

《心理学报》审稿意见与作者回应

题目：高善良特质在情绪调节行动控制中的内隐优势

作者：孙俊才；寻凤娇；刘萍；张文海

第一轮

审稿人 1 意见：该文章通过三个行为实验考察了善良特质对情绪加工的影响。研究发现，与低善良被试相比，高善良被试对情绪性人际关系词表现出显著的正性注意偏向，更倾向于采用控制型情绪调节方式，而且高善良被试能更好地克服“愤怒优势效应”（表现为更快地从愤怒面孔中搜索出高兴面孔），反映了其更强的情绪调节能力。研究问题具有较好的科学意义和应用价值，研究设计思路清楚。但有如下问题需要进一步澄清：

（备注：在文稿中，根据第一位审稿专家进行的修改见绿色字部分）

意见 1：实验 1 的结果描述有错误，如 p9：“简单效应分析发现，高善良被试在人际积极词上的反应时显著长于低善良被试， $F(1,48)$ ”

回应：感谢专家的审阅！实验 1 的结果改为：简单效应分析发现，在积极人际词上，高善良被试的反应时显著长于低善良被试的反应时，见“2.3 结果”部分的绿字。

意见 2：实验 2 的研究假设是：善良特质优化内隐控制型情绪调节态度，那么我们会预期高善良被试在相容任务阶段的反应时小于不相容任务阶段的反应时，可是作者却说高善良被试在相容任务阶段的反应时大于不相容任务阶段的反应时支持了其研究假设，这里的逻辑是什么？另外，统计结果的报告也不完整，如“简单效应分析发现，高善良组在两种任务上的反应时没有显著差异，但是表现出相容任务反应时长于不相容任务反应时的倾向， $F(1, 54) = 3.36, p = 0.07, \eta_p^2 = 0.06, 95\%CI$ 为 $[-12.71, 284.98]$ 。但低善良组的结果没有报告。

回应：感谢专家的质疑！

实验中为了平衡试次顺序影响，设计了两个实验程序，一个程序是相容任务在第四部分，不相容在第七部分；另一个程序是不相容在第四部分，相容在第七部分。由于我们的失误，在统计分析时没有将数据调换过来，报告了错误的结果，我们为此表示深深的歉意！

我们重新整理了数据，报告了正确的实验结果，结果显示，高善良被试在相容任务上的反应时显著短于不相容任务的反应时，符合经典 IAT 效应的假设，而且高组的 D 值显著大

于低组的 D 值，这与预期结果一致，说明高善良被试持有更为积极的情绪调节控制性内隐态度，而低善良被试持有更为消极的情绪调节表达性内隐态度。具体修改见“3.4 结果”和“3.5 小结”的绿字部分。

意见 3: 实验 3 在 $p = 0.097$ 的情况下把交互作用定义为边缘显著是否合适？之后为什么对两组被试在三种面孔组合上的反应时进行独立样本 t 检验，而不是简单效应检验？

回应: 感谢专家的质疑！

(1) 实验 3 在 $p = 0.097$ 的情况下把交互作用定义为边缘显著确实不合适。 $p = 0.097$ 表示交互作用不显著。

(2) 虽然交互作用不显著，为了进一步探讨两组被试在三种面孔组合上是否存在显著差异，因此我们借鉴以往研究（何宁，朱云莉，2016）采用独立样本 t 检验，考察高善良被试从愤怒面孔中搜索到高兴面孔的反应时是否短于低善良被试，从而证明高善良被试的情绪恢复效率更高。

综上，我们在“4.4.2 面孔表情视觉搜索任务的反应时”部分改为：面孔组合方式和组别之间的交互作用不显著， $F(2, 110) = 2.38$ ， $p = 0.097$ 。进一步对两组被试在三种面孔组合上的反应时进行独立样本 t 检验，见绿字部分。

参考文献:

何宁，朱云莉. (2016). 自爱与他爱:自恋、共情与内隐利他的关系. *心理学报*, 48(2), 199-210.

以人际反应指标为因变量进行 2(自恋水平: 自恋、非自恋)×2(实验处理条件: 启动组、控制组)方差分析。结果显示，自恋水平主效应不显著， $F(1,132) = 0.58$ ， $p = 0.447$ ， $\text{partial } \eta^2 = 0.004$ ；实验处理条件主效应显著， $F(1,132) = 8.84$ ， $p = 0.003$ ， $\text{partial } \eta^2 = 0.063$ ；自恋水平与实验条件的交互作用边缘显著， $F(1,132) = 3.80$ ， $p = 0.053$ ， $\text{partial } \eta^2 = 0.028$ 。进一步独立样本 t 检验结果发现，启动条件下自恋组在人际反应指标问卷上的得分显著高于控制条件下自恋组被试($t(56) = 3.233$ ， $p = 0.002$ ， $\text{Cohen } d = 0.850$)，而非自恋组被试人际反应指标问卷得分在启动与控制条件下没有差异(见图 1)。共情启动后自恋组与非自恋组被试在人际反应指标问卷上的得分差异不显著， $t(70) = 0.847$ ， $p = 0.400$ ， $\text{Cohen } d = 0.203$ 。

