究中带出来的，占据较少的篇幅且位置不明显。读者可能会较难发现亮点。

回应：非常感谢审稿专家的建议，使文章结构更加合理同时也突出了文章的新意。作者已对问题提出部分的结构做了调整，由原来的“先论述语素意识对读写的影响，后论述读写发展的研究”调整为“先论述读写发展的研究，后论述语素意识对读写的影响”，这样也使得文章结构与文章标题更加一致。由于改动较大，因此请详见文中第二段的红色字体部分。

审稿人 2 意见：

《小学低段汉字识别和听写的发展轨迹：语素意识的预测作用》一文采用追踪研究设计就儿童汉字识别和听写的发展轨迹进行了探索，并检验了语素意识的预测作用。文章有一定的理论价值，但文章的统计分析部分存在显著的缺陷。

意见 1：“有 5 名被试在一般认知能力测验上有缺失，相关分析中采用了平均数替代的处理方式”这种做法是不妥当的，相关分析可以使用对删法。

回应：非常感谢审稿专家的建议。根据专家意见，一般认知能力测验上存在的缺失值在相关分析中已使用了 spss 软件默认的对删法(exclude cases pairwise)处理了缺失值。在文章表 1 中，相关数据结果发生了轻微变化，已做了修改，请见表 1 红色字体部分。

意见 2：样本量的估计方法不妥，本研究采用的是 LGCM，而 Gpower 并不适合估计该类型的样本量。可以采用蒙特卡洛模拟的方法估计。

回应：非常感谢审稿专家的意见。受作者统计知识的局限，作者先前在询问了擅长实验设计的同行后，错误运用了 Gpower 来估计样本量。审稿专家提出了蒙特卡洛模拟的方法后，作者认真查询了相关的知识(Muthén & Muthén, 2002; 王济川, 王小倩, 姜宝法, 2011)。目前在应用统计实践中，基于马尔科夫链蒙特卡洛算法的贝叶斯统计逐渐得到社会科学领域学者的重视（王孟成, 毕向阳, 2018）。运用 Mplus 软件实现样本量的估计过程，当统计功效为 0.80 时，采用蒙特卡洛模拟法估计的样本量为 100；当统计功效在 0.83 时，采用蒙特卡洛模拟法估计的样本量为 110。此外，LGM 也可采用 Sattaro-Saris 法估计样本量，当统计功效为 0.79 时，采用 Sattaro-Saris 法估计的样本量为 100；当统计功效在 0.82 时，采用 Sattaro-Saris 法估计的样本量为 110。通过这一问题的回答，作者扩大了统计知识面，非常感谢审稿专家！作者在自检报告中已做了修改。

参考文献：

Muthén, L. K., & Muthén, B. O. (2002). How to use a Monte Carlo study to decide on sample size and determine power. *Structural Equation Modeling*, 9(4), 599-620.

王济川, 王小倩, 姜宝法. (2011). *结构方程模型: 方法与应用*. 北京: 高等教育出版社, 249-276.

王孟成, 毕向阳. (2018). *潜变量建模与 Mplus 应用进阶篇*. 重庆: 重庆大学出版社, 289-319.

意见 3：在 LGCM 建模时作者采用如下策略的依据是什么？“在自由估计时，第一时间点斜

率的载荷为“0”，而第二、三时间点中斜率的载荷改为自由估计，第四时间点斜率的载荷改为“1”。

回应：非常感谢审稿专家的意见。已有文献（刘红云，张雷，2005；Meredith, & Tisak, 1990；王孟成，毕向阳，2018）提出了不定义曲线类型的潜变量增长模型，在估计斜率因子时，将其两个因子载荷分别固定为 0 和 1（模型识别的限定），其他因子载荷自由估计，这一模型常用于曲线形状不确定的情况，通过对数据的分析来确定曲线的类型，然后根据定义变化参数的因子载荷解释数据的发展特点，这一方法往往具有探索性价值。

