

# 《心理学报》审稿意见与作者回应

题目：小学一年级数感不良儿童的筛查与动态干预

作者：张丽锦 毕远 梁熠 刘敏红

---

## 第一轮

### 审稿人 1 意见：

作者首先编制了儿童数感能力测验，然后利用测验筛查出数感不良儿童进行干预，发现这些干预措施有积极的效果。整个研究思路清晰，实验设计良好，结果有效。但有一些地方需要作者再考虑：

**意见 1：**对“数感”的提法需要更多证据。例如，提供英文关键词，提供五个任务的样例；作者似乎讨论的是基本数字加工能力。

**回应：**数感（number sense）是关于数字间关联的意识和直觉能力，是对数、数量关系和数量操作的直觉，它从幼儿早期开始显现并一直发展到小学中年级初步成熟。所以数感可以说是初级数学能力（Geary, Bailey, & Hoard, 2009），其核心成分包括对小数量的快速直觉、数型识别、数量比较、数量估计、数数以及简单的数量转换等（见 Berch, 2005; Case, 1998）。幼儿早期就表现出来的对小数量的识别、比较和操作上的不同敏感度和差异体现了不同的数感水平。对数感的概念和解释我们在前言部分有所阐述，根据审稿专家的意见又作了进一步补充，有关数感与数学困难和数学成绩的关系在前言和对审稿专家 2 的第 4 个问题的答复中作了详细说明。我们挑选了数感测验各成分样例附到了附录中（见附录 1 表 7）。不过鉴于版面的关系，以及样例而非完整测验实际上不能体现数感测验所涉及的广泛的数学加工领域，同时应当注意到对数感的测量关键在于数学问题（相对于课堂教学内容）的新异性和问题情境的陌生程度，而不是所涉及到的数学加工过程，所以我们不建议把样例作为附录，不知审稿专家是否认可。当然，如果需要，我们可以把样例附上，我们已挑选了数感测验各成分样例附到了附录中。

数字加工（numerical processing）是指个体加工数字刺激，形成相应的数字心理表征，并运用其进行相应的认知加工过程，如数量比较、数学运算、数字估计等（张红川, 2005, 2007）。数感的具体成分与基本数字加工能力可能涉及到类似的数学操作，但两个概念所关注的心理过程是不同的：数字加工能力关注的是在对抽象的数字和运算形式的加工过程中的能力结构问题，包括对视觉、听觉刺激和类比表征的编码、对数字事实等静态记忆的提取（和提取速度）、基于工作记忆的过程性数字（加数、进位等）的保持等认知能力，而数感关注的是儿童对数与数、数与现实世界之间联系的直觉或感觉，以及在新情境中主动运用这种直觉“以灵活而有创造性的方式”（Sowder, 1992）解决新的数学问题的能力。两个概念应当是从所涉及的心理成分而不是（也无法从）数学成分上加以区分。

数感测验内容比同一时期的数学测验所涉及的数字范围要广，除了在题型设计上保证相对的新异性外，还要尽量满足儿童无法依靠已经学到的算术知识或数字事实直接解决问题的要求。

**意见 2：**研究 1 部分可以简写，避免文章太长；

**回应：**接受审稿专家的建议，我们已对研究一进行了必要删减。

**意见 3:** 表 5 中的数据组织不合理, 比较凌乱;

**回应:** 表 5 中的信息比较丰富, 既有干预组、控制组和正常组三者之间在数感五成分和总分的差异性比较 (F 检验), 又有各组内的前、后测比较 (t 检验), 所以显得表中的信息内容不那么可读。但如果不将 F 值和 t 值放到表里, 在文中由于数值太多而显得过于琐碎。考虑到审稿专家的意见, 我们在表 5 下方“表注”中对该表进行了解释, 希望能增加可读性, 请审稿专家指教。

**意见 4:** 根据表 6, 三组不同数感水平被试在“计算流畅性测验”前后测的比较中没有差异 (即都有类似的进步)。似乎说明干预没有特定的效果;

**回应:** 表 6 的结果显示的是有差异, 首先, 三组被试在计算流畅性前测上数感不良干预组和控制组差异不显著, 但二者均显著低于正常组。其次, 也是最能说明差异的是, 在后测上干预组显著好于控制组, 且与正常组差异不显著, 这体现了干预组经干预后好于控制组的进步。我们在文中对此又作了进一步的强调和说明, 请审稿专家指正。

