

## 《心理学报》审稿意见与作者回应

题目：工作记忆刷新训练改善抑郁倾向大学生情绪调节能力的 HRV 证据

作者：彭婉晴；罗伟；周仁来

---

### 第一轮

**审稿人 1 意见：**该研究旨在研究是否可以通过工作记忆训练，改善抑郁倾向的大学生的情绪调节能力。研究采用抑郁量表区分抑郁和正常被试，采用情绪调节量表的得分，以及心率变异性作为指标。结果发现，通过 20 天的工作记忆训练该研究，情绪评分和调节量表得分，表现为不同组别的被试在认知重评上有显著差别。就 HRV 的结果，训练组被试在后测时与健康对照组的差别变小，与抑郁对照组的差别变大。这一研究在改善抑郁倾向的被试的情绪调节能力方面做了有益的探索，具有一定的创新性，结果具有实践意义。但该文存在以下主要问题：

**意见 1：**工作记忆的刷新能力与改善抑郁被试的情绪调节能力之间的关系是本研究的中心问题，因此在前言中应重点论述这一方面。而就该研究的结果而言，这两者的关系未得到显著印证。实验结果表明，无论是行为评分还是心率变异性，均未发现组别和前后测之间的交互作用，这提示通过训练并未引起显著的情绪调节的改善和心率变异性的变化。从图 6 和图 8 来看，图中的 error bars 均很大，这可能和被试量，以及测验指标的敏感性等有关。

**回应：**首先，我们依照审稿人的建议在引言增加了刷新功能和情绪调节能力之间关系的文献支持。其次，对于本实验的结果，虽然在主观量表得分上没有显著变化，但心率变异性的结果对两者之间的关系提供了支持：抑郁倾向训练组的 HF-HRV 在后测有显著的提高，并与抑郁倾向对照组产生差异。关于心率变异性的结果，我们补充了重复测量方差分析的结果，时间 $\times$ 组别的交互作用是显著的 ( $p=0.017$ )，简单效应检验显示，与两个对照组相比，抑郁倾向训练组在后测 HF-HRV 有显著的提高，并与抑郁倾向对照组拉开差距。表明工作记忆刷新训练提高了抑郁倾向个体的副交感神经活性，进而改善了他们的情绪调节能力。图 6 和图 8 里的指标是 HF-HRV 和 LF/HF-HRV 的净改变量，是采用了做差法，为考察被试情绪的纯净变化量，消除个体本身在观看视频时 HRV 基线水平的高低本身的影响，由于其本身的数值较小，个体间差异较大所导致。它们是实验的辅助性指标。

**意见 2：**由于抑郁对照组没有任何训练，因而它和训练组之间的差别可能是由于被试的主观动机等造成。这也反映在 HRV 的数据分析中。如需排除这一因素，可加入健康训练组或抑郁其它训练组。

**回应：**本研究是一个典型的双对照组的前后测三组设计，包括实验组（抑郁倾向训练组）、对照组 1（抑郁倾向对照组）和对照组 2（健康对照组）。审稿人提出，对照组没有进行其他的任何训练，因此无法排除训练组的提高是由于除工作记忆训练外的其他因素造成的，例如安慰剂效应。有的研究为解决这一问题会设置积极对照组(active control group)，即让对照组也进行某种训练。但其实关于这种前后测设计是否设置 active control group 一直也是存在争议的。第一，就工作记忆刷新训练而言，若让 active control group 做另外一种类型的训练，如数学训练或者简单的按键反应 (Gray et al., 2012)，客观上等于比较了这种训练与工作记忆刷新训练的效果差异，这会导致偏离研究的目标；

第二, 若让 active control group 做一些与比所训练的工作记忆内容减少的低版本训练, 如非适应性的工作记忆训练任务 (Hotton, Derakshan, & Fox, 2018; Holmes, Gathercole, & Dunning, 2009), 会由于训练内容的难度水平较低或相对缺乏吸引力会导致被试对持续较长时间的训练失去兴趣, 不仅未能达到控制的效果, 还会产生因无聊或动机减弱导致的反向效果, 这种反向效果客观上造成与实验组的差异, 并非是训练带来的, 恰是这种 active control group 带来的。并且像这种非适应性的工作记忆任务即使再简单也并非对被试的工作记忆毫无影响, 正如 Hotton 等人的研究发现的, active control group 和实验组均在工作记忆的某些方面发现提高, 但仍无法确定是由于安慰剂效应导致, 还是训练本身所导致的。因此, 不可排除简单的任务也会对个体的工作记忆以及中央执行功能产生影响。

第三, 其他一些发表的工作记忆训练的文章大多也都采用的是空白对照组 (Xiu, Zhou, & Jiang, 2016; Zhao, Zhou, & Fu, 2013; Gropper, Gotlieb, Kronitz, & Tannock, 2014)。在我们的这个设计中, 对抑郁倾向训练组加以训练, 采用另一组抑郁倾向的被试作为对照组, 同时也用另一组健康的被试作为对照, 也就是说有两个对照组。该实验目的是看工作记忆刷新训练是否能改善被试的情绪调节能力, 采用的是心率变异性 (HRV) 作为因变量的指标, 该指标已经有来自我们的和其他研究者的多项证据可以作为反映情绪调节的一个指标 (Xiu, Zhou, & Jiang, 2016)。这种设计的目的一方面是要看经过工作记忆刷新训练的抑郁倾向训练组的情绪调节水平是否好于未经训练的抑郁倾向对照组, 同时也要看经过训练的抑郁倾向训练组的情绪调节水平的改善是否达到了非抑郁倾向对照组 (健康组) 的水平, 等于从两个角度来检验训练的效果。应该说这种设计还是比较严格的。该研究设计借鉴了 Zhang 等人发表在《Frontiers in Human Neuroscience》杂志上的类似实验设计 (Zhang, Chang, Chen, Ma, & Zhou, 2018)。

此外, 也一直有研究者对动机、期待等因素可能在工作记忆训练中产生“安慰剂效应”给予关注, 本课题组研究生张心在其学位论文中专门针对“安慰剂效应”的问题进行了连续实验检验, 先后操纵了公开实验目的的招聘方式、个体智力信念测量的方式, 采用与本实验相同的训练方式进行了研究, 均否定了采用本实验的工作记忆训练方式存在“安慰剂效应”的可能 (见南京大学心理系 2018 届张心心理学硕士学位论文“工作记忆训练中的安慰剂效应”)。

最后, 在实验的过程中, 我们也采取了许多方式来控制其他影响因素, 一是从招募被试到实验过程, 都没有向被试透露真正的实验目的, 实验结束时被试的反馈也是没有猜到实验的目的。因为使用的是工作记忆训练, 如果有安慰剂效应, 被试的期望也会是在智力等方面获得提升, 而不会期待在情绪调节能力上得到提高。第二, 由于我们考察的是 HRV 这个生理指标, 它主要受被试的生理状况 (如年龄, 是否剧烈运动等) 以及任务本身的影响较大。在本研究中, 被试的 HRV 都是在平静的状态下记录的, 实验任务相同, 都考察了被试在五个状态下 HF-HRV 的变化。可以看到被试在不同状态下的 HF-HRV 会有波动, 但发现抑郁倾向训练组相比抑郁倾向对照组都有明显的提高, 说明训练组 HF-HRV 的提高是较为稳定的。

#### 参考文献:

- Gray, S. A., Chaban, P., Martinussen, R., Goldberg, R., Gotlieb, H., Kronitz, R., et al. (2012). Effects of a computerized working memory training program on working memory, attention, and academics in adolescents with severe LD and comorbid ADHD: a randomized controlled trial. *Journal of Child Psychology and Psychiatry*, 53, 1277–1284.
- Gropper, R. J., Gotlieb, H., Kronitz, R., & Tannock, R. (2014). Working memory training in college students with ADHD or LD. *Journal of Attention Disorders*, 18, 331–345.
- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor

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

Hotton, M., Derakshan, N., & Fox, E. (2018). A randomised controlled trial investigating the benefits of adaptive working memory training for working memory capacity and attentional control in high worriers. *Behaviour Research and Therapy*, 100, 67-77.