意见 4.所有的结果图都不规范，应该标注标准差或标准误，以及差异显著性。而且重复的结果表和结果图只保留一种就可以。

回应: 感谢专家的建议！

我们重新绘制了三个实验的结果图。考虑到原稿实验 3 结果图 7 与结果表 5 的重复度较高，在修改中只保留了结果表 5，删除了结果图 7。请专家审阅！

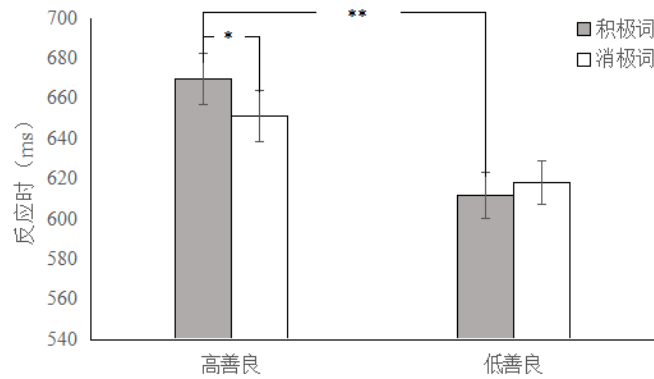


图2 两组被试在积极词和消极词上的反应时
注：误差线为标准误；*为 $p < .05$ ，**为 $p < .01$

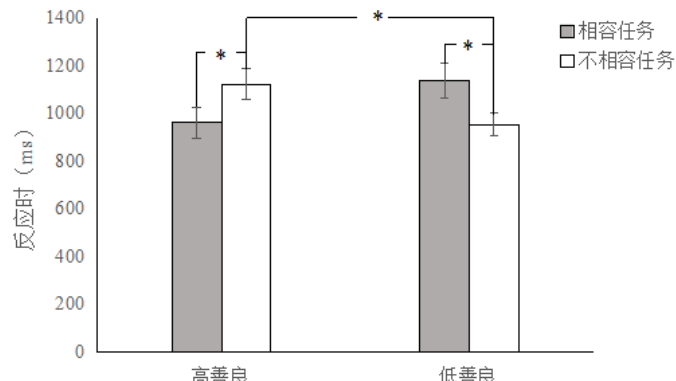


图4 两组被试在不同任务上的反应时
注：误差线为标准误；*为 $p < .05$

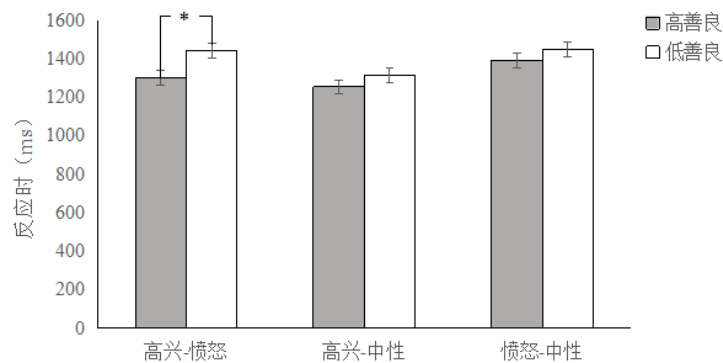


图7 两组被试在不同面孔组合上的反应时
注：误差线为标准误；*为 $p < .05$

.....

审稿人 2 意见：该研究以情绪调节的行动控制理论为基础，通过三个实验，考察了高低善良特质对内隐情绪调节各项子任务的影响，结果发现，与低善良被试相比，高善良被试对情绪

性人际关系词表现出显著的正性注意偏向；高善良被试更倾向于采用控制型情绪调节方式调节情绪；高善良被试在搜索“靶目标为高兴面孔，干扰目标为愤怒面孔”时的反应时显著较快。研究设计合理，逻辑清晰，数据分析合理，结果能够较好的支持结论。但仍存在以下问题。

（备注：根据第二位审稿专家进行的修改，见文稿红色字部分）。

意见 1：文中假设部分缺乏相应的参考文献支撑。例如，“多项研究证实了视觉搜索任务中存在愤怒优势效应(徐展，李灿举，2014)。基于上述研究，从愤怒面孔背景中搜索高兴面孔的反应时指标可以作为内隐调节效果的指标之一。”多项研究证实了....，作者却只列了一个研究，是否应该多引用几个？

回应：感谢专家的建议！

我们在文中添加了相关的引用，见“1.2.3 善良特质影响内隐情绪恢复效果”的红字部分，并将相应参考文献补充完整。

新增参考文献为

Lipp, O. V., Price, S. M., & Tellegen, C. L. (2009). Emotional faces in neutral crowds: detecting displays of anger, happiness, and sadness on schematic and photographic images of faces. *Motivation and Emotion*, 33(3), 249-260.

Öhman, A., Juth, P., & Lundqvist, D. (2010). Finding the face in a crowd: Relationships between distractor redundancy, target emotion, and target gender. *Cognition and Emotion*, 24(7), 1216-1228.

Schmidt-Daffy, M. (2011). Modeling automatic threat detection: development of a face-in-the-crowd task. *Emotion*, 11(1), 153.

意见 2：实验一的材料与程序部分，“所选的积极人际关系情绪词和消极人际关系词在情绪效价等级评定上得分差异显著，”那么积极情绪词和消极情绪词的唤醒度如何控制？作者没有进行说明。

回应：非常感谢审稿专家指出这个问题。在实验过程中，由于汉语情感词系统库的词汇不能完全满足我们的实验需要，因此，我们参照王一牛，周立明和罗跃嘉(2008)评定情绪词愉悦度和唤醒度的方法，进一步评定了这两类词的愉悦度和唤醒度，防止情绪词的唤醒度对效价

产生干扰作用。在第一稿中考虑到篇幅没有报告这部分结果。现把这部分结果报告到文稿中。请专家审阅。

结果表明，积极情绪词的愉悦度 ($M = 7.33, SD = 1.42$) 显著高于消极情绪词的愉悦度 ($M = 2.32, SD = 1.60$)， $t(29) = 10.14, p < 0.0001$ ；积极情绪词的唤醒度 ($M = 6.09, SD = 1.52$) 与消极情绪词的唤醒度 ($M = 5.24, SD = 1.70$) 差异不显著， $t(29) = 1.79, p = 0.083$ 。

参考文献

王一牛，周立明，罗跃嘉. (2008). 汉语情感词系统的初步编制及评定. *中国心理卫生杂志*, 22(8), 608-612.