由于已有文献（刘红云，张雷，2005；Meredith, & Tisak, 1990；王孟成，毕向阳，2018）把斜率因子的前两个载荷固定为 0 和 1，因此作者在文中更改为“第一、二时间点斜率的载荷分别为 0 和 1，第三、四时间点斜率的载荷为自由估计”，其在本质上与“第一时间点斜率的载荷为 0，第二、三时间点斜率的载荷自由估计，第四时间点斜率的载荷为 1”相似。作者已按照第一种方法对数据进行了重新分析，四个时间点的载荷数字大小发生了变化，由原来的“0、0.56、0.75、1”变为“0、1、1.33、1.77”，两种分析方法的数据结果趋势一致，作者已在文中对具体的数字大小进行了修改，也添加了相关的参考文献，详见 3.2 和 3.3 中的红色字体。

参考文献：

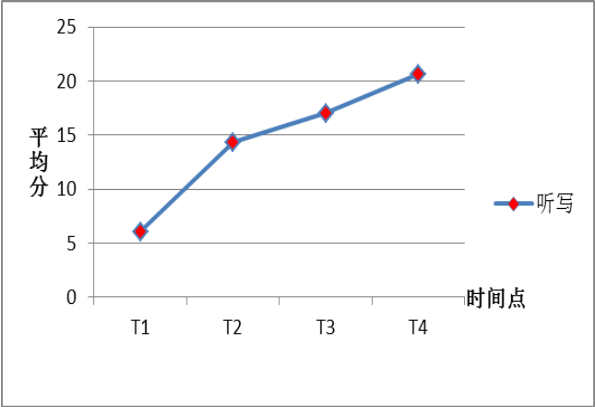
刘红云，张雷. (2005). *追踪数据分析方法及其应用*. 北京：教育科学出版社，P192-193.

Meredith, W., & Tisak, J. (1990). Latent curve analysis. *Psychometrika*, 55, 107-122.

王孟成，毕向阳. (2018). *潜变量建模与 Mplus 应用进阶篇*. 重庆：重庆大学出版社，P89.

意见 4：作者在拟合听写模型时，发现听写的线性模型不正定，作者得出听写为非线性关系，这个推论并不成立，不正定的因素有很多，样本量，测量次数等都可能造成不正定。建议将听写的发展轨迹通过可视化的形成呈现出来，用以简单评价增长趋势。

回应：非常感谢审稿专家的意见。诚如专家所言，作者在论文撰写初期的确进行了四个时间点平均数的折线图，如下图，发现折线图中第一到第二个时间点之间，听写的发展速率较大，因此判定听写的发展可能是曲线性关系，这样采用曲线发展来拟合。的确样本量、测验次数均会影响模型，但这些因素在研究结束的情况下已经无法改变，因此本研究选择了调整模型来拟合数据，作者在今后的研究中会注意这些因素。



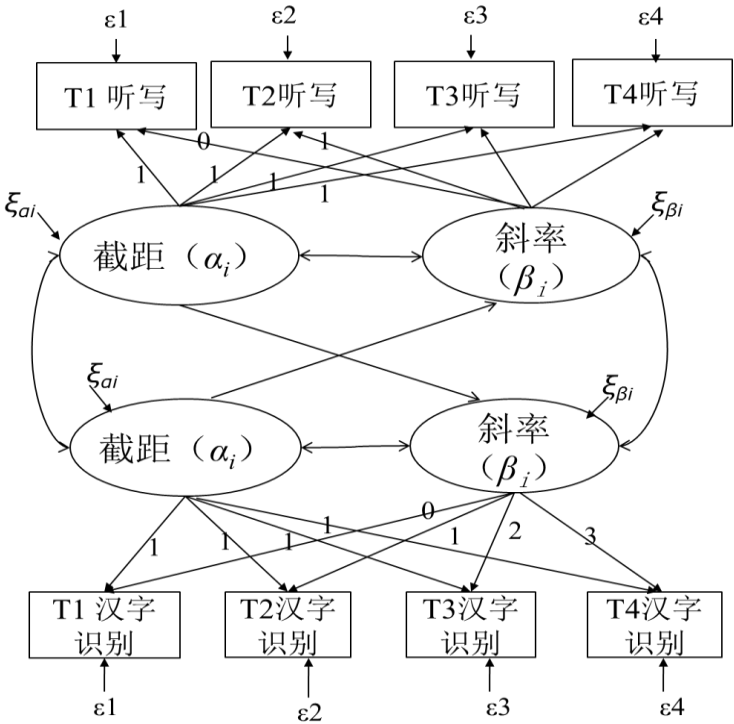
意见 5：关于方法还存在另外一个重要的问题，儿童的识别和听写通常是同时发展变化的，进行 LGCM 建模时应该将两者同时构建平行发展模型并探索两者之间的关系。