Xiu, L., Zhou, R., & Jiang, Y. (2016). Working memory training improves emotion regulation ability: evidence from hrv. *Physiology & Behavior*, 155(6), 25–29.

Zhang, H., Chang, L., Chen, X., Ma, L., & Zhou, R. (2018). Working memory updating training improves mathematics performance in middle school students with learning difficulties. *Frontiers in Human Neuroscience*, 12, 154.

Zhao, X., Zhou, R., & Fu, L. (2013). Working memory updating function training influenced brain activity. *Plos One*, 8(8), e71063.

意见 3: HRV 是自主神经系统平衡性的指标,但这一部分是否可快速地反映情绪调节,尚需更多的实验支持。建议作者采用皮肤电等更为普遍应用的指标,将 HRV 作为探索性指标。

回应:关于 HRV 是否能较好地反映个体的情绪调节能力的方面上,本实验主要使用了 HF-HRV 和 LF/HF-HRV 这两个指标,其中前者反映的是副交感神经,也就是迷走神经的活性,后者反映的是自主神经系统的平衡性。与情绪调节能力更为相关的是 HF-HRV,也是本研究中主要采用的指标。根据已有的研究,静息状态下的 HF-HRV 越高的个体情绪调节能力越好,自我报告的情绪调节困难程度越低 (Williams et al., 2015);另外,当个体主动调节自己的情绪时,也会伴随着 HF-HRV 的升高 (Libby et al., 2012)。HF-HRV 和情绪调节能力的关系还有神经的基础,研究表明 HF-HRV 和情绪唤起和情绪调节的脑区血流量的变化有显著相关 (Lane et al., 2009)。因此,近些年来,有不少研究将 HF-HRV 看作是反映研究情绪调节能力的重要参考指标,由来自本课题组另一位博士生采用本实验一样的工作记忆训练方法的研究,证实了工作记忆训练能够明显提高 HRV 的水平,进而反映了情绪调节能力的提高,该研究发表在 2015 年的《Physiology and Behavior》杂志上,本研究是这项研究的扩展,再次验证了该研究的结论。对于皮肤电等指标在其他的研究中使用的比较多,也能一定程度上反映情绪调节能力的变化,但在实验过程中我们并未采集皮肤电等指标,因此很遗憾无法提供这方面的数据。

#### 参考文献:

Lane, R. D., McRae, K., Reiman, E.M., Chen, K., Ahern, G. L., & Thayer, J. F. (2009). Neural correlates of heart rate variability during emotion. *NeuroImage*, 44, 213-222.

Libby, D. J., Worhunsky, P. D., Pilver, C. E., & Brewer, J. A. (2012). Meditation-induced changes in high-frequency heart rate variability predict smoking outcomes. *Frontiers in Human Neuroscience*, 6(12), 54.

Williams, D. P., Cash, C., Rankin, C., Bernardi, A., Koenig, J., & Thayer, J. F. (2015). Resting heart rate variability predicts self-reported difficulties in emotion regulation: a focus on different facets of emotion regulation. *Frontiers in Psychology*, 6, 261.

意见 4: 被试通过工作记忆训练,是否其刷新能力得到提高?这一部分的成绩变化需报告。提高的结果也可和行为评分和生理指标作相关分析,这可以部分回答上述第一个问题。

回应:根据审稿人的要求,我们把这部分数据专门进行了整理,三组被试的刷新能力成绩如下所示:

		抑郁倾向训练组	抑郁倾向对照组	健康对照组
2-back 正确率 (%)	前测	77.82±15.95	78.69±15.70	74.85±14.90
	后测	84.13±9.06	89.56±7.90	84.70±9.38
	提高	6.30±14.13	10.86±14.82	9.85±14.37
2-back 反应时 (ms)	前测	1036.29±204.40	1049.52±182.21	1121.83±206.04
	后测	980.61±207.58	935.94±182.18	978.98±169.21
	提高	-55.68±221.96	-113.58±194.04	-142.84±138.54
3-back 正确率 (%)	前测	62.04±11.82	66.59±14.47	64.70±14.06
	后测	75.90±12.94	75.45±13.46	72.70±12.63
	提高	13.86±16.25	8.86±17.60	8.00±15.18
3-back 反应时 (ms)	前测	1095.30±247.68	1086.36±189.90	1084.86±199.58
	后测	1021.73±241.30	984.06±201.24	1015.10±163.16
	提高	-73.57±224.06	-102.30±204.90	-69.76±173.54

可以看到，三组在 2-back 和 3-back 任务的后测中都有明显的进步，表现在正确率的提高和反应时的缩短上面。对三组的提高量进行方差分析发现，组别的主效应均不显著。可能的原因如下：对于成年人，我们在编制工作记忆测验时，都经过反复的预实验，将难度水平设定在中等稍偏上一点（也可见本课题组发表在 2017 年 PLOS ONE 上的文章“Working memory test battery for young adults Computerized working memory assessment”）的程度，对于多数被试，稍加训练（两次测验本身就是练习），都会提高，但提高的幅度有限，差不多接近天花板效应。根据 Zhao(2013)和 Xiu(2016)等人的研究可以发现，大学生群体在 2-back 的正确率都是较高，反应时较短的 (Zhao, Zhou, & Fu, 2013; Xiu, Zhou, & Jiang, 2016)，因此本实验中三组的提高量没有组别上的差异也是正常的。

没有在文章中报告刷新能力的成绩是由于本研究中工作记忆训练的重点不在于工作记忆本身提高多少，主要在于引起功能的改变，即 HRV 的变化。本研究使用的工作记忆刷新训练已有很多文章证明过是有效的，如 Zhao(2013)等人的研究表明，该工作记忆刷新训练能引起个体明显的脑电活动如 P2、P3 等的改变；Xiu(2016)等人的研究表明此训练能引起正常大学生 HF-HRV 的改变。因此，本研究主要想证明的也是工作记忆刷新训练对抑郁倾向大学生情绪调节能力的影响，主要表现在 HF-HRV 这个指标上。

#### 参考文献：

- Levens, S. M., & Gotlib, I. H. (2010). Updating positive and negative stimuli in working memory in depression. *Journal of Experimental Psychology General*, 139(4), 654.
- Xiu, L., Zhou, R., & Jiang, Y. (2016). Working memory training improves emotion regulation ability: evidence from hrv. *Physiology & Behavior*, 155(6), 25–29.
- Zhao, X., Zhou, R., & Fu, L. (2013). Working memory updating function training influenced brain activity. *Plos One*, 8(8), e71063.
- Ma, L., Chang, L., Chen, X., & Zhou, R. (2017) Working memory test battery for young adults: Computerized working memory assessment. *Plos One*, 12(3): e0175047. <https://doi.org/10.1371/journal.pone.0175047>

意见 5: 图 5 和图 7 需加入 error bars。

回应: 已按照要求改进文章图表的规范性问题, 加入 error bars。

意见 6: 建议正文中少用或不用黑体标注文字。

回应: 已按要求改正。

.....

