·儿童学习能力与家庭养育环境·

基于 Rasch 模型的普通话高年级小学生听写测试字表的编制

冯琳，王梦寰，梁丹丹

南京师范大学文学院，江苏 南京 210097

摘 要：目的 通过 Rasch 模型构建适用于五、六年级学生的普通话听写能力评估工具，并经测量学优化提升读写障碍和书写障碍儿童的筛查准确性。**方法** 从部编版语文教材中低学段生字库中选取 180 个汉字构建初始测试题库，2023 年 11 月对江苏省南通市某小学 595 名五、六年级学生进行测试，采用 Rasch 模型进行题项筛选与模型优化。**结果** 经筛选后保留 163 个拟合度优良的题项。正式字表信度优异（题项信度 0.975，样本信度 0.956），分离度高（题项分离度 6.32，样本分离度 4.66）。单维性检验（第一/第二成分特征根比值 9.545，Smith's *t*-test 显著差异比例 5.88%）与局部独立性假设（MADaQ3=0.042）均成立；怀特图显示题项难度与目标群体能力匹配良好，测验信息函数在低能力区间（$\theta \approx -2$）达到峰值（信息量>25），提示其对读写障碍和书写障碍儿童筛查效能更优。字表内部一致性良好（Cronbach's $\alpha=0.975$），评分者间一致性高（ICC=0.998）。ROC 分析显示，该字表在本研究样本中对书写困难风险儿童具有较高区分能力（AUC>0.99）；以标准分 37 为截断值时，敏感度为 100%，特异度为 98.8%。**结论** 普通话高年级小学生听写测试字表具有良好的信效度和筛查敏感性，为读写障碍和书写障碍的早期识别提供了客观测量框架，并为计算机自适应测验开发奠定基础。

关键词： 高年级小学生；Rasch 模型；听写能力；语言筛查

中图分类号：R179 **文献标识码：**A **文章编号：**1008-6579(2026)08-0890-07 doi:10.11852/zgetbjzz2025-0839

Development of a Mandarin Dictation Test Word List for Upper-grade Primary School Students based on the Rasch model

FENG Lin, WANG Menghuan, LIANG Dandan

School of Chinese Language and Literature, Nanjing Normal University, Nanjing, Jiangsu 210097, China

Corresponding author: LIANG Dandan, E-mail: ldd233@163.com

Abstract: Objective To construct a mandarin dictation ability assessment tool for fifth- and sixth-grade students using the Rasch model, and to enhance the screening accuracy for children with developmental dyslexia and dysgraphia through psychometric optimization. **Methods** A preliminary item bank of 180 Chinese characters was constructed from the lower grade character database of the Ministry of Education's Chinese textbooks. In November 2023, the test was administrated to 595 students in Nantong, Jiangsu Province. Rasch model analysis was used for item selection and model optimization. **Results** A total of 163 well-fitting items were retained after the screening process. The final word list demonstrated excellent reliability (item reliability=0.975, person reliability=0.956) and high separation indices (item separation=6.32, person separation=4.66). The assumptions of unidimensionality (ratio of the first to second component eigenvalues=9.545; percentage of significant differences in Smith's *t*-test=5.88%) and local independence (MADaQ3=0.042) were robustly supported. The Wright Map indicated good alignment between item difficulty and person ability distribution. The test information function peaked in the low-ability range ($\theta \approx -2$, information value > 25), suggesting superior screening efficacy for children with dyslexia and dysgraphia. The word list showed excellent internal consistency (Cronbach's $\alpha=0.975$) and high inter-rater reliability (ICC=0.998). ROC analysis indicated high discriminative performance for identifying children at risk of writing difficulties in this sample (AUC>0.99). Using a standard score of 37 as the cutoff, the sensitivity and specificity of this test were 100% and 98.8%, respectively. **Conclusions** The Mandarin Dictation Test Word List for Upper-grade Primary School Students demonstrates good performance in terms of reliability, validity, and screening sensitivity. It provides an objective framework for early identification of dyslexia and dysgraphia, and lays the foundation for computerized adaptive testing.