**意见 5:** 关于数集测验作为计算流畅性测验, 作者需要作出说明。

**回应:** Geary 等人 (Geary, Bailey, & Hoard, 2009) 开发的数集测验是考查儿童识别和加工表征阿拉伯数字与几何符号数量关系的速度和准确性的测验, 它本质上是一种限时加法测验, 而限时加法和/或限时减法测验考查的正是计算流畅性。该数集测验的材料设计和施测步骤对计算流畅性的测查具有良好的表面效度。我们将把这些解释补充进正文中。谢谢审稿专家的建议。

---

**审稿人 2 意见:**

研究编制了针对小学一年级学生数感发展水平的测验, 并对数感不良儿童进行了干预。研究结果表明所编制的测验达到了心理测量学要求, 干预的效果良好。整篇文章的研究思路清晰, 操作规范, 写作流畅。

在引言和讨论的阐述中, 有些地方显得比较模糊, 造成本研究和所编制测验的意义不够清晰。

**意见 1:** 对于鉴别小学低年级数学困难的必要性, 作者只提到了能力-成绩差异模型适用于更高年级的学生, 而数学困难需要早发现早干预等。但没有提到我国已有对幼儿数学困难进行鉴别和干预的相当多的工作 (华东师范大学周欣老师进行了系统研究)。既然已有针对更小年龄的鉴别和干预方法, 为何还需要针对更大年龄的? 针对小学低年级数学困难的研究当然是有必要的, 但似乎不应简单以“早发现早干预”为理由, 忽略了针对学前期的已有工作。

**回应:** 我们在文中所说的“早发现早干预”原本是针对小学四年级而言的, 因为国内对“数学学习困难”的研究筛查标准大都用的是能力-成绩差异模型 (见 毕远, 张丽锦, 2014), 而小学低年级儿童的成绩分数普遍偏高, 四年级才开始表现出分化。所以对于低年级儿童使用能力-成绩差异模型进行筛查在数学成绩方面的标准不好确定, 严了增加漏报、松了增加误报, 会造成取样偏差。所以我们以为研究者若使用能力-成绩差异模型对数学困难儿童进行筛查来选取被试的话, 最好在小学四年级以上。当然, 我们一直都不否认在小学低年级乃至学前期也可以对儿童进行数学水平评估、甄别和筛查, 因为数感发展及其差异性在幼儿期就表现十分明显。但是, 是否可以直接使用“数学学习困难”的概念还需谨慎。根据 DSM-V 的诊断标准, 学习困难出现在学龄期, 当对儿童学业能力的要求超过他们有限的能力时 (例如, 限时测验、在规定时间内完成阅读或撰写一定长度和复杂度的文章, 超负荷的学习负担等), 学习困难才会明显表现出来 (Association, 2013)。所以对学前儿童使用“数学学习困难”应当

谨慎，对小学低年级儿童因为学习成绩的普遍偏高也要谨慎。国外研究者通常使用的是“数学困难风险儿童”（children at risk for mathematics difficulties）的概念，我们用的是“数感不良”，为的也是准确使用概念和诊断标准。

事实上，我们一直都在关注周欣老师及其同事的一系列研究，他们关于“3~6岁儿童学习和发展指南（草案）：认知领域发展水平测查工具（R）”中对儿童的数概念（包括数数策略、数量守恒、书面数符号表征、加减运算和估算）进行了界定和评估（周欣，黄瑾，赵振国，杨宗华，2009），并应用于大、中、小班儿童。他们曾在了一项研究（周欣等，2013）中使用了“早期数学学习困难”的概念，认为早期数学学习困难儿童是智力发展正常，但数学发展水平显著低于同龄儿童（他们的被试是5岁）的儿童。他们使用的是教师评定加标准化测验（“早期数学能力测验（III），TEMA）。该研究对数学困难组进行筛选的标准是TEMA在29分以下（比正常儿童低约2个标准差），智商在80以上。作者认为，这是严格意义上的早期数学学习困难儿童。研究还提到，已有研究发现仅仅根据一次数学得分的考察来确定数学学习困难儿童可能不完全准确，更为理想的鉴别方法是在连续两年的考察中都是低分（见Mazzocco & Thompson, 2005）鉴于此，我们认为，直接提出对“数学困难”进行诊断和筛查对年龄小的儿童都需谨慎。尽管如此，这些并不妨碍研究者对数学认知落后的儿童进行诊断、筛查和干预，周欣老师的工作如此、我们的工作也是如此。

