

# 《心理学报》审稿意见与作者回应

题目：执行功能在不同年级儿童数学能力中的作用

作者：祝孝亮 赵鑫

## 第一轮

### 审稿人 1 意见：

本研究对执行功能各个子成分对算数，逻辑推理，空间能力的影响进行了研究，并且分析了在不同年级两者关系的变化。该研究的优点是包括了相对大的样本量，涵盖 3-6 年级。采用了结构方程模型，研究了各个子成分的独立贡献。写作较为清晰，表述比较准确，且结合文献对结果进行了比较详细的讨论。但该研究也存在一些重大问题。

### 意见 1：横断研究方法难以解释变化和因果关系。

**回应：**感谢审稿专家的意见。如您所言，横断研究是发展心理学研究中常用的研究设计，其仅测查不同年龄被试在特定时间点各心理属性间的关系，因而无法获取个体心理属性间关系的连续发展趋势，难以对变量之间的因果联系做出推论。可见，这也是本研究所存在的一大研究局限。对此，我们已在文章的展望部分对其进行了说明，并提出了未来研究可以使用纵向研究方法来进行进一步考察因果关系的相应展望。虽然本研究所使用的横断研究方法存在难以解释变化和因果关系的局限，但纵观已有探究执行功能与儿童数学能力关系的相关研究，本横断研究仍存在一些明显优势。具体如下：**第一**，相比于考察执行功能与数学能力关系的类似横断研究(Cragg et al., 2017; Gómez et al., 2015)，本研究具有较大的样本量(本研究中被试总数为 812; Cragg 等人的研究中被试总数为 293; Gómez 等人的研究中被试总数为 450)，因而这种基于大样本所获得的研究结果更具有可信性与稳定性，也更加符合结构方程模型对于大样本量的要求。**第二**，本研究中不同年级组被试的年龄差异较小，这大大规避了横断研究中存在的代群效应(唐细容等, 2013)，即得出年龄/年级差异在很大程度上是造成不同年级组被试两者关系发生变化的主要原因。因此，本横断研究可以在一定程度上揭示各执行功能子成分与不同数学能力间关系的发展趋势，其不仅回答了以往研究中的部分争议，还为未来研究者进一步探究其因果联系提供了一定研究思路的参考。**第三**，相较于已有研究，本研究考虑到了执行功能子成分的再分性(即抑制控制分为干扰抑制与反应抑制，工作记忆分为工作记忆广度和工作记忆刷新)、数学能力的多维性(运算、逻辑思维、空间想象)、年龄效应(横断视角)以及使用了结构方程模型，这不仅评估了各执行功能成分在不同数学能力中的作用，也探究了这种预测模式随年级升高而变化的趋势。其得出的重要研究结论(低阶执行功能成分在低年级儿童数学能力的发展中所起的作用较大，高阶执行功能成分支持高年级儿童数学能力的重要认知因素)是对于已有研究结果的重要补充，即高阶执行功能成分除支持解决更复杂数学问题外(Magalhães et al., 2020)，也会在大龄儿童简单数学问题的加工中发挥重要作用。这说明执行功能成分在数学能力中的作用是与个体各执行功能成分的发展顺序相一致，即低阶执行功能成分成熟较早，对早期数学能力的影响较大；高阶执行功能成分成熟较晚，在后期数学能力的发展中发挥的作用较大。综上，虽然本横断研究并未揭示出执行功能与数学能力的因果关系，但对于明晰各变量间的关系及指引未来研究者进一步探究两者之间的实质联系具有重要意义，因而具有较大理论价值与实证意义。

## 参考文献:

- 唐细容, 蒋莉, 曾慧, 姚树桥. (2013). 成人认知功能的发展与老化:一项控制同辈效应的横断研究. *心理发展与教育*, 29(3), 262–267.
- Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, 162, 12–26.
- Gómez, D. M., Jiménez, A., Bobadilla, R., Reyes, C., & Dartnell, P. (2015). The effect of inhibitory control on general mathematics achievement and fraction comparison in middle school children. *Zdm*, 47(5), 801–811.
- Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934–952.

**意见 2:** 关于执行功能以及各个子成分对数学的影响已经存在大量研究, 创新性并不高。

**回应:** 非常感谢审稿专家的意见。如专家所言, 已有大量研究探讨了执行功能/执行功能各成分与数学能力的关系。但以往绝大多数研究对两者关系的考察主要存在以下三点局限: **第一**, 缺少考虑执行功能子成分的可再分性(抑制控制分为干扰抑制与反应抑制, 工作记忆分为工作记忆广度和工作记忆刷新), 其需要再分的必要性在于这两个成分的不同方面与数学能力的关系存在差异。如研究发现, 学龄儿童干扰抑制与数学能力之间的相关程度较低或学龄儿童的干扰抑制无法有效预测其数学能力(Cragg et al., 2017; Lee et al., 2018), 而反应抑制可以显著预测学龄儿童的数学能力(St Clair-Thompson & Gathercole, 2006; Yazdani et al., 2021)。**第二**, 缺少考虑数学能力的多维性(如多数研究仅以数学运算来衡量儿童的数学能力), 其需要划分为不同数学能力的必要性在于各执行功能成分在不同数学能力中的作用存在差异。如研究发现, 抑制控制对儿童数学推理不起独立的预测作用(Lee et al., 2018), 但可以显著预测儿童在计数任务上的表现(Fernández Cueli et al., 2020); 工作记忆与数学运算间的联系较低, 而与应用问题解决之间的关系更为紧密(Spiegel et al., 2021); 相比于数学运算, 认知灵活性和数学推理间的联系更加密切(Georgiou et al., 2020)。**第三**, 缺少考虑执行功能/执行功能各成分与数学能力的发展性联系, 即年龄/年级因素对二者关系的影响。这是由于已有研究发现执行功能与数学能力的联系会随个体年龄的增长而发生变化(Magalhães et al., 2020; Stipek & Valentino, 2015)。综上, 这三点局限往往导致已有研究结论的不一致, 因而在考察执行功能与数学能力的关系时, 明晰具体的执行功能成分、数学能力的不同方面及考虑年龄因素对于揭示执行功能与数学能力间的实质联系显得极其必要。然而, 虽然以往也有少数几项研究较为细致的探讨过不同执行功能成分(工作记忆广度、工作记忆刷新、反应抑制、干扰抑制及认知灵活性)对儿童数学能力的影响, 但也仅以学校标准化数学测验(St Clair-Thompson & Gathercole, 2006)或数学运算和数学应用问题(Yang et al., 2019)来评估儿童的数学能力, 其不仅存在缺乏考虑数学能力多维性与较小样本量的问题, 同时这种以低龄儿童为被试的研究依旧存在是否可以真正测量到认知灵活性和刷新能力的疑惑。如研究发现, 学龄儿童执行功能的三因素模型(抑制、转换和刷新)在三至六年级儿童中拟合最好(Lehto et al., 2003)。但相对比而言, 本横断研究的样本量较大、涉及多年龄段儿童、控制了部分混淆变量(如, 年龄、性别)、系统考虑了执行功能各子成分与数学能力的多维性, 还使用结构方程模型来验证研究的理论假设, 其所具有的低误差和针对理论模型验证等特点使之更能揭示变量间的实质联系。因此, 相比于探讨执行功能与数学能力关系的同类研究, 本研究虽无法揭示其二者间的因果关系, 但可以対以往研究结果不一致的争议做出一定回答, 其所得出的研究结论对于未来两者关系的进一步探索具有较好的指引作用, 即在探究执行功能与儿童数学能力的关系时不仅需要考虑执行功能子成分的可再分性和数学能力的多维性, 还需要考虑年龄对二者关系的影响。

## 参考文献:

- Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, 162, 12–26.
- Fernández Cueli, M. S., Areces Martínez, D., García Fernández, T., Alves, R. A. T., & González Castro, P. (2020). Attention, inhibitory control and early mathematical skills in preschool students. *Psicothema*.
- Georgiou, G. K., Wei, W., Inoue, T., Das, J. P., & Deng, C. (2020). Cultural influences on the relation between executive functions and academic achievement. *Reading and Writing*, 33(4), 991–1013.
- Lee, K., Ng, S. F., & Bull, R. (2018). Learning and solving algebra word problems: The roles of relational skills, arithmetic, and executive functioning. *Developmental Psychology*, 54(9), 1758–1772.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21(1), 59–80.
- Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934–952.
- Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., & Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychological bulletin*, 147(4), 329–351.
- St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology*, 59(4), 745–759.
- Stipek, D., & Valentino, R. A. (2015). Early childhood memory and attention as predictors of academic growth trajectories. *Journal of Educational Psychology*, 107(3), 771–788.
- Yang, X., Chung, K. K. H., & McBride, C. (2019). Longitudinal contributions of executive functioning and visual-spatial skills to mathematics learning in young Chinese children. *Educational Psychology*, 39(5), 678–704.
- Yazdani, S., Soluki, S., Arjmandnia, A. A., Fathabadi, J., Hassanzadeh, S., & Nejati, V. (2021). Spatial Ability in Children with Mathematics Learning Disorder (MLD) and Its Impact on Executive Functions. *Developmental Neuropsychology*, 46(3), 232–248.

意见 3：对于为什么要研究运算能力，逻辑推理能力，和空间能力缺乏理论支撑。

回应：非常感谢您指出的这一重要问题，这是本文写作中的不足之处。经分析初稿中(引言第四段)对相关部分的阐述，我们的确缺少了将数学能力划分为数学运算能力、逻辑思维能力与空间想象能力的理论支撑与推导。进一步文献整合发现，研究者们对于学龄儿童数学能力的具体组成部分仍存在较大争议，主要显示为国内外差异。如国外研究者指出，学龄儿童的数学能力主要包括运算、代数、几何与测量(Spiegel et al., 2021)，或运算、代数和几何(Gilmore et al., 2015)，而国内研究者则将学龄儿童数学能力划分为数学运算能力、逻辑思维能力与空间想象能力(林崇德, 2011; 吴汉荣, 李丽, 2005)。如林崇德(2011)重点强调数学思维能力，即思维的深刻性、灵活性、独创性、批判性和敏捷性，认为数学能力是运算能力、空间想象能力和逻辑思维能力与五种思维品质相互交叉构成的统一整体。同时，根据我国历次小学数学教学大纲，教育界长期将学龄儿童的数学能力划分为数学运算能力、逻辑思维能力与空间想象能力(或空间观念)三大核心能力(杜文平, 2013)。因此，基于本研究是针对国内儿童执行功能与数学能力关系的探讨，所以我们在本研究中仍沿用国内学者对数学能力的划分，即将学龄儿童的数学能力分为：数学运算能力、逻辑思维能力与空间想象能力。根据专家的建议，我们在修改稿(1.1 部分, 第二段)中进行了相应的补充。具体内容如下：迄今为止，关于数学能力的组成部分，研究者们观点仍未统一。如国外研究者指出，学龄儿童的数学

能力主要包括运算、代数、几何与测量(Spiegel et al., 2021), 或运算、代数和几何(Gilmore et al., 2015)。而国内研究者及小学数学教学大纲通常将学龄儿童的数学能力划分为: 数学运算能力、逻辑思维能力和空间想象能力(杜文平, 2013; 林崇德, 2011; 吴汉荣, 李丽, 2005)。基于此, 本研究继续沿用国内学者对学龄儿童数学能力的分类, 将小学儿童的数学能力分为: 数学运算能力、逻辑思维能力和空间想象能力。

#### 参考文献:

- 杜文平. (2013). 小学数学教学大纲或课程标准中数学能力内涵的演变. *中小学数学(小学版)*, 12, 43-45.
- 林崇德. (2011). 智力发展与数学学习(第四版). 北京: 中国轻工业出版社
- 吴汉荣, 李丽. (2005). 小学生数学能力测试量表的编制及信效度检验. *中国公共卫生*, 21(4), 473-475.
- Gilmore, C., Keeble, S., Richardson, S., & Cragg, L. (2015). The role of cognitive inhibition in different components of arithmetic. *Zdm*, 47(5), 771-782.
- Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., & Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychological Bulletin*, 147(4), 329-351.