回应：非常感谢审稿专家的意见。诚如专家所言，读和写是两个关系密切的技能，二者关系的探讨是阅读领域的重点和热点。建构两者同时发展模型，可以探讨读的起始水平、发展速

度与写的起始水平、发展速度的作用关系。本研究没有应用平行发展模型探索两者之间的关系是基于以下几点考虑：

首先，目前已有较多研究（Fletcher-Flinn, Shankweiler, & Frost, 2004; Martin-Chang, Ouellette, & Madden, 2014）关注并得出了有意义的结果，已经初步达成一致，认为母语学习者的读会更多地影响到写，但写对后期读的影响作用甚微。作者所在的团队研究（***, Wu X. C., Cheng Y. H., & Nguyen, T. P. 2019）也发现汉语儿童的读影响后期的写。

其次，运用 LGCM 构建读和写两者平行发展模型，在一个整合模型中探讨读的起始水平、发展速度与写的起始水平、发展速度的关系，此模型（见下图）比较复杂，模型拟合指数为： $\chi^2(df=80)=271.29$, $p<0.001$, CFI=0.78, TLI=0.75, RMSEA=0.13 (90%CI=0.11~0.14), SRMR=0.18, 在模型拟合指数上不太理想。



最后，本研究重点探讨字层面的读写发展轨迹是什么、它们的发展是马太效应还是补偿效应，以及它们的发展轨迹是否受语素意识的影响，因此二者之间的相互关系不是本研究重点探讨部分。

参考文献：

Fletcher-Flinn, C.M., Shankweiler, D. & Frost, S.J. (2004). Coordination of reading and spelling in early literacy development: An examination of the discrepancy hypothesis. *Reading and Writing*, 17(6), 617–644.

Martin-Chang, S., Ouellette, G. & Madden, M. (2014). Does poor spelling equate to slow reading? The relationship between reading, spelling, and orthographic quality. *Reading and Writing*, 27(8), 1485–1505.

***, Wu X. C., Cheng Y. H., & Nguyen, T. P. (2019). The relationship of character reading and spelling: A longitudinal study in Chinese. *Journal of Research in Reading*, 42(1), 18-36.

意见 6：2.4 部分对 LGCM 的解释是错误的。尽管可以用多水平模型的思路分析和理解纵向数据，但用多水平的模型设置解释 LGCM 显然不妥。

回应：非常感谢审稿专家的意见。多水平模型设计之初是用于处理嵌套数据，重复测量存在

嵌套的特征,因此多水平模型也可处理重复测量数据(王孟成,毕向阳,2018)。统计学家(刘红云,张雷,2005)认为潜变量增长曲线模型(LGCM)通过因子(潜变量:截距和斜率)来描述重复测量变量的发展特征,其处理技术类似于含有均值的结构方程模型,用类似于验证性因子分析的方法来定义潜变量,用时间的函数确定因子载荷,斜率因子的载荷显示了变化的轨迹。文中的模型解释是基于已有著作(刘红云,张雷,2005)和一些实证研究(侯桂云,黎光明,谢晋艳,杨栋,2018;刘俊升,周颖,李丹,2013;杨立军,张小青,张薇,2016)而进行的。一稿中的模型描述、解释可能比较突兀,因此作者已在文中补充了相应的参考文献,详见2.4的红色字体部分。

参考文献:

- 侯桂云,黎光明,谢晋艳,杨栋.(2018).老年人认知功能的变化轨迹:基于潜变量增长模型的分析.《心理科学》,41(04),835-841.
- 刘红云,张雷.(2005).追踪数据分析方法及其应用.北京:教育科学出版社,P190-193.
- 刘俊升,周颖,李丹.(2013).童年中晚期孤独感的发展轨迹:一项潜变量增长模型分析.《心理学报》,45(02),179-192.
- 杨立军,张小青,张薇.(2016).大学期间学生学习成果增值轨迹:一项潜变量增长模型分析.《教育发展研究》,36(17):34-42.
- 王孟成,毕向阳.(2018).潜变量建模与Mplus应用进阶篇.重庆大学出版社,85-87.