鉴于我们的上述考虑和以往丰富的对学前和小学低年级儿童数学认知的评估与干预研究，我们认真思考了评审专家的批评建议，认为“早发现早干预”的关于研究目的的提法不够严谨、妥当，我们已在文中做了修改。在引言中对学前儿童的有关研究进行了补充，并论述了数感研究的特殊意义和本研究选取一年级被试的理由。进一步地，我们将原题目“数感不良儿童的早期筛查与动态干预”改为“一年级数感不良儿童的筛查与动态干预”，不知是否恰当，敬请审稿专家批评指正。

**意见 2：**对于采用数感测量来鉴别数学困难的理论基础和优势等，在文中的体现并不充分。作者提到，由于小学低年级的数学成绩尚不够稳定，为避免标签效应，我国研究者对四年级以下儿童数学困难倾向的鉴别相当谨慎。那么，使用数感测量来鉴别低年级数学困难是否弥补了之前所考虑到的风险？是否以牺牲这种谨慎为代价？如果作者能更加清晰、有力地对这一问题进行阐述，研究的意义将得到进一步凸显。

**回应：**首先感谢审稿专家的细致评审，结合对审稿专家2第1个问题的回答，我们承认在文中的表述不够严谨恰当。事实上，国内很多研究者对四年级以下儿童数学困难的评估采用的是能力-成绩差异模型，而且以年级水平离差法居多（见毕远，张丽锦，2014）。文中是我们表述有误，接受审稿专家的批评！综合考虑审稿专家2的几个主要疑点，对引言的修改我们将以问题解决的思路为线索进行重新整合，以便突出本研究的合理性与研究意义。在相关表述上，也强调数感鉴别是以及早补充教学降低风险为目的，意图对存在潜在数学困难风险儿童进行筛查以便干预，而非提前对低年级儿童作“数学学习困难”的诊断。

**意见 3：**由于文中多次提到数感测量成绩和学校数学成绩的关系问题，并以数学成绩作为效标，建议作者进一步明确这两者之间的界线。

**回应：**该问题与“问题 4.”有共同之处，合并到“问题 4.”中回答。

**意见 4：**关于计算流畅性的阐述，引言和讨论部分似乎都没有很好地明确其地位。数感、计算流畅性、数学成绩之间是什么关系？本研究又希望用计算流畅性指标代表哪一变量？特别是在讨论中，提到“将计算流畅性作为评估数学困难，尤其是低龄儿童的数学困难倾向的指标有助于早诊断、早发现、早干预”。是说直接用计算流畅性来鉴别指标就可以了吗？那么

本研究所做的数感测量又将扮演什么角色呢？

回应：首先来看数感，数感是对数、数量关系和数量操作的直觉，是关于数字间关联的意识和直觉能力，它从幼儿早期开始显现并一直发展到小学中年级。对幼儿和小学低年级儿童来讲，数感的核心成分包括对小数量（小于3或5——视不同任务而定）的快速直觉、数型识别、数量比较、数量估计、数数以及执行简单的数量转换等（见 Berch, 2005; Case, 1998）。这些基本数感是独立于一般记忆能力、语言和空间知识的特定的认知功能（Gleman & Butterworth, 2005; Landerl, Bevan, & Butterworth, 2004）。幼儿早期就表现出来的对小数量的识别、比较和操作上的不同敏感度和差异体现了不同的数感水平。数字系统有语言和非语言之分，婴儿期就表现出的对非语言数字系统的数量操作的敏感和差异提示个体对数量的敏感有着神经基础，位于顶内沟（intraparietal sulcus, IPS）。数感的这种对数量和大小的直觉能力可能是先天的并且是早期数学学习的基础（Geary, 2006），但它也与儿童早期数概念的家庭教育和经验有关（Case & Griffin, 1990）。即数感不仅涉及非语言数字系统，也与受经验和学习影响的语言数字系统有关。因为数感概念的难以确定统一，以及数感涉及到数量加工的多个方面，使数感究竟包括哪些成分一直众说纷纭。尽管 Jordan 的数感五成分理论及其测验未必详尽涵盖了被研究确认的和与小学数学课程有关的数学认知内容，但对基本的数学能力都有涵盖，而且也并不是对基本认知能力（如，一般工作记忆）的考查。五成分模型强调数感是对儿童关于数、数量关系和数量操作的知觉，包括数数、数知识、数量转换、数估计和数型。数感的基本成分对儿童日后数学学习中的更高层级（higher-order）的数学思维（如数量操作与数量程序中的流畅性与灵活性）非常重要，所以数感对数学成绩有显著的预测作用。不仅如此，数感各成分之间、各成分与其他数学能力和数学成绩之间也存在着复杂的联系，比如，研究证实：