审稿人 2 意见: 该文具有一定的理论意义和应用价值。审稿人认为, 下列问题需要修改:

意见 1: 建议对工作记忆刷新训练的后效, 即长期效应进行讨论, 尽管该问题不属于该研究内容, 但有助于提升本研究的应用价值。

回应: 已在讨论中加入对工作记忆刷新训练的长期效应的讨论。

意见 2: 对全文的标点符号、乘号 (\*) 等进行修改。

回应: 已对使用不当的标点符号进行改进。

意见 3: 细读全文, 对表述不清的内容精心修改。如“要求被试依次记住最后出现的三个字母, 即一直保持记住最近出现的三个字母。如出现的字母依次为 S-D-F-G-H-J-K, 被试要依次记住 S, S-D, S-D-F, D-F-G, F-G-H, G-H-J, H-J-K。最后, 屏幕中央会出现一个方框, 被试按照顺序依次输入最后出现的三个字母”, 怎样理解?

回应: 已对表述不清的内容进行修改, 如以上的内容改为: 首先, 在屏幕中央会出现一个“+”, 提示任务马上开始, 接着屏幕正中央会依次、逐个、随机地出现字母, 且出现的字母个数并不是固定的, 一共有 5、7、9、11 四种长度, 每种长度随机出现。要求被试依次记住最后出现的三个字母, 即一直保持记住最近出现的三个字母。如出现的字母依次为 S-D-F-G-H, 被试需要回答 F-G-H; 若依次出现的字母为 S-D-F-G-H-J-K, 被试则要回答 H-J-K。

意见 4: 修改图标, 注意文字布局, 做到文图自清。

回应: 已对文中的图表进行修改。

意见 5: 控制心脏的交感神经和迷走神经, 建议直接使用“心交感神经”, “心迷走神经”。

回应: 已对此进行修改。

意见 6: 其他问题, 见审稿文本。

回应: 已一一按照审稿文档进行修改。

---

## 第二轮

审稿人 1 意见:

意见 1: 建议对图中的误差线的长短、图注字体字号等做进一步修改, 以保持一致。

回应: 已对文中的图进行了改进, 图的标题统一改为五号字体, 图的注解为 10 号字体。

审稿人 2 意见：作者虽然回复了前两位评审人的意见，但文章还存在一些重要的问题：

意见 1：关于抑郁组分组的问题。“以自愿参加训练的方式将有抑郁倾向被试分为工作记忆刷新训练组和对照组”，自愿而不是随机分配，这会带来比较大的问题。可能存在一个动机效应，即参加训练组的被试，都是有较强参与和改变动机的，这可能会影响本实验的结果。此外，文中的表述“随机分成抑郁倾向训练组和抑郁倾向对照组”，与自愿分组的安排不一致。回应：关于被试自愿分组的问题，首先，我们在筛选被试时并未告知被试具体的实验内容，所筛选的抑郁倾向被试都是达到事先确定的量表筛选标准的大学生，也就是说他们都达到了实验所需要的标准。由于工作记忆的训练时间较长，分组时原则上都是随机分配的，但有些同学因课程安排等原因，在被分配到训练组后就明显表示无法保证全程参加，但可以参加时间较短的控制组，因此，把部分愿意参加训练组的控制组被试调到了训练组。从实际发生的情况看，我们不能说这是严格意义上的随机分配。因此可能存在动机效应的干扰，这个问题在文中也进行了讨论。表达不一致的地方也进行了修改。

原稿：最终筛选出 40 名抑郁倾向被试（随机分成抑郁倾向训练组和抑郁倾向对照组）和 20 名健康被试。

修改稿：最终筛选出 40 名抑郁倾向被试（自愿分成抑郁倾向训练组和抑郁倾向对照组）和 20 名健康被试。

意见 2：对照组没有任何训练，这也是一个问题。

回应：关于工作记忆训练的对照组是否也要进行训练，或者进行何种训练一直也是存在争议的。首先，就工作记忆刷新训练而言，若让对照组做另外一种类型的训练，如数学训练或者简单的按键反应（Gray et al., 2012），客观上等于比较了这种训练与工作记忆刷新训练的效果差异，这会导致偏离研究的目标；其次，若让对照组做一些与原本所训练的工作记忆内容减少的低版本训练，如非适应性的工作记忆训练任务（Hotton, Derakshan, & Fox, 2018; Holmes, Gathercole, & Dunning, 2009），会由于训练内容的难度水平较低或相对缺乏吸引力而导致被试对持续较长时间的训练失去兴趣，不仅未能达到控制的效果，还会产生因无聊或动机减弱导致的反向效果，这种反向效果客观上造成与实验组的差异，并非是工作记忆训练带来的，恰是这种对照训练带来的。并且像这种非适应性的工作记忆任务即使再简单也并非对被试的工作记忆毫无影响，正如 Hotton 等人的研究所发现的，积极对照组和实验组均在工作记忆的某些方面得到提高，但仍无法确定是由于安慰剂效应导致，还是训练本身所导致的。因此，不可排除简单的任务也会对个体的工作记忆以及中央执行功能产生影响。第三，其他一些已发表的工作记忆训练的文章大多也都采用的是空白对照组（Xiu, Zhou, & Jiang, 2016; Zhao, Zhou, & Fu, 2013; Gropper, Gotlieb, Kronitz, & Tannock, 2014）。第四，在我们的这个设计中，对抑郁倾向训练组加以训练，采用另一组抑郁倾向的被试作为对照组，同时也用另一组健康的被试作为对照，也就是说有两个对照组。该实验目的是看工作记忆刷新训练是否能改善被试的情绪调节能力，采用的是心率变异性（HRV）作为因变量的指标，该指标已经有来自我们的和其他研究者的多项证据可以作为反映情绪调节的一个指标（Xiu, Zhou, & Jiang, 2016）。这种设计的目的一方面是要看经过工作记忆刷新训练的抑郁倾向训练组的情绪调节水平是否好于未经训练的抑郁倾向对照组，同时也要看经过训练的抑郁倾向训练组的情绪调节水平的改善是否达到了非抑郁倾向对照组（健康组）的水平，等于从两个角度来检验训练的效果。应该说这种设计还是相对比较严格的。此外，该研究设计也借鉴了 Zhang 等人发表在《Frontiers in Human Neuroscience》杂志上的类似实验设计（Zhang, Chang, Chen, Ma, & Zhou, 2018）。