**意见 4:** 分析部分, 对于不同的年级的模型有不同的路径设定, 无法判断两个年级的模型的结果差异是由于路径设定的不同导致, 还是由于年级不同导致。

**回应:** 感谢审稿专家细致的思考。如您所言, 初稿中 3~4 年级阶段与 5~6 年级阶段的结构方程模型确实存在修正路径不同的问题。具体为: 3~4 年级阶段的修正路径为刷新成绩与目测长度误差项; 5~6 年级阶段的修正路径为年龄与目测长度误差项。在初稿模型修正时, 我们仅以使模型拟合系数更好为出发点, 并未考虑到支持此修正的相应理论, 这是我们考虑不周之处。因此, 这种模型修正路径的差异将致使不同年级阶段的结果无法直接进行对比。然而, 需要强调的一点是除上述修正路径不同之外, 我们对于文中所呈现的三个结构方程模型的路径设定是完全基于其假设理论模型(图 1), 其缘由为这些自变量之间本身就存在其相关性。在 Amos 中构建每个模型时, 我们画出了所有自变量间的相关关系(除性别与年龄之间的相关路径), 但由于部分变量间的相关关系并不显著, 因此在所示的结构方程模型图中并未画出不显著相关的路径, 从而造成了表面上各模型中自变量路径不一致的现象。所以, 初稿中 3~4 年级阶段与 5~6 年级阶段的结构方程模型除修正路径存在差异外, 其余所有路径的设定均是一致的(未呈现自变量间相关不显著的路径)。基于专家的意见 5 与意见 20, 我们在修改稿的假设理论模型(图 1)中也将性别作为了一个控制变量。基于已有研究基础, 即性别与执行功能各子成分和不同数学能力、年龄与执行功能各子成分和不同数学能力、执行功能各子成分与不同数学能力、执行功能各子成分间均存在一定的联系, 本文提出了新的假设理论模型(图 1)。在此, 我们仍需说明的是, 修改稿中三个结构方程模型自变量间的路径均为一致, 所有模型的初始模型已完全达到模型拟合指标的要求, 因而对所有模型未做任何修正。为简化模型及提升模型结果的清晰表达性, 各年级段模型图中仍不呈现自变量相关未达到显著水平的路径, 但我们在结果部分(图 2 下方)统一做出了说明, 这样可有效规避专家和读者可能存在的疑惑。因此, 修改稿中不同年级阶段模型的路径设定均是一致, 可通过比较不同年级阶段模型结果的差异来获得各执行功能子成分与不同数学能力关系的发展趋势。

**意见 5:** 只有年龄作为控制变量, 没有其他的控制变量。

**回应:** 非常感谢审稿专家的意见。如您在详细审稿意见(意见 20)中提到的一样, 本文初稿中所构建的结构方程模型仅控制了年龄, 未有其他控制变量的纳入。如专家所言, 年龄、性别、父母受教育水平、家庭社会经济地位、语言能力和非言语智力等均会对个体的执行功能和数学能力产生一定影响(Fung et al., 2020; Georgiou et al., 2020; Stipek & Valentino, 2015; Yang et

al., 2019), 我们也确实欠缺对部分混淆变量的控制。根据专家的意见, 我们在修改稿中将性别因素也纳入了新的假设理论模型, 并对所有的结构方程模型进行了重新构建。新构建的结构方程模型图以及对模型结果的相关表述已在文章(3.3、3.5)中使用红色字体作出了标识。

#### 参考文献:

- Fung, W. K., Chung, K. K. H., & Lam, C. B. (2020). Mathematics, executive functioning, and visual-spatial skills in Chinese kindergarten children: Examining the bidirectionality. *Journal of Experimental Child Psychology*, 199, 1–10.
- Georgiou, G. K., Wei, W., Inoue, T., Das, J. P., & Deng, C. (2020). Cultural influences on the relation between executive functions and academic achievement. *Reading and Writing*, 33(4), 991–1013.
- Stipek, D., & Valentino, R. A. (2015). Early childhood memory and attention as predictors of academic growth trajectories. *Journal of Educational Psychology*, 107(3), 771–788.
- Yang, X., Chung, K. K. H., & McBride, C. (2019). Longitudinal contributions of executive functioning and visual-spatial skills to mathematics learning in young Chinese children. *Educational Psychology*, 39(5), 678–704.

#### [续: 详细审稿意见]

该研究对小学阶段的儿童的执行功能与数学之间的关系进行了深入探讨, 以及两者的关系在低年级和高年级之间的区别进行了详细的分析。该研究的优点是拥有较大的样本量, 且涵盖的年级范围较广, 从三年级到六年级。另外, 该研究对于执行功能的测查也较为全面, 包括了执行功能的主要部分, 有助于了解执行功能各个部分对数学的独立贡献。作者的写作也较为流畅, 讨论部分对各个研究结果进行了详细的解释, 且结合了以往研究对各个解释提出了理论的支撑。但该研究还存在很多问题, 主要体现在文献综述, 方法部分和结果分析部分, 具体意见如下:

**意见 6:** 作者在文献综述中提到“较少有研究细致的探讨执行功能各子成分在不同的数学能力中起到的作用是否相同”。这个说法是不对的, 以往有很多关于执行功能各个子成分对数学能力的影响的研究和文献综述(Cragg & Gilmore, 2014; Friso-van den Bos et al., 2013; Fuchs et al., 2010; Schmerold, Bock, Peterson et al., 2016.....)。而且“不同数学能力”的描述太过模糊, 应该直接点出具体是哪些数学能力。

**回应:** 非常感谢审稿专家所指出的易于误解的表述。**首先**, 我们阐释一下初稿引言部分第四段中所提到的“较少有研究细致的探讨执行功能各子成分在不同的数学能力中起到的作用是否相同”这句话的本意, 即现有研究较少考察各执行功能子成分(即本文中所提到的五个成分)对不同数学能力(基于现有理论的分类)的影响。这样探究的必要性在于尽管干扰抑制与反应抑制同属抑制控制, 工作记忆广度和工作记忆刷新同属工作记忆, 但这些成分的不同方面往往具有不同的认知基础(如工作记忆刷新的神经基础为双侧额叶和右侧顶叶、工作记忆广度的神经基础为右前额叶)(McKenna et al., 2017), 其在数学加工中也扮演中不同的作用。然而, 目前绝大多数研究均是未考虑这些成分的可再分性(Sulik et al., 2018, 2020), 从而导致现有执行功能与数学能力的关系出现研究结果不一致的争议。同时, 绝大多数研究者也未考虑数学能力的多维性, 即缺少基于理论来划分不同的数学能力(如类似本文对数学能力的分类), 而是较为直观的以计数、计算、应用问题或标准化数学测验等来直接衡量个体的数学能力(e.g., Magalhães et al., 2020; Pellizzoni et al., 2020; Yang et al., 2019)。其划分的必要性在于各执行功能成分在不同数学能力中的参与度存在差异, 缺少考虑数学能力的多维性将导致难以全面揭示执行功能与数学能力的关系。**其次**, 经对专家所列举的部分文献进行分析, 发现本文与这些研究所关注的问题仍存在明显差异。具体而言, 虽然 Cragg 和 Gilmore(2014)



在综述中从三个视角(相关研究、实验研究和学习研究)论述了执行功能和数学能力的关系,但仍是未详细区分执行功能的不同方面以及数学能力的多维性; Friso-van den Bos 等人(2013)和 Fuchs 等人(2010)的研究是基于 Baddeley 工作记忆模型,探讨了 phonological loop, central executive, visuospatial sketchpad 和数学技能的关系, 仍是未详细区分执行功能的不同方面以及考虑数学能力的多维性; Schmerold 等人(2017)的研究也仅以 74 名儿童为研究对象,考察了工作记忆广度和干扰抑制与标准化数学测验之间的关系。相比之下,本研究是专门基于执行功能理论,使用大样本来详细探讨各执行功能成分在不同数学能力中的作用,并检验了年龄对其关系的影响。因此,某种程度上是对以往研究的高度整合。**最后**,我们对“不同数学能力”的论述确实比较模糊,非常认可专家的建议。根据专家的意见和建议,我们对引言部分的推论逻辑重新进行了调整,在 **1.1 和 1.2 部分**中阐述清楚了本研究所指的各执行功能成分和不同数学能力,并在**引言 1.3 部分(第 1 段)**对所指表述重新进行了改写,具体修改为:如前所述,执行功能在支持儿童数学能力发展的过程中发挥着重要作用。但目前较少有研究系统探究各执行功能成分在儿童不同数学能力(数学运算能力、逻辑思维能力与空间想象能力)中的作用,以及这种影响模式是否会随儿童年级阶段的不同而显现出差异。

#### 参考文献:

- Cragg, L., & Gilmore, C. (2014). Skills underlying mathematics: The role of executive function in the development of mathematics proficiency. *Trends in Neuroscience and Education*, 3(2), 63–68.
- Friso-Van den Bos, I., Van der Ven, S. H., Kroesbergen, E. H., & Van Luit, J. E. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, 10, 29–44.
- Fuchs, L. S., Geary, D. C., Compton, D. L., Fuchs, D., Hamlett, C. L., & Bryant, J. D. (2010). The contributions of numerosity and domain-general abilities to school readiness. *Child development*, 81(5), 1520-1533.
- Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934–952.
- McKenna, R., Rushe, T., & Woodcock, K. A. (2017). Informing the structure of executive function in children: a meta-analysis of functional neuroimaging data. *Frontiers in human neuroscience*, 11, 154.
- Pellizzoni, S., Apuzzo, G. M., De Vita, C., Agostini, T., Ambrosini, M., & Passolunghi, M. C. (2020). Exploring EFs and math abilities in highly deprived contexts. *Frontiers in psychology*, 11, 383.
- Schmerold, K., Bock, A., Peterson, M., Leaf, B., Vennergrund, K., & Pasnak, R. (2017). The relations between patterning, executive function, and mathematics. *The Journal of Psychology*, 151(2), 207-228.
- Sulik, M. J., Finch, J. E., & Obradović, J. (2020). Moving beyond executive functions: Challenge preference as a predictor of academic achievement in elementary school. *Journal of Experimental Child Psychology*, 198, 104883.
- Sulik, M. J., Haft, S. L., & Obradović, J. (2018). Visual-motor integration, executive functions, and academic achievement: Concurrent and longitudinal relations in late elementary school. *Early Education and Development*, 29(7), 956-970.
- Yang, X., Chung, K. K. H., & McBride, C. (2019). Longitudinal contributions of executive functioning and visual-spatial skills to mathematics learning in young Chinese children. *Educational Psychology*, 39(5), 678–704.