意见 7: 文章中有很多标点符号错误。

回应: 非常感谢审稿专家的严谨和认真,作者已对文中的标点符号进行了仔细核查和修改。

意见 8: 英文摘要需要润色。

回应: 非常感谢审稿专家的建议,我们已经找了专业的润色公司对英文摘要进行了润色服务,摘要已按照润色内容进行了修改,润色证明已附在附录部分。

意见 9: 总的来说,文章使用的样本量并不大,统计方法存在不妥之处,研究的创新性也欠佳。

回应: 非常感谢审稿专家的意见。诚如专家所言,文章的样本量的确不是大样本,仅符合建立模型的最低数据要求,在今后的研究中作者会多加注意此问题。

在研究创新性方面,本研究首次运用追踪数据模型考察低年级儿童在字层面的读写两项能力发展的轨迹,发现汉字识别和听写的发展轨迹不同并且呈现补偿效应,在字层面的读写发展上,本研究数据结果并未支持人们经验中的“不让儿童输在起跑线上”的观点,为教育实践提供了一定的数据支持。

正如审稿人1的意见8,可能受限于作者的表述能力,有些地方表述欠妥,一稿文中逻辑结构欠佳,致使审稿专家误解,在此作者表示深深的歉意,作者会根据两位审稿专家的意见进一步对文章仔细修改和完善,使其重点和创新性更加突出。

第二轮

审稿人1 意见:

文章经过修改后,统计方法的使用更加严谨,文章架构也更加清晰。但是仍然存在部分表述不严谨;问题提出中缺乏理论基础;对结果的讨论不够充分,相关的实证研究、理论支

撑不足等问题。

意见 1: 问题提出中只是罗列了“语素意识对某一时间点的字词识别以及听写的影响”的相关研究，并未阐述其背后的理论基础。此外，为何提出研究假设 2“语素意识影响汉字识别和听写的发展轨迹”，也未能在文中找到相应的理论支撑。

回应: 非常感谢审稿专家的建议，作者在问题提出部分增加了相应的理论阐述，使文章的理论性更强。首先，在语素意识对某一时间点的字词识别以及听写影响部分，增加了语素分解理论，具体内容修改为：“尽管字词识别和听写有着不同的加工过程(Verhoeven & Carlisle, 2006)，但二者均存在语素分解。语素分解理论(Verhoeven & Perfetti, 2003)认为在心理词典中存在语素的表征，字词的识别和提取过程中存在对组成字词的最小语义单元——语素的分解加工，语素分解有助于提高字词识别和提取能力的发展”及“听写中语素及语素结构的分解加工(Verhoeven & Perfetti, 2003)促进了正确字形的输出”。详见问题提出中的第三段蓝色字体部分。

其次，在假设 2 的提出中，作者增加了“基于语素分解理论中语素分解促进读写能力的发展(Verhoeven & Perfetti, 2003)及语素意识对汉语儿童读写发展的重要性(李利平等, 2016; McBride-Chang et al, 2003; Tong et al, 2009)”来提出研究假设 2，增强了假设 2 的理论性。详见问题提出中的最后一段蓝色字体部分。

意见 2: 提出假设时使用的表述是“鉴于实践中人们的感知经验”，为什么不是基于已有的研究结论？此处表述不严谨。

回应: 非常感谢审稿专家的意见。作者已把原有表述“鉴于实践中人们的感知经验认为儿童的起始点对后期发展起着重要作用”修改为“鉴于拼音文字的实证研究结果(Carreker et al., 2007; Kempe et al., 2011; Pfost et al., 2012)”，详见问题提出部分中最后一段的蓝色字体部分。

意见 3: 讨论中对于补偿效应的第二点解释中的“起始点落后的儿童在非认知因素上准备较好”是否有相关研究进行支撑？第一点解释也略显牵强，不能很好地解释为什么起始水平低的儿童后期发展速度较快。