意见 3: 实验任务的表述还需要更加细致，如，“相容任务是指被试把情绪调节控制词和积极属性词联合并按“F”键反应，把情绪调节表达词和消极属性词联合并按“J”键反应；不相容任务是指被试把情绪调节表达词和积极属性词联合并按“F”键反应，把情绪调节控制词和消极属性词联合并按“J”键反应。”受试者按键有没有时长限制？

回应: 感谢专家的建议！我们细化了表达。

(1) 相容任务是指被试把属于“情绪控制”和“积极属性”的刺激词都视为一类并按“F”键，把属于“情绪表达”和“消极属性”的刺激词都视为一类并按“J”键。不相容任务是指被试把属于“情绪表达”和“积极属性”的刺激词都视为一类并按“F”键，把属于“情绪控制”和“消极属性”的刺激词都视为一类并按“J”键。见文稿“3.3 实验程序”部分红字修改。

(2) 被试的按键没有时长限制，在结果统计分析前，我们首先对数据进行预处理，对反应时小于 300ms 和大于 10000ms 的数据进行剔除，对错误的反应时用该任务的平均反应时加上 600ms 来替换（刘俊升，桑标，2011）。

意见 4: “情绪控制目标词”和“情绪调节控制词”要统一；“情绪表达目标词”和“情绪调节表达词”要统一。

回应: 感谢专家的建议！

我们统一表述为“情绪控制目标词”与“情绪表达目标词”。

意见 5: 所有结果的柱状图都只有平均值，应该加上标准差，并且注明显著的差异。

回应：第一位审稿专家也指出了这个问题，非常感谢两位专家的建议！

我们重新绘制了结果柱状图，另外，根据第一位审稿专家意见 4 所指出的结果表与结果图只保留一种的建议，我们在正文文稿中删除了图 7，详见文稿与第一位审稿专家意见 4 的回应。

意见 6：文中参考文献部分出现了 Mauss, Bunge, & Gross(2010)，应该改成 Mauss, Bunge 和 Gross(2010)，建议规范引用；其次，在文后参考文献部分，出现了期刊名缩写的情况。

回应：感谢审稿专家提出宝贵意见，在认真校对的基础上，修改了（见文稿红字部分）三个方面的内容：

- （1）正文中只写第一作者的引用，根据规范要求补充了全部作者。
- （2）修改了引用时的英文括号，采用中文括号，符合引用规范。
- （3）修改了参考文献，将期刊名改为了全称（“&”改为“and”）并斜体，修改了作者名不规范的文献，补充了文献的卷次和页码信息，将页码范围符号改为“-”，删除了一条重复的文献,增添了三条新文献。

意见 7：文中出现英文单词错误，比如“bebefits”等；其次出现若干中文语言错误；英文摘要存在一些语法问题，比如，“At last, the participants’ emotion were tested by PANAS once more.”和“showed a significant positive attention”等等，建议作者认真修改英文摘要。

回应：感谢审稿专家明察秋毫！

我们根据《心理学报》的摘要要求重新写了英文摘要，并请英语专业的同事进行了细致修改，请专家审阅。

第二轮

审稿人 1 意见：审阅了作者的修改和回复，我没有进一步的意见，同意发表。

回应：感谢审稿专家提出的宝贵修改意见，提升了文稿的质量。

审稿人 2 意见：该研究以情绪调节的行动控制理论为支撑，设计 3 个行为实验考察了高、低善良特质与情绪调节三个子任务之间的关系。结果发现，高（vs.低）善良被试对积极的情

情绪性人际关系词更敏感；高善良被试更倾向于采用控制型情绪调节方式调节情绪；高善良被试搜索愤怒背景下的高兴面孔的速度更快。该选题较有新意，实验设计较为合理，但仍有较多的地方需要继续完善。

回应：非常感谢审稿专家对本研究的肯定！您的支持和建议极大地促进文稿的完善。

备注：第二轮修改和补充文字采用蓝色字标记。

主要问题

意见 1：摘要对实验结果的概述不全面，建议分别阐述两组被试对应的结果。

回应：非常感谢专家建议，完善了摘要写作。修改后的摘要为：

摘要 基于善良特质在中国人人格修养中的重要价值，本研究以情绪调节的行动控制理论为模型参照，设计了三项行为实验检验善良特质的导向性与内隐情绪调节三项子任务的关系。研究 1 采用 Stroop 任务，比较高低善良被试判断人际关系词颜色时，是否受到色词内容的干扰，结果发现，高（vs.低）善良被试对积极的情绪性人际关系词更敏感；研究 2 采用情绪调节内隐联想测验任务，比较情绪调节内隐态度差异，发现高（vs.低）善良被试更倾向于采用控制型情绪调节方式调节情绪；研究 3 在诱发被试的负性情绪条件下，采用面孔表情视觉搜索任务，间接考察内隐情绪修复效果。结果显示高（vs.低）善良被试搜索愤怒背景下的高兴面孔的速度更快。三项实验结果一致地表明善良特质利他重感情的行动目的性在情绪调节的行动控制中具有内隐优势，具有积极的心理健康促进作用。

意见 2：在问题提出部分，作者介绍了善良特质的作用以及研究善良的必要性，也介绍了内隐情绪调节的相关研究，并将其与外显情绪调节进行了初步比较，但从头至尾都没有给出“善良”、“内隐情绪调节”和“外显情绪调节”的定义。对行动控制理论和内隐情绪调节之间的关系阐述不够。

回应：非常感谢专家的建议。在修改稿中，分别引用了这三个概念的基本定义并从目的达成的角度进一步解释了行动控制理论与内隐情绪调节的关系。

（1）善良特质：善良特质是中国人格结构的重要维度之一（王登峰，崔红，2003），代表了中国文化中的“好人”特质，即在人际交往中所表现出的对人友好、真诚，能够顾及他人，关心、宽容他人，重视感情生活，遇事公平、正直等品质（王登峰，崔红，2005）。

(2) 外显情绪调节与内隐情绪调节：我们引用 Gyurak, Gross 和 Etkin (2011) 从意识水平对外显情绪调节与内隐情绪调节的界定，完善了这两个概念。Gross 和 Thompson (2007) 把情绪调节定义为目标指向的心理加工过程，其目的在于影响所体验情绪的强度、持续时间与类型等。目标指向性的特点，使得情绪调节具有灵活性，以应对情境的独特性并实现自我的长期目标。基于这些观点，Gyurak, Gross 和 Etkin (2011) 从情绪调节的意识水平，提出情绪调节的双加工模型，把外显情绪调节 (explicit emotion regulation) 定义为“需要意识努力启动，在执行过程中具有一定的程度的意识监控，而且与一定水平的内在觉察和领悟有关联”的调节加工过程；与之相对应，把内隐情绪调节 (implicit emotion regulation) 定义为“由刺激自动激活，并且调节的执行与完结不需要意识监控，也不需要觉察与领悟”。