**意见 7:** 关于为什么要研究数学运算能力, 逻辑思维能力和空间想象能力的理论支撑不足。如何从文献综述得出, 数学能力由作者提的这三个成分构成? 另外, 文章中引用的 Spiegel(2021)提到数学的构成部分包括运算, 代数, 几何与测量, 为什么没有研究这些部分? 作者提到了几种关于数学的分类, 为什么采用(吴汉荣, 李丽, 2005)的这种分类方式, 没有明

确描述。

**回应：**非常感谢审稿专家的宝贵意见，此意见同**上述意见 3**。如您所言，本文初稿中确实存在此问题。经分析已有文献，发现现有研究者对于学龄儿童数学能力的具体组成部分仍存在较大争议，如国外研究者指出，学龄儿童的数学能力主要包括运算、代数、几何与测量(Spiegel et al., 2021)，而国内研究者则将学龄儿童数学能力划分为数学运算能力、逻辑思维能力与空间想象能力(林崇德, 2011; 吴汉荣, 李丽, 2005)。如林崇德(2011)重点强调数学思维能力，即思维的深刻性、灵活性、独创性、批判性和敏捷性，认为数学能力是运算能力、空间想象能力和逻辑思维能力与五种思维品质相互交叉构成的统一整体。同时，根据我国历次小学数学教学大纲，中国教育界长期认为学龄儿童的数学能力包含数学运算能力、逻辑思维能力与空间想象能力(或空间观念)三大核心能力(杜文平, 2013)。因此，鉴于本研究的研究对象为中国儿童，为更贴合中国小学数学教育的现状(国内小学数学教学大纲培养学龄儿童数学能力的目标)，所以我们仍沿用国内研究者对数学能力的分类，即将学龄儿童的数学能力分为：数学运算能力、逻辑思维能力与空间想象能力。根据专家的建议，我们在修改稿**(1.1 部分，第二段)**中进行了相应的补充。具体内容如下：**迄今为止，关于数学能力的组成部分，研究者的观点仍未统一。如国外研究者指出，学龄儿童的数学能力主要包括运算、代数、几何与测量(Spiegel et al., 2021)，或运算、代数和几何(Gilmore et al., 2015)。而国内研究者及小学数学教学大纲通常将学龄儿童的数学能力划分为：数学运算能力、逻辑思维能力和空间想象能力(杜文平, 2013; 林崇德, 2011; 吴汉荣, 李丽, 2005)。基于此，本研究继续沿用国内学者对学龄儿童数学能力的分类，将小学儿童的数学能力分为：数学运算能力、逻辑思维能力与空间想象能力。**

#### 参考文献：

- 杜文平. (2013). 小学数学教学大纲或课程标准中数学能力内涵的演变. *中小学数学(小学版)*, 12, 43–45.
- 林崇德. (2011). 智力发展与数学学习(第四版). 北京：中国轻工业出版社
- 吴汉荣, 李丽. (2005). 小学生数学能力测试量表的编制及信效度检验. *中国公共卫生*, 21(4), 473–475.
- Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., & Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychological Bulletin*, 147(4), 329–351.

**意见 8：P3：“如前所述，执行功能在数学.....”**这一段在描述以往研究局限，应该放到文献综述的最后部分。后面的关于执行功能的子成分对不同数学能力的影响应该作为文献综述放在前面，假设部分放到研究问题后面的部分提出。

**回应：**非常感谢审稿专家宝贵的意见，我们非常认可您的建议。根据您所提出的意见 10，我们使用小标题将文章引言部分(综述部分)进行了划分，具体为：**1.1 执行功能子成分与不同数学能力之间的关系；1.2 执行功能子成分与数学能力关系的变化；1.3 研究假设与内容。**将“如前所述，执行功能在支持.....”放置在了引言 1.3 部分的第一段；将“执行功能子成分对不同数学能力的影响”部分放置在了 1.1 部分的第三段；将本文的研究假设在 1.3 部分第一段中全部提出。具体修改内容，我们在文中已用**红色字体**进行了标注。

**意见 9：**关于每个执行功能的子成分包括哪些能力，为什么要研究这些能力没有清晰的描述。比如，以往研究发现 Baddeley 的工作记忆的模型中的三个成分(phonological loop, central executive, visuospatial sketchpad)对数学有影响(e.g., Harvey & Miller, 2017; Rasmussen & Bisanz, 2005; Swanson & Kim, 2007)，为什么没有研究这三个成分？而是区分成工作记忆广度和刷新度？

**回应：**非常感谢审稿专家严谨细致的思考及所提出的宝贵意见。**首先**，回顾本文初稿，我们

在初稿中的确欠缺将执行功能子成分进行再次划分的理由描述,即未明确指出其划分的必要性。在修改稿中(引言 1.1 部分第一段),我们较为细致的阐述了执行功能子成分的不同方面(如,抑制控制分为干扰抑制和反应抑制,工作记忆分为工作记忆刷新和工作记忆广度),及其为何要区分开来探讨对数学能力的影响。具体修改的逻辑为:已有研究发现,执行功能与数学能力的关系存在研究结论不一致的争议(定位于引言第二段后半部分),而造成这种争议的一大原因就在于已有研究并未细致考虑执行功能子成分(抑制控制和工作记忆)的可再分性(1.1 部分中的第一段)。因此,我们需要系统且细致的评估各执行功能子成分在数学能力中的作用,以进一步解释已有研究结果不一致的争议以及系统构建各执行功能成分在不同数学能力中的影响模式。具体修改内容详见引言第二段后半部分和 1.1 部分中的第一段。其次,本文重点探讨了执行功能在儿童数学能力中的作用,而未探究 Baddeley 工作记忆三/四成分对数学能力的影响,其原因在于前者是目前绝大多数研究者的重点关注之处,且现有的研究结论也存在不一致。基于此,本研究的出发点为对已有研究结果不一致的争议做进一步解释,以及用大样本的横断研究来探究其关系变化的大致趋势。与 Baddeley 工作记忆三成分相比,本研究所考虑到的执行功能缺少视觉工作记忆,因此我们非常肯定专家所提的建议,即未来研究需进一步探究 Baddeley 工作记忆三成分在不同数学能力中的作用(以往研究已考察过 Baddeley 工作记忆三成分对儿童数学能力的影响,但较少有研究对数学能力进行划分,并探究年龄对二者关系的影响)。最后,我们简述将工作记忆区分为工作记忆广度和工作记忆刷新的原因。其原因如下:在以往研究中,部分研究者认为执行功能的核心成分,即工作记忆是指工作记忆的容量(即工作记忆广度)(e.g., Ellefson et al., 2020; Fung et al., 2020),而另外一部分研究者认为工作记忆则是指更新信息的能力(即工作记忆刷新)(e.g., Gashaj et al., 2019)。还有研究者同时使用了衡量刷新和广度的任务去评估个体的工作记忆(e.g., Toll et al., 2011)。然而,鉴于工作记忆广度和工作记忆刷新具有可分离性(Cirino et al., 2018; McKenna et al., 2017),两者在数学加工中发挥的作用不同(Friso-Van den Bos et al., 2013)。因此,本研究探讨了工作记忆的刷新和广度在不同数学能力中的作用,以进一步厘清两者对不同数学能力的影响。

#### 参考文献:

- Cirino, P. T., Ahmed, Y., Miciak, J., Taylor, W. P., Gerst, E. H., & Barnes, M. A. (2018). A framework for executive function in the late elementary years. *Neuropsychology, 32*(2), 176–189.
- Ellefson, M. R., Zachariou, A., Ng, F. F. Y., Wang, Q., & Hughes, C. (2020). Do executive functions mediate the link between socioeconomic status and numeracy skills? A cross-site comparison of Hong Kong and the United Kingdom. *Journal of Experimental Child Psychology, 194*, 104734.
- Friso-Van den Bos, I., Van der Ven, S. H., Kroesbergen, E. H., & Van Luit, J. E. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review, 10*, 29–44.
- Fung, W. K., Chung, K. K. H., & Lam, C. B. (2020). Mathematics, executive functioning, and visual-spatial skills in Chinese kindergarten children: Examining the bidirectionality. *Journal of Experimental Child Psychology, 199*, 1–10.
- Gashaj, V., Oberer, N., Mast, F. W., & Roebers, C. M. (2019). The relation between executive functions, fine motor skills, and basic numerical skills and their relevance for later mathematics achievement. *Early Education and Development, 30*(7), 913–926.
- McKenna, R., Rushe, T., & Woodcock, K. A. (2017). Informing the structure of executive function in children: a meta-analysis of functional neuroimaging data. *Frontiers in Human Neuroscience, 11*, 154.
- Toll, S. W., Van der Ven, S. H., Kroesbergen, E. H., & Van Luit, J. E. (2011). Executive functions as predictors of math learning disabilities. *Journal of Learning Disabilities, 44*(6), 521–532.



**意见 10：**文献综述部分可以用小标题把内容区分开来。尤其是，需要加入为什么要研究运算，空间能力和逻辑思维能力这三种数学能力，以及各个成分与不同数学能力之间的关系的文献综述(虽然在 p4 有部分综述，但是比如抑制控制并没有分成后面所测的干扰抑制和反应抑制去讨论他们对数学的影响，所以为什么要区分二者并不清晰)。

**回应：**非常感谢审稿专家宝贵的建议。根据专家的建议，我们已将文献综述部分(1 引言)划分为了三个小标题，分别为：1.1 执行功能子成分与不同数学能力之间的关系(包含了为什么要研究这三个数学能力以及有关执行功能各成分与不同数学能力关系的文献综述)；1.2 执行功能子成分与数学能力关系的变化；1.3 研究假设与内容。此外，我们在修改稿中进一步明晰了抑制控制与数学能力的关系。具体修改思路为：在 1.1 部分第一段中指出了区分为反应抑制和干扰抑制的必要性；在 1.1 部分第三段中阐述了抑制控制的两个方面在本研究所评估的数学能力中发挥的作用较低(基于已有研究)；在 1.2 部分论述了抑制控制对数学能力的影响将随儿童年级的升高而降低(未详细说明干扰抑制、反应抑制与数学能力的发展性关系的原因在于 1.1 部分已经阐明了这两个方面在本研究所测数学能力中的作用较小)。

**意见 11：**假设部分和前面的文献综述以及后面的任务不一致。文献综述部分提到要把这些执行功能做更精细的分类，去研究他们对数学的影响，但是在提出假设的时候抑制控制并没有区分。如果按照假设的意思，抑制控制应该是一个变量，或者用一个潜变量来代替，但是后面又分成了反应抑制和干扰抑制来分析，原因是什么？

**回应：**非常感谢审稿专家细致的思考。**首先**，我们对初稿中此部分写作模糊之处进行说明。在初稿中我们提出假设时以“抑制控制”整体涵盖其两个方面(初稿引言部分第五段末)，而未详细阐明干扰抑制和反应抑制在数学能力中的作用究竟怎样，这是一处较容易引起误解的表述。基于以往研究结果，即工作记忆广度是数学运算能力和空间想象能力最为显著的预测因素(e.g., Friso-Van den Bos et al., 2013; Passolunghi et al., 2008)，而工作记忆刷新则更能预测儿童的逻辑思维能力(e.g., Holmes et al., 2009)。我们当时的写作思路是侧重凸显执行功能子成分对特定数学能力具有最大预测的一对一关系(如，工作记忆广度是数学运算能力和空间想象能力最为显著的预测因素，而工作记忆刷新则更能预测儿童的逻辑思维能力)。但需要承认的是，如果以整体的“抑制控制”来论述时确实会造成文章前后逻辑上的矛盾。**其次**，基于专家的意见 6、8、10，我们对文章的撰写逻辑重新进行了梳理，在 1.1 和 1.2 两部分中分别论述了各执行功能成分与不同数学能力的关系，及其这种关系的发展趋势。基于 1.1 和 1.2 的论述，我们在 1.3 部分第一段中提出了本研究的研究假设，具体内容为：**如前所述，执行功能在支持儿童数学能力发展的过程中发挥着重要作用。但目前较少有研究系统探究各执行功能成分在儿童不同数学能力(数学运算能力、逻辑思维能力与空间想象能力)中的作用，以及这种影响模式是否会随儿童年级阶段的不同而显现出差异。基于上述分析，我们假设不同执行功能成分在不同数学能力中起到的作用存在差异，具体来说，工作记忆广度是数学运算能力和空间想象能力最为显著的预测因素；工作记忆刷新更能预测儿童的逻辑思维能力；干扰抑制、反应抑制和认知灵活性在数学运算、空间想象和逻辑思维中的作用相对较低。同时，鉴于年龄在执行功能与数学能力关系中的作用，我们假设不同执行功能成分对不同数学能力的影响会随儿童年级的升高而发生改变。具体而言，随着年级水平的升高，低阶执行功能成分(干扰抑制、反应抑制、工作记忆广度)在数学能力中的作用会逐渐降低，而高阶执行功能成分(工作记忆刷新、认知灵活性)对数学能力的预测作用将逐渐增大。**

#### **参考文献：**

- Friso-Van den Bos, I., Van der Ven, S. H., Kroesbergen, E. H., & Van Luit, J. E. (2013). Working memory and mathematics in primary school children: A meta-analysis. *Educational Research Review*, 10, 29–44.
- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor

working memory in children. *Developmental Science*, 12(4), F9–F15.