回应: 非常感谢专家的建议。针对讨论中补偿效应的解释，作者再次查阅了相关文献后，对原来的解释进行了修改和补充，由原来的“第一，系统的小学教育加速了起始点读写落后儿童的发展，他们通过学校教育得到更多的促进和提升；第二，起始点落后的儿童虽然在客观知识的掌握上初始水平较低，但是他们在非认知因素上准备较好，如他们可能比起始水平高的儿童有更强的学习动机、更浓的学习兴趣，这些非认知因素促进了后期读写的快速发展”修改补充为现在的“这可能是因为：第一，语言技能不同层面的高低限制程度不同，字词的读写属高限制技能(Paris, 2005)，因字词的范围、数量有限，儿童可以在短时间内掌握，起点低的儿童后期的提升速度快，而起点高的儿童后期提升空间较小，在字词读写技能方面个体间差异随时间会不断缩小。第二，学校教育环境和资源加速了起始点读写落后儿童的学习和成长。已有研究(张启睿, 边玉芳, 王烨晖, 苑春勇, 2012)发现学校教育环境和资源是学生学业成就的重要预测变量。起始点落后的儿童在接受了学校教育后，潜在的学习能力得到开发和发展，进而学习技能得到迅速提升。第三，起始点落后的儿童虽然在客观知识的掌握上初始水平较低，但是他们在非认知因素上可能准备较好，如他们比起始水平高的儿童有更强的学习动机、更浓的学习兴趣，在同伴关系、生活习惯上处理较好，能快速适应小学生活方式(马健生, 陈元龙, 2019)，这些非认知因素促进了后期学业的快速发展。”详见 4.1 中蓝色字体部分。

参考文献：

- Paris, S. G. (2005). Reinterpreting the development of reading skills. *Reading Research Quarterly*, 40, 184–202.
- 张启睿, 边玉芳, 王烨晖, 苑春勇. (2012). 学校教育环境与资源对青少年学业成就的影响. *教育研究*, 33, 32–40.
- 马健生, 陈元龙. (2019). 学前教育小学化:困惑与澄清——基于“儿童发展中心”的分析. *北京师范大学学报(社会科学版)*, (04), 5–14.

意见 4: 讨论中提到的“本研究结果与已有的研究结果(Apel et al., 2012; Deacon et al., 2009; Pittas & Nunes, 2014)并不一致。可能的原因是这些研究均是对某个时间点上静态的听写水平的预测, 语素意识预测了听写水平但是并不预测听写的发展轨迹”, 既然本研究针对的是发展轨迹, 而引用的研究针对的是听写水平, 因变量不一样, 不存在结果不一致, 文中的表达不严谨, 直接对出现的两种结果进行解释即可。此外, 文中的讨论不能很好地解释为什么语素意识预测了听写水平但是并不预测听写的发展轨迹。

回应: 非常感谢审稿专家的意见, 使得文章表述更加严谨。根据专家意见, 作者针对静态和动态的两种结果分别进行了解释, 修改为“此外, 虽然已有研究(Apel et al., 2012; Deacon et al., 2009; Pittas & Nunes, 2014)显示了语素意识对某个时间点上静态的听写水平的预测, 但是本研究结果显示语素意识并不预测听写的发展轨迹, 听写的发展轨迹可能受其他因素(如记忆、动机、阅读量等)的影响”。详见 4.2 中蓝色字体部分。

此外, 可能由于作者的表述问题, 致使专家误认为“语素意识预测了听写水平但是并不预测听写的发展轨迹”。本研究结果显示第一时间点的语素意识对听写的起始水平和发展速度均没有显著影响。作者已对表述方式进行了修改, 详见 4.2 中蓝色字体部分。

.....