## 参考文献：

- Gray, S. A., Chaban, P., Martinussen, R., Goldberg, R., Gotlieb, H., Kronitz, R., et al. (2012). Effects of a computerized working memory training program on working memory, attention, and academics in adolescents with severe LD and comorbid ADHD: a randomized controlled trial. *Journal of Child Psychology and Psychiatry*, 53, 1277–1284.
- Gropper, R. J., Gotlieb, H., Kronitz, R., & Tannock, R. (2014). Working memory training in college students with ADHD or LD. *Journal of Attention Disorders*, 18, 331–345.
- Holmes, J., Gathercole, S. E., & Dunning, D. L. (2009). Adaptive training leads to sustained enhancement of poor working memory in children. *Developmental Science*, 12, F9–F15.
- Hotton, M., Derakshan, N., & Fox, E. (2018). A randomised controlled trial investigating the benefits of adaptive working memory training for working memory capacity and attentional control in high worriers. *Behaviour Research and Therapy*, 100, 67–77.
- Xiu, L., Zhou, R., & Jiang, Y. (2016). Working memory training improves emotion regulation ability: evidence from hrv. *Physiology & Behavior*, 155(6), 25–29.
- Zhang, H., Chang, L., Chen, X., Ma, L., & Zhou, R. (2018). Working memory updating training improves mathematics performance in middle school students with learning difficulties. *Frontiers in Human Neuroscience*, 12, 154.
- Zhao, X., Zhou, R., & Fu, L. (2013). Working memory updating function training influenced brain activity. *Plos One*, 8(8), e71063.

意见 3：作者回复了评审人 1 的意见 4，建议将相关结果在文章中予以报告和讨论。

回应：关于工作记忆刷新功能测试（2-back 和 3-back）的部分已经在文章的方法、结果、讨论中进行了补充。

修改稿：

## 2.2 主要研究工具

### 2.2.1 刷新功能测试

前后测采用 2-back 和 3-back 任务来测量被试的工作记忆刷新能力，在 2-back 任务中，要求被试比较当前出现的字母是否和此前刚呈现过的前面第二个字母相同。3-back 任务则是比较当前字母和此前呈现过的前面第三个字母是否相同。分别统计抑郁倾向训练组、抑郁倾向对照组和健康对照组被试在前后测的 2-back 和 3-back 任务的反应时和正确率。

## 3 结果

### 3.1 2-back 和 3-back 结果

三组被试在前后测的 2-back 和 3-back 成绩如表 2 所示。对三组的 2-back 正确率和反应时进行重复测量方差分析，发现时间的主效应显著 ( $F(1,57)=23.31, p<0.001, \eta^2_p=0.290$ ;  $F(1,57)=18.36, p<0.001, \eta^2_p=0.244$ )，时间和组别的交互作用均不显著 ( $F(2,57)=0.55, p=0.580, \eta^2_p=0.019$ ;  $F(2,57)=1.11, p=0.336, \eta^2_p=0.038$ )，三组在后测的正确率都有显著的提高，反应时都有显著的缩短。对三组的 3-back 正确率和反应时进行重复测量方差分析，同样发现时间的主效应显著 ( $F(1,57)=23.46, p<0.001, \eta^2_p=0.292$ ;  $F(1,57)=9.86, p=0.003, \eta^2_p=0.148$ )，时间和组别的交互作用均不显著 ( $F(2,57)=0.74, p=0.478, \eta^2_p=0.026$ ;  $F(2,57)=0.15, p=0.857, \eta^2_p=0.005$ )，三组在 3-back 后测的正确率都有显著的提高，反应时都有显著的缩短。对三组



的提高量进行方差分析发现组别的主效应均不显著。

表 2 2-back 和 3-back 任务反应时和正确率

		抑郁倾向训练组	抑郁倾向对照组	健康对照组
2-back 正确率 (%)	前测	77.82±15.95	78.69±15.70	74.85±14.90
	后测	84.13±9.06	89.56±7.90	84.70±9.38
	提高	6.30±14.13	10.86±14.82	9.85±14.37
2-back 反应时 (ms)	前测	1036.29±204.40	1049.52±182.21	1121.83±206.04
	后测	980.61±207.58	935.94±182.18	978.98±169.21
	提高	-55.68±221.96	-113.58±194.04	-142.84±138.54
3-back 正确率 (%)	前测	62.04±11.82	66.59±14.47	64.70±14.06
	后测	75.90±12.94	75.45±13.46	72.70±12.63
	提高	13.86±16.25	8.86±17.60	8.00±15.18
3-back 反应时 (ms)	前测	1095.30±247.68	1086.36±189.90	1084.86±199.58
	后测	1021.73±241.30	984.06±201.24	1015.10±163.16
	提高	-73.57±224.06	-102.30±204.90	-69.76±173.54

## 4 讨论

### 4.1 实验结果讨论

……就刷新功能而言，三组在 2-back 和 3-back 任务的后测中都有明显的进步，表现在正确率的提高和反应时的缩短上面。对三组的方差分析发现，组别的主效应均不显著，即没有发现训练组在刷新任务上有更大的提高。一个可能的原因是在编制对于成年人的工作记忆测验时会将难度水平设定在中等偏上一点的程度，对于多数被试，稍加训练（两次测验本身就是练习）都会提高，但提高的幅度有限，差不多接近天花板效应。根据 Zhao (2013)和 Xiu (2016)等人的研究可以发现，大学生群体在 2-back 的正确率都是较高，反应时较短的，因此本实验中三组在 2-back 和 3-back 任务的成绩上没有组别上的差异也是正常的。本研究中工作记忆训练的重点不在于工作记忆本身提高多少，主要在于引起功能的改变，即 HRV 的变化。本研究使用的工作记忆刷新训练已有很多文章证明过是有效的，如 Zhao(2013)等人的研究表明，该工作记忆刷新训练能引起个体明显的脑电活动如 P2、P3 等的改变；Xiu(2016)等人的研究表明此训练能引起正常大学生 HF-HRV 的改变。本研究主要想证明的也是工作记忆刷新训练对抑郁倾向大学生情绪调节能力的影响，主要表现在 HF-HRV 这个指标上。

意见 4：抑郁倾向组和健康组被试组的区分以 15 分（BDI）或 20 分（CES）为界，中间没有间隔区间，是否一高一低更好，而不是采用二分法。

回应：实验中抑郁倾向组被试的纳入标准是 BDI-II-C 得分≥15 分，且 CES-D 得分≥20 分；健康组被试的纳入标准为 BDI-II-C 得分<15 分，且 CES-D 得分<20 分。因此被试要同时满足两个抑郁量表的得分才通过筛选，相比其他采用单一量表进行筛选的，我们的标准会更加严格。虽然采用的是二分法，但在实际筛选健康组被试时会更多地选择在抑郁量表中得分较低的被试，一般 BDI-II-C 得分均低于 10，因此健康组的平均 BDI-II-C 得分为 3.90±3.86，平均 CES-D 得分为 11.55±5.07，抑郁倾向组的平均 BDI-II-C 得分为 20.57±5.33，平均 CES-D