(3) 情绪调节的行动控制理论与内隐情绪调节的关系。从内隐情绪调节的研究资料来看，研究者明确了内隐情绪调节没有意识决策性但是具有目的导向性的特点，并且这种目的导向性具有自动激活的特点。而情绪调节行动控制的理论，恰恰在于解释为了达到情绪调节目的，应该采用的行动控制序列。文稿围绕这个核心桥梁，进一步解释了情绪调节的行动控制理论与内隐情绪调节的关系。在修改过程中，调整了原稿的表达顺序，使得观点更为明确清晰。

意见 3：P16“善良特质影响内隐情绪恢复效果”下面第一段逻辑混乱，建议按照参考文献的年份顺序（内隐情绪调节领域的发展历程）来组织本部分。

回应：感谢专家的建议，我们重新调整了“善良特质影响内隐情绪恢复效果”的表达顺序。具体修改如下：

由于内隐情绪调节的目标可能有助于人们决定自身是否需要进行情绪调节 (Koole et al., 2015)，一旦情绪调节的目标被激活了，相应的内隐情绪调节过程就会自动启动。因此，如果假设 2 成立，即高善良被试更倾向于采用控制性的调节目标，那么，从逻辑上来说他们在内隐情绪调节的效果应该更好。基于威胁优势效应原理，从愤怒面孔中识别高兴面孔成为内隐情绪调节效果任务测量的典型方法。

Öhman, Lundqvist 和 Esteves (2001) 提出的表情面孔搜索范式获得了较多的研究支持，他们的研究表明，识别威胁性愤怒面孔的速度和准确率优于其他负性面孔（如悲伤），表现为威胁优势效应，即潜在的威胁可以快速捕获注意，并且不受目标以及个体的意向控制。后续研究表明，无论使用真实的情绪面孔图片或面孔简图作为实验材料，研究结果都证实视觉

搜索任务中存在愤怒优势效应 (Schmidt-Daffy, 2011; Öhman, Juth, & Lundqvist, 2010; Lipp, Price, & Tellegen, 2009)。

Jostmann, Koole, Van 和 Fockenberg (2005) 指出, 被试搜索愤怒表情面孔背景中高兴面孔的反应时, 能够作为个体情绪调节能力高低的指标。这种测量方法由于可以克服自我报告测量方法的局限性, 因而受到研究者的青睐 (Uhlmann, Leavitt, Menges, Koopman, Howe, & Johnson, 2012)。后续进一步变化表情面孔搜索任务的组合类型, 并用来测量情绪调节效果。基于上述研究, 从愤怒面孔背景中搜索高兴面孔的反应时指标可以作为内隐调节效果的指标之一。

意见 4: P17 (1) “其中男生 102 人, 女生 297 人, 平均年龄 19.32 岁 (SD = 3.71)。按善良人格得分从高到低排序, 以 27% 为标准, 筛选出高分者 30 名, 低分者 30 名”, (102+297)*27%=108,30 人是怎么算出来的? “参加实验 1 被试为 54 人(高分组 27 人, 低分组 27 人)” 与前文说的高低组各 30 人矛盾。

回应: 感谢专家的质疑。由于我们的表述不当, 产生了疑问。修改为: 按善良人格得分从高到低排序, 以上下 27% 为标准, 并根据被试的意愿, 请 30 名高分者和 30 名低分者参加后续的实验。由于在实验过程中, 个别被试没有完成所有实验任务, 在实验 1 中, 有 54 名被试参加了实验。特此说明。

(2) 自检报告中用 Gpower 预估样本量时, 设置的 η^2 值为 0.4, 0.4 这一数值是怎样确定的? 实际上本研究中所得到的最大 η^2 值才为 0.25, 是否可以理解为本实验没有达到预期的实验效果? Gpower 中的 Number of measurements 值应该为 4 (2 善良组 \times 2 词语类型) 而不是 2。按照 $\eta^2=0.25$, Number of measurements=4 计算出的样本量为 72, 说明本实验中的样本量 (54) 不够, 还需继续补充实验。

回应: 感谢专家的质疑! 已有研究 (卢谢峰, 唐源鸿, 曾凡梅, 2011) 指出 effect size 的大小判断标准, 对于 η^2 这个效应量指标来说, 0.04 为小效应, 0.25 为中等效应, 0.64 为大效应。因此, 在第一稿中, 以 0.4 为标准设定效应量不够严谨。这次, 我们按照该文献讨论的大小判断标准, 将 effect size 值设定为 0.25, 并进一步对本研究的三个实验进行了样本量估计。结果如下。

实验一采用: 2 (组别: 高善良、低善良) \times 2 (词语类型: 积极词、消极词) 的混合实验设计, 其中组别为被试间变量, 词语类型为被试内变量。在 G-power 中, Number of measurements 指的是被试内变量的水平数, 因此 Number of measurements 的值为 2, 计算

出的样本量为 54 人。实验二也是 2（组别：高善良、低善良）*2（实验任务：相容、不相容）的混合实验设计，样本估计量的算法同实验一。实验三为 2（组别：高善良、低善良）*3（面孔组合方式：高兴-愤怒、高兴-中性、愤怒-中性）的混合实验设计，Number of measurements 的值为 3，计算出的样本量为 44 人。

The screenshot shows the G*Power software interface. The 'Test family' is set to 'F tests' and the 'Statistical test' is 'ANOVA: Repeated measures, within-between interaction'. The 'Type of power analysis' is 'A priori: Compute required sample size - given alpha, power, and effect size'. The 'Input Parameters' section includes: 'Determine =>', 'Effect size f' (0.25), 'alpha err prob' (0.05), 'Power (1-beta err prob)' (0.95), 'Number of groups' (2), 'Number of measurements' (2), 'Corr among rep measures' (0.5), and 'Nonsphericity correction epsilon' (1). The 'Output Parameters' section includes: 'Noncentrality parameter lambda' (13.5000000), 'Critical F' (4.0266314), 'Numerator df' (1.0000000), 'Denominator df' (52.0000000), 'Total sample size' (54), and 'Actual power' (0.9500773).