Passolunghi, M. C., Mammarella, I. C., & Altoè, G. (2008). Cognitive abilities as precursors of the early acquisition of mathematical skills during first through second grades. *Developmental Neuropsychology*, 33(3), 229–250.

**意见 12：**假设部分只提出了工作记忆广度和工作记忆刷新对于不同的数学能力的预测作用。没有对抑制控制，认知灵活性对数学的几种能力有何不同的预测作用，以及抑制控制中的反应抑制和干扰抑制有何不同的作用也没有作出假设。另外针对不同的年级这些关系有何变化，也没有对应的假设。应该把研究问题和针对每个研究问题的所有的假设都单独列出来。

**回应：**非常感谢审稿专家的意见和建议。在此，我们先回答专家的第一个问题，即为何只突显出了工作记忆广度和工作记忆刷新对不同数学能力的预测作用，而对抑制控制和认知灵活性未作出详细的假设。具体来说，由于执行功能是促使个体完成目标导向行为的一系列认知能力(Diamond, 2013)，其核心成分包含了工作记忆、抑制控制和认知灵活性。已有研究也证实了这些不同成分之间存在部分重叠(Himi et al., 2021; McKenna et al., 2017)，因而较难精准衡量单个成分在数学能力中的作用。基于此，已有绝大多数研究均是在同一研究中考察这三个成分在数学能力中的作用(e.g., Ahmed et al., 2019)，即同时比较所获得的相对作用大小。根据已有研究，我们可清晰推论出工作记忆广度和工作记忆刷新相比抑制控制(干扰抑制、反应抑制)和认知灵活性，各自在这三个数学能力中发挥的作用较大。因此，我们将其论述为相比于抑制控制(干扰抑制、反应抑制)和认知灵活性，工作记忆广度是数学运算能力和空间想象能力最为显著的预测因素；工作记忆刷新更能预测儿童的逻辑思维能力。其理由为：既然已经阐明了这两者(抑制控制、认知灵活性)的作用较小，那么继续分别做出假设会显得较为赘述。但我们以“抑制控制”来直接涵盖两个成分，确有不妥，我们已在修改稿假设部分做出了相应的修改(1.3 部分第一段)。同时，我们在修改稿 1.2 部分中论述了年龄对执行功能子成分与数学能力关系的影响，阐明了各个子成分与数学能力关系的发展趋势。此外，基于 1.1 与 1.2 的相关论述，我们在修改稿 1.3 部分的第一段中提出了本研究的研究假设。具体修改内容(1.1、1.2、1.3)，我们在文中已用红色字体进行了标注。

#### 参考文献：

Ahmed, S. F., Tang, S., Waters, N. E., & Davis-Kean, P. (2019). Executive function and academic achievement: Longitudinal relations from early childhood to adolescence. *Journal of Educational Psychology*, 111(3), 446–458.

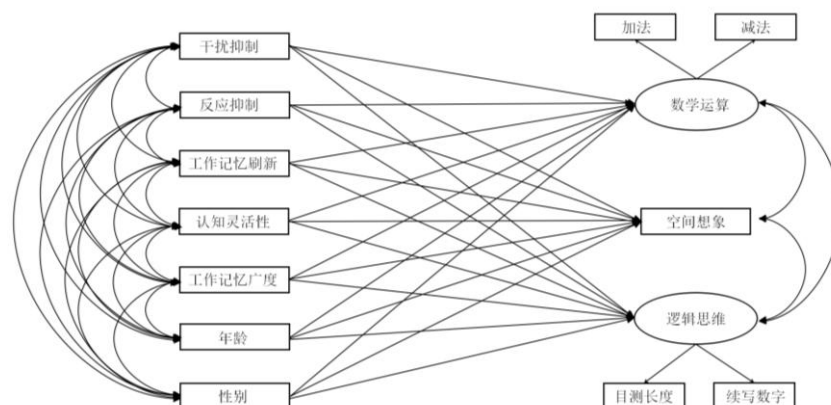
Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.

Himi, S. A., Bühner, M., & Hilbert, S. (2021). Advancing the understanding of the factor structure of executive functioning. *Journal of Intelligence*, 9, 16.

McKenna, R., Rushe, T., & Woodcock, K. A. (2017). Informing the structure of executive function in children: a meta-analysis of functional neuroimaging data. *Frontiers in Human Neuroscience*, 11, 154.

**意见 13：**图 1 应该标注出来每个执行功能的成分，而不是用任务/正确率来构建理论模型。

**回应：**感谢审稿专家细致的思考和建议。在修改稿中，我们使用执行功能各个子成分替代了原来的任务/正确率。新构建的理论模型图(图 1)如下：



**意见 14：**被试部分提到 13 名学生未完成所有的测验，予以剔除。如果是数据不全的情况，可以用最大信息量(full information maximum likelihood, FIML)或者其他插补的方法去处理缺失值的，不应该直接删除。另外应该做缺失值分析，是否是随机缺失。

**回应：**非常感谢审稿专家的意见。诚如专家所述，在结构方程模型统计分析中，缺失值的处理是一个关键的问题，处理缺失值的常用方法主要包括删除法(Listwise Deletion, LD)、多重插补(Multiple Imputation, MI)、全息极大似然估计法(Full Information Maximum Likelihood, FIML)等(邓居敏等, 2018; 王孟成, 邓俏文, 2016; Graham, 2009)，其后两种方法更受研究者的推崇。然而，本文中所提到有 13 名学生未完成所有测验，其本意是指这些学生只有一个或至多两个观测变量的数据(如，只有反应抑制和正背)，其余变量的相应观测值因学生未参加测试(请假错过测试时间)或研究助手遗漏拷贝数据等原因均未收集到相应数据。因此，在后期的数据整理中，我们考虑到本研究的样本量较大，这些被试中单个被试的缺失数据较多，采用删除法应是保证研究结果可靠的有效途径，从而对这几名被试的有限数据进行了永久性删除。同时，从缺失值处理的方式来看，由于本文中缺失数据既有可恢复的缺失值(完全随机缺失)和不可恢复的缺失值(未测量，不存在)，因而采用删除法是降低估计系数出现偏误的有效方式。

#### 参考文献：

- 邓居敏, 陈羽, 关颖. (2018). 基于结构方程模型的缺失数据填补方法比较. *数理医药学杂志*, 31(2), 159–161.
- 王孟成, 邓俏文. (2016). 缺失数据的结构方程建模: 全息极大似然估计时辅助变量的作用. *心理学报*, 48(11), 1489–1498.
- Graham, J. W. (2009). Missing data analysis: Making it work in the real world. *Annual Review of Psychology*, 60(1), 549–576.

**意见 15：**为什么用目测长度来表示逻辑思维能力？关于为什么要这样计分的依据应该要进行描述。目测长度涉及到数学中的测量(measurement)能力，为什么这个任务是测量思维逻辑能力而不是测量能力？

**回应：**非常感谢审稿专家所提出的意见。逻辑思维能力是指对于数学现象的比较分析，是个体通过比较、分类、概括、类比、归纳与演绎、分析与综合等来解决数学问题的能力(谭瑞, 李汉, 1989)。在本文中，我们使用了目测长度分测验作为衡量儿童逻辑思维能力的一项指标，其原因包括以下三点：**第一**，基于编制该量表作者本人的分类。吴汉荣和李丽(2005)引进并修订了小学生数学基本能力测试量表，并认为目测长度是衡量儿童逻辑思维能力的一项分测验。**第二**，该任务的确符合逻辑思维能力的定义。如附录(目测长度)所示，该测验是给出了三条标准长度，要求被试根据所示的三个标准去估计题干中的长度为多少。这要求儿童根据

题干去快速选定所参照的标准,通过比较分析来灵活的变换思路,从而快速解决当前的问题,与逻辑思维能力紧密相关(孙敦甲,1987)。**第三**,就目测长度测验的任务要求而言,确实涉及到数学中的测量能力,即将测定的量与一个作为标准的同类量进行比较的能力(张华等,2003)。但研究者指出,数学测量任务的解决需要个体使用大量的逻辑思维能力(史亚娟等,2007)。综上,本研究以目测长度分测验作为衡量儿童逻辑思维能力的一项指标。此外,根据专家的意见,我们在 2.2 测验工具部分(**2.2.1 数学能力中第一段末尾**),从整体上补充了各测验的计分方式。具体增加内容为:各测验的计分方式为每对一题,记为 1 分,总分越高,代表被试相应的数学能力越高(吴汉荣,李丽,2005)。

#### 参考文献:

- 孙敦甲.(1987). 学生数学能力结构. *心理发展与教育*, 4, 42-46.
- 史亚娟, 韩小雨, 张华, 庞丽娟. (2007). 4-6 岁儿童对空间测量中逻辑关系的理解研究. *教育理论与实践* 27(2),28-30.
- 谭瑞, 李汉. (1989). 小学生数学能力结构探讨. *心理发展与教育*, 2, 41-46.
- 吴汉荣, 李丽. (2005). 小学生数学能力测试量表的编制及信效度检验. *中国公共卫生*, 21(4), 473-475.
- 张华, 庞丽娟, 陶沙, 陈瑶, 董奇. (2003). 儿童早期数学认知能力的结构及其特点. *心理学报*, 6,810-817.

**意见 16:** 为什么选了 6 个分测验来考察,而不是原始的 12 个分测验? 选择的依据是什么? 比如图形计数是否也算是空间想象能力? 为什么没有包括进来。以及其他的运算能力。

**回应:** 非常感谢审稿专家严谨细致的思考及意见。本文仅选用了《小学生数学基本能力测试量表》中的六个分测验(包含一个热身任务)来评估儿童的数学能力。其原因如下:**第一**,参照以往研究中衡量特定数学能力(数学运算能力)的常用工具。在已有研究中,研究者们最常使用加、减法测验来评估个体的运算能力(e.g., Best et al., 2011; Fung et al., 2020)。因此,为使研究结果更具有可比性,我们只使用加法和减法分测验来评估个体的数学运算能力。**第二**,基于儿童实际掌握的数学知识和准确评估所考察的数学能力。详细来说,该量表最初用来评估数学运算能力的分测验包括加法、减法、乘法、除法、比较大小和填空六个分测验,但由于三年级儿童还未系统学习乘法和除法知识,其分数相比其他高年级肯定出现地板效应,吴汉荣和李丽(2005)在问卷修订中也指出对三年级儿童不进行乘法和除法测验。因此,在衡量数学运算能力时排除乘法、除法分测验。同时,完成该量表中比较大小和填空两个分测验的数学基础几乎均为加、减法,即需要熟练的加、减法技能来充当解决相应问题的重要辅助,因而所考察的数学技能的重点还是在于加、减法运算的熟练度。因此,本文未再选用这两个测验来考察儿童的运算能力。此外,在剩余的五个任务中,我们选用目测长度分测验与续写数字分测验来衡量个体的逻辑思维能力(**关于选用目测长度分测验来评估逻辑思维能力的理由已在意见 15 中做出了详细说明**),选取方块计数分测验来评估个体的空间想象能力。我们未选用图形计数和数字连接分测验的原因为:图形计数是一项衡量儿童数数能力的测验,即儿童需要快速的数出图示(如下图 1)中的简单形状总共有多少个,因而是一项计数测验。并且在吴汉荣和李丽(2005)的研究中,图形计数与其他测验间均呈现出较低的相关( $r$  的范围为 0.206~0.325),这也在一定程度上说明其所测量的数学能力与其他测验存在差异。我们未选用数字连接分测验的原因为:该测验也是一项类似计数的测验,即儿童需要通过快速的视觉搜索来找寻数字并按照其顺序来连接数字链(如下图 2),因此主要衡量了儿童的视觉跟踪能力(吴汉荣,李丽,2005)。并且在吴汉荣和李丽(2005)的研究中,数字连接与其他测验间也均呈现出较低的相关( $r$  的范围为 0.209~0.285),但与图形计数测验间呈现最高的相关度  $r=0.325$ 。这说明,图形计数与连接数字所衡量的数学能力具有相似性,主要是以衡量个体的计数能力为主。综上,本研究未选取图形计数与连接数字分测验来衡量个体的数学能力。为在文中做进一步澄清,我们在 2.2.1 中加入了补充描述:基于儿童实际掌握的数学知识和