得分为  $29.77 \pm 5.65$ 。独立样本 T 检验显示，健康组和抑郁倾向组的 BDI-II-C 得分有显著差异 ( $t(58)=12.42, p<0.001, d=3.58$ )，两组在 CES-D 得分上也有显著差异 ( $t(58)=12.15, p<0.001, d=3.39$ )。因此，实际进行的筛选接近于审稿人所说的“一高一低”，并非简单的二分法。

意见 5：“CES-D 及 BDI-II-C 量表得分情况均表现为抑郁倾向组显著高于健康被试组”，要有相关统计报告，请作者举一反三。

回应：相关统计报告：对三组的 CES-D 和 BDI-II-C 得分进行方差分析，发现在 CES-D 得分上，组别的主效应显著 ( $F(2,57)=74.65, p<0.001, \eta^2 p=0.724$ )，且健康对照组的得分显著高于抑郁倾向训练组和抑郁倾向对照组 ( $p<0.001; p<0.001$ )，抑郁倾向训练组和对照组没有显著差异 ( $p=0.289$ )。在 BDI-II-C 得分上，组别的主效应也显著 ( $F(2,57)=78.44, p<0.001, \eta^2 p=0.734$ )，健康对照组的得分同样显著高于抑郁倾向训练组和抑郁倾向对照组 ( $p<0.001; p<0.001$ )，抑郁倾向训练组和对照组没有显著差异 ( $p=0.236$ )。其余地方也进行了修改。

修改稿：

### 2.3.1 被试的分组

以自愿参加训练的方式将有抑郁倾向被试分为工作记忆刷新训练组和对照组，每组 20 例，分组后的被试基本情况见表 1。对三组的 CES-D 和 BDI-II-C 得分进行方差分析，发现在 CES-D 得分上，组别的主效应显著 ( $F(2,57)=74.65, p<0.001, \eta^2 p=0.724$ )，且健康对照组的得分显著高于抑郁倾向训练组和抑郁倾向对照组 ( $p<0.001; p<0.001$ )，抑郁倾向训练组和对照组没有显著差异 ( $p=0.289$ )。在 BDI-II-C 得分上，组别的主效应也显著 ( $F(2,57)=78.44, p<0.001, \eta^2 p=0.734$ )，健康对照组的得分同样显著高于抑郁倾向训练组和抑郁倾向对照组 ( $p<0.001; p<0.001$ )，抑郁倾向训练组和对照组没有显著差异 ( $p=0.236$ )。

意见 6：全文图太多，一些并不必要（如图 1）和不显著的结果（如图 6），可以不用放图。图可以做得更精致、紧凑一些。图中显著的对比如用\*,\*\*等标识出来。建议作者仔细思考，对全文中的图予以精简、修改、美化。

回应：已对文中的大部分图进行了改进和美化，原来的图 1 和图 6 已经删除，图中的显著对比也用星号标出，如图 5，其余的图不需要单独加显著星号。

修改稿：

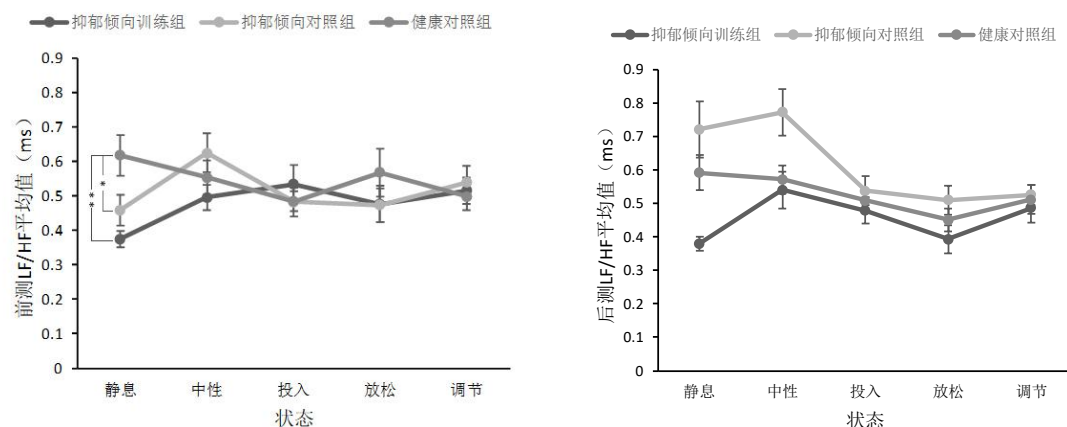


图 5 三组被试在前后测的情绪调节任务中五个条件下的平均 LF/HF-HRV 值

注：误差线为标准误 (SE)，\*  $p<0.05$ ，\*\*  $p<0.01$ 。其中上图是三组在前测的 LF/HF-HRV 值，差异检验表明，在静息条件下，健康对照组的 LF/HF-HRV 的平均水平显著高于抑郁倾向训练组和抑郁倾向对照组的平均水平 ( $p=0.001, p=0.043$ )。下图是三组后测的 LF/HF-HRV 结果，抑郁倾向训练组的 LF/HF-HRV 显著低于抑郁倾向对照组 ( $p=0.007$ )。

意见 7：“在各阶段的转换之间都会有两分钟的放松时间去帮助被试恢复平静并填写视频主观评分量表。”，“每段视频开始前有 3 分钟的静息时间让被试的生理指标恢复到基线水平”，到底是 2 分钟还是 3 分钟，依据是什么？

回应：这里应该是原文没表述清楚。两种状态之间转换的休息时间是 2 分钟，例如看完两段中性视频后被试需要静息两分钟，然后再看后面的两段投入性视频。在看两段同样性质的视频之间被试不做休息，仅对刚刚看完的视频进行主观评分。3 分钟是每段情绪视频的呈现时间。被试在两种状态转换之间休息的 2 分钟，目的是为了被试从前一段视频的情绪中脱离出来，减少顺序效应的影响。这是依据 Xiu(2016), Goldin(2005), Gyurak(2012)等人的实验设计。例如 Xiu(2016)和 Schweizer(2013)等人所使用的情绪调节任务流程与本实验是基本一致的，前者在不同情绪状态的视频之间设置了 2 分钟的休息时间，而后者也在负性视频后设置了 45 秒的恢复时间。另外，Gyurak 等人在 2012 年的情绪调节研究的实验设计里，两段不同的视频之间有 1 分钟的恢复时间。还有，在 Goldin(2005)等人的研究里，他们在悲伤和愉悦视频中间也插入了 2 分钟的中性视频，这本身也是一种恢复。因此，总的来说，2 分钟是一个比较合适的时间让被试从前一个情绪状态里恢复。表述不一致的地方已经在方法的叙述部分做了修改。

原稿：每段视频开始前有 3 分钟的静息时间让被试的生理指标恢复到基线水平

修改稿：每种条件的视频观看结束后有 2 分钟的静息时间让被试的生理指标恢复到基线水平。

参考文献：

- Goldin, P. R., Hutcherson, C. A. C., Ochsner, K. N., Glover, G. H., Gabrieli, J. D. E., & Gross, J. J. (2005). The neural bases of amusement and sadness: A comparison of block contrast and subject-specific emotion intensity regression approaches. *Neuroimage*, 27(1), 26-36. doi:10.1016/j.neuroimage.2005.03.018
- Gyurak, A., Goodkind, M. S., Kramer, J. H., Miller, B. L., & Levenson, R. W. (2012). Executive functions and the down-regulation and up-regulation of emotion. *Cognition & Emotion*, 26(1), 103-118. doi:10.1080/02699931.2011.557291
- Schweizer, S., Grahn, J., Hampshire, A., Mobbs, D., & Dalgleish, T. (2013). Training the Emotional Brain: Improving Affective Control through Emotional Working Memory Training. *Journal of Neuroscience*, 33(12), 5301-5311. doi:10.1523/jneurosci.2593-12.2013
- Xiu, L., Zhou, R., & Jiang, Y. (2016). Working memory training improves emotion regulation ability: evidence from hrv. *Physiology & Behavior*, 155(6), 25-29.