早期的数数困难可以预测日后的算术运算困难（Geary, Hoard, & Hamson, 1999）；

数知识是一年级算术成绩的强预测变量（其零阶系数达0.73）（Baker et al., 2002）；

数量转换如果涉及到与语言相关的内容，属于言语计算的范畴，和社会阶层、和共有文化知识（conventional knowledge）有关；如果涉及非言语计算的数量转换，则依赖于个体把握工作记忆中的数量表征能力。所以数感基于先天神经机制，又受后天经验学习的影响；

数估计与要求精确数知识的算术运算技能有显著相关（Dowker, 1997; Rubenstein, 1985）；

数型的再认知和使用可以提高数字组合（属于数量转换）的技能（Threfall & Frobisher, 1999），数型的掌握可以使数字组合变得更流畅（Gray & Tall, 1994; Jordan et al., 1994）。

数感不良能够导致计数加工不足、算术事实提取缓慢、精确计算不足等几乎所有数学困难的特征（Grary et al., 2000; Jordan et al., 2003a, 2003b），所以数感是数学困难的强预测变量。另外，数感不良儿童从数学教学中获益困难（Baroody & Rosu, 2006）又进一步证实数感与数学成绩的相关关系。

数学困难除了有单纯的数学困难和伴阅读困难的数学困难的不同类别，就单纯的数学困难本身其表现也存在跨领域的不均衡性（Ginsburg, 1997），因为数学的不同方面包括不同的认知能力（Geary et al., 2000），而非只是笼而统之的数学能力（mathematical competence）。比如，有些数学困难儿童表现出算术事实提取的不足，但对数数原则和数学概念把握良好；有些数学困难儿童有良好的计算技能但却不能很好理解数概念（Jordan & Hanich, 2000; Jordan & Montani, 1997; Russell & Ginsburg, 1984）。所以 DSM-V 提出：诊断数学困难，要列出受损的所有领域和子领域。如果受损领域超过一个，应该根据下列分类具体指明究竟是哪些领域受损——数感（number sense）、算术事实记忆（memorization of arithmetic facts）、精确或流畅的计算（accurate or fluent calculation）、精确数学推理（accurate math reasoning）。

再来看计算流畅性，计算流畅性是指快速、准确、轻松地进行基本运算并灵活应用的能

力 (Council, 2001), 是解决数学问题的重要能力。计算流畅性是低年级数学学习的特点, 与核心计算原理的基础知识相互促进, 对数学学习的很多方面 (从解决简单的整数问题到计算分数、小数、百分数, 再到代数方程、甚至基本的几何问题) 都有促进作用 (Jordan et al., 2008)。如前所述, 数感不良能够导致儿童在算术事实提取上的困难。而算术事实 (arithmetic fact) 是已经转化为陈述性知识的数量转换关系 (如,  $4 + 3 = 7$ ), 其特点是在解决数学问题时能够快速、轻松、准确地进行提取 (Hasselbring, Goin, & Bransford, 1988), 即不用计算直接从记忆中获取准确答案 (并不是仅靠死记硬背积累算术事实), 这显然和计算流畅性密切相关。有研究者主张, 计算流畅性的前提性基础包括数感、记忆广度、阅读/语言能力、空间能力 (Locuniak & Jordan, 2008), 数感既然作为计算流畅性的基础性前提, 它的发展不良必然会导致儿童在计算方面的缺陷, 并容易导致数学困难 (Gersten, Jordan, & Flojo, 2005; Mazzocco & Thompson, 2005)。许多数学困难儿童存在计算流畅性不足的事实 (Bryant, 2005; Locuniak & Jordan, 2008; Jordan, 2007; Reigosa-Crespo et al., 2012; Fuchs et al., 2013; Cowan & Powell, 2014) 就是对数感、计算流畅性、数学困难三者之间相互关联关系的支持性证明。