准确评估所考察的数学能力,我们选取了该量表中的 6 个分测验用以评估儿童数学能力的三个方面。

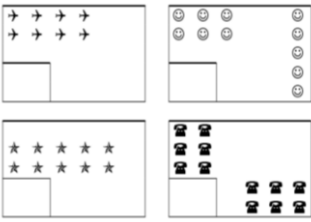


图 1 图形计数测验示例

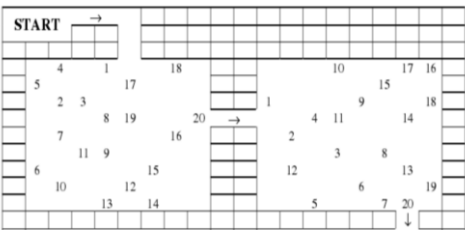


图 2 数字连接测验示例

意见 17: 数学运算能力用了一个计时的测试, 考察的是学生的运算的流畅度(arithmetic fluency), 为什么用运算流畅度而不是用非计时计算测试来考察学生的运算能力?

回应: 感谢审稿专家的意见。本文选用《小学生数学基本能力测试量表》(中国修订版)中的加法测验和减法测验来评估儿童的数学运算能力, 其理由如下: **首先**, 《小学生数学基本能力测试量表》是对《德国海德堡大学小学生数学基本能力测试量表》的引进与修订, 其适宜测试我国小学生的基本数学能力, 具有较好的信、效度(吴汉荣, 李丽, 2005), 在以往研究中使用广泛(e.g., 吴甲旺等, 2022; 曾桐奥等, 2020)。同时, 该量表中的各分测验由于具有限时和题量较大的特点, 大大降低了高年级儿童测试分数出现天花板效应的风险, 从而使各年级儿童的测试分数更具有可比性。**其次**, 加法与减法分测验(如附录所示)的限时性及项目难度递增性是一个典型的速度测验和难度测验的结合, 因而可以更加准确的评估个体的数学运算能力, 并且以往研究也显示加、减法测验是评估儿童数学运算能力的重要方式(e.g., Best et al., 2011; Fung et al., 2020)。**最后**, 诚如专家所述, 本文评估儿童数学运算能力的加、减法测验主要衡量了学生的运算流畅度(吴甲旺等, 2022), 但以往研究也证实了这种限时测验可以有效甄别学生的数学运算能力(American Psychiatric Association, 2013), 是衡量个体数学运算能力的重要方式(Zhou et al., 2015)。

参考文献:

吴汉荣, 李丽. (2005). 小学生数学能力测试量表的编制及信效度检验. *中国公共卫生*, 21(4), 473-475.

吴甲旺, 李红霞, 司继伟. (2022). 小学低年级儿童的计算流畅性与数学焦虑: 基于变量为中心与个体为中心的分析. *心理发展与教育*, 38(1), 72-80.

曾桐奥, 卜晓鸥, 王庭照, 马海林. (2020). 听障儿童抑制控制能力特点及其与数学基本能力的关系. *心理与行为研究*, 18(4), 517-523.

American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders* (5th ed.). Arlington: American Psychiatric Publishing.

Best, J. R., Miller, P. H., & Naglieri, J. A. (2011). Relations between executive function and academic achievement from ages 5 to 17 in a large, representative national sample. *Learning and Individual Differences*, 21(4), 327-336.

Fung, W. K., Chung, K. K. H., & Lam, C. B. (2020). Mathematics, executive functioning, and visual-spatial skills in Chinese kindergarten children: Examining the bidirectionality. *Journal of Experimental Child Psychology*, 199, 1-10.

Zhou, X., Wei, W., Zhang, Y., Cui, J., & Chen, C. (2015). Visual perception can account for the close relation between numerosity processing and computational fluency. *Frontiers in Psychology*, 6, 1364.

意见 18: 描述统计和差异检验和相关分析的篇幅可以减少。因为, 本研究的主要目的是探讨执行功能对与数学能力的影响, 年级差异并不是主要关心的问题。而且, 讨论部分也没有



针对这部分的结果作出讨论，这部分结果可以精简一些。

**回应：**非常感谢审稿专家的建议。根据您的意见，我们已对结果部分(3.1、3.2、3.4)的有关内容进行了相应精炼，具体修改内容已在正文中使用红色字体进行标注。请专家详见结果部分中的3.1、3.2及3.4。

**意见 19：**各个结构方程模型，都做了哪些修正应该呈现出来，不能单纯为了模型拟合好，就把修正路径都加上。需要呈现为什么要加这些修正的路径。在对各个年级阶段的模型拟合中，也都分别修正了模型，要清楚的解释为什么要做这些修正，且在不同年龄阶段的模型中的修正的路径是否相同？如果不同怎么解释？

**回应：**我们非常赞同审稿专家的意见。**首先**，我们对初稿三个模型中的修正路径做进一步明晰，分别为：3~6 年级阶段的模型中修正路径为年龄与目测长度误差项；3~4 年级阶段的模型中修正路径为工作记忆刷新与目测长度误差项；5~6 年级阶段的模型中修正路径为年龄与目测长度误差项。**其次**，经进一步查阅结构方程建模的相关文献，我们意识到初稿中对三个模型的修正仅考虑了使模型的拟合指数更佳，而缺乏从理论角度考虑做出相应修正的理论依据，因此忽视了修正后模型结果的现实意义和理论价值。同时，对于不同年级阶段模型所作出的不同修正导致了不同年级阶段的模型结果无法直接进行比较，从而难以揭示各执行功能子成分与不同数学能力关系的发展模式。这的确是本文初稿建模时欠缺考虑之处。**最后**，结合专家意见及更符合现有理论依据，在修改稿中将性别因素作为控制变量纳入模型后，发现三个模型的初始模型已具有较好的模型拟合度，即证实了假设模型的合理性，因而在修改稿中未做任何修正。因此，无需在修改稿中阐释修正路径的有关问题，以及可以对修改稿中不同年级阶段模型结果进行直接比较。

**意见 20：**只有年龄作为一个控制变量是一个重要的缺陷。同时影响数学能力和执行功能的变量有很多，通常需要排除年龄，性别和父母教育水平或者家庭 ses 的影响。以及需要排除语言能力与智力的影响。这些因素都会影响工作记忆，和数学能力。虽然作者在研究局限中提到缺乏对这些变量的考虑，但是性别至少可以加入到模型中作为一个控制变量。

**回应：**非常感谢审稿专家提出的宝贵意见和建议，此意见同上述意见 5。如您所言，年龄、性别、父母受教育程度、家庭社会经济地位、语言能力和智力等均会对个体的数学能力和执行功能有一定影响，本研究在施测时也确实存在欠缺对部分混淆变量的考虑。基于专家的建议，我们将性别因素作为控制变量纳入了本研究的假设理论模型。并在修改稿的各模型中对性别这一变量进行了控制，新构建的结构方程模型图以及对模型结果的相关表述已在文章(3.3、3.5)中使用红色字体作出了标识。

**意见 21：**为什么要分成三四年级，五六年级来进行分析结构方程模型的分析？前面的文献综述并没有提供这样分析的依据。这样分开分析的假设是三四年级的执行功能与数学的关系和五六年级阶段的执行功能与数学的关系是不一致的。虽然文献综述有提到随着年龄增长，工作记忆的影响会减弱，但是并没有提出为什么三四年级和五六年级是不同的，这样分的依据是什么？为什么不每个年级单独做一个模型呢？

**回应：**非常感谢审稿专家所提出的意见和建议。我们在初稿中确实缺少划分为这两个年龄阶段的有力描述。结合前人对儿童数学思维发展阶段的探讨、满足建构复杂结构方程模型样本量及参照以往研究者在横断研究中对年龄段的划分，我们在修改稿中仍将儿童的年龄段划分为三~四年级阶段与五~六年级阶段。具体理由如下：**第一**，在以往研究中，研究者们对学龄儿童数学能力的发展阶段做了一些探讨，如刘范等(1981)考察了学龄儿童数学概念的发展阶段，发现一~二年级儿童只能解决简单的计算问题，空间观念和逻辑思维还未有明显发展；

三~四年级儿童的数概念体系逐步形成，具有了一定的空间观念与抽象思维；五~六年级儿童的数概念体系已经成熟，空间观念与抽象思维获得了较大的发展。于萍和左梦兰(1996)探讨了三~六年级小学生数学能力与认知结构的发展，结果显示，三~六年级儿童数学能力会随年级的增长而提高，即三~四年级儿童处于具体运算阶段，对数学问题的加工是一种双重协调的过程；而五~六年级儿童的抽象推理能力获得较大发展，对数学问题的加工则转向精致协调。由此可见，三~四年级儿童与五~六年级儿童的数学加工及所具备的数学思维存在明显差异，因而分别对这两个年级阶段儿童进行考察将更有助于了解儿童加工不同数学能力的实质。这也是划分为不同年级阶段的最关键原因。**第二**，由于本研究的假设理论模型较为复杂，模型中的预测路径较多，因而考虑到使模型的参数估计更为准确以及降低模型出现过拟合问题的风险，将三~四年级与五~六年级阶段数据分别合并是保证模型具备大样本量支持的重要方式。**第三**，结合前人研究，如 Magalhães 等人(2020)使用横断研究方法考察了学龄儿童执行功能对学业成绩的预测作用，儿童的年龄阶段分别为 2、4、6 年级。虽然作者并未阐明为何要以 2、4、6 年级学生为研究对象(而非 1、3、5)，但我们认为这或许与执行功能成分的发展速度有关。已有大量研究表明，尽管执行功能各成分的成熟是一个线性增长的过程(e.g., Andrews et al., 2021; Luciana et al., 2005)，但认知能力的显著提升始终较为缓慢，因此鉴于执行功能在支持儿童数学能力发展中的重要作用(e.g., Wang et al., 2019)，考察具有较大年龄间隔儿童的执行功能与数学能力的关系将更能清晰勾勒出两者关系变化的趋势。基于以上表述，本文将儿童划分为三~四年级与五~六年级两阶段分别建立结构方程模型，以刻画出各执行功能子成分与不同数学能力关系的发展趋势。根据专家的意见，我们在文献综述部分(1.3 部分第二段)补充了为何要划分为三~四年级阶段与五~六年级阶段的相关论据，具体增添与修改内容如下：与前人研究相反，本研究主要关注于小学中高年级的儿童，其原因主要分以下三点：第一，选取学龄中高年级的儿童将降低由于发育迟缓和缺乏数学指导所导致的数学成绩较差的风险。第二，这些儿童在数学能力的发展上已经超越了简单计算，其逻辑思维和空间想象等数学能力获得了较大发展，因而兼具了不同的数学能力。第三，儿童各执行功能子成分已获得了进一步发展(Lehto et al., 2003)。基于于萍和左梦兰(1996)的研究，三~六年级儿童的数学能力会随年级的增长而逐渐提高，即三、四年级儿童处于具体运算阶段，对数学问题的加工是一种双重协调的过程；而五、六年级儿童的抽象推理能力获得较大发展，对数学问题的加工则转向精致协调。由此，对学龄中高年级儿童进行考察将更清晰的刻画出各执行功能成分与不同数学能力间的实质联系。因此，本研究将通过一系列任务考察三至六年级儿童在各执行功能子成分和不同数学能力上的表现，通过对各任务间关系的分析来验证本研究的假设。在控制了年龄和性别的影响后，我们构建了本研究的假设理论模型，如图 1 所示。