Keywords: upper-grade primary students; Rasch model; dictation ability; language screening

基金项目：国家社会科学基金重点项目（22AYY013）

作者简介：冯琳（2002—），女，硕士研究生在读，主要研究方向为中文读写障碍。

通信作者：梁丹丹，E-mail：ldd233@163.com

汉字书写能力是读写素养的核心要素，也是学龄儿童学业发展的基础。部分儿童因神经发育障碍出现听写困难，属于发展性书写障碍（developmental dysgraphia）；若同时存在阅读困难，则诊断为发展性读写障碍（developmental dyslexia）。二者听写表现均显著落后于同龄人，提示听写任务可作为有效筛查指标。现有中文听写筛查工具存在明显局限：多数标准化测验聚焦中低年级[1-2]，缺乏高年级分年级设计；非标准化工具系统性不足[3]；而覆盖高小阶段的香港小学生读写障碍测验（The Hong Kong Test of Specific Learning Difficulties in Reading and Writing for Primary School Students, HKT-P）[4]听写词汇与内地存在文化差异[5]。此外，当前评估多采用经典测量理论的简单计分方式，其结果依赖特定样本与测验，缺乏客观性。相比之下，Rasch 模型能实现能力与难度参数的分离估计，确保测量的样本独立性和测验不变性[6]。因此，本研究基于 Rasch 模型对单字听写测验进行测量学优化，旨在构建适用于高年级小学生、具有良好心理计量特性的书写能力筛查工具。

1　对象和方法

1.1　*对象*　本研究于 2023 年 11 月选取江苏南通某小学五、六年级全体学生为研究对象，采用自编听写测验和同字认读测试进行施测。障碍组纳入标准（满足以下任一条）：1）书写障碍与读写障碍　听写正确字数低于常模均值 1.5s；2）听写障碍　认读正确率＞95%且读写差异率高于平均水平 1.5s（改自丁杨等[7]）。正常组：除障碍组外的其余正常发展儿童。本研究经过南京师范大学生物医学研究伦理审查委员会审批（NNU202310026），所有参与者均知情同意。

1.2　*量表编制*　依据《义务教育语文课程标准（2022 年版）》第三学段要求，采用分层抽样法从小学语文教材（2020 版）一、二学段生字表中抽取 10%（共 180 个，其中第一学段 81 个，第二学段 99 个）作为目标字构建题库。对目标字以下 4 个维度进行控制：1）笔画复杂度　一～四年级生字均值 8.53 画，目标字均值 9.63 画；2）字频　参考 CCLOOW 数据库[8]，一～四年级平均字频为 5.14×10^{-4} 次/10^6，目标字均值为 3.50×10^{-4} 次/10^6；3）结构　上下、左右及其他结构占比分别为 26.10%、53.89%、20.00%，分布特征与一、二学段生字表一致；4）形声字规则性　依据舒华等[9]的标准，控制规则、半规则、不规则及非形声字占比分别为 23.9%、16.7%、30.0%、29.4%。在语文教材中选取目标字组成的双字词（排除同音词）并伪随机排列。由普通话一级乙等的专业人员进行标准化录音。

1.3　*施测方法*　采用班级集体施测法。所有主试均接受规范化培训，确保指导语及流程高度统一。通过多媒体设备播放标准化录音。正式施测前设置 2 例练习题以熟悉作答规范。测试时，学生根据语音指令（如‘谐’，‘和谐’的‘谐’）在答题纸指定方框内书写，并对无法作答的题项进行标记，录音间隔 15 秒。

1.4　*评分标准与质量控制*　随机抽取 30 份试卷进行预评分，并制定以下评分细则：零反应、部件错误、笔画增减以及替代性错误（如形近替代、音近替代）等均判定为错误（记 0 分）；笔画及结构不够美观但符合字形拓扑结构的判定为正确（记 1 分）。评分过程采用“盲评法”（密封卷头）以减小主观偏倚。随机抽取 20% 样本由两名评分者独立评分，计算得 Cohen′s Kappa 系数为 0.94；争议项由专家最终裁定。

1.5　*统计学方法*　鉴于测验条目为二元计分，采用 Rasch 二元模型（Rasch Dichotomous Model, RDM）进行项目分析，剔除拟合不良题项，并对正式量表进行心理测量学检验。

1.5.1　Rasch 模型分析　通过 Jamovi 2.4.8 以及 R 4.2.0 的 TestAnaAPP 工具包[10]进行 Rasch 模型分析。Rasch 模型分析流程采用 Aryadoust 等[11]的建议，步骤如下：1）初步分析，基于原始数据进行 Rasch 建模。2）模型优化，采用 Infit Mean Square（Infit MNSQ）和 Outfit Mean Square（Outfit MNSQ）对题项拟合度进行评价，删除拟合度欠佳的题项。Infit MNSQ 检测受试者能力水平附近的不稳定反应，而 Outfit MNSQ 对远离受试者能力水平的极端反应更为敏感。Infit 和 Outfit 的 MNSQ 值介于 0.5～1.5 之间表示拟合良好[12]。3）模型校准，重复步骤（2）循环运行 Rasch 分析，直至无新增不达标题项，题项的估计参数稳定。4）基于最终剩余的题项建立常模题库，并汇报参数。Rasch 模型评价包括 3 个方面：1）模型的假设检验，包括局部独立性检验、单维性检验、模型拟合度检验等；2）模型的参数估计，包括题项的难度估计值、受试者的能力估计值及其分布范围；3）模型的可靠性，包括题项和受试者的信度和分离系数，以及测验信息量。采用 Yen′s Q3 残差相关系数反映各个题项间的局部独立性[13]，用中心化 Q3 绝对值的平均值（mean of absolute values of centered Q3, MADaQ3）反映此模型局部独立性的整体特征[14]；采用主成分分析（principal component analysis, PCA）对模型的单维性进行度量；采用 M2 统计量、近似均方根误差（root mean square error of approximation, RMSEA）、标准化均方根误差（standardized root mean squared