意见 8：“5（条件）”具体是什么，要明确指出来。

回应：具体是指 5（条件：静息、中性、投入、放松、调节）。已在文中改进。

意见 9：本文设计比较复杂，结果也很多，建议作者增加结果报告的结构性，注意分段的逻辑性，如一个 ANOVA 结果一段（含简单效应分析等）。

回应：已针对结果报告的结构性和分段的逻辑性进行了修改。例如，在“3.4.2 三组 HF-HRV 的 2（时间）×3（组别）×5（条件）混合方差分析”的结果里，从原来的三段改为一段。在“3.5.2 三组 LF/HF-HRV 的 2（时间）×3（组别）×5（条件）混合方差分析”的结果中，由原先的五段，修改为一段。在“3.4.3  $\Delta$  HF 的重复测量方差分析”中，由原来的一段改为两段。以及在“3.5.3  $\Delta$  LF/HF 的重复测量方差分析”，由一段改为两段。

修改稿：

### 3.4.2 三组 HF-HRV 的 2（时间）×3（组别）×5（条件）混合方差分析

对三组在前后测的 HF-HRV 数据进行 2（时间：前测、后测） $\times$ 3（组别：抑郁倾向训练组、抑郁倾向对照组、健康对照组） $\times$ 5（条件：静息、中性、投入、放松、调节）的重复测量方差分析。发现组别的主效应不显著 ( $F(2,57)=2.558, p=0.086, \eta^2_p=0.082$ )，时间的主效应也不显著 ( $F(1,57)=0.247, p=0.621, \eta^2_p=0.004$ )，条件的主效应显著 ( $F(4,228)=2.419, p=0.049, \eta^2_p=0.041$ )。五个条件两两进行比较发现，中性状态和投入状态下的 HF-HRV 之间存在显著差异 ( $p=0.048$ )，投入状态下的 HF-HRV 显著高于中性状态的 HF-HRV，其余差异不显著。另外，还发现时间 $\times$ 组别的交互作用显著 ( $F(2,57)=4.410, p=0.017, \eta^2_p=0.134$ )，条件 $\times$ 组别的交互作用也显著 ( $F(8,228)=3.202, p=0.002, \eta^2_p=0.101$ ) 时间 $\times$ 条件、时间 $\times$ 条件 $\times$ 组别的交互作用均不显著 ( $F(4,228)=0.413, p=0.799, \eta^2_p=0.007; F(8,228)=1.806, p=0.077, \eta^2_p=0.060$ )。简单效应分析显示，抑郁倾向训练组和抑郁倾向对照组被试在前测 HF-HRV 没有显著差异，在后测抑郁倾向训练组和抑郁倾向对照组差异边缘显著 ( $p=0.052$ )，后测抑郁倾向训练组的 HF-HRV 高于抑郁倾向对照组。还发现，抑郁倾向训练组在前后测的 HF-HRV 有显著差异 ( $p=0.042$ )，后测的 HF-HRV 要显著高于前测，抑郁倾向对照组和健康对照组在前后测均没有显著差异 ( $p=0.316$ )。抑郁倾向训练组在调节状态下的 HF-HRV 要显著高于静息状态 ( $p=0.002$ )，其余条件没有显著差异，抑郁倾向对照组和健康对照组在各个条件下均没有发现显著差异。

### 3.5.2 三组 LF/HF-HRV 的 2（时间） $\times$ 3（组别） $\times$ 5（条件）混合方差分析

对前后测的 LF/HF-HRV 数据进行 2（时间：前、后测） $\times$ 3（组别：抑郁倾向训练组、抑郁倾向对照组、健康对照组） $\times$ 5（条件：静息、中性、投入、放松、调节）的重复测量方差分析。发现组别的主效应不显著 ( $F(2,57)=2.945, p=0.061, \eta^2_p=0.094$ )，时间的主效应也不显著 ( $F(1,57)=0.728, p=0.397, \eta^2_p=0.013$ )，条件的主效应显著 ( $F(4,228)=8.494, p<0.001, \eta^2_p=0.130$ )。五个条件两两进行比较发现，中性状态和投入状态、放松状态、调节状态下的 LF/HF-HRV 均有显著性差异 ( $p<0.001, p<0.001, p=0.002$ )，中性状态下的 LF/HF-HRV 均显著高于投入状态、放松状态和调节状态。另外，还发现时间 $\times$ 组别的交互作用边缘显著 ( $F(2,57)=3.088, p=0.053, \eta^2_p=0.098$ )，条件 $\times$ 组别、时间 $\times$ 条件的交互作用也显著 ( $F(8,228)=5.623, p<0.001, \eta^2_p=0.165; F(4,228)=3.693, p=0.006, \eta^2_p=0.061$ )，时间 $\times$ 条件

×组别的交互作用不显著 ( $F(8,228)=1.600, p=0.126, \eta^2_p=0.053$ )。简单效应分析显示, 三组被试在前测 LF/HF-HRV 没有显著差异, 在后测抑郁倾向训练组和抑郁倾向对照组差异显著 ( $p=0.007$ ), 在后测抑郁倾向训练组的 LF/HF-HRV 要显著低于抑郁倾向对照组。在投入状态下, 抑郁倾向训练组的 LF/HF-HRV 显著低于抑郁倾向对照组 ( $p=0.011$ ), 抑郁倾向训练组和健康对照组没有显著差异 ( $p=1.000$ )。简单效应分析显示, 静息状态的 LF/HF-HRV 前后测差异显著 ( $p=0.036$ ), 后测的 LF/HF-HRV 显著高于前测。

意见 10: “5 局限性与展望”也是讨论的一部分, 建议整合到讨论中, 在讨论中也要更加结构化, 分小标题分别论述。重点论述一条主线, 结合各方面的发现来论证。

回应: 已将原来的“局限与展望”整合到讨论中。将原来的“4 讨论”与“5 局限与展望”修改为 4 讨论 (4.1 实验结果讨论; 4.2 HRV 差值意义; 4.3 研究局限; 4.4 未来展望) 几个部分。

修改稿:

## 4 讨论

### 4.3 研究局限

特别指出的是, 心率也许在投入和调节状态下都会显著高于中性条件, 但心率变异性则不是, HRV 与情绪范式相结合时需要考虑其特殊性, 通过作差法、取差值这一方式来平衡基线更要慎重, 因为基线本身 (比如这里指的中性视频的 HRV 值) 就表征着波动, 甚至已经反映了自主神经功能的状况, 所以未必适于作为被减数。另外, 考虑到神经细胞兴奋的不应期, 在今后此类研究中, 应该要注意设计专门适用于 HRV 的情绪诱发和调节范式。

其次, 由 3.2 中的结果可以看出, 自愿接受 20 天工作记忆刷新训练的抑郁倾向被试, 其情绪调节量表的得分情况与健康对照组没有显著差异, 只有抑郁倾向对照组在情绪调节量表得分上体现出了显著低于健康被试的现象, 考虑愿意参加训练的抑郁倾向被试或许有在潜意识中调节自己的情绪。以自愿参加为原则进行分组难免产生混淆因素, 但由于心理学科发展和实验条件的限制, 此类问题目前尚没有得到很好的解决。虽然在实际实验中, 采取自愿训练的分组方法可以保证较低的被试脱离率, 但难免也降低了整个实验分析的效能。在参与率与随机化中进行取舍, 依然是目前在工作记忆刷新训练的相关实验中需面对的问题。随着心理学实验方法的发展, 希望能在实验招募上有更好的处理措施, 以尽可能减少或消除被试在前测由于实验组和对照组非双盲的分组和处理而导致的心理各方面的差异。另外, 为了避免昼夜节律对人体心率变异性的影响, 此类涉及生理指标的研究均应在固定季节固定时段进行。

### 4.4 未来展望

在关于工作记忆刷新训练的长期效应上, 本研究没有进行长期的追踪, 因此无法得知训练的后续效果如何, 这是今后可以进一步探究的问题。目前还没有关于工作记忆训练在情绪调节方面远迁移长期效应的讨论, 但我们注意到相关的一些工作记忆训练研究对个体自身工作记忆的提高以及在智力方面迁移并维持方面进行了讨论, 如 Gropper 等人对有注意缺陷多动障碍 (ADHD) 的大学生进行为期五周的工作记忆训练, 发现训练的效果可以持续到两个月后 (Gropper R. J., Gotlieb H., Kronitz R., & Tannock R., 2014); Chen (2017) 等人对儿童进行工作记忆刷新训练, 发现在训练完六个月后受训者在数学成绩上仍有明显提高 (Chen, X., Ye, M., Chang, L., Chen, W., & Zhou, R., 2017)。因此, 在未来研究可以进一步讨论工作记忆刷新训练的维持效应的问题。

其次是关于 HRV 正常取值范围的国内常模建立。本研究发现 HRV 有可能因人种的不

同而有不同的波动范围，本研究所测的 HF 值大约为  $160\text{ms}^2/\text{Hz}$ ，在  $0.15\text{-}0.4\text{Hz}$  的波段内积分，并开方，得到的结果在  $6\text{ms}$  左右，这和巴西学者 Antelmi 等在欧洲开展的实验结果稍有差别 (Antelmi, Paula, Shinzato, Peres, Mansur, & Grupi, 2004)，其研究表明 20 岁左右的巴西人群的正常 HF 的 24 小时心电图记录下的 HF-HRV 值大概在  $16\text{-}18\text{ms}$  范围内 (见下图 7)，这一方面也许是人种差异，二方面也许是记录时长的差异导致的，有文献表明 24h 的记录结果会高于 5min 的结果，而国内还没有大样本研究或是常模出现，这是未来值得研究的方向。另外，考虑短时程 HRV 与长时程 HRV 相比，能较好控制各种影响因素，稳定程度及真实性更好 (柯素颖, 赵娣, & 杨芳, 2015)，欧洲心血管病学会及北美心脏起搏和电生理学会专题委员会也建议以 5 min 作为标准 (Listed, 1996)。故本实验采取短时程 HRV 有较好的稳定性，但关于可比性，仍需要大样本调查研究来论证。

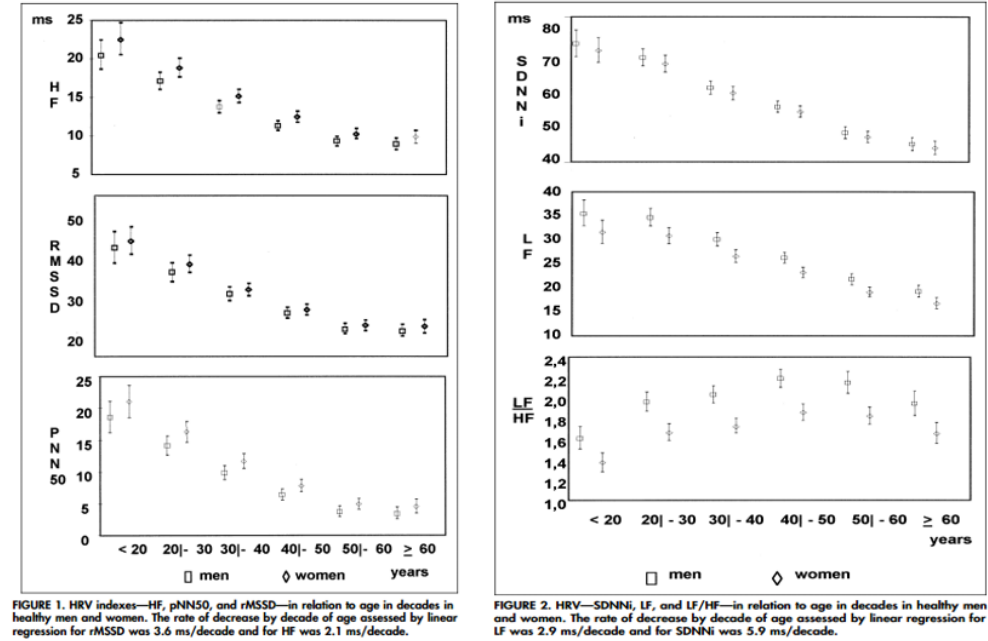


图 7 巴西大样本 HRV 频域、时域指标正常范围图<sup>1</sup>

最后，在抑郁的干预中，也有研究者尝试了使用 HRV 生物反馈调节的手段 (李欣, 张通, & 宋鲁平, 2013)，这不失为直接作用于人类的情绪调节能力的一项措施，但目前国际上还没有公认的可用于压力识别研究的 HRV 或心电信号数据库 (刘振, 2016)。HRV 的参数是丰富的，包括频域参数、时域参数以及非线性参数，每项参数里还划分许多子参数，可以考虑将一系列参数值作为识别特征，以随机森林、贝叶斯等算法为建模基础 (陈彦会, 2017)，建立压力或情绪的识别系统。希望本项研究可以为今后的实验提供参考。

意见 11：一些文字表达问题。（1）表达错误，如并不存在“图 6 左和右”，而是上和下；（2）表达不够精炼，如“是有一个分支起主导作用的”；（3）一些标点符号使用不当，如有的地方应用中文符号的用成了英文符号（如“，”和“，”）。请作者举一反三，认真修改。

<sup>1</sup> 摘自 Antelmi, I., Paula, R. S. D., Shinzato, A. R., Peres, C. A., Mansur, A. J., & Grupi, C. J. (2004). Influence of age, gender, body mass index, and functional capacity on heart rate variability in a cohort of subjects without heart disease. *American Journal of Cardiology*, 93(3), 381-5.