#### 参考文献：

- 刘范, 吕静, 沈家鲜, 赖昌贵, 张增杰, 刘静娴, 丁松年, 徐秀嫦, 曹子方, 王鹏飞. (1981). 国内十个地区 7—12 岁儿童数学概念和运算能力发展的初步研究. *心理学报*, 2, 135–149.
- 于萍, 左梦兰. (1996). 三~六年级小学生数学能力及认知结构的发展. *心理发展与教育*, 3, 30–36.
- Andrews, K., Atkinson, L., Harris, M., & Gonzalez, A. (2021). Examining the effects of household chaos on child executive functions: A meta-analysis. *Psychological Bulletin*, 147(1), 16–32.
- Luciana, M., Conklin, H. M., Hooper, C. J., & Yarger, R. S. (2005). The development of nonverbal working memory and executive control processes in adolescents. *Child Development*, 76(3), 697–712.
- Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934–952.
- Wang, C., Jaeggi, S. M., Yang, L., Zhang, T., He, X., Buschkuhl, M., & Zhang, Q. (2019). Narrowing the

**意见 22:** 在三四年级的模型中, 刷新成绩与目测长度有一个显著的负相关, 如何解释?

**回应:** 感谢审稿专家的意见。非常抱歉! 我们对此问题的回答较为有限。在初稿中, 我们在三~四年级阶段的结构方程模型中构建了一条工作记忆刷新与目测长度误差项之间的修正路径, 其本意只为使模型的各拟合指标更好, 欠缺从理论层面分析构建此修正路径的必要性。同时, 这两个变量间的残差系数为-0.23\*\*\*, 与皮尔逊相关系数 0.15\*\*(三~四年级阶段中工作记忆刷新与目测长度的皮尔逊相关)的相关关系方向刚好相反, 经查阅大量有关结构方程模型建模的书籍和文献, 认为最合理的解释应为估计值(拟合值)可能大于了实际观测值, 从而导致残差系数为负。但需要进一步澄清的是, 由于缺乏做出相应修正的理论依据和保障模型结果的现实意义和理论价值, 以及基于新构建的各个模型的各项拟合指标已完全达到统计要求, 因此我们在修改稿中对三个模型均未作任何修正, 因而不存在上述问题。这不仅符合了建模的理论基础要求, 也使不同模型的结果更具可比性。

**意见 23:** 结构方程模型在不同的年龄阶段, 模型的路径存在大量差异。比如在执行功能各个变量的相关路径有大量差异。所以这两个模型造成的结果差异如何判定是因为年级的差异还是因为各个路径的设定不同的原因?

**回应:** 非常感谢审稿专家的意见, 此意见同**上述意见 4**。如您所说, 我们在初稿中确实存在对不同年级阶段模型进行不同修正的问题, 但此问题在修改稿中不存在(**详见所回应的意见 19**)。此外, 我们需要进一步澄清的是不论在初稿还是修改稿中, 三个结构方程模型中自变量之间的相关路径均为一致, 即与本研究的假设理论模型(图 1)保持一致。在 Amos 中构建每个模型时, 我们画出了所有自变量间的相关关系(除性别与年龄之间的相关路径), 但由于部分变量间的相关关系并不显著, 因此在所示的结构方程模型图中并未画出不显著相关的路径, 从而造成了表面上各模型中自变量路径不一致的现象。基于专家的意见 5 与意见 20, 我们在修改稿的假设理论模型(图 1)中也将性别作为了一个控制变量。基于已有研究基础, 即性别与执行功能各子成分和不同数学能力、年龄与执行功能各子成分和不同数学能力、执行功能各子成分与不同数学能力、执行功能各子成分间均存在一定的联系, 本文提出了新的假设理论模型(图 1)。在此, 我们仍需说明的是, 修改稿中三个结构方程模型自变量间的相关路径均为一致, 所有模型的初始模型已完全达到模型拟合指标的要求, 因而对所有模型未做任何修正。为简化模型及提升模型结果的清晰表达性, 模型图中仍不呈现自变量相关未达到显著水平的路径, 但我们在结果部分(**图 2 下方**)统一做出了说明, 这样可有效规避专家和读者可能存在的疑惑。因此, 修改稿中不同年级阶段模型的路径设定均是一致, 可通过比较不同年级阶段模型结果的差异来获得各执行功能子成分与不同数学能力关系的发展趋势。

**意见 24:** 讨论部分 p21 “工作记忆广度对三种基本数学能力的预测将随儿童年级的增高而降低。”这里对于结果的描述并不准确。因为作者只比较了三四年级和五六年级, 而且是横断研究, 并不能说明随儿童年级的增高而降低, 因为并不知道从三年级到四年级的情况是如何, 从五年级到六年级的情况如何。只能说明工作记忆广度与数学能力的关系在高年级儿童和低年级儿童呈现出不同的模式。类似的问题也出现在其他的表述中“认知灵活性对儿童基本舒徐嗯能力的影响将随儿童的年级的增高而变大”, 和结论中。

**回应:** 非常感谢审稿专家细致的思考和建议, 我们对这几部分的表述确实存在问题。如专家所言, 横断研究无法揭示因果关系, 以及文中模型并非包含所有年级。基于专家的建议, 我们在修改稿中尽量抹去了因果关系的类似表述, 以及对您所提到的需修改的几处进行了重新

表述，具体修改之处请您详见文中(4.2、5 结论)的红色字体。

.....

## 审稿人 2 意见：

**意见 1：**作者提到执行功能结构的三因素模型在三至六年级的儿童拟合最好，为何作者又进一步将执行功能细分为 5 个子成分，有何特殊的理论或实践价值吗？

**回应：**非常感谢审稿专家的意见。在初稿引言部分的最后一段中，我们阐明了本研究关注年龄较大儿童的有关依据，其第三个原因为“第三，尽管已有研究证实了学龄儿童执行功能结构的三因素模型，但这种结构在三至六年级的儿童中拟合最好(Lehto et al., 2003)”。**首先**，我们先解释一下做此论证的原因。在以往研究中，研究者们对于个体在何时真正具备执行功能的三个成分(工作记忆、抑制控制、认知灵活性)仍存在争议。如有研究者认为，学前儿童就已经具有执行功能的三因素成分(e.g., Fung et al., 2020)，而另外一部分研究者则指出，学龄中高年级儿童才完全具备执行功能的三因素成分(Karr et al., 2018)尤其是三至六年级儿童(Lehto et al., 2003)。因此，初稿中作此论证的本意是突出当前研究的合理性，即在儿童各执行功能成分相对分离时考察其与不同数学能力的关系将更能揭示两者间的实质联系。**其次**，需要承认的是我们在初稿中确实欠缺将工作记忆和抑制控制进行再分的相关描述。虽然我们在修改稿(1.1 部分和 1.2 部分)中已经补充了相关内容，但我们在此对再分的必要性再作以简述，即以以往研究所得出的执行功能子成分与数学能力的关系并不一致。而导致这种结果不一致的主要原因有以下三点：(1)多数研究并未系统考虑执行功能子成分的可再分性(抑制控制分为干扰抑制与反应抑制，工作记忆分为工作记忆广度和工作记忆刷新)；(2)已有研究较少考虑数学能力的多维性；(3)以往研究较少考虑年龄/年级对执行功能各成分与不同数学能力关系的影响。因此，细致的探究各执行功能子成分与不同数学能力的关系将不仅可以解答以往研究中所存在的部分争议，同时这种基于横断视角的研究还可以揭示其关系变化的大致趋势。**最后**，尽管三至六年级儿童的执行功能是一个三因素结构，但是干扰抑制与反应抑制、工作记忆广度和工作记忆刷新分别代表了抑制控制与工作记忆的不同方面(Bryce et al., 2011; Yang et al., 2019)，因此，在考察执行功能与儿童数学能力的关系时对这两个成分进行细分将更能揭示各执行功能成分与不同数学能力间的实质联系。综上，本横断研究系统探究了各执行功能成分在不同数学能力中的作用，以及年龄/年级对两者关系的影响。其有助于进一步明晰执行功能与数学能力的关系，所得出的研究结论对于未来两者关系的进一步探索具有较好的指引作用。

## 参考文献：

- Bryce, D., Szűcs, D., Soltész, F., & Whitebread, D. (2011). The development of inhibitory control: an averaged and single-trial lateralized readiness potential study. *NeuroImage*, 57(3), 671–685.
- Fung, W. K., Chung, K. K. H., & Lam, C. B. (2020). Mathematics, executive functioning, and visual-spatial skills in Chinese kindergarten children: Examining the bidirectionality. *Journal of Experimental Child Psychology*, 199, 1–10.
- Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, 21(1), 59–80.
- Karr, J. E., Areshenkoff, C. N., Rast, P., Hofer, S. M., Iverson, G. L., & Garcia-Barrera, M. A. (2018). The unity and diversity of executive functions: A systematic review and re-analysis of latent variable studies. *Psychological Bulletin*, 144(11), 1147–1185.
- Yang, X., Chung, K. K. H., & McBride, C. (2019). Longitudinal contributions of executive functioning and visual-spatial skills to mathematics learning in young Chinese children. *Educational Psychology*, 39(5),



**意见 2:** 横断研究无法揭示执行功能的不同成分与数学成就之间的关系是如何发展变化的, 并且无法得出认知能力与数学成就之间的因果关系, 目前有研究表明, 数学成就也可以影响儿童认知能力的发展。

**回应:** 非常感谢审稿专家严谨细致的思考及意见。如您所言, 横断研究是发展心理学研究中常用的研究设计, 其仅测查不同年龄被试在特定时间点各心理属性间的关系, 因而无法获取个体心理属性间关系的连续发展趋势, 难以对变量之间的因果联系做出推论。可见, 这也是本研究所存在的一大研究局限。对此, 我们已在文章的展望部分对其进行了说明, 并提出了未来研究可以使用纵向研究方法来进行进一步考察因果关系的相应展望。虽然本研究所使用的横断研究方法存在难以解释变化和因果关系的局限, 但纵观已有探究执行功能与儿童数学能力关系的相关研究, 本横断研究也存在一些明显优势。具体如下: **第一**, 相比于考察执行功能与数学能力关系的类似横断研究(Cragg et al., 2017; Gómez et al., 2015), 本研究具有较大的样本量(本研究中被试总数为 812; Cragg 等人的研究中被试总数为 293; Gómez 等人的研究中被试总数为 450), 因而这种基于大样本所获得的研究结果更具有可信性与稳定性, 也更加符合结构方程模型对于大样本量的要求。**第二**, 本研究中不同年级组被试的年龄差异较小, 这大大规避了横断研究中存在的代群效应(唐细容等, 2013), 即得出年龄/年级差异在很大程度上是造成不同年级组被试两者关系发生变化的主要原因。因此, 本横断研究可以在一定程度上揭示各执行功能子成分与不同数学能力间关系的变化趋势, 其不仅回答了以往研究中的部分争议, 还为未来研究者进一步探究其因果联系提供了一定研究思路的参考。**第三**, 相较于已有研究, 本研究考虑到了执行功能子成分的再分性(即抑制控制分为干扰抑制与反应抑制, 工作记忆分为广度和刷新)、数学能力的多维性(运算、逻辑思维、空间想象)、年龄效应(横断视角)以及使用了结构方程模型, 这不仅评估了各执行功能成分在不同数学能力中的作用, 也探究了这种影响模式随年级升高而变化的趋势。其所得出的重要研究结论(低阶执行功能成分在低年级儿童数学能力的发展中所起的作用较大, 高阶执行功能成分支持高年级儿童数学能力的重要认知因素)是对于已有研究结果的重要补充, 即高阶执行功能成分除支持解决更复杂数学问题外(Magalhães et al., 2020), 也会在大龄儿童简单数学问题的加工中发挥重要作用。这说明执行功能成分对数学能力的影响是与个体各执行功能成分的发展顺序相一致, 即低阶执行功能成分成熟较早, 对早期数学能力的影响较大; 高阶执行功能成分成熟较晚, 在后期数学能力的发展中发挥的作用较大。**第四**, 如专家所指, 现有少数研究发现儿童的数学能力也可以预测随后的执行功能(Watts et al., 2015; Welsh et al., 2010), 但也有研究结果并未证实数学能力对执行功能具有预测作用(e.g., Fung et al., 2020)。由此可见, 执行功能与数学能力间或许存在双向预测, 但这些结果仍无法判定二者间存在因果关系, 对因果关系的考察还需要更严格的实验设计来进行考证。因此, 我们在研究展望部分阐述了未来研究可以使用追踪和实证研究(干预研究)来进一步揭示二者间因果关系的相应展望。综上, 虽然本横断研究并未揭示出执行功能与数学能力的因果关系, 但对于明晰各变量间的关系及指引未来研究者进一步探究两者之间的实质联系具有重要意义, 因而具有重要的理论与实证意义。