residual，SRMSR）、比较拟合指数（comparative fit index，CFI）与 Tucker-Lewis 指数（Tucker-Lewis index，TLI）对模型的拟合度进行检验，至少一项满足拟合要求的指标即认为模型拟合检验通过。

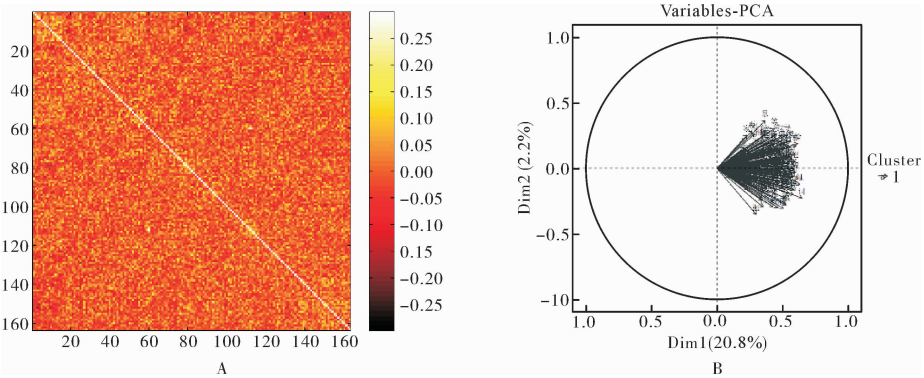
1.5.2 心理测量学检验 基于《选择健康测量工具的共识标准》（consensus-based standards for the selection of health measurement instruments，COSMIN）框架，对正式字表的心理测量学属性进行考察。1）采用 Cronbach′s α 系数评估字表的内部一致性；2）随机抽取约 50% 的样本（$n=310$），由第二位评分者对测试进行二次打分，计算字表总分的组内一致性系数（intra-class correlation coefficients，ICC）以评估评分者间信度；3）受试者个人能力与总分之间的相关性采用 Pearson 相关分析，以检验字表的同时效度；4）用 PCA 分析、Yen′s Q3 残差相关系数和 MADaQ3 检验字表的结构效度（Rasch 建模已涉及）；5）用双因素方差分析对不同年级典型发展儿童表现和书写障碍与读写障碍儿童表现进行比较，以考察字表的预测效度；6）绘制正式字表总分的受试者工作特征（receiver operating characteristic，ROC）曲线，计算曲线下面积（area under curve，AUC）以分析字表的评估准确性；依据 ROC 曲线分析筛查被试者的阳性截断值以评价字表的反应度。

2 结 果

2.1 研究对象的一般资料 共发放问卷 603 份，回收有效问卷 595 份，有效回收率 98.7%；其中五年级 310 份，有效率 98.7%；六年级 285 份，有效率 98.6%。两组性别与年级分布具体为：正常组 537 人，其中五年级女生 138 人、男生 141 人，六年级女生 131 人、男生 127 人；障碍组 58 人，其中五年级女生 8 人、男生 23 人，六年级女生 10 人、男生 17 人。