审稿人 2 意见:

《小学低段汉字识别和听写的发展轨迹: 语素意识的预测作用》一文经过作者修改后整体质量有了显著提升, 但在一些细节上还需要作者进一步完善, 具体如下:

意见 1: 汉字识别测验的在本研究中的信度系数需要报告, 文章只引用了别人的结果, 这样会产生信度引入的问题。

回应: 非常感谢审稿专家的意见。这一点的确有我们考虑不周的地方, 特此解释如下:

本研究是一项汉语儿童阅读追踪研究项目中的一部分。该项目第一次测量时间点(T1)收集了 1 到 6 年级的横断数据, 并随后对一年级进行了为期 6 年 8 个时间点的长程追踪。在该项目中, 涉及到诸多变量, 包括一般认知能力、家长问卷、阅读相关认知技能以及阅读水平等, 其中汉字识别测验(Li et al., 2012)由 150 个汉字构成, 这 150 个汉字中, 其中前 40 个来自已有幼儿汉字识别的测验(Shu et al., 2008), 并邀请了两位幼儿教师进行评定, 可以适用于一年级儿童, 另外邀请了两位小学语文教师对其余 110 个汉字进行了评定, 其中 100 个汉字选自二至六年级语文教科书, 每个年级 20 个(Shu et al., 2003), 最后 10 个汉字是教科书中没有出现过的。这 150 个汉字依难度水平(以预研究为基础)进行排列。在 Li 等人的研究中, 其信度达到 0.99。我们 T1 的六个年级横断数据的信度分析中, 汉字识别测验在 6 个年级的信度范围在 0.97~0.99 之间, 与已有研究基本相同。因此, 我们认为汉字识别测验的信度在该项目的测量中是可以保证的。

由于我们在追踪数据的录入过程中发现, 每次都对 150 个汉字进行录入来计算其信度, 工作量巨大, 也很繁琐, 因而采用了计入总成绩的录入方法, 虽然避免了录入的繁琐, 但却带来了信度引用的问题。我们基于这是一个比较成熟的测验, 并且 T1 的数据分析也发现其信度可以保证, 在后期的数据录入过程中进行了简化处理, 我们已经认识到这种处理方式

过于轻率，非常感谢审稿专家细致而严谨的提问，在今后的研究中我们一定注意此类问题。
参考文献：

- Li, H., Shu, H., McBride-Chang, C., Liu, H., & Peng, H. (2012). Chinese children's character recognition: Visuo-orthographic, phonological processing and morphological skills. *Journal of Research in Reading*, 35, 287–307.
- Shu, H., Chen, X., Anderson, R.C., Wu, N. & Xuan, Y. (2003). Properties of school Chinese: Implications for learning to read. *Child Development*, 74, 27–47.
- Shu, H., Peng, H. & McBride-Chang (2008). Phonological awareness in young Chinese children. *Developmental Science*, 11, 171–181.

意见 2：对第一轮意见 6 的回复稍有不妥，尽管在处理追踪数据时多水平模型和 LGCM 等价，但使用多水平的语言去描述 LGCM 有点别扭，建议修改。

回应：非常感谢审稿专家的意见。作者已对 LGCM 的描述进行了修改，例如，增加“参数”的表述、删去了“方程分别为 $\alpha_i = \mu_\alpha + \gamma_{a1}\chi_{i1} + \gamma_{a2}\chi_{i2} + \gamma_{a3}\chi_{i3} + \gamma_{a4}\chi_{i4} + \gamma_{a5}\chi_{i5} + \zeta_{ai}$ ； $\beta_i = \mu_\beta + \gamma_{\beta1}\chi_{i1} + \gamma_{\beta2}\chi_{i2} + \gamma_{\beta3}\chi_{i3} + \gamma_{\beta4}\chi_{i4} + \gamma_{\beta5}\chi_{i5} + \zeta_{\beta i}$ ，其中 χ_{i1} 、 χ_{i2} 、 χ_{i3} 、 χ_{i4} 、 χ_{i5} 分别为被试 i 的一般认知能力、语音意识、正字法意识、同音语素意识和复合语素意识”的表述。详见 2.4 蓝色字体部分。

第三轮

主编意见：第二位审稿人提出一些方法问题，作者修改是否达到了审稿人的要求？请审稿人再看一下。

审稿人 2 回应：已对我提出的问题进行了认真回复和修改，谢谢。

第四轮

主编意见：稿件经过修改，已经达到可刊发的水平，同意发表。