因为数感测量涉及到了多种数学加工成分, 可以相对广泛地说明数学困难的多种表现, 所以以数感测验甄别潜在数学困难风险儿童是可靠和可行的, 我们将这些儿童称为“数感不良儿童”。因为数感是早期发展的, 对学前及小学低年级数学困难风险儿童的甄别不适合采用年级水平离差法, 而数感测验则是一个相对合适的选择。至于计算流畅性, 尽管即使在小学低年级, 考查和甄别儿童在计算方面是否快速准确很容易操作, 计算流畅性可以评估小学低年级儿童的数学能力, 但它并不能作为诊断和筛查量表来使用。我们前面说过, 数学困难的表现有多种, 而计算流畅性只是其中相对更核心的一个特征, 就像心理数线与儿童的数学能力、数学成绩有很强的相关, 可以预测数学成绩一样 (Booth & Siegler, 2008) (其实心理数线与计算流畅性之间也有很强的相关)。这样与数学能力有很高相关的数感成分当然不止流畅性和数线, 它们只是良好的预测指标之一, 还不能广泛地代表数学认知的能力范畴。而我们依据五成分理论编制的数感测验, 包含数感的多种成分, 有心理数线、有精确计算和数量估计等, 以此评估潜在数学困难风险儿童, 比任何单一成分的数学能力测验都具有更好的代表性和更广泛的解释性。

上述是对数感、计算流畅性、数学困难三者关系的阐述和说明, 相关观点我们在前言和讨论中进行了补充、梳理和厘清。

**意见 5:** 建议作者关注文中“数量”和“数字”的用法, 更准确地使用这两个概念。现在文中的有些“数字”似乎应改为“数量”更加合适 (如“数字间关联”、“数字估计”)。

**回应:** 感谢并接受审稿专家的建议。经过进一步斟酌, 我们将相关概念“数字组型”、“数字间关联/关系”改为“数量组合”、“数量间关联/关系”。对于“数字估计”, 因原作者 (周广东, 莫雷, 温红博, 2009) 使用的是“数字估计” (numerical estimation) 的概念, 所以我们在文中没有更改。请审稿专家批评指正。

**意见 6:** 对于方法的介绍有一些模糊之处

(1) 方法 2.1 部分指出“另选 89 人为考查重测信度被试”。这里似乎不应该是“另选”, 若作为重测信度被试, 应是从之前的样本中选择?

**答:** 这的确是我们表述上的疏忽, 是从初测的 202 名被试中再次抽取的 89 名被试进行的重测。感谢审稿专家的批评指正, 我们已在文中作了修改。

(2) 3.3 部分筛选数感不良儿童时, 指出“言语分量表和记忆分量表达到常模水平”。这里的“常模水平”意义很含糊, 最好能具体指出。

**答:** 接受审稿专家的建议。已将麦卡锡常模数据 (5.25 岁~8.5 岁) 和本研究对被试测查的数

据在文中呈现（见“3.3 程序（1）前测阶段”）。

（3）3.3 部分介绍干预措施时，提到“以 4~5 人为一小组进行干预”，但是从后面的操作来看，如进行对错反馈和询问策略等，似乎只能一对一进行，并不适于小组活动。这里具体是如何操作的？

回应：4~5 人的小组干预是以情景游戏的形式设计的，游戏（测量、数字板竞赛等）是以“回合制”（每名儿童轮流操作）形式出现，因而主试可以记录每名儿童在游戏中的表现。在自由游戏环节之后对每名儿童在游戏过程中的操作进行反馈和询问，这一环节确实是一一对一进行的，是在儿童轮流作答环节之后，因而反馈不会影响到其他儿童作答。在 3.3 程序中干预部分作了更为准确的表述。

意见 7：文中所出现的引用与文后的参考文献不能一一对应，建议作者仔细检查。

回应：这个错误非常令我们汗颜！修改稿中对此进行了仔细修改，在 Endnote 插入文献的基础上对“见 作者名”和“作者名（年）”这样与 Endnote 格式不符的文献引用进行了再检查和修改。

建议修改后再审。

---

## 第二轮

审稿人 1 意见：

意见 1：建议删掉标题中“动态”二字；

回应：因为本项目组的一系列研究（包括本研究）都是在维果斯基动态测验（dynamic test 或 dynamic assessment）理念和测量范式下进行的干预研究，我们觉得保留“动态”二字比较合适。不知我们这样考虑是否合适。

意见 2：19 页“见附录 2 的表 7” 有误。附录 3 文章中没有引用。

回应：感谢审稿专家的细致评审，很抱歉这是我们粗心所致。现已按照出现的先后顺序重新调整了 3 个附录，在正文里：15 页—附录 1 表 7，19 页—附录 2，20 页—附录 3 表 8，用蓝色字体标识。

审稿人 2 意见：

意见 1：修改后的稿件较好地解决了之前存在的问题，或者对之前的问题做了比较恰当的回答。建议采用。

回应：感谢审稿专家。已将文稿再次进行校对和捋顺。