2.2 Rasch 建模结果

2.2.1 Rasch 模型的假设检验 Rasch 模型经 4 轮迭代剔除 16 个拟合度不良项，最终保留 163 项题目形成正式字表，正式字表的模型检验参数结果：模型拟合度检验结果显示，M2=16 459.89（$P<0.001$），SRMSR 为 0.075，RMSEA 为 0.02，且 TLI 与 CFI 均为 0.99，表明模型具有较好的拟合度。局部独立性检验结果发现，测验整体 MADaQ3 = 0.042（<0.10），maxaQ3=0.312，符合 Romaní 的标准[13]；除“危”和“险”字对受高频共现影响导致其 Yen′s Q3 残差相关系数略高于 0.3（0.312）外，其余题项间均不存在局部依赖性。见图 1A。单维度分析显示，模型的第一个成分解释了总方差的 20.834%，与第二个成分特征值之比为 9.545（>5），可认为测验主要由单一潜在特质所主导[15]。残差主成分分析的变量图显示，各题项向量几乎均指向第一残差维度（Dim1）的同一方向，整体呈单向扇形分布，无明确题项聚类或方向性分化，且未发现与语义类别、字形特征或词频属性等内容维度相对应的稳定分化模式，提示残差中未形成具有可解释意义的次级维度。见图 1B。进一步采用 Smith′s t-test 对单维性进行验证。根据第一残差主成分载荷，选取载荷最高与最低各 25% 的题项构成两个对立题项子集，并分别估计被试在两组题项上的 Rasch 能力参数。通过对个体能力估计差异进行配对 t 检验，结果显示仅有 5.88% 的被试在两组题项上的能力估计差异有统计学意义（$|t|>1.96$），处于 Smith 提出的 5%～10% 判据范围[16]。综上，该字表满足 Rasch 模型单维性假设，具有很好的结构效度。



注：A 为 Q3 残差相关矩阵；B 为 PCA 变量相关图

图 1 Rasch 模型的假设检验分析

Fig. 1 Analysis of Rasch model hypothesis tests

2.2.2　Rasch 模型的参数估计　Rasch 模型分析结果显示，题项难度的平均值为－1.95（范围－5.504～1.401）logit，而受试者能力的平均值为 0.019（范围－4.323～3.039）logit。由怀特图可见，除极少数高能力受试者外，题项难度水平与受试者能力分布基本匹配且均接近正态分布，提示量表内容覆盖度良好，符合筛查工具要求。其中，题项“国”难度最低，“庭”和“秩”难度最高。见图 2。

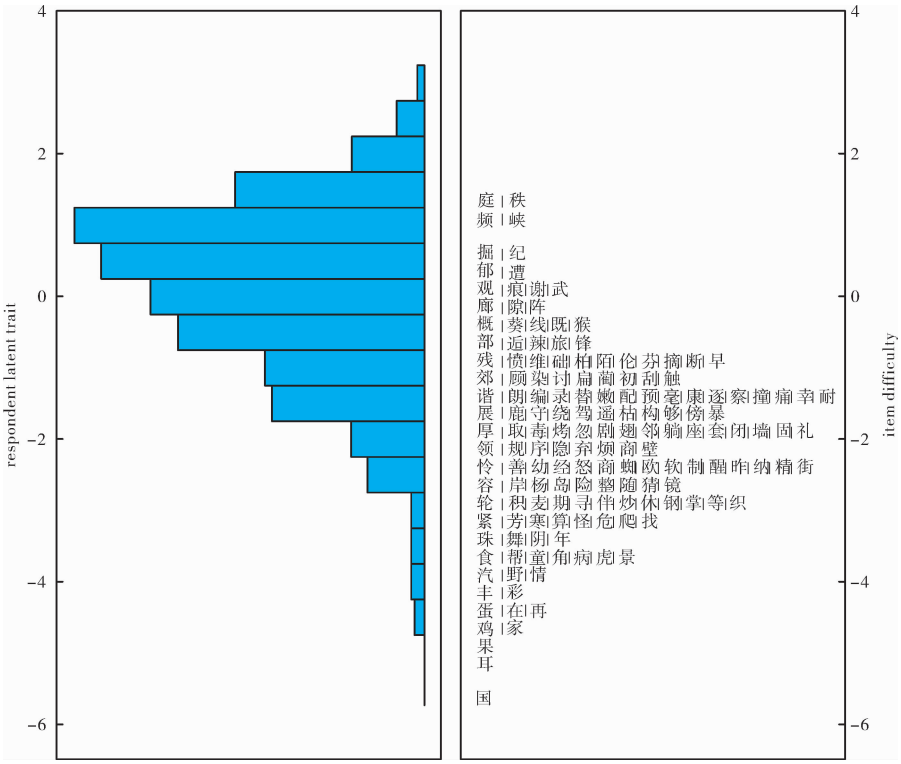


图 2　听写字表的怀特图分析

Fig. 2　Wright map analysis of the dictation word list

2.2.3　Rasch 模型的可靠性检验　听写测验题项与样本的信度分别为 0.975 和 0.956，均在 0.9 以上，测验的可信程度较高；题项与样本的分离度分别为 6.32 和 4.66，均满足 0.9 置信水平对应的临界值（被试分离度≥2、测验分离度≥3）[12]，表明字表区分不同能力水平被试的效能良好。见表 1。

表 1　样本及题项的信度、分离度、难度/能力值拟合结果

Tab. 1　Fitting results of reliability, separation, difficulty/ability values for samples and items