参考文献：

- (1) Richardson, J. T. E. (2011). Eta squared and partial eta squared as measures of effect size in educational research. *Educational Research Review*, 6(2), 135-147
- (2) 卢谢峰, 唐源鸿, 曾凡梅. (2011). 效应量:估计、报告和解释. *心理学探新*, 31(3), 260-264.
- (3) Gpower 手册第 27 页解释了 η_p^2 的理论期望与 SPSS 输出结果之间的差异。

We first calculate the three means for factor A: $\mu_{i**} = \{-0.722231, 1.30556, -0.583331\}$. Due to the fact that we have first subtracted the grand mean from each cell we have $\sum_i \mu_{i**} = 0$, and we can easily compute the variance of these means as mean square: $\sigma_m^2 = 1/3 \sum_i \mu_{i**}^2 = 0.85546$. With these values we calculate $f = \sqrt{\sigma_m / \sigma} = \sqrt{0.85546 / 1.71296} = 0.7066856$. The effect size drawer in G*POWER can be used to do the last calculation: We choose From Variances and insert 0.85546 in the Variance explained by special effect and 1.71296 in the Error Variance field. Pressing the Calculate button gives the above value for f and a partial η^2 of 0.3330686. Note that the partial η^2 given by G*POWER is calculated from f according to the formula $\eta^2 = f^2 / (1 + f^2)$ and is not identical to the SPSS partial η^2 , which is based on sample estimates. The relation between the two is "SPSS $\eta_0^2 = \eta^2 N / (N + k(\eta^2 - 1))$ ", where N denotes the total sample size, k the total number of groups in the design and η^2 the G*POWER value. Thus $\eta_0^2 = 0.33306806 \cdot 108 / (108 - 36 + 0.33306806 \cdot 36) = 0.42828$, which is the value given in the SPSS output.

意见 5: P18 (1)“整个实验程序有 80 个实验试次”实际上本研究共包含 2 (善良程度) × 2 (效价) = 4 种条件, 那么每种条件只包含 10 个试次, 且还要剔除极端值, 这么少的试次如何保证结果的可靠性?

回应: 感谢专家的疑问。在本研究中, 实验一采用: 2 (组别: 高善良、低善良) * 2 (词语类型: 积极词、消极词) 的混合实验设计, 其中组别为被试间变量, 词语类型为被试内变量。本实验采用 Stroop 色词判断任务, 由于本实验的目的是比较高 (Vs.低) 善良被试对词汇颜色判断的反应时是否与词汇类型 (积极、消极) 具有关联。因此, 在实验刺激材料的设计上, 分别筛选出 10 个积极词和 10 个消极词, 为了平衡颜色的判断, 每个词汇分别呈现为 4 种颜色。由此, 组成 80 个实验试次。即积极词 40 个试次, 消极词 40 个试次, 每个被试都完成对积极词和消极词共 80 个试次的判断。

在研究文献中, 有些 Stroop 任务的判断任务试次较多, 如 Orem 和 Bedwell (2010) 的试次为 119 个试次, Lamers 和 Roelofs 的三个实验也分别有 72/72/60。但是这些实验任务中, 都蕴含相符类型这个自变量, 即相符、不相符、中性 (congruent, incongruent, and neutral), 研究者藉此进一步比较 Stroop 任务的一致效应与不一致效应。如果考虑到这个变量, 每个实验条件的试次为 40 次 (Orem & Bedwell, 2010)。Lamers 和 Roelof 实验的试次也需要除以 3。

本实验研究, 不探讨 Stroop 任务的一致效应与不一致效应, 采用 40 个实验试次。在已有实验中, 例如白学军, 刘湍丽和沈德立 (2014) 的采用 Stroop 任务研究认知抑制, 实验任务为 36 个试次。因此, 本实验的对积极词与消极词每类 40 个试次的设计是合理的。

参考材料

白学军, 刘湍丽, 沈德立. 部分线索效应的认知抑制过程: 情绪 Stroop 任务证据. *心理学报*, 2014, 46(02): 143-155.

Orem, D. M., & Bedwell, J. S. (2010). A preliminary investigation on the relationship between color-word stroop task performance and delusion-proneness in nonpsychiatric adults. *Psychiatry Research*, 175(1-2), 27

Orem 和 Bedwell (2010) 的文献摘录

2.2.2. The Stroop Task

The current study employed a three-condition (congruent, incongruent, and neutral) computerized Stroop task with four color choices—red, green, blue, and yellow. E-Prime software (Psychology Software Tools) was used to build and run the Stroop task. The neutral condition, during which a row of four “X”s appeared in one of the four colors, served as the baseline measure of reaction time of simply responding to the font color. During the congruent condition, a color-word appeared in matching, or “congruent,” font color (i.e. the word RED appears in a red font). During the incongruent condition, the color-word and the actual color of the font differ (i.e. the word “GREEN” appearing in a red font). A total of 119 trials were administered, with an approximately equal number of trials from each of the three conditions (neutral, congruent, and incongruent). All trials were presented in a fixed semi-random order of condition type, with the restriction that following an incongruent trial, the answer (font color) of the next trial was not that of the color-word that the participant just ignored during the preceding incongruent trial. This restriction was added to avoid a

(2) “剔除反应时低于 200 ms 和超过 1000 ms 的极端反应时试次，以及错误率超过 10% 的被试数据”：首先，该文所用的实验范式和 Lamers 和 Roelofs 的实验中范式并不完全相同；其次，Lamers 和 Roelofs 实验中不同颜色的词语呈现的时间是固定的（1500 ms），且注视点和空屏（1000 ms）呈现的时间也远长于本研究，因而不宜直接照搬该文中剔除极端值的标准，建议按照 $M \pm 3SD$ 的标准剔除极端值后重新分析数据；

回应：感谢专家的质疑。虽然本实验与 Lamers 和 Roelofs 在刺激呈现的时间设定上不完全一致，但是，整理的数据都是被试在判断色词颜色时的反应时，从统计分析上来讲，色词判断反应时应该与刺激呈现时间应是相分离的。从现有研究来看“剔除反应时低于 200 ms 和超过 1000 ms 的极端反应时试次，以及错误率超过 10% 的被试数据”是较为常用的方法，例如白学军，刘湍丽和沈德立（2014）的研究也采用这种方法。从本实验结果来看，这个标准也比较适合。

参考文献：

白学军,刘湍丽,沈德立.部分线索效应的认知抑制过程:情绪 Stroop 任务证据.《心理学报》, 2014,46(02):143-155.

(3) Lamers 他们这篇论文中一共包含了 3 个实验，3 个实验中每种条件下的试次数依次为 72/72/60，这也从侧面表明本研究中每种条件下的试次数太少了。

回应：感谢专家的严谨求证。参见对 5（1）的回应。

意见 6：P24“本实验选择三种目标-干扰格式的面孔表情组合矩阵（高兴-愤怒，高兴-中性和愤怒-中性）”分别在愤怒和中性背景下选用高兴作为目标刺激，只在中性背景下选用愤怒

作为目标刺激，这样设置的理由是什么？因为高兴作为目标刺激出现的次数比愤怒多，可能导致高兴目标更容易被识别。

回应：感谢专家全方位关注实验控制。本实验的目的主要是针对高、低两组被试的情绪调节与恢复的对比，对于两组来说实验试次的影响是相同；第二，从效果指标的设计来看，根据已有研究，从面孔矩阵中识别高兴面孔是效果指标。Koole 和 Jostmann（2004）认为识别出愤怒表情面孔群中的高兴面孔，和中性表情面孔群中的高兴面孔的反应时与情绪调节能力有关。他们认为要对表情面孔组合中的高兴面孔进行识别，需要将注意从表情面孔背景中转移到目标表情面孔上，也就是说把注意从负性信息中解离出来，对注意的解离越迅速，那么情绪调节能力就越强。Jostmann 和 Koole（2005）根据其研究结果进一步指出，被试对愤怒表情面孔背景中高兴面孔的识别反应时能够作为个体情绪调节能力高低的指标，并且此任务中的情绪调节加工主要是指内隐情绪调节；第三，刺激材料的试次设计和反应方式，能够较好平衡高兴面孔与愤怒面孔出现次数。面孔识别实验总共是 108 个试次，其中男性面孔 54 张，女性面孔 54 张。具体设计方式如下：男生面孔 54 个试次：包括差异面孔组合 27 张，高兴-愤怒，高兴-中性，愤怒-中性三种面孔组合各 9 张；相同表情面孔组合 27 张，高兴、愤怒、中性，各 9 张；女性面孔与男性面孔的设计方式一样。请被试判断“矩阵如果有不相同的面孔表情图片”，则按“Q”键反应，如果“矩阵中都是相同面孔表情图片”，则按“P”键。屏幕上首先呈现“+”注视点 2s，接着呈现面孔表情矩阵，且直到被试按键反应才会消失，之后会有 1.5s 的空屏，提示进入下一试次。第四，本研究与同类研究的实验材料及任务相同，例如，Quirin, Bode 和 Kuhl (2011)的研究。

Quirin, M., Bode, R. C., & Kuhl, J. (2011). Recovering from negative events by boosting implicit positive affect. *Cognition & Emotion*, 25(3), 559-570. 在这篇文献 562 页，介绍了实验刺激材料。

Face in the crowd task. To provide an additional measure of affect regulation, participants were presented with a series of 54 (plus 9 practice) matrices in random order with each matrix consisting of nine (3×3) schematic faces, which were taken from Öhman et al. (2001). The size of each individual face was 3.5×4.0 cm. Twenty-seven matrices, that is, half of the matrices presented, were composed of faces with one of three identical expressions (happy, angry, or neutral). The other matrices consisted of eight identical distractor faces and one target face with an expression different from the distractors. Since we used three target-distractor combinations (happy in angry, happy in neutral, and angry in neutral) and since in each combination the target appeared once at every one of the nine possible positions, the target matrices account for the remaining 27 (3 combination \times 9 positions) matrices presented to participants.

在 Koole, S.L., Jostmann, N.B. (2004). Getting a grip on your feelings: Effects of action orientation and external demands on intuitive affect regulation. *Journal of Personality and Social Psychology*, 87, 974-990, 这篇文献的 984 页, 也介绍了实验刺激材料的制作。

Dependent Measures

Face discrimination task. Participants were shown a series of matrices that were composed of nine schematic faces that were taken from Öhman et al. (2001). The size of each individual face was 3.5×4.0 cm. Each set of nine faces was arranged in a 3×3 matrix. Half of the matrices were composed of faces that all had the same emotional expression (i.e., happy, angry, or neutral). In the other matrices, one target face had a different emotional expression from that of the background distractors. The target could occur at any of nine positions in the matrix. Thus, there were 54 different matrices containing a target. In addition, there were 3 different distractor matrices without targets (happy, angry, or neutral), each of which was shown 18 times to create an equal number of distractor matrices. The instructions explained to the participants that their task was to detect discrepant faces among the crowds of distractor faces that appeared on the computer screen. Participants responded by pushing either a “one face is different” or an “all faces are the same” button on the keyboard and were asked to keep their hands near the buttons throughout the task. The matrices always appeared on the center of the screen and remained on screen until participants pressed one of the response keys. After making a response, the screen went blank for 1.5 s before the next matrix appeared. With 54 target matrices and 54 distractors, each participant was exposed to 108 randomly ordered matrices.

参考文献:

Jostmann, N.B., Koole, S.L., van der Wulp, N.Y., Fockenberg, D.A. (2005). Subliminal affect regulation: The moderating role of action vs. state orientation. *European Psychologist*, 10, 209-217

Koole, S.L., Jostmann, N.B. (2004). Getting a grip on your feelings: Effects of action orientation and external demands on intuitive affect regulation. *Journal of Personality and Social Psychology*, 87, 974-990

Quirin, M., Bode, R. C., & Kuhl, J. (2011). Recovering from negative events by boosting implicit positive affect. *Cognition & Emotion*, 25(3), 559-570.

意见 7: 讨论部分较单薄, 讨论的深度不够。例如, P28“控制型的内隐情绪调节与被试的健康及有意识地使用认知重评策略密切相关”这个观点说明什么呢? 与本文有何直接关系? 习惯使用认知重评策略的优点、对个体的积极影响有哪些呢?

回应: 非常感谢专家的质疑和建议。我们对总讨论部分进行了修改和完善, 重点讨论了人格特质与成功情绪调节的关系, 以及善良人格特质在成功内隐情绪调节中所发挥的作用及可能的机制。

新增参考文献如下:

Kokkonen, M., & Pulkkinen, L. (2001). Examination of the paths between personality, current mood, its evaluation, and emotion regulation. *European Journal of Personality*, 15(2), 83-104

Morawetz, C., Alexandrowicz, R. W., & Heekeren, H. R. (2016). Successful emotion regulation is predicted by amygdala activity and aspects of personality: a latent variable approach. *Emotion*, 17(3).

Paschke, L. M., Dörfel, D., Steimke, R., Trempler, I., Magrabi, A., & Ludwig, V. U., et al. (2016). Individual differences in self-reported self-control predict successful emotion regulation. *Social Cognitive and Affective Neuroscience*, 11(8), 1193-1204.

Steel, P., Schmidt, J., & Shultz, J. (2008). Refining the relationship between personality and subjective well-being. *Psychological Bulletin*, 134(1), 138-161.

细节问题

意见 1: 作者在回复审稿人 2 的第二个问题时, 提到“..... $t(29) = 10.14$, $p < 0.0001$ ”, p 值不对, P17 存在同样的问题。

回应: 感谢专家指出我们的错误。对精确 p 值的表达格式进行了修改。

意见 2: 参考文献

(1) “其中Thman, Lundqvist 和 Esteves (2001) 提出的表情面孔搜索范式获得了较多的研究支持, 他们的研究表明.....并且不受目标以及个体的意向控制 (徐展, 李灿举, 2014)”一句中的文献引用不当, 直接引用一手文献即可, 徐展的文献应删掉。

回应: 感谢专家的指正, 删除了该文献。

(2) P19“从信息加工的特点来看, 情绪意义偏向将会调节对应于 EPN 和 LPP 波的激活增强, 并且还影响右半球的 N170 波的激活幅度 (Frühholz, Jellinghaus, & Herrmann, 2011)”与本研究并无直接关系, 建议删掉。

回应: 感谢专家的指正, 删除了该引用文献。

(3) P23“(Greenwald et al., 2003; Cai, Sriram, Greenwald, & McFarland, 2004)”两篇文献的顺序应互换。

回应: 感谢指正, 调整了顺序。

(4) P23 文献引用错误: “负性情绪诱发材料: 采用阅读法诱发被试的负性情绪”应给出参考文献。

回应: 感谢指正, 补充了参考文献. Pu, Z., Liu, C. H., & Yu, G. L. (2012). An overview of mood-induction methods. *Advances in Psychological Science*, 20(1), 45-55.[郑璞, 刘聪慧, & 俞国良. (2012). 情绪诱发方法述评. *心理科学进展*, 20(1), 45-55.]

(5) P27“Quirin 和 Lane (2012) 指出, 内隐积极情绪的增强可以提高对愤怒面孔群中的高兴面孔的识别能力, 他们认为个体在体验到情绪唤醒的恢复阶段中, 其内隐积极情绪的增加与内隐情绪调节加工有关”该文并没有这个观点。

回应: 非常感谢审稿专家非常严谨的审稿。由于我们的失误, 这个观点的参考文献的引用出现了错误。正确的引用文献为: Quirin, M., Bode, R. C., & Kuhl, J. (2011). Recovering from negative events by boosting implicit positive affect. *Cognition & Emotion*, 25(3), 559-570.

我们引用的这个结论性观点在本文 567 页, 原文如下。

and a delayed increase in implicit PA. This increase in implicit PA was related to a measure that has previously been found to be associated with affect regulation, namely detection of a happy face in an angry crowd (Study 1). Finally,

意见 3: 用语的一致性。例如, P16“高兴面孔”/“高兴面孔”、“积极的情绪调节控制性内隐态度”/“控制型的积极情绪调节内隐态度”、“消极的情绪调节表达性内隐态度”/“表达型的消极

情绪调节内隐态”。

回应：感谢专家严谨细致的指正。我们进行了修改。统一为：高兴面孔、控制型的情绪调节内隐态度，表达型的情绪调节内隐态度。

意见 4：用语准确性。P17“假设（3）：诱发负性情绪后，高善良被试从愤怒面孔背景中搜索高兴面孔的反应时更快”，反应时应用长/短来形容，快/慢——反应速度。P27“搜索高兴面孔任务的反应时显著快于低善良特质者”存在同样的问题。

回应：感谢专家严谨细致的指正。我们进行了修改。

意见 5：图 3 中词语对应的方框后面应该为“按键消失”，而不是“F 键”、“J 键”。

回应：感谢专家的指正，我们修改了流程图。修改结果如下。

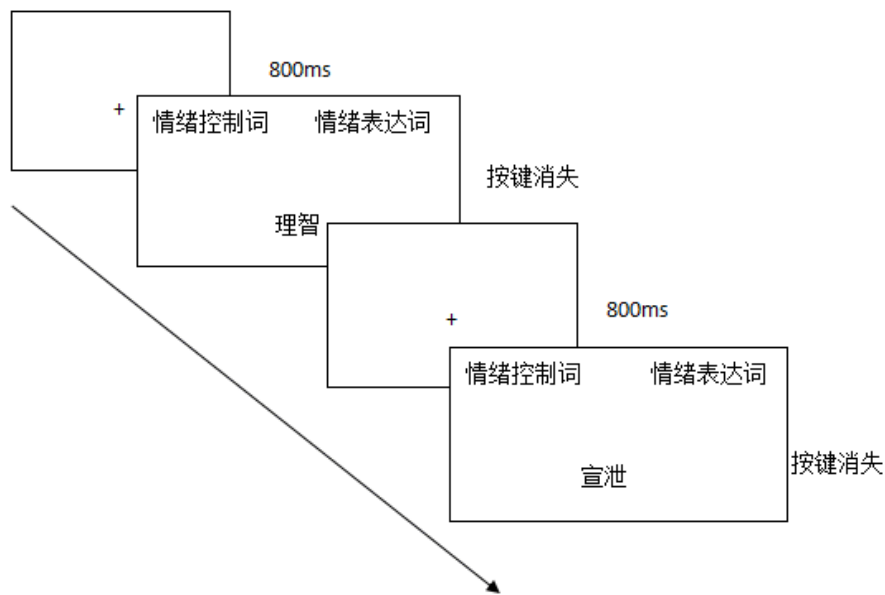


图 3 情绪调节版内隐联想测验（ER-IAT）程序的部分示例

最后再次感谢两位审稿专家严谨、细致地审稿，很好地指导了我们修改稿件，并对我们今后从事更加科学规范的心理实验研究提供了重要帮助。

第三轮

审稿人 2 意见：作者已进行了修改和回复，在完善以下细节问题后同意发表。

非常感谢审稿专家细致地审稿，并提出宝贵的修改意见！

备注：第三轮修改或补充文字，采用紫色字。

意见 1：正文中引用的文献与文后的文献列表不完全一致。例如，Gross 和 Thompson（2007）和 Gyurak, Gross 和 Etkin（2011）在文后的文献列表中没有列出；

回应：感谢专家细致审校。补充文献如下：

Gross, J. , & Thompson, R. . (2007). Emotion regulation: conceptual foundations.In. J. J. Gross Emotion Regulation Handbook, 21(3), 431-441.

Gyurak, A., Gross, J. J., & Etkin, A. (2011). Explicit and implicit emotion regulation: A dual-process framework. *Cognition and Emotion*, 25(3), 400-412.

意见 2：正文中的文献引用有误。例如，刘文文，江琦，任晶晶，李树芳，徐雅珮（2016）。

回应：感谢专家细致审校，修改了相关引用错误。

正文：

刘文文，江琦，任晶晶，李树芳，徐雅珮（2015）；

面孔表情视觉搜索任务参照 Öhman, Lundqvist 和 Esteves（2001）的设计程序；

视觉搜索任务中存在愤怒优势效应（Schmidt-Daffy, & Martin., 2011; Öhman, Juth, & Lundqvist, 2010; Lipp, Price, & Tellegen, 2009）。

参考文献：

Liu, W., Jiang, Q., Ren, J., L, S., & X, Y. (2015). The Impact of Trait Anger on Aggressive Behavior: Moderated Mediating Effect of Hostile Cognition and Impulsivity Level. *Psychological Development and Education*, 31(4), 485–493.

Webb, T. L. , Gallo, I. S. , Miles, E., Gollwitzer, P. M. , & Sheeran, P. (2012). Effective regulation of affect: an action control perspective on emotion regulation. *European Review of Social Psychology*, 23(1), 143–186.

意见 3：文中还存在个别语句不完整通顺的问题。例如，1.1 内隐情绪调节与行动控制理论部分，把内隐情绪调节（implicit emotion regulation）定义为“由刺激自动激活，并且调节的执行与完结不需要意识监控，也不需要觉察与领悟”，这句缺少“的调节加工过程”。

回应：感谢专家细致审校，修改了个别语句，见正文。

意见 4：英文错误。例如，“1.2.2 善良特质优化内隐控制型情绪调节态度”部分，外显-自动（implicit-automatic）和内隐自动（explicit-automatic）出现错误。

回应：感谢专家细致审校，修改为：外显-自动（explicit-automatic）和内隐自动（implicit-automatic）。

意见 5：讨论部分有些地方逻辑性不够强。例如，“5.2 善良特质促进控制型内隐情绪调节”部分的第一段。

回应：感谢专家细致审校，5.2 部分的讨论修改为：

5.2 善良特质促进控制型内隐情绪调节，增进健康

本研究发现，高善良特质被试更倾向于控制型情绪调节内隐态度，即他们更倾向于适度控制情绪而不是基于个人喜好宣泄情绪。Sperduti 等（2017）研究了内隐情绪调节与执行功能的关系，发现内隐情绪调节与工作记忆的信息更新能力（updating abilities）这一复杂执行功能指标相关联。从这个意义上来说，内隐控制具有实现的可能性。同时，研究还表明，控制型相比宣泄型内隐调节态度更有利于个体健康。例如，Mauss 等（2006）发现，控制型情绪调节内隐态度与成功的、自动的和生理适应性的愤怒情绪减弱调节显著相关；而且，从整体适应上来看，这种调节态度与感情性反应的适应性显著相关，即对情绪控制的内隐评估越积极，个体在自动的、成功的、生理层面上的情绪调节的适应性越高（Mauss et al., 2010）。刘俊升和桑标（2009）研究结果也表明，在负性情绪诱发情境中，对情绪调节持积极内隐控制型调节态度的个体，会自动化的调节和控制自己的情绪。

因此，可推论善良特质通过控制型情绪调节内隐态度促进健康。这一研究结果与以往研究具有一致性，王登峰和崔红（2007）的研究表明，善良人格间接负向预测心身症状，即善良特质与身心健康相关联。控制型内隐调节与健康的关联性也已经获得实证证据，如 Hopp 等人（2011）采用情绪调节内隐联想测验（ER-IAT）进行的研究也表明，控制型的内隐情绪调节与被试的健康及有意识地使用认知重评策略调节情绪密切相关。因此，实验 2 的结果表明，善良特质通过促进控制型的情绪调节态度优化内隐情绪调节，进而影响身心健康。