#### 参考文献:

- 唐细容, 蒋莉, 曾慧, 姚树桥. (2013). 成人认知功能的发展与老化: 一项控制同辈效应的横断研究. *心理发展与教育*, 29(3), 262-267.
- Cragg, L., Keeble, S., Richardson, S., Roome, H. E., & Gilmore, C. (2017). Direct and indirect influences of executive functions on mathematics achievement. *Cognition*, 162, 12-26.
- Fung, W. K., Chung, K. K. H., & Lam, C. B. (2020). Mathematics, executive functioning, and visual-spatial skills in Chinese kindergarten children: Examining the bidirectionality. *Journal of Experimental Child Psychology*.



*Psychology*, 199, 1–10.

- Gómez, D. M., Jiménez, A., Bobadilla, R., Reyes, C., & Dartnell, P. (2015). The effect of inhibitory control on general mathematics achievement and fraction comparison in middle school children. *Zdm*, 47(5), 801–811.
- Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934–952.
- Watts, T. W., Duncan, G. J., Chen, M., Claessens, A., Davis-Kean, P. E., Duckworth, K., ... & Susperreguy, M. I. (2015). The role of mediators in the development of longitudinal mathematics achievement associations. *Child Development*, 86(6), 1892–1907.
- Welsh, J. A., Nix, R. L., Blair, C., Bierman, K. L., & Nelson, K. E. (2010). The development of cognitive skills and gains in academic school readiness for children from low-income families. *Journal of Educational Psychology*, 102(1), 43–53.

意见 3：研究者称：“我们假设随着儿童年级的增长，执行功能在完成数学任务中的作用将逐渐降低”，但是有另一种可能：随着儿童年龄的增长，是否不同的执行功能子成分对数学能力的作用会发生改变。

回应：非常感谢审稿专家的意见。本研究的核心目的是从横断的视角初步揭示各执行功能子成分与不同数学能力之间的发展性关系。确如专家所述，不同执行功能子成分与数学能力的关系在不同年龄/年级儿童中表现模式存在差异。如相关研究发现，低阶的执行功能成分(干扰抑制、反应抑制、工作记忆广度)与数学能力的相关程度在低龄儿童中更大(e.g., Bryce et al., 2015; Stipek & Valentino, 2015)，而高阶执行功能成分(工作记忆刷新、认知灵活性)与数学能力的相关程度在大龄儿童中更强(e.g., Lee et al., 2018; Magalhães et al., 2020)。这或许是由于低阶执行功能成分发展较早，因此更早参与儿童的数学学习和加工；而高级执行功能成分的发展速度缓慢、成熟较迟，因而在大龄儿童数学能力的发展中起的作用更大。但仍需进一步的实证研究去论证这一假说的合理性。尽管已有研究从横向或纵向的视角考察了执行功能与数学能力关系的发展模式，但这些研究要么欠缺考虑执行功能子成分的可再分性(抑制控制分为干扰抑制和反应抑制、工作记忆分为工作记忆广度和工作记忆刷新)，要么缺少考虑数学能力的多维性(如类似本文的基于理论的数学能力类别划分)。这将导致难以明晰各执行功能子成分与不同数学能力关系的发展模式。因此，本研究对执行功能的各子成分进行了较为全面的考察，揭示了其在不同数学能力中的作用，并从横断的视角探讨了二者关系随年级/年龄而变化的趋势。基于专家的建议，我们对本文综述部分的逻辑重新进行了梳理与改写(详见 1.1 部分、1.2 部分)，并进一步完善了本研究的假设(1.3 部分)。具体修改之处请专家详见修改稿(1.1 引言)中用红色字体标注的部分。

#### 参考文献：

- Bryce, D., Whitebread, D., & Szűcs, D. (2015). The relationships among executive functions, metacognitive skills and educational achievement in 5 and 7 year-old children. *Metacognition and Learning*, 10(2), 181–198.
- Lee, K., Ng, S. F., & Bull, R. (2018). Learning and solving algebra word problems: The roles of relational skills, arithmetic, and executive functioning. *Developmental Psychology*, 54(9), 1758–1772.
- Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934–952.
- Stipek, D., & Valentino, R. A. (2015). Early childhood memory and attention as predictors of academic growth trajectories. *Journal of Educational Psychology*, 107(3), 771–788.

意见 4: 单个任务的年级差异在表格里呈现出来之后, 需要在文字中再赘述了。因为这些并不是研究关心的主要问题。

回应: 非常感谢审稿专家的意见。根据专家的建议, 我们已对文中相关内容进行了精简(3.1 部分)。具体修改为: 表 2 显示的是各年级儿童在数学运算、逻辑思维、空间想象等数学能力及中央执行功能各任务上的平均数与标准差。同时, 对不同年级被试在各测验上的成绩进行了差异检验。由表 2 可知, 四个年级的儿童在衡量三个数学能力的各测验上均存在显著差异( $F_s > 7; p_s < 0.001$ )。在中央执行功能方面, Stroop 干扰效应在不同年级间存在显著差异,  $F(3, 808) = 2.79, p < 0.05, \eta_p^2 = 0.01$ , 四年级儿童的成绩显著高于三年级儿童的成绩( $p < 0.05$ )。NOGO 正确率在年级上差异显著,  $F(3, 808) = 5.55, p < 0.01, \eta_p^2 = 0.02$ , 事后比较结果显示, 四、五和六年级儿童在 NOGO 正确率上的成绩显著高于三年级儿童。在衡量工作记忆刷新、工作记忆广度和认知灵活性的各任务上, 四个年级之间均存在显著差异( $F_s > 19; p_s < 0.001$ )。事后比较结果显示, 五年级和六年级儿童在各任务上的表现显著优于三年级和四年级儿童( $p_s < 0.05$ )。

意见 5: 4.2 第一段: 不同模型之间不能直接通过比较系数的大小来说明影响的大小, 需要进一步提供解释量或其他指标来进行效应之间的比较。

回应: 非常感谢审稿专家的意见。在修改稿中, 我们所建构的不同年级阶段的结构方程模型均为一致, 因而可以对高、低年级阶段的模型结果进行比较。但由于横断研究无法揭示其变量关系的连续变化模式, 我们也对讨论部分(4.2)的相关内容进行了修改, 具体修改内容请您详见 4.2 中用红色字体所标注的部分。

意见 6: “工作记忆刷新对儿童的空间想象能力和逻辑思维能力具有重要重要影响”, 此处多了一个“重要”。

回应: 非常感谢审稿专家的意见。我们已将文中重现的“重要”二字删除。

---

## 第二轮

### 审稿人 1 意见:

在修改之后, 文章的质量大幅度提高, 且逻辑更为顺畅, 之前不清楚的地方作者都做了进一步的澄清和更为详尽的阐述。作者关于审稿意见的回复十分详尽, 且在解决和回答审稿人的疑虑方面做的很好。仅有几个小问题需要作者注意。

意见 1: 关于审稿意见 14 的回复, 我指出的缺失值处理是希望作者提供一个统计上的缺失值分析, 比如 little's MCAR test (1988), 来说明缺失值是完全随机的, 可以删除。

回应: 非常感谢审稿专家对该问题的进一步说明。首先, 我们在回复第一轮审稿意见(意见 14)时已指出了本研究缺失数据的来源, 即这 13 名学生只有一个或至多两个观测变量的数据(如, 只有反应抑制和正背), 其余变量的相应观测值因学生未参加测试(请假错过测试时间)或研究助手遗漏拷贝数据等原因均未收集到相应数据。因此, 从数据缺失的内在机制来看, 本研究中的数据缺失类型符合完全随机缺失, 即目标变量的缺失与自身或其他观察变量的取值无关(Little & Rubin, 2002), 因而更宜采用删除法(Clarke & Hardy, 2007)。其次, 根据您的意见, 我们从最原始的数据库中核查并查找出了存在数据缺失的 13 个被试的部分数据, 将其补充在总数据库中并使用 SPSS 25.0 执行缺失值分析。独立方差  $t$  检验的结果显示, 几乎

所有变量的  $t$  值均小于 1.963(本样本中  $p = 0.05$  对应的  $t$  值为 1.963)。可见, 现有变量的数据缺失不会明显影响其他变量, 这说明本研究中变量的缺失是一种完全随机缺失。随后, 继续使用 little's MCAR test 进一步对缺失值进行分析。结果显示, 卡方值为 9.924, 自由度为 142, 显著性为 1.000, 即卡方检验的显著性明显大于 0.05, 因此我们接受缺失值为完全随机缺失的假设。综上, 本研究中这 13 个被试的数据缺失类型为完全随机缺失, 说明我们对这 13 名被试的数据进行删除是合理的。**最后**, 经与已有类似研究对比, 在**本文 2.1 部分**中阐述出这 13 名被试的数据缺失为完全随机缺失可能显得有些赘述。现先不将此内容在修改稿中呈现, 但若专家老师觉得极有必要呈现, 我们再进行补充。

**参考文献:**

Clarke, P., & Hardy, R. (2007). Methods for handling missing data. In A. Pickles, B. Maughan, & M. Wadsworth (Eds.), *Epidemiological methods in life course research*(Vol. 1, pp. 157–197). New York: Oxford University Press.

Little, R. J., & Rubin, D. B. (2002). *Statistical analysis with missing data* [M]. 2nded. New York: John Wiley & Sons, 1–10.

**意见 2:** 关于审稿意见 21 的回复和正文中“第一, 选取学龄中高年级的儿童将降低由于发育迟缓和缺乏数学指导所导致的数学成绩较差的风险。”这一点的表述存在问题, 低年级儿童就是发育迟缓吗? 应该是避免低年级儿童在数学测试上出现地板效应? 或者是这个测试不适合低年级的学生做?

**回应:** 非常感谢审稿专家严谨细致的思考及所提出的宝贵意见, 此处的表述的确存在歧义。我们所写的“降低由于发育迟缓”是一种错误的表述。我们最初想表达的意思为: 中高年级儿童由于已经接受了较多基础数学运算的训练, 数学运算能力获得较大发展, 同时儿童的数概念体系也逐渐建立, 空间观念和抽象思维获得了明显发展(刘范, 1981; 于萍, 左梦兰, 1996), 因而此年级阶段儿童的数学能力相比学龄低年级儿童较高也较为稳定。同时, 低年级儿童或许由于缺乏数学指导而存在数学成绩较差的风险(即数学测验分数出现地板效应), 而中高年级儿童则出现这种风险的概率相对较低。因此, 对中高年级儿童进行考察将更能准确评估各执行功能子成分在不同数学能力中的作用。综合专家的意见和建议, 我们将该句中的“发育迟缓”四字进行删除, 并进一步规范了相应表述。具体修改为: **第一, 选取学龄中高年级的儿童将降低由于缺乏数学指导而导致数学成绩较差的风险。**

**参考文献:**

刘范, 吕静, 沈家鲜, 赖昌贵, 张增杰, 刘静娴, 丁松年, 徐秀娣, 曹子方, 王鹏飞. (1981). 国内十个地区 7—12 岁儿童数学概念和运算能力发展的初步研究. *心理学报*, 2, 135–149.

