

# 《心理学报》审稿意见与作者回应

题目：中国情绪调节词语库的初步编制与试用

作者：袁加锦 张祎程 陈圣栋 罗利 茹怡珊

## 第一轮

### 审稿人 1 意见：

**意见 1：**从实验一的统计方法来讲，作者在选择情绪调节词语时候用了相关的方法，所选的词语与被试理解的几种策略高度排他性相关则入选。但是，这些词语与某类情绪调节策略的意思相关却并不代表这些词语可以诱发出被试相应的情绪调节策略，或者说作者并没有证明某类策略的词语就真的诱发了该类的情绪调节策略，有很大的可能性这些所选的词语只是语义上与这些情绪调节策略的语义理解相关，而并不是该词语诱发了该情绪调节的策略。

**回应：**通过查阅文献可知，以往许多研究已证明形成情绪调节词语的语义表征，可以内隐地启动情绪调节目标从而达到情绪调节的目的。比如，以句子组织任务为例，对接受目标的内隐启动可以成功调节个体的主观情绪体验与情绪相关的生理活动（如脑电与心率，Yuan et al., 2019; Ding et al., 2015）；对一般性情绪控制的内隐目标启动可以降低主观负面情绪体验（Mauss et al., 2007）；对注意分散和认知重评的内隐目标启动可以降低应激引起的主观情绪或情绪相关的生理活动（Williams et al., 2009; Yuan et al., 2015; 李红, 袁加锦, 2018）。

在本研究中，情绪调节启动组与控制组在其他情绪有关的个体差异因素上得到了匹配，所接受的实验处理上唯一的差别来自于启动目标的不同。然而，注意分散、表达抑制与认知重评三个情绪调节启动组相比控制组（即中性启动组）其主观情绪评价均出现显著下降，可以证明在本研究中被试的情绪调节效果是来自于情绪调节启动的作用。

然而需要承认的是，如果纳入操作有效性检验的程序，比如内隐联想测验和残句补全等方式，以检验被试经过启动以后对情绪调节目标的表征确实建立起来了会更加严谨。本研究未纳入该检验程序，已在讨论部分承认其局限，并将此作为未来研究方向（p.24; line 23-24）。

**意见 2：**从研究结果来讲，作者发现认知重评和表达抑制的效果要明显于其他几个情绪调节的策略，我们知道 Gross 的认知重评和表达抑制是比较有名的情绪调节策略，那其他的几个注意分散，接受策略，表情宣泄，作者能否解释的更加清楚一些在文章中，或注明引用出处，或在讨论部分重点讨论这几个策略的结果差异。

**回应：**非常感谢审稿专家的意见。本文在前言部分完善了五种情绪调节策略的解释和定义，在讨论部分重新阐述了几种情绪调节策略的结果差异，并注明了引用出处，请您审阅（具体见 p.5, line 3-13）。

**意见 3：**实验二中，作者从负性情绪中选取了厌恶图片作为负性情绪的诱发，缘由是什么呢？厌恶情绪最具代表性？文章提到“由于不同情绪调节策略对不同情绪强度的调节效果不同”请注明出处。

**回应：**首先，本研究选择以血腥场景为代表的厌恶刺激作为负面情绪材料，是由于先前研究（如 Chen et al., 2018; Ma et al., 2019）发现血腥图片可以很稳定的诱发厌恶情绪，而厌恶情

绪本身也是令人不愉快的负面情绪的代表(Rozin, Haidt, & McCauley, 2000); 有证据表明厌恶情绪会诱发负性情绪的典型脑区 (黄好, 罗禹, 冯廷勇, & 李红, 2010), 因而厌恶情绪的调节对认知神经科学研究有着重要意义。

此外, 以往有研究表明情绪调节的效果会因为情绪类别和情绪强度而有所不同 (Shafir, Thiruchselvam, Suri, Gross, & Sheppes, 2016; 李红 & 袁加锦, 2018)。我们选择以厌恶刺激为对象的另一个考虑是避免情绪调节的效果随着情绪类别或强度而发生变化(见修改后的讨论 p.14, line 11-14)。但是, 近以厌恶材料为调节对象也可能导致情绪调节效果的代表性不足, 例如本研究中我们发现接受目标启动未能对厌恶图片产生情绪调节效果。对于这一局限我们在讨论部分进行了相应的阐述 (见 p14, line11-16)。

“由于不同情绪调节策略对不同情绪强度的调节效果不同”此处已注明出处, 请您审阅 (见 p.10, line 22)。

**意见 4:** 实验二中, 作者选取了负性情绪图片分为高低唤醒度的情绪图片, 而中性图片却并没有唤醒度。负性或中性图片只是情绪的效价, 而这些图片都会有唤醒度之间的差异。

**回应:** 已有基于情绪维度理论的研究证明效价和唤醒是定义情绪刺激的最核心的两个维度, 且情绪刺激的效价维度与唤醒度之间存在着倒 U 型的相关关系 (Lang et al.,1998; Bai et al.,2005)。与此相一致, 本研究也发现随着负性刺激情绪强度的增强, 刺激的效价更低 (更加负性) 而唤醒度更高; 另外情绪材料的效价低于中性刺激而唤醒度高于中性刺激。我们已在文中补充高低强度的情绪图片和中性图片的效价与唤醒度。

**意见 5:** 本研究旨在建立一个能够满足内隐情绪调节需要的、具有多个属性维度的、标准化的词语库。但总体来说, 词语库每个策略可以选取作为实验刺激的词汇量不足, 还需要完善。

**回应:** 在制作情绪调节词语库的第一阶段, 我们尽量均等地考虑了每一种调节策略相应词语的数量 (情绪调节词语数 Min=32; 情绪调节词语数 Max =43)。然而, 由于不同情绪调节目标相应词语的丰富度不同, 尤其是所纳入的情绪调节词语需要具有情绪调节策略的特异性, 在开发情绪调节词语库的过程中我们确实发现某些策略符合纳入标准的词语数量较少, 如注意分散和接受。对于这一问题, 需要随着语言丰富度的进一步发展留待以后解决。

本研究建立的词语库最终包含了符合标准的接受策略 19 词, 注意分散策略 13 词, 表情宣泄 30 词, 表达抑制 30 词, 认知重评 25 词, 中性 32 词。通过查阅国内外隐情绪调节的文献发现, 本词语库能基本满足使用句子整理范式进行内隐情绪调节研究的材料要求。

.....

**审稿人 2 意见:**

**意见 1:** 研究方法: 句子整理范式是否具有代表性? 可否推广到其他任务范式? 是否是最具代表性的任务? 依据是什么?

**回应:** 在内隐情绪调节中有多种研究范式, 句子整理范式是其中一种常见的研究范式。以 "implicit emotion regulation" or "unconscious emotion regulation" or "automatic emotion regulation"为关键词在 web of science 上检索得到相关文献, 在结果中检索使用关键词启动被试情绪调节目标的研究, 发现大部分研究均采用句子整理范式来启动内隐情绪调节(Ajaya, Peckham, & Johnson, 2016; Ding, Yang, Liu, & Yuan, 2015; Mauss, Cook, & Gross, 2007; Williams, Bargh, Nocera, & Gray, 2009; Yuan, Ding, Liu, & Yang, 2015), 也有部分研究使用词语配对任务(Chen et al., 2017; Liu, Wang, & Li, 2018; Wang & Li, 2017; Yang, Tang, Gu, Luo, & Luo, 2015)。

**意见 2:** 研究一: 男女比例统计上是否有差异。年龄范围跨度较大, 被试如何分布? 尤其 17 岁还是未成年, 低年龄段和高年龄段之间是否有差异?

**回应:** 在研究一中, 男女分别为 57 人和 71 人, 经过卡方检验  $p=0.453$ , 男女比例无显著差异。对于年龄范围, 本研究年龄范围如下图, 虽然年龄范围较广, 但是 90% 以上的人群年龄分布在 18-23 岁之间, 仅有一人为 17 岁。

**意见 3:** 1.1.2 实验材料的筛选采用三个标准, 其中熟悉度怎么定义? 为什么不用词频进行控制? 问卷二进行 6 个维度的评定, 这 6 个维度的选取依据是什么?

**回应:**

- 1、在三个标准中, 熟悉度的定义为阅读该词时感觉熟悉的程度。
- 2、由于汉字的特殊性, 不能像英文单词一样简便地对词频进行统计, 在统计汉字词频时, 分词是一个困难的工作, 因此目前可查的官方资料包含《现代汉语频率词典》(1986)、《现代汉语常用词词频词典》(1990)。但由于其为三十余年前的资料, 而现代人语言习惯早有不同。作者也查阅了“语料库在线”等网站资料, 发现大部分四字词语在其中的频率 $<0.0005$ , 出现次数 $<50$ , 参考价值有限, 而且国内的词语库也未将词频纳入其指标。故综合考虑, 本文采取了熟悉度作为相应的标准。
- 3、六个维度的选取是基于外国理论和本国实际确定的, 例如, 在英文情绪词库中包含了愉悦度、唤醒度、优势度三个指标; 在汉语情感词系统中愉悦度、唤醒度、优势度、趋向度、熟悉度五个指标被用到, 本研究在五个指标基础上, 考虑到某些四字词的笔画数差异较大, 故将其拼写复杂度也作为一个维度。

**意见 4:** 2.1.2“一般性的厌恶图片”是什么? IAPS 和 CAPS 本身有评分, 为什么用自评的? 可靠性有多少?

**回应:** 非常感谢审稿专家的意见。

- 1、本文修改了“一般性的厌恶图片”的说法, 将其更改为“典型的厌恶图片(包含血腥内容)”。
- 2、由于我们的图片包含了来自网络的图片, 需要在一个研究中以同一批被试来对图片的情绪属性进行评价; 如果来自网络的图片是我们自评而其他图片直接使用 IAPS/CAPS 的数据, 则两批数据可能缺乏可比性, 因此也就无法用同一个标准来进行分类和评价。
- 3、IAPS 的刺激应用于中国被试已被证明存在文化差异(黄宇霞, 罗跃嘉, 2005), 而在同一个研究中 CAPS 与 IAPS 的刺激材料其情绪属性数据也需要经过同一批被试的评定, 其数据才具有可比性。为了验证我们所评定数据的准确性, 我们计算了不包含文化差异等混淆因素的 CAPS 图片的数据库得分与自评得分之间的相关系数, 结果发现 CAPS 原始数据与我们的自评数据之间的相关性为效应  $r=0.536^*$ , 唤醒度  $r=0.50^*$ (Pearson 相关,  $*p<.05$ ), 两组数据之间的高相关性佐证了我们的自评数据具有较高的可靠性。

**意见 5:** 2.1.2 情绪调解组为什么选 10 个情绪词和 5 个中性词? 需要内部比较? 或者其他原因。

**回应:** 非常感谢审稿专家的意见。为了避免被试通过连续地加工情绪调节词汇而猜测到实验的目的, 本研究没有全部选择情绪调节词汇, 而是在情绪调节词汇中穿插中性词汇(Mauss et al., 2007; Williams et al., 2009; Yuan et al., 2015; Ding et al., 2015)。

---

## 第二轮

审稿人 1 意见：

**意见 1：**作者在引言部分介绍了外显与内隐的两种情绪调节的方式，意见一作者的回应中强调情绪调节词语的语义表征是可以内隐地启动情绪调节的目标，既然词语库是作为内隐情绪调节的启动，题目叫内隐情绪调节词语库是不是更妥当？

**回应：**非常感谢审稿专家的意见。本文如此命名的原因是词语库本身的内容就是情绪调节词语。本词语库可用于内隐情绪调节，但不限于内隐情绪调节，也可用于其他需要用到情绪调节词语的外显情绪调节的研究中。例如，在外显情绪调节特别是涉及多策略转换的研究中，需要在实验过程中用一个词语提醒被试当前试次使用的或被试选择的情绪调节方法（e.g. Shafir, Schwartz, Blechert and Sheppes, 2015）。另外，通过构建执行意图实现外显情绪调节目标的研究，也通常需要使用表达特定情绪调节意义的词语来建立情境与反应方式的关联(e.g. Gallo et al.,2009; Ma et al., 2019)，例如“如果我遭遇失败，我将以顺其自然的方式看待它”。综上可知，情绪调节词语也可以用于外显情绪调节的研究中，因而我们保留当前文章的命名方式。

在修改稿的讨论部分（p.15, paragraph 3），我们进一步阐述了情绪调节词语库在内隐和外显情绪调节中的用途，具体如下：

“需要指出的是，虽然情绪调节词语库的典型用途是有效启动内隐情绪调节，但它也可以用于外显情绪调节目标的实现。例如，在情绪调节的实验流程中，研究者经常需要使用特定情绪调节目标相关的词语来提醒被试当前需要使用的情绪调节策略，这一点尤其在涉及情绪调节策略选择的研究中更为明显（Shafir, Schwartz, