《心理学报》审稿意见与作者回应

题目：工作记忆训练提升幼儿流体智力表现
作者：彭君 黄平 周莹 王靖 莫雷

第一轮

审稿人 1 意见：
意见 1：摘要和全文中的用语请尽量简洁、规范，不要使用空泛的词句；
回应：感谢老师的意见，已经对原摘要进行修改。（修改见 摘要）

意见 2：作者在前文没提到要对控制组 2 的流体智力进行后试，后面呈现数据时是测试了，是不是检查确认一下？
回应：之前表述可能模糊，已经在文中进行修改。（修改见 2.2 实验设计 红字部分）

意见 3：作者是不是可以对前、后测、延续效果等三个时间段的数据进行一下连续分析？从数据看，控制组 1 前测数据与延续效果的数据差值接近 5 分，而实验组同样的数据差值才 3 分多一点。不知作者是如何看待这两组之间在这两个时间点上的差异？
回应：
（1）前测、后测、延时测验所用的工具不同，原始分数并不具有可比性，直接用原始分数进行连续分析并不合适。
（2）为进行连续分析，将各个测验的分数进行标准化转换，Ti (标准分) = 50 + 10 Zi (Z 分数以所有组别被试于该测验的成绩计算所得)，继而进行 3 × 2 非重复测量方差分析。（修改见 3.2 流体智力测验结果）

意见 4：请作者将实验程序至少放在附录中；
回应：实验程序示意图已根据老师意见放入附录。（修改见 附录）

意见 5：n-back 任务本身比较枯燥，是如何吸引幼儿持续进行训练的？2-back 和 3-back 尤其 3-back 对一般的成人难度都挺大，对幼儿的难度就更大，其训练的正确率如何？幼儿真正掌握的情况如何？每次训练实际花费多长时间？在每天的什么时间训练的？
回应：在附录中的实验程序示意图可以看到：一方面，n-back 任务在实验人员的加工下较为美观，加入了平板游戏的元素；另一方面在训练中分数和等级有进步的幼儿将得到贴纸奖励。经过这样的处理，n-back 任务吸引了幼儿持续训练。2-back 和 3-back 在幼儿中确实有较大难度，从成绩上来看，幼儿在 14 天的训练中绝大部分可以达到 2-back，只有少数幼儿可以达到 3-back。程序使用的是 JAVA 编程，因此正确率难以搜集，但程序设置 2-back 有两个关卡（第 3、4 关），3-back 难度也有两个关卡（第 5、6 关），因此若幼儿凑巧达到 3-back，也很难再从第 5 关晋级到第 6 关，但观察幼儿训练的过程是可以从 3、4 关（2-back）通过，并且有少数人可以达到第 6 关。说明在适当的指导下，幼儿可以做到 3-back 任务。每天训练程序运行的时间为 15min，若加上老师组织幼儿、提醒幼儿注意游戏规则等时间大约每天为 20min。训练在每天的 9 点至 10 点进行，幼儿在这个时间段精神状态较好。（修改见 2.3.1 工作记忆训练程序）

意见 6：训练的实施要交代清楚，如：是在什么性质的幼儿园，是如何对幼儿园老师培训的，

如何保证幼儿园老师实施质量的？是否对幼儿园老师进行了物质奖励？
回应：实验在民办的市一级幼儿园中进行（修改见 2.1 被试），由实验员配合实验程序对老师进行游戏规则讲解，明确对幼儿园老师的实施要求，每次训练都有实验员在旁监督。实验项目组给予参与实验的幼儿园老师相应的劳务报酬。

意见 7：智力测验的施测过程和环境是怎样的？
回应：智力测验是单独施测，在幼儿园每个班的休息室中逐个进行测验，由一位实验员对幼儿进行测试，幼儿园教师配合实验员组织幼儿逐个进入休息室进行测试。瑞文标准推理测验使用的是打印的测验材料，被试需在答案下做记号，不限定时间，所有被试均能在 20min 内完成。TONI 使用的是类似台历的测试材料，被试需要用手指出答案，不限定时间，所有被试均能在 15min 内完成。休息室隔音效果较好，室温和光照都适宜进行测试。

意见 8：训练的环境是什么样的？幼儿是在集体环境中进行训练的吗？
回应：幼儿在该幼儿园的一个活动室中进行训练，活动室中配备桌椅，通风、光照、隔音效果良好，一次可以对 16 名幼儿进行训练，幼儿有固定座位，且每人一台平板机和一副耳机，幼儿园教师和实验员在旁监督，训练过程中幼儿秩序良好，没有出现相互打扰的情况。（修改见 2.3.1 工作记忆训练程序）

意见 9：F 和 t 检验的自由度没有写上，p 值小于 0.05 的也要把具体值写上；
回应：感谢您的建议，已进行修改。（修改见 3 结果）

意见 10：在统计时把效应值（partial Eta 在 SPSS 可以做出来，或者汇报 Cohen’s d）和统计检验效力也写上；
回应：感谢您的建议，已根据您的建议进行修改。（修改见 3 结果）

意见 11：该研究设计不是随机分组的，对数据的统计方法需要谨慎思考一下：作者对前测后测分别进行了统计，得出前测两组被试差异不显著，后测两组被试差异显著，但这并不能推论实验效应显著，可考虑使用协方差分析，将前测数据作为协变量进行协方差分析(只针对 SPM 的来做吗？)；或者对后测减去前测的增值分数进行显著性检验；
回应：前后测所用题目均来自 SPM，除题目呈现顺序外，测评过程完全相同，可以进行分数相减。修改稿中将后测与前测的分数差异作为因变量，使用单因素（分组）方差分析证明了训练的即时效果。（修改见 3.2 流体智力测验结果）

意见 12：工作记忆训练程序的结果应该放在前面，这个结果的描述也有些简略；
回应：根据您的建议，将训练程序的结果调整至前面，并细化了描述。（修改见 3.1 工作记忆训练成绩）
意见 13：态度和动机水平是怎么测的，测试题目没有交待；
回应：使用了教师测评、幼儿自评和实验员测评三份问卷。（修改见 2.3.4 态度及动机问卷 及附录）

意见 14：前言里面提到 3-6 岁是记忆力发展的关键时期，在讨论部分建议从儿童注意和记忆能力发展的角度进一步讨论；
回应：非常感谢您的建议。经过文献查阅和思考，发现：
（1）前言中没有提到“3~6 岁是记忆力发展的关键时期”，可能原因是作者的原文表述不清

晰，所以误导了您和其他读者，原句为“学术界公认促进智力发展的最佳期是婴幼儿时期，该时期是脑对环境和经验的可塑性和敏感性很强的时期，3~6 岁幼儿阶段是人生最重要的启蒙时期……”，现已修改为“……3~6 岁幼儿阶段是智力的启蒙时期……”。

（2）在讨论第一部分提到的注意是由于曾有人质疑过游戏训练对注意力有提升作用，进而导致被试在训练过后的测试中表现良好，因此本研究中控制 1 组也进行了游戏，是为了平衡游戏对被试注意力的影响，进而说明注意因素没有影响被试流体智力测试的表现，因此注意不是本研究关注的要点。

（3）我们思考过将注意和记忆展开讨论，但是在讨论中第一部分提到心理机能的发展规律，其中也包括了注意和记忆能力，因此若展开讨论，讨论的逻辑会变得不清晰。为了在讨论中也关注到注意和记忆力，作者对讨论进行了文字改动。（修改见 4.1 个体心理技能短期训练的效果问题）

意见 15：参考文献作者虽然用了 endnote 插入，但也存在个别错误，作者也需要仔细检查每一篇文献

回应：感谢您的意见，由于写作修改过程出现文献错乱，现已检查过每一篇参考文献的插入。

意见 16：英文摘要翻译痕迹明显，需要进一步加工润色，以更好地理解。

回应：感谢您的建议，已进行进一步的加工润色。（修改见 英文摘要）

审稿人 2 意见：

意见 1：作者在摘要中认为“工作记忆训练可以提高幼儿的流体智力成绩，并且这种训练效果具有持续的有稳定性”，需要辨析 的是“持续的稳定性”之含义，作者在进行即时流体智力测验后六个月再做了次流体智力测验，发现训练组成绩仍然高于控制组。但应注意到，6 个月后实验组和控制组成绩均有下降趋势（见表 3）。试问，假如一年后再测流体智力，结果又如何呢？ 仅仅做了一次半年的后测就能证明“持续的稳定性”吗？这是一个最重要的问题。请作者回答。

回应：您的建议提得很好，对我们很有帮助。

（1）虽然两个测验所测试的都是流体智力，但由于使用的测试不同，所以不能直接用原始分来进行比较，成绩的下降不能说明流体智力的退步，为了在两个不同测验的三次测验成绩进行横向比较，因此我们将三次测验的分数进行了标准化，Ti (标准分) = 50 + 10 Zi (Z 分数以所有组别被试于该测验的成绩计算所得)，继而进行 3 × 2 非重复测量方差分析，发现训练主效应显著，且在 6 个月测试中，实验组的优势保持。

（2）本研究中“持续的稳定性”可以解释为训练效果能持续并且稳定保持，前人有许多考察效果保持的研究，例如， 2011 年 Jaeggi 等人发表在 PNAS 上的文章对小学生的训练结束 3 个月后进行考察，结果发现 3 个月后训练组流体智力成绩依然保持优势，他们在研究基础上提出工作记忆训练有效且具有长期持续性。其次，Li 等人（2008）对老年人进行工作记忆训练的研究中，研究者在训练结束后 3 个月也做了追踪测试，结果发现老年人的训练效果在 3 个月后出现回落，而成人年的训练效果依然保持。该研究表明若训练效果不能保持，那么在 3 个月内会出现回落。另外，Holmes 等人(2009)对工作记忆缺陷儿童进行训练，在训练结束 6 个月测试发现迁移效果依然能保持，他们认为这类测验可以作为克服工作记忆缺陷的手段，等等。综合前人的研究，考察训练效果保持的间隔时间鲜少超过半年，因此在本研究中我们将 6 个月作为考察期，流体智力成绩优势依然保持，因此使用 “持续的稳定性”的说法。

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081–10086.

Li, S. C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychology and aging*, *23*(4), 731.

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.

意见 2：文章存在一些需要解释的地方：

（1）在 2.3.3 中，在流体智力测验中，“前测与后测题目相同”，虽然“相同难度题目顺序不同”，但是前测与后测的题目毕竟相同，练习效应（原文为“顺序效应”）并没有排除，如何解释这个问题？

回应：感谢您的提问，由于 3 个组的被试都接受同样的测试，可以平衡练习效应。

（2）在 2.4.2 中，作者说“训练结束后三个月使用 TONI-4 的 A 部分进行测试”，三个月后也进行了测试吗？

回应：感谢您的提问，这里是作者书写疏漏，在校对时没能及时发现，已经进行修改并全篇再次审校。

（3） 数据的统计分析中，既然设立了三个组，那么不进行方差分析而只做 t 检验的理由何在？我认为，应该先通过方差分析考察组间差异，只有实验组与控制组的组间差异主效应显著，这才有说服力。请作者解释。

回应：感谢您的意见，现已将三个组的成绩进行了标准化，然后进行方差分析，结果显示实验组与控制组的组建差异主效应显著。（修改见 3.2 流体智力测验成绩）

（4）本研究用平板电脑测试，不知平均年龄 4.9 岁的幼儿使用平板电脑测试是否合适？

回应：感谢您的提问，研究采用纸质版测试，用平板电脑进行训练。训练程序被制作成类似平板电脑游戏，便于幼儿理解和操作，程序的介绍添加在附录中，烦请您查阅。（修改见 附录）

第二轮

审稿人 1 对研究结果与统计部分提出的意见及作者回应

意见 1：表 2 是作者的主要结果，实验组经过前测、后测、延迟后测，标准 SPM 分数先上升（差值 7），然后略下降（差值 5）；控制 1 是一直下降（前后测差值约 6.6；前测与延迟后测差值约 9.6.这两组的对比的差异既可能是训练后质量提高，更可能是控制 1 不知原因的一直下降，导致两组的反向结果造成的。现有的数据很难证明作者最有价值的“工作记忆训练提升流体智力”的结论。

意见 2：控制组 2 前后测差值基本为零，且前测值为 47，与控制组 1 一开始就有几乎 6 分的差值。作者在讨论中提到两个控制组前后测没差异，只有实验组前后测有差异，但没看到这个结果。

意见 3：（针对表 2 提出）：控制组 1 三次智力测试标准分数的结果是直线下降，差值约 10，即使前后测差值也有 6，控制组 2 差值基本为零，而实验组是上升的，这就是说，实验组大于控制组 1，不一定是仅仅实验组提高的原因，同时还可能有控制组 1 下降的原因。作者如何进行解释？

意见 4：（针对“而在即时后测与延缓后测中，实验组得分均比控制 1 组显著更高”提出）：

这与上面是同样的问题，控制 1 在直线下降，实验组是先上升再略有下降，结果的解释无法完全归因于提高的结果。

意见 5：（针对“两个控制组前后测的成绩没有显著差异”提出）：不清楚作者如何理解这个表述。你在前面表 2 结果分析时采用的是前后测差值，控制组 1 和 2 之间差异是显著的，没有其他的专门针对每个组前后测差值的其他检验，这里说的不显著指什么？

对以上意见的回应：

感谢老师提出以上问题，这些问题是作者在第一稿与第二稿中用了不同的方式对研究结果进行整理分析并统计检验而造成。由于本研究的设计是：即时测试是三个组的比较，延缓测试是两个组的比较，两次测试所用的量表也不同，前面两稿都没有清晰地进行结果的分析比较。第二稿改变了一稿的分数形式（将原始分转换为标准分）进行统计，将表格中原始分数和转化后的标准分放在一个表格内，又没有表述清楚，显得混乱，导致了数据结果有不少疑问，审稿人提出这些疑问是非常中肯的。

本次修改稿，作者经过仔细考虑，调整了实验结果的呈现方式与分数的统计检验，力求分析条理清晰。本次修改，首先，对三个组（实验组与两个控制组）的即时测试结果进行统计与分析，第二步再对两个组（实验组与控制 1 组）的即时测试、延时测试的分数（转化为标准分）进行统计分析。

首先，对三个组（实验组与两个控制组）的即时测试分数进行统计与分析，结果见文中表 2：

表 2 两组被试训练后即时测试的流体智力测验平均成绩

	前测（SPM）	即时后测（SPM）
实验组	21.64 (4.88)	26.96 (3.14)
控制1组	23.04 (6.85)	20.66 (6.12)
控制2组	20.42 (5.34)	21.85 (3.74)

统计分析了三个组所有被试前测和即时后测的成绩，前测三个组成绩差异不显著（F（2，83）=1.704，p = 0.188），即时后测三个组成绩差异显著（F（2，83）=13.695，p < 0.001），实验组即时后测成绩显著高于两个控制组（p < 0.001），两个控制组即时后测成绩差异不显著（p = 0.388）；实验组前测和即时后测成绩之间差异显著（p < 0.001），控制 1 组前测和即时后测成绩差异不显著（p = 0.093），控制 2 组前测和即时后测成绩差异不显著（p = 0.185）。

第二步，对两个组（实验组与控制 1 组）的即时测试、延时测试的分数（转化为标准分）进行统计分析。结果见文中表 3：

表 3 两组被试三次流体智力测验成绩

	前测（SPM）	即时后测（SPM）	延缓后测（TONI-4）
实验组	48.64 (8.64)	55.94 (5.20)	55.34 (7.52)
控制1组	51.13 (11.02)	45.05 (10.00)	45.55 (9.71)

统计分析了两个组（实验组和控制 1 组）在前测、即时后测、延缓后测三次测验上的连续比较。由于延缓后测使用的是非语言智力测验，因此要和前两次瑞文智力测试相比较，必须先将被试的成绩转化为 T 分，转化方法为 Ti = 50 + 10 zi（值得注意的是，计算某次测验的 z 值时使用的平均分和标准差来自于该次测验两个组被试原始分的平均分和标准差）。结果显示两组前测成绩不显著（p = 0.365），实验组即时后测成绩显著高于控制 1 组（p < 0.001），实验组延缓后测成绩也显著高于控制 1 组（p < 0.001）。进一步，分别对实验组与控制 1 组训练前、训练后即时后测与延缓后测的成绩进行简单效应分析，结果显示，对于实验组，即时后测成绩显著高于前测成绩（p < 0.001），延缓后测成绩均显著高于前测标成绩（p < 0.001），而即时后测与延缓后测成绩差异不显著（p = 0.770）；对于控制 1 组，前测成绩显著优于即时后测成绩（p = 0.026），也显著优于延缓后测成绩（p = 0.041），而即时后测成绩分与延缓

后测差异不显著（$p = 0.852$），此处表现出外审老师提出的关于控制 1 组的成绩下降的现象，实际上这是由于将原始分转化为标准分而造成，因为，标准分是标示原始分的相对位置的分数，控制 1 组的前测和两次后测的 T 分出现差异（而原始分没有差异），是由于实验组成绩在即时后测中出现了上升，而控制 1 组的被试水平没有显著变化，这样，一旦转化为有相对位置意义的标准分，由于实验组后测的实际成绩显著提高了，而控制 1 组由于实际成绩没有进步就会造成 T 分比前测下降，T 分数这种下降并非由于控制组实际成绩的退步导致，而是由于实验组实际成绩的进步造成的。

在这里特别要说明的是，本文第一稿只是分别对实验组与控制组的即时后测的成绩（原始分）与延缓后测的成绩（原始分）进行比较，这样表述结果时虽然不容易造成误解，但是，不能更好地展示三次测试结果的连续变化，因此，根据审稿人第一次审稿意见，我们将前测、即时后测与延缓后测的结果统一进行比较，由于即时后测与延缓后测所用的量表不同（使用不同量表是为了提高研究效度，国外研究多有采用这种方法），为了将三次结果进行连续比较（连续比较的建议是外审老师在第一次修改意见中第 3 条提出），我们统一将三次测试的原始分数转为标准分，这样就会形成由于实验组训练后实际成绩的进步而造成控制组后测的标准分的下降的情况。

审稿人 1 针对文章其他部分提出的意见与作者回应

意见 1：（针对前言中“第二个是 Bergman 等人(2011)对 101 名 4~6 岁幼儿进行工作记忆训练”）：从服务于你的研究目的出发，你应该在这里着重叙述他使用的训练方法。你后面直接说他没使用刷新任务，显得没说服力。

回应：感谢老师的指出，根据再次查阅文献的结果，Bergman 等人（2011）的这篇文章与前面提到的 Thorell 等人（2009）使用的训练任务相同，是 Cogmed system 中的工作记忆任务和抑制任务，Cogmed system 中的工作记忆任务的核心规则是：按难度呈现出若干刺激，被试按显示顺序和要求（顺序或逆序）用鼠标在画面中点击刺激作答。该核心规则属于工作记忆广度任务，不属于工作记忆刷新任务(刘春雷 & 周仁来, 2012)。（修改见 1 引言 蓝色字体）

意见 2：（针对“学术界公认促进智力发展的最佳期是婴幼儿时期”）请给出文献支持。

回应：感谢老师的指出，根据思考与确认，此类说法出自于一般教材，放在此处确实不合适，现将此句删除。（修改见 1 引言）

意见 3：（针对图 1 实验组训练表现提出）：作者提到实验组“成绩”使用的是反应时，原因是程序记录不了正确率，作为记忆任务，正确率更能说明问题，这应该是常识。不知作者这里检验“组次”是什么指标？

回应：感谢老师提出疑问，这里的平均成绩不是反应时，而是实验组被试每天在游戏上达到的关卡的平均值，这里的“组次”是关卡的意思，即组次 1 是 3 个位置的 1-back，组次 2 是 4 个位置的 1-back 等等。被试只有在一轮的训练中达到正确数 12 个（总数 15 个）以上才能够在下一轮晋级一个关卡，晋级规则见 2.3.1 工作记忆训练程序。之前描述不清，现已纠正，感谢老师的指出。（修改见 图 1 实验组训练表现 及蓝色文字）

意见 4：（针对“流体智力提高量差异显著”）：这里有提高吗？根据你的数据，不是提高，而是下降，如何解释呢？

回应：感谢老师的提问，这里的提高量是用各个组前测和即时后测的 SPM 原始分相减得到的，每个组的每个被试都得到一个差值，实验组的差值显著大于控制 1 组和控制 2 组。控制

1 组成绩没有显著的下降。我们在之前的统计中使用了更明确的统计方法来说明控制 1 组和控制 2 组不存在差异，因此该部分数据统计略显多余并可能误导外审老师，经过思考我们将这部分删除。（修改见 3.2 流体智力测验结果）

第三轮

审稿人 1 意见：此研究对理解工作记忆及流体智力的发展有重要的理论意义，而且对幼儿基本认知能力的培养有现实意义。文章思路清晰，实验设计合理，实验结果可靠。以下几点还请思考。

意见 1：6 个月的延时测量是否可称为“长期保持”？

回应：感谢老师的提问。前人有许多考察效果保持的研究，例如， 2011 年 Jaeggi 等人发表在 PNAS 上的文章对小学生的训练结束 3 个月后进行考察，结果发现 3 个月后训练组流体智力成绩依然保持优势，他们在研究基础上提出工作记忆训练有效且具有长期持续性。其次，Li 等人（2008）对老年人进行工作记忆训练的研究中，研究者在训练结束后 3 个月也做了追踪测试，结果发现老年人的训练效果在 3 个月后出现回落，而成人年的训练效果依然保持。该研究表明若训练效果不能保持，那么在 3 个月内会出现回落。另外，von Bastian 和 Oberauer(2013)对成年人进行工作记忆训练的迁移效果研究，结果发现训练停止后 6 个月依然能探测到训练的迁移效果。Holmes 等人(2009)对工作记忆缺陷儿童进行训练，在训练结束 6 个月测试发现迁移效果依然能保持，他们认为这类测验可以作为克服工作记忆缺陷的手段，等等。综合前人的研究，考察训练效果保持的间隔时间鲜少超过半年，因此在本研究中我们将 6 个月作为考察期，流体智力成绩优势依然保持，因此使用 “长期保持”的说法。外审老师的提的建议非常宝贵，我们会根据这个建议做其他追踪研究。（修改见 4 讨论，绿色字体）

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081–10086.

Li, S. C., Schmiedek, F., Huxhold, O., Röcke, C., Smith, J., & Lindenberger, U. (2008). Working memory plasticity in old age: practice gain, transfer, and maintenance. *Psychology and aging*, *23*(4), 731.

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.

von Bastian, Claudia C, & Oberauer, Klaus. (2013). Distinct transfer effects of training different facets of working memory capacity. *Journal of Memory and Language*, *69*(1), 36-58.

意见 2： 结果部分 T 分数的处理可以增加组间及测量时间点之间的可比性，但也掩盖了原始分数的直接信息，是否可以考虑多加一个表或者在表中加一列，这样读者可以看到三个组的智力测验分数都是随时间增加的。

回应：感谢老师的建议，需向审稿老师说明两点原因：1、原始分数的直接信息在表 2 呈现，并且对原始成绩进行了统计，若把分数合并入表格，可能会造成结果描述的混乱；2、在 6 个月的追踪测试中，使用的测验（TONI-4）与前两次测验（SPM）的测验目的相同——均为流体智力测验，但计分原则不同，因此两种测试得出的原始分数并不能直接做对比，因此我们将 3 次分数转化为了 T 分，较方便进行比较。若将两种测验的原始分合并在一个表内，可能会造成读者的误解，认为两次测验间被试水平有差异。综上两个原因，我们经过认真反复地思考，认为分两个表格描述结果会比较合适。我们也在文字部分注重了对统计结果的说明，以便读者理解。

意见 3：控制 1 组的流体智力成绩并没有在训练后提高，是否可以说明注意力训练对流体智力的促进无效，最好有相关文献支持。

回应：感谢老师的提问。当前存在各种各样的注意力训练，其中有一些可以迁移至流体智力，但另一些并没有导致流体智力的提高，这是一个引人争议的问题。有研究(Ren, Altmeyer, Reiss, & Schweizer, 2013)指出目前注意力的研究可以主要分为两种：一种是知觉注意，集中在心理加工的知觉层面，例如警觉、聚焦和选择性注意，知觉注意和高水平认知能力之间的研究有各种各样的结果，例如和智力相关(Schweizer, Moosbrugger, & Goldhammer, 2005)；另一种是更高级的心理加工层面的执行注意，执行注意与工作记忆相关紧密，有研究者认为他们的功能是相似的，研究结果表明工作记忆能力的高低也可以影响认知控制、Stroop 任务和双耳分听任务的成绩(Engle, 2002)。知觉注意任务和执行注意任务涉及的加工过程有不同的偏重，Ren 等人的研究(2013)也提出，解码加工对于知觉注意任务十分重要，而刷新加工和转换加工则在执行注意任务中更为重要。虽然知觉注意和执行注意与流体智力都存在相关，但执行注意可以解释更多的流体智力的变异。当前，执行注意方面存在许多于智力相关的训练研究，例如 Rueda 等人(2012)使用 4 种执行注意任务训练学前幼儿，训练进行 5 周，结果发现训练组和非训练组的注意力没有显著提高，但反应时和错误率显示两组被试反应更熟练，研究者将其解释为测试任务本身的练习效应；该结果还显示训练组流体智力出现了提升，并且显著高于控制组，在停止训练两个月后依然保持提升效果，说明了流体智力可以通过干预提高，与前人的研究结果相符(Jaeggi, Buschkuehl, Jonides, & Perrig, 2008)。对于注意训练，其他分类方法也有类似的结果，例如 Tang(2009)将注意相关的认知训练归纳为注意训练和注意状态训练：注意状态是指类似于冥想和暴露于自然的训练；而注意训练是指冲突相关任务、工作记忆任务和其他与执行控制机制有关的任务的训练，此类训练可以提高注意力和流体智力，对幼儿以后的学业成绩也有积极的效果。

综合前人的研究可以发现，在注意训练中出现流体智力迁移的研究，其训练内容都涉及执行注意训练。同时，执行注意与工作记忆密切相关，难以分割，且刷新功能在执行注意任务中起到了重要作用。值得注意的是，前人研究显示刷新功能与流体智力紧密相关(Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008)，n-back 任务是针对刷新功能进行的训练，且有证据证明训练对提高流体智力有效(Jaeggi et al., 2008; Jaeggi, Buschkuehl, Jonides, & Shah, 2011)，因此，执行注意训练导致流体智力提升的原因，极可能是训练了工作记忆的刷新功能而产生的迁移。

本研究中，实验组使用 n-back 训练程序进行训练，我们对控制 1 组的任务选择有如下考虑：首先，有人发现通过工作记忆训练后，被试的注意力表现有所提升(Colom et al., 2013; Thorell, Lindqvist, Bergman Nutley, Bohlin, & Klingberg, 2009)；其次，动作电脑游戏也被证明可以提高视觉注意(Boot, Kramer, Simons, Fabiani, & Gratton, 2008; Green & Bavelier, 2003)，我们使用的 n-back 工作记忆训练程序被制作成了类似电脑游戏，被试进行 n-back 程序任务时，与进行动作电脑游戏有类似的要求，例如眼手协调操作、对全屏幕进行关注、短时间内作出快速的决定等(Green & Bavelier, 2008)。最后，Green 等人(2003)的研究表明，即使之前没有电脑游戏的经验，在 10h 的动作电脑游戏训练以后，他们的视觉注意力表现也提高了。因此，综合以上原因，我们为避免因 n-back 训练程序产生的注意力提高影响实验结果，对控制 1 组采用了“水果忍者”游戏以平衡可能产生的知觉注意力提升。实验结果发现，控制 1 组没有出现流体智力的提升，该实验结果虽然不能直接证明注意训练不能提高流体智力，但与前人的研究结果相符——没有足够的证据证明知觉注意训练显著提高流体智力，涉及执行注意的训练才有可能出现流体智力提升，并且提升的可能原因是任务中包含了刷新功能的训练。感谢老师提出的具有研究意义的问题，我们已在文中进行相应内容的添加和解释（见 4 讨论，绿色字体）。

Boot, W. R., Kramer, A. F., Simons, D. J., Fabiani, M., & Gratton, G. (2008). The effects of video game playing on attention, memory, and executive control. *Acta psychologica*, *129*(3), 387–398.

Colom, R., Abad, F. J., Quiroga, M. Á., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why?. *Intelligence*, *36*(6), 584–606.

Colom, R., Román, F. J., Abad, F. J., Shih, P. C., Privado, J., Froufe, M., et al. (2013). Adaptive n-back training does not improve fluid intelligence at the construct level: Gains on individual tests suggest that training may enhance visuospatial processing. *Intelligence*, *41*(5), 712–727.

Engle, R. W. (2002). Working memory capacity as executive attention. *Current directions in psychological science*, *11*(1), 19–23.

Green, C. S., & Bavelier, Daphne. (2003). Action video game modifies visual selective attention. *Nature*, *423*(6939), 534–537.

Green, C. S., & Bavelier, D. (2008). Exercising your brain: a review of human brain plasticity and training-induced learning. *Psychology and aging*, *23*(4), 692.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Perrig, W. J. (2008). Improving fluid intelligence with training on working memory. *Proceedings of the National Academy of Sciences*, *105*(19), 6829–6833.

Jaeggi, S. M., Buschkuehl, M., Jonides, J., & Shah, P. (2011). Short-and long-term benefits of cognitive training. *Proceedings of the National Academy of Sciences*, *108*(25), 10081–10086.

Ren, X., Altmeyer, M., Reiss, S., & Schweizer, K. (2013). Process-based account for the effects of perceptual attention and executive attention on fluid intelligence: An integrative approach. *Acta psychologica*, *142*(2), 195–202.

Rueda, M. R., Checa, P., & Cómbita, L. M. (2012). Enhanced efficiency of the executive attention network after training in preschool children: immediate changes and effects after two months. *Developmental Cognitive Neuroscience*, 2, S192–S204.

Schweizer, K., Moosbrugger, H., & Goldhammer, F. (2005). The structure of the relationship between attention and intelligence. *Intelligence*, *33*(6), 589–611.

Tang, Y. Y., & Posner, M. I. (2009). Attention training and attention state training. Trends in cognitive sciences, 13(5), 222–227.

Thorell, L. B., Lindqvist, S., Bergman Nutley, S., Bohlin, G., & Klingberg, T. (2009). Training and transfer effects of executive functions in preschool children. *Developmental science*, *12*(1), 106–113.

意见 4：讨论部分请再加强，尤其要结合前人文献。

回应：感谢老师宝贵的建议，我们已对前人研究进行适当添加和介绍，增强了文章的逻辑完整性（修改见 1 引言 和 4 讨论，绿色字体部分）。

意见 5：此实验是否经过伦理委员会审核，家长是否知情并同意,请提供相应信息。

回应：感谢老师宝贵的意见，实验前家长和老师都知情并同意，我们已在文中进行说明（修改见 2.1 被试，绿色字体）

意见 6：训练组中有四名儿童因“不明白规则”而退出，考虑到被试人数，这个比例不低。请作者对此现象进行解释，而且此研究结果的推广也会因此受限。

回应：感谢老师的宝贵意见。我们经过对该现象的慎重思考，结合幼儿园老师以及当时的实验员对程序的意见，得出幼儿不明白规则的可能原因是：程序中没有设置适合幼儿理解动画说明。我们就老师提出的问题询问了该幼儿园的老师和当时的实验员对程序的看法。他们认为程序的规则对于幼儿并不难理解，因为在老师的简单讲解之后，绝大部分幼儿正确地理解了程序的规则。他们提出如果在程序之前能够设置配合以更生动的说明动画，会让幼儿更好理解程序的规则。我们针对程序的说明动画，对国外认知训练程序进行了调查和体验，例如适用于幼儿的工作记忆训练 Cogmed 系统，它在进入正式训练之前出现了制作精良的说明动画，引导幼儿理解程序规则。我们也就动画问题询问了制作程序的人员，他们表示动画的制

作可以实现，并且可以令幼儿更好地理解。因此，我们认为幼儿“不明白规则”而退出的原因是在程序之前缺乏生动的引导说明动画。我们会汲取当前的宝贵经验，更完善未来的研究，同时也为研究结果的推广提出有用的建议。