于萍, 左梦兰. (1996). 三~六年级小学生数学能力及认知结构的发展. *心理发展与教育*, 3, 30–36.

**审稿人 2 意见:**

**意见 1:** 文稿有明显改善。

**回应:** 非常感谢审稿专家对我们修改工作的肯定。

**意见 2:** 引言第二段为什么不按照执行功能三个成分的排列顺序来展开相关研究, 而把抑制控制的研究放在工作记忆相关研究的前面?

**回应:** 非常感谢您严谨细致的思考及所提出的宝贵意见, 此处的写作逻辑确有不当之处。工作记忆、抑制控制和认知灵活性作为执行功能的三个核心成分, 在文章论述中的排列顺序并非固定。如多数研究报告中的排列顺序为抑制控制、工作记忆、认知灵活性(e.g., Ellefson et

al., 2020; Spiegel et al., 2021)。我们在引言第二段句首指出执行功能主要包括工作记忆、抑制控制和认知灵活性三个核心成分,但在该段后半部分论述已有研究结果的不一致时又以抑制控制开始,这种前后顺序的不一致使该处的写作逻辑显得不连贯。基于专家的意见,我们对引言第二段句首中执行功能三个成分的排列顺序进行了调整,交换了工作记忆和抑制控制的顺序。具体修改为:作为一种高阶的认知功能,执行功能是指对完成目标导向行为至关重要的一系列认知能力(Diamond, 2013),主要包括抑制控制、工作记忆、和认知灵活性三个核心成分。

#### 参考文献:

- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.
- Ellefson, M. R., Zachariou, A., Ng, F. F. Y., Wang, Q., & Hughes, C. (2020). Do executive functions mediate the link between socioeconomic status and numeracy skills? A cross-site comparison of Hong Kong and the United Kingdom. *Journal of Experimental Child Psychology*, 194, 104734.
- Spiegel, J. A., Goodrich, J. M., Morris, B. M., Osborne, C. M., & Lonigan, C. J. (2021). Relations between executive functions and academic outcomes in elementary school children: A meta-analysis. *Psychological bulletin*, 147(4), 329–351.

意见 3: 引言第二段中执行功能的三个成分的排列为“工作记忆、抑制控制和认知灵活性”,而 1.1 中第二行则换成了“(抑制控制、工作记忆和认知灵活性)”,请保持一致。

回应: 非常感谢您的细致审阅。结合专家的意见 2, 我们已将引言第二段中执行功能三个成分的排列顺序修改为: 抑制控制、工作记忆和认知灵活性。因此, 不再存在前后顺序不一致的问题。

意见 4: “一般可将其划分为干扰抑制和反应抑制(Diamond, 2013), 其分别侧重于抑制不恰当的行为与保持良好的注意力控制(赵鑫等, 2016; Diamond, 2013)。”这里的对应顺序也不对。

回应: 非常感谢审稿专家的宝贵建议, 此处前后的对应顺序确实不一致。根据专家的意见, 我们在修改稿中(引言部分第 3 段)对该句进行了修改, 具体修改为: 对于抑制控制而言, 一般可将其划分为反应抑制和干扰抑制(Diamond, 2013), 其分别侧重于抑制不恰当的行为与保持良好的注意力控制(赵鑫等, 2016; Diamond, 2013)。

#### 参考文献:

- 赵鑫, 贾丽娜, 咎香怡. (2016). 干扰控制的训练: 内容、效果与机制. *心理科学进展*, 24(6), 900–908.
- Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.

意见 5: “二者的可分离性代表了工作记忆的不同方面”这句话是什么意思?

回应: 非常感谢审稿专家所提出的宝贵意见。此句话的意思为: 工作记忆广度和工作记忆刷新代表了工作记忆的不同方面, 两者虽同属工作记忆这一能力, 但又彼此独立。这已获得已有研究的大量论证(e.g., Cirino et al., 2018; Himi et al., 2021; McKenna et al., 2017; Yang et al., 2019)。经专家指出, 此句话的写作确实略显晦涩。结合专家的意见, 我们将其修改为: 但二者又彼此独立, 代表着工作记忆的不同方面。

#### 参考文献:

- Cirino, P. T., Ahmed, Y., Miciak, J., Taylor, W. P., Gerst, E. H., & Barnes, M. A. (2018). A framework for executive function in the late elementary years. *Neuropsychology*, 32(2), 176–189.
- Himi, S. A., Bühner, M., & Hilbert, S. (2021). Advancing the understanding of the factor structure of executive functioning. *Journal of Intelligence*, 9(1), 16.
- McKenna, R., Rushe, T., & Woodcock, K. A. (2017). Informing the structure of executive function in children: a



meta-analysis of functional neuroimaging data. *Frontiers in Human Neuroscience*, 11, 154.

Yang, X., Chung, K. K. H., & McBride, C. (2019). Longitudinal contributions of executive functioning and visual-spatial skills to mathematics learning in young Chinese children. *Educational Psychology*, 39(5), 678–704.

**意见 6:** 1.1 的倒数第二段和最后一段之间缺少逻辑连接。

**回应:** 非常感谢审稿专家严谨细致的思考, 我们非常认可您的意见。根据您的建议, 我们在 **1.1 部分倒数第二段末尾**补充了与后一段连接逻辑的相应表述, 具体修改为: **基于此, 本研究继续沿用国内学者对学龄儿童数学能力的分类, 将小学儿童的数学能力分为: 数学运算能力、逻辑思维能力与空间想象能力, 以期探究执行功能各子成分在不同数学能力中的作用。**做出此修改的具体思路为: 1.1 部分第一段论述了对执行功能和儿童数学能力关系的考察应考虑抑制控制和工作记忆的不同方面(抑制控制可分离为反应抑制和干扰抑制; 工作记忆可划分为工作记忆广度和工作记忆刷新); 1.1 部分第二段论述了对执行功能和儿童数学能力关系的考察应考虑数学能力的多维性(如本研究沿用了国内学者对学龄儿童数学能力的分类, 将小学儿童的数学能力分为数学运算能力、逻辑思维能力和空间想象能力); 1.1 部分第三段综述了执行功能各子成分与不同数学能力关系的已有研究结果。因此, 在 1.1 部分倒数第二段末尾补充 **“以期探究执行功能各子成分在不同数学能力中的作用”**, 其将不仅起到对 1.1 部分中第一段和第二段内容进行有机融合, 同时也开启了后段阐述执行功能各子成分与不同数学能力关系的相应综述, 起到一个承上启下的作用。

**意见 7:** 认知灵活性对数学能力的影响似乎是突然出现的, 引言的前面两段都没提到相关的研究, 需要补充。

**回应:** 非常感谢审稿专家的意见和建议。**首先**, 需要陈述的是我们在引言第一段尾句中确有描述认知灵活性对数学能力影响的相关表述。在引言第一段的尾句中, 写道“在各种操作和解题策略间灵活转换”, 这则是认知灵活性在数学加工中起作用的集中体现。但确如专家所述, 我们在一审修改稿(**引言第二段**)中缺少阐释认知灵活性对数学能力的影响。结合专家的建议, 并考虑到使文章的行文逻辑更为精炼, 我们在修改稿中(**引言第二段后半部分**)补充了相关内容。具体修改为: **同时, 工作记忆和认知灵活性与儿童数学能力的关系也显现出研究结论的不一致。部分学者发现, 工作记忆(Lan et al., 2011; Monette et al., 2011)或认知灵活性(Magalhães et al., 2020)是执行功能成分中预测儿童数学能力的唯一因子。然而, 也有研究表明, 相比于其他执行功能成分, 工作记忆和认知灵活性与数学能力之间的联系较小(Cantin et al., 2016; Ellefson et al., 2020; Zhang et al., 2018)。**这说明, 执行功能与儿童数学能力间的实质联系仍需进一步探究。

#### **参考文献:**

- Cantin, R. H., Gnaedinger, E. K., Gallaway, K. C., Hesson-McInnis, M. S., & Hund, A. M. (2016). Executive functioning predicts reading, mathematics, and theory of mind during the elementary years. *Journal of Experimental Child Psychology*, 146, 66–78.
- Ellefson, M. R., Zachariou, A., Ng, F. F. Y., Wang, Q., & Hughes, C. (2020). Do executive functions mediate the link between socioeconomic status and numeracy skills? A cross-site comparison of Hong Kong and the United Kingdom. *Journal of Experimental Child Psychology*, 194, 104734.
- Lan, X., Legare, C. H., Ponitz, C. C., Li, S., & Morrison, F. J. (2011). Investigating the links between the subcomponents of executive function and academic achievement: A cross-cultural analysis of Chinese and American preschoolers. *Journal of Experimental Child Psychology*, 108(3), 677–692.
- Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics



achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934–952.

Monette, S., Bigras, M., & Guay, M. C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology*, 109(2), 158–173.

Zhang, J., Cheung, S. K., Wu, C., & Meng, Y. (2018). Cognitive and affective correlates of Chinese children's mathematical word problem solving. *Frontiers in Psychology*, 9, 2357.

**意见 8:** 正如作者在局限性指出, 横断研究无法明确地回答因果关系和影响机制, 因此, 请在文中谨慎使用诸如“预测”、“导致”、“影响”这样的字眼。

**回应:** 非常感谢审稿专家的意见和提示。确如专家所述, 本研究的横断性质使其无法明确因果关系和影响机制。但鉴于本研究在分析数据时使用了结构方程模型, 其作为多元数据分析的重要工具, 可以通过同时观测多个变量来分析他们之间潜在的因果关系(MacCallum & Austin, 2000)。且以往研究多以预测关系来表述基于结构方程模型所获得的结果(e.g., 刘馨元等, 2017; Georgiou et al., 2020; Hawes et al., 2019)。基于专家建议并参考已有研究中的有关论述, 我们进一步对文中的相关表述进行了规范, 在文章多处以“预测”替代“影响”、“导致”, 以期实现更为严谨表述本研究所得结论的目的。具体修改已在文中使用绿色字体进行了标注。

#### 参考文献:

刘馨元, 李铭心, 张志杰. (2017). 中央执行子功能对预期式时距估计的作用——基于结构方程模型的结果. *心理科学*, 40(6), 1316–1321.

Georgiou, G. K., Wei, W., Inoue, T., Das, J. P., & Deng, C. (2020). Cultural influences on the relation between executive functions and academic achievement. *Reading and Writing*, 33(4), 991–1013.

Hawes, Z., Moss, J., Caswell, B., Seo, J., & Ansari, D. (2019). Relations between numerical, spatial, and executive function skills and mathematics achievement: A latent-variable approach. *Cognitive Psychology*, 109, 68–90.

MacCallum, R. C., & Austin, J. T. (2000). Applications of structural equation modeling in psychological research. *Annual Review of Psychology*, 51, 201–226.

---

### 第三轮

#### 审稿人 2 意见:

文稿改善显著, 无意见。

#### 编委意见:

作者全面详细回应了两位外审意见, 文章做了相应大修, 实验操作正确, 样本量大, 结果显著, 文字通畅, 建议接收和发表此修改稿。

#### 主编意见:

本论文以三至六年级儿童为研究对象, 借助相关分析和结构方程模型等方法, 对各执行功能成分在不同年级儿童基本数学能力中的作用进行了考察。本论文研究思路清晰, 数据处理过程科学规范, 得到的研究结论较为可信。