项目	信度	分离度	困难度/能力	
			平均值	范围
被试（n=595）	0.956	4.66	0.019	－4.323～3.039
题项（n=163）	0.975	6.32	－1.953	－5.504～1.401

测验信息曲线（TIC）及测量标准误差（SEM）曲线结果显示，信息曲线呈右偏态，在被试能力值约为－2logit 处达到峰值，提示该能力段的测量精度最高。依据判定标准（测验信息量＞25 为理想[17]），本测验在能力值（－3.025～－0.699）logit 区间内的信息量均＞25，表明该工具对低能力水平受试者具有极高的测量灵敏度，符合筛查书写困难儿童的预设目标。见图 3。

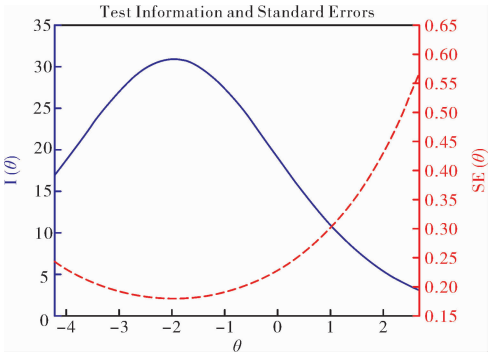


图 3　听写字表的测验信息函数

Fig. 3　Test information function of the dictation word list

2.3　心理测量学结果　字表的 Cronbach′s α 系数为 0.975（＞0.8），提示该字表的内部一致性极高。采用双向随机效应模型的绝对一致性定义计算 ICC，结果为 0.998（95% CI：0.933～0.999，P＜

0.001)，提示评估者间存在非常高的相关性。

以受试者字表原始得分与受试者能力水平进行相关性分析，显示字表得分与受试者能力相关性极其显著($r=0.98$，$P<0.001$)，表明该字表具有很高的同时效度。见图 4。

双因素方差分析结果显示组别主效应显著($F=801.898$，$P<0.001$，$\eta^2=0.575$)，而年级主效应($F=1.237$，$P=0.266$，$\eta^2=0.001$)和组别与年级的交互效应($F=1.409$，$P=0.236$，$\eta^2=0.001$)不显著。事后比较结果显示，五、六年级典型发展儿童得分均显著高于书写障碍与读写障碍儿童($t=21.588$，$P_{Tukey}<0.001$，$d=4.087$；$t=18.580$，$P_{Tukey}<0.001$，$d=3.758$)，说明字表能够有效反映正常组和障碍组儿童听写能力上的差异，具有良好的预测效度。

将字表原始得分转换为标准分，绘制书写障碍与读写障碍儿童诊断 ROC 曲线。结果显示 AUC>0.99(95%CI：0.999～1.000，$P<0.001$)，提示字表具有较高的筛出准确性。由于该字表以筛查为目的，选择敏感度最大(100%)时的标准分 37 作为截断值(转换原始分 90，样本占比 10.8%，$Z=-1.402$，$T=35.979$)，此时特异度为 98.8%，阳性预测值(PPV)为 90.5%，阴性预测值(NPV)为 100%。见图 5。

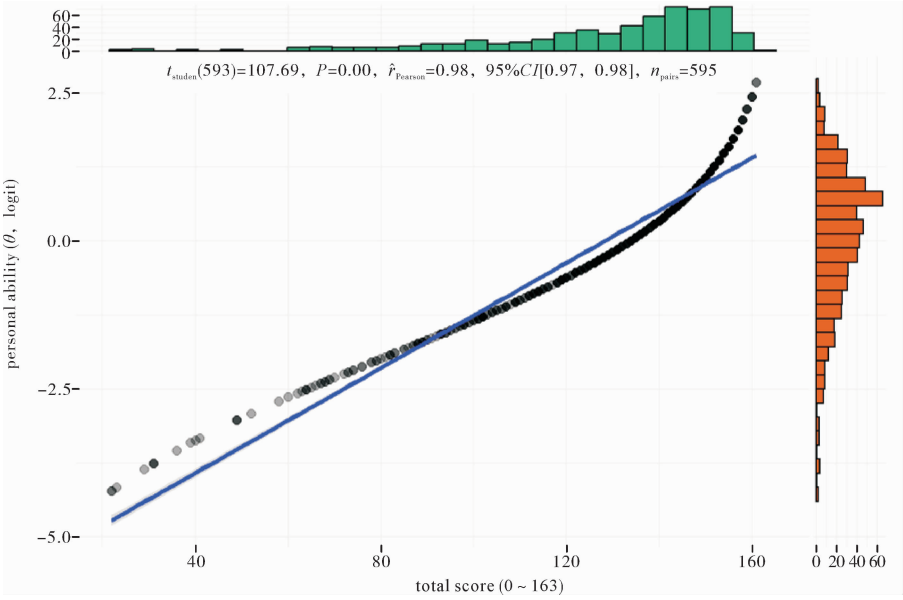


图 4 受试者字表总分与能力水平的相关性

Fig. 4 Correlation between total score and ability level

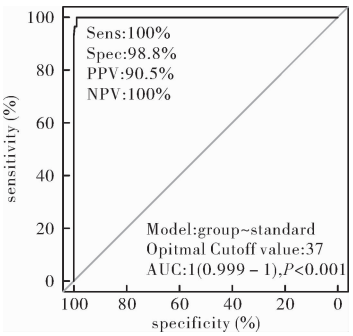


图 5 字表标准分的 ROC 分析和截断值

Fig. 5 ROC analysis and cutoff value for standardized word list scores

3 讨 论

3.1 听写测试的信效度及其结构合理性 本研究基于 Rasch 模型编制并验证了一套适用于高年级小学生的听写能力评估工具，听写测验题项选自小学语文教材中低学段生字库，通过笔画复杂度、字频、结构、形声字规则性四维度控制，确保测评内容与课程标准要求一致。最终结果表明，该字表在信度、结构效度和筛查敏感性方面均表现优异。听写任务是一种以语音-正字法映射和字形输出稳定性为核心的语言产出活动，其成功完成主要依赖于个体内部正字法表征的精确性与稳固程度[7]。与阅读理解或书面表达等多加工通道并行的复杂语言任务不同，听写任务中不同认知成分(如语音加工、语义激活)的变异在一定程度上被任务形式所压缩，最终主要体现在字形生成这一单一输出维度上。Rasch 模型分析结果显示，本测验的题项符合单维性和局部独立性假设，提示个体差异主要由单一潜在能力维度

主导，这一结构特征与正字法加工模型及既有关于汉字书写能力的实证研究结果一致。理想的测量工具应确保题项的平均难度与其目标群体的平均能力水平相匹配[18]。怀特图与测验信息函数的结果表明，测验题项难度与被试能力分布匹配良好，但难度较低的项目较多，且在低能力区间的信息量更为集中，说明该字表在能力水平较低人群中的区分精度更高。这一特征与本工具“筛查型”定位一致，即优先区分典型发展儿童与存在书写或读写障碍风险的个体，而非对正常群体内部进行精细分层。此外，测验信息函数在低能力区间（$\theta \approx -2$）的最大信息量高于理想阈值（25），进一步验证了其对书写障碍与读写障碍儿童的高敏感性，符合筛查工具的定位需求。

3.2　*听写测试对书写障碍与读写障碍儿童的筛查效用及其适用范围*　目前针对读写障碍儿童的书写能力筛查多为常模参照测验[19]，HKT-P[4]和普通话发展性阅读障碍量表（The Developmental Dyslexia Scale for Standard Mandarin，DDSSM）[2]的听写测验设置了连续错误达阈值即终止测试的机制，判定标准依赖题项难度设置而定。尽管 Lee 等[20]指出听写测验中的笔画、部件等错误分析对读写障碍识别具有诊断价值，但受限于编码的时间与人力成本，学校场景下的大规模筛查仍以正误判断的二元计分为主。针对书写障碍儿童的筛查常聚焦于手写运动特征[21]或依赖读写差异率，尚未建立独立的听写测验截断标准。鉴于此，本研究尝试通过 ROC 曲线分析探索听写测验的截断值，旨在建立统一标准以同步识别书写障碍与读写障碍风险群体。

基于这一目标，本研究通过 ROC 分析发现 AUC 值高于 0.99，表明该听写测验在本研究样本中具有极高的筛查准确性[22]。在以筛查为目的的应用情境下，本研究优先保证敏感度，将标准分 37 作为截断值，以最大限度识别潜在风险个体。在此阈值下，测验不仅保持了 100%的敏感度，同时特异度亦维持在较高水平，显示出良好的实际应用价值。需要指出的是，该截断值基于江苏地区样本及部编版教材体系确定，外推至不同地区、教材版本或教学进度的学生群体时仍需谨慎。Rasch 模型所提供的参数独立性为未来通过再校准或等值分析实现跨样本应用奠定了基础，但当前结果仍有必要在独立样本中进一步验证，以确保其稳定性并避免对高 AUC 结果的过度解读。在适用范围上，本字表最适合用于五、六年级学生在学校情境下的快速筛查，尤其适用于识别书写能力明显落后的个体。对于低年级学生，由于书写能力尚处于快速发展阶段，其测量结果更易受到教学进度和发展波动的影响，需结合纵向随访与其他评估工具进行谨慎判定。同样，对于存在明显方言影响或特殊教育需求的儿童，也建议将该工具作为综合评估体系中的一部分，而非单独的诊断依据。在此基础上，未来可通过扩充题库、覆盖各学段课程标准，进一步构建全学段书写能力测评体系，并引入计算机自适应测验技术（computerized adaptive testing，CAT），在保持筛查精度的同时提高测评效率。

利益冲突声明　本文不存在利益冲突

数据可用性声明　本研究的原始数据因涉及参与者隐私保护与伦理限制，无法公开共享。经脱敏处理后的汇总分析数据（包括样本分布、量表信效度及组间成绩差异等）已在论文中全面呈现；如需复现研究方法或验证结论，可向通信作者（邮箱等联系方式）提出合理申请，获取脱敏后的数据集。

AIGC 工具使用声明　在本研究的准备过程中，作者使用了豆包和 Gemini 来进行英文摘要翻译、文本润色及参考文献格式整理。使用此工具/服务的时间为 2025 年 7 月—2026 年 3 月，并通过“请将中文摘要翻译为规范英文摘要”“请润色论文讨论部分的语言表达”“请按期刊要求格式整理参考文献”等提示生成了相关文本。使用此工具/服务后，作者根据需要对内容进行了必要的审查和编辑，并对出版物的内容承担全部责任。

参考文献

[1] 张琳丹. 正字法意识、阅读能力与汉字书写能力关系研究[D]. 深圳：深圳大学，2020.
Zhang LD. The research of relationship between orthographic awareness, reading ability and Chinese character writing ability[D]. Shenzhen: Shenzhen University, 2020. (in Chinese)

[2] Liu Z, Wang L, Wang T, et al. The developmental dyslexia scale for standard mandarin: A study among early primary students[J]. Res Dev Disabil, 2024, 154: 104841.

[3] 孟祥芝. 65 名儿童中文书写评估及其类型初步分析[J]. 中国心理卫生杂志，2004，18(12)：833-835.
Meng XZ. Assessing and subtypes of writing in Chinese school children[J]. Chinese Mental Health Journal, 2004, 18(12): 833-835. (in Chinese)

[4] Ho CSH, Chan DWO, Chung KKH, et al. The Hong Kong test of specific learning difficulties in reading and writing for primary school students-Third Edition (HKT-P(III))[M]. Hong Kong: Hong Kong Specific Learning Difficulties Research Team, University of Hong Kong, 2015.

[5] Tsang YK. Dictation and vocabulary knowledge tests for adult native Chinese readers[J]. Behav Res Methods, 2025, 57: 151.

[6] 杨慷，贺文洁，王海龙. 单参数单维度 Rasch 模型的优势与意义[J]. 心理科学，2021，44(6)：1491-1498.
Yang Q, He WJ, Wang HL. Advantages and significance of the one-parameter and unidimensional Rasch model[J]. Journal of Psychological Science, 2021, 44(6): 1491-1498. (in

Chinese)

[7] 丁杨,张傲雪,杨双,等.空间信息加工对听写困难儿童字形记忆的影响[J].中国临床心理学杂志,2021,29(1):14-18.
Ding Y, Zhang AX, Yang S, et al. Effects of spatial information processing on glyph memory in children with spelling difficulties[J]. Chinese Journal of Clinical Psychology, 2021, 29(1):14-18. (in Chinese)

[8] Luan L, Wentao Z, Ming S, et al. CCLOOW: Chinese children's lexicon of oral words[J]. Behav Res Methods, 2023, 56(2):846-859.

[9] 舒华,武宁宁,郑先隽,等.小学汉字形声字表音特点及其分布的研究[J].语言文字应用,1998,2:65-70.

[10] Jiang Y, Zeng Q, Wen H. TestAnaAPP: An interactive R-shiny application for various test analysis methods[J]. SoftwareX, 2024, 28:101967.

[11] Aryadoust V, Ng YL, Sayama H. A comprehensive review of Rasch measurement in language assessment: Recommendations and guidelines for research[J]. Language Testing, 2021, 1:6-40.

[12] Linacre JM. A user's guide to WINSTEPS® MINISTEP Rasch-model computer programs. Program Manual5. 9. 2. [EB/OL]. (2020-06-25)[2025-01-10]. https://www.winsteps.com/winman/copyright.htm.

[13] Ha HT, Nguyen DTB, Stoeckel T. A comparison of Yen's Q3 coefficient and Rasch testlet modeling for identifying local item dependence: Evidence from two vocabulary matching tests[J]. Lang Assess Q, 2025, 22(1):56-76.

[14] Romaní-Romaní F. Cultural adaptation to Peruvian Spanish and psychometric validation of an instrument to measure scientific literacy in medical students[J]. Anales de la Facultad de Medicina, 2024, 85(1):34-42.

[15] 武慧多,刘巧云,张艳丽,等.基于 Rasch 模型的普通话学前儿童词语命名测试词表的编制[J].听力学及言语疾病杂志,2022,30(5):495-500.
Wu HD, Liu QY, Zhang YL, et al. Development of Mandarin vocabulary naming test for preschoolers based on the Rasch model[J]. Journal of Audiology and Speech Pathology, 2022, 30(5):495-500. (in Chinese)

[16] Smith EV Jr. Detecting and evaluating the impact of multidimensionality using item fit statisticsand principal component analysis of residuals[J]. J Appl Meas, 2002, 3(2):205-231.

[17] 温红博,刘先伟,姜有祥. K-means 聚类方法在中考标准设定中的信度分析[J].中国考试,2024,8:69-78.
Wen HB, Liu XW, Jiang YX. Reliability analysis of the K-means clustering method in standard setting for junior high school academic proficiency examinations[J]. Journal of China Examinations, 2024, 8:69-78. (in Chinese)

[18] Tesio L, Caronni A, Kumbhare D, et al. Interpreting results from Rasch analysis 1. The "most likely" measures coming from the model[J]. Disabil Rehabil, 2024, 46(3):591-603.

[19] 王久菊,孟祥芝,李虹,等.汉语发展性阅读障碍诊断与干预的专家意见[J].中国心理卫生杂志,2023,37(3):185-191.
Wang JJ, Meng XZ, Li H, et al. Expert advice on the diagnosis and intervention of Chinese developmental dyslexia[J]. Chinese Mental Health Journal, 2023, 37(3):185-191. (in Chinese)

[20] Lee MK, Liu HW, Tong SX. Identifying Chinese children with dyslexia using machine learning with character dictation[J]. Sci Stud Read, 2023, 27(1):82-100.

[21] 韩平,杨文艺,谢溢洋,等.智能书写分析与识别平台用于6～12岁儿童书写能力评估的研究报告[J].康复学报,2024,34(3):251-261.
Han P, Yang WY, Xie YY, et al. Research reports on assessment of handwriting ability in children aged 6-12 years by using smart handwriting analysis recognition platform[J]. Rehabil Med, 2024, 34(3):251-261. (in Chinese)

[22] Shune S, Gray LT, Perry S, et al. Validation of the caregiver analysis of reported experiences with swallowing disorders (CARES) screening tool for neurodegenerative disease[J]. Am J Speech Lang Pathol, 2025, 34(2):633-645.

收稿日期:2025-07-30 **修回日期**:2026-03-09 **本文编辑**:李少闻

(上接第 878 页)

[18] 马飞飞,赵国辉,宋文静,等.水中动作技能习得对学龄前儿童基本动作技能的影响[J].中国学校卫生,2025,46(7):1014-1018.
Ma FF, Zhao GH, Song WJ, et al. Impact of preschool children's aquatic motor skill acquisition on their fundamental motor skill development[J]. Chin J Sch Health, 2025, 46(7):1014-1018. (in Chinese)

[19] Ma G, Liu Y, Wang L, et al. Hierarchy in sensory processing reflected by innervation balance on cortical interneurons[J]. Sci Adv, 2021, 7(20):eabf5676.

[20] Shahbazi M, Mirzakhani N. Assessment of sensory processing characteristics in children between 0 and 14 years of age: A systematic review[J]. Iran J Child Neurol, 2019, 15(1):29-46.

[21] 刘如强,徐子纯,张如磊,等. 4～5岁儿童静态平衡中的感觉融合特征研究[J].中国儿童保健杂志,2024,32(4):372-376,394.
Liu RQ, Xu ZC, Zhang RL, et al. Characteristics of sensory integration in static balance of 4 - to 5-year-old children[J]. Chin J Child Health Care, 2024, 32(4):372-376, 394. (in Chinese)

[22] 黄春香,郝增明,王健.学龄前儿童静态站姿平衡及其姿势控制策略[J].中国学校卫生,2021,42(11):1708-1711,1716.
Huang CX, Hao ZM, Wang J. Static balance and control strategy in preschool children[J]. Chin J Sch Health, 2021, 42(11):1708-1711, 1716. (in Chinese)

收稿日期:2025-09-23 **修回日期**:2026-05-08 **本文编辑**:李少闻