回应：谢谢审稿人的详细意见，已把这几处进行了改进。如：

原稿：在通常情况下，心交感神经和心迷走神经这两个之中，是有一个分支起主导作用的。

修改稿：在通常情况下，心交感神经和心迷走神经这两个之中，是其中一个在起主导作用的。  
其他具体修改见文中红色部分。

### 第三轮

审稿人 1 意见：作者较好回答了评审人的意见。还有几个小问题：

意见 1：中文摘要太琐细、太长，建议只写最重要的结果，即能够支持结论的核心结果。

回应：已对中文摘要进行修改使其更加简洁，并且突出主要的结果。

修改稿：

**摘要** 根据流调中心用抑郁量表（CES-D）和贝克抑郁量表（BDI-II-C）的得分情况招募健康被试 20 例，抑郁倾向被试 40 例，以自愿参加的分组方式将抑郁倾向被试分为工作记忆刷新训练组和对照组，每组 20 例。对训练组进行为期 20 天的工作记忆刷新训练，对照组不处理。记录各组被试在前后测的刷新功能以及在情绪调节任务中量表的得分情况，并收集各组被试在 5 种实验条件下的心率变异性（HRV）的频域指标值，结果发现：前测，抑郁倾向个体的 HF-HRV 显著低于健康个体的平均水平。经过工作记忆训练后，抑郁倾向训练组在情绪调节任务中的 HF-HRV 水平有显著的提高，贴近健康对照组的水平，并与抑郁倾向对照组分离。研究表明，工作记忆刷新训练能够使抑郁倾向大学生的 HRV 活动更贴近于健康被试的 HRV 活动，使抑郁倾向大学生的情绪调节能力得到改善。

意见 2：建议多个显著性不要同时报告（如  $F(2,285)=2.37, p=0.095, \eta^2_p=0.016$ ;  $F(4,285)=1.03, p=0.392, \eta^2_p=0.014$ ）；另外，不显著的不需要提供效应量值。

回应：已将文中同时报告的结果分开来报告，并且去掉了不显著结果的效应量值，例如：

原稿：

#### 3.1 2-back 和 3-back 结果

三组被试在前后测的 2-back 和 3-back 成绩如表 2 所示。对三组的 2-back 正确率和反应时进行重复测量方差分析，发现时间的主效应显著（ $F(1,57)=23.31, p<0.001, \eta^2_p=0.290$ ； $F(1,57)=18.36, p<0.001, \eta^2_p=0.244$ ），时间和组别的交互作用均不显著（ $F(2,57)=0.55, p=0.580, \eta^2_p=0.019$ ； $F(2,57)=1.11, p=0.336, \eta^2_p=0.038$ ），三组在后测的正确率都有显著的提高，反应时都有显著的缩短。对三组的 3-back 正确率和反应时进行重复测量方差分析，同样发现时间的主效应显著（ $F(1,57)=23.46, p<0.001, \eta^2_p=0.292$ ； $F(1,57)=9.86, p=0.003, \eta^2_p=0.148$ ），时间和组别的交互作用均不显著（ $F(2,57)=0.74, p=0.478, \eta^2_p=0.026$ ； $F(2,57)=0.15, p=0.857, \eta^2_p=0.005$ ），三组在 3-back 后测的正确率都有显著的提高，反应时都有显著的缩短。对三组

的提高量进行方差分析发现组别的主效应均不显著。

修改稿:

### 3.1 2-back 和 3-back 结果

三组被试在前后测的 2-back 和 3-back 成绩如表 2 所示。对三组的 2-back 正确率进行重复测量方差分析,发现时间的主效应显著 ( $F(1,57)=23.31, p<0.001, \eta^2_p=0.290$ ), 时间和组别的交互作用不显著 ( $p=0.580$ )。对 2-back 反应时进行方差分析,同样发现时间的主效应显著 ( $F(1,57)=18.36, p<0.001, \eta^2_p=0.244$ ), 时间和组别的交互作用不显著 ( $p=0.336$ )。表明三组在后测的正确率都有显著的提高, 反应时都有显著的缩短。

对三组的 3-back 正确率进行重复测量方差分析,发现时间的主效应显著 ( $F(1,57)=23.46, p<0.001, \eta^2_p=0.292$ ), 时间和组别的交互作用不显著 ( $p=0.478$ )。对 3-back 反应时进行分析,同样发现时间的主效应显著 ( $F(1,57)=9.86, p=0.003, \eta^2_p=0.148$ ), 时间和组别的交互作用不显著 ( $p=0.857$ )。三组在 3-back 后测的正确率都有显著的提高, 反应时都有显著的缩短。

对三组的提高量进行方差分析发现组别的主效应均不显著。

此为其中一处修改, 其他类似的也都一一进行了修改。

意见 3: “4.1 实验结果讨论”标题不太合适, 理论上所有讨论都是实验结果的讨论。建议进一步细分小标题进行分别讨论。

回应: 已将 4.1 的标题改为“4.1 抑郁倾向个体情绪调节能力”, 更符合讨论的内容。

---

## 第四轮

审稿人 1 意见:

意见 1: 不显著的结果的 F 值或 T 值还是需要报告, 只是不用报告效应量, 请参考《心理学报》其他文章。

回应: 已重新修改, 只去掉了不显著结果的效应量值, 保留了其 F 值。例如:

修改稿:

### 3.1 2-back 和 3-back 结果

三组被试在前后测的 2-back 和 3-back 成绩如表 2 所示。对三组的 2-back 正确率进行重复测量方差分析,发现时间的主效应显著 ( $F(1,57)=23.31, p<0.001, \eta^2_p=0.290$ ), 时间和组别的交互作用不显著 ( $F(2,57)=0.55, p=0.580$ )。对 2-back 反应时进行方差分析,同样发现时间的主效应显著 ( $F(1,57)=18.36, p<0.001, \eta^2_p=0.244$ ), 时间和组别的交互作用不显著 ( $F(2,57)=1.11, p=0.336$ )。表明三组在后测的正确率都有显著的提高, 反应时都有显著的缩短。

对三组的 3-back 正确率进行重复测量方差分析,发现时间的主效应显著 ( $F(1,57)=23.46, p<0.001, \eta^2_p=0.292$ ), 时间和组别的交互作用不显著 ( $F(2,57)=0.74, p=0.478$ )。对 3-back 反应时进行分析,同样发现时间的主效应显著 ( $F(1,57)=9.86, p=0.003, \eta^2_p=0.148$ ), 时间和



组别的交互作用不显著 ( $F(2,57)=0.15$ ,  $p=0.857$ )。三组在 3-back 后测的正确率都有显著的提高, 反应时都有显著的缩短。

对三组的提高量进行方差分析发现组别的主效应均不显著。

此为其中一处修改, 其他类似的也都一一进行了修改。

---

#### 第五轮

编委复审: 稿件已符合心理学报要求, 建议发表。

主编终审: 同意外审和编委意见, 建议录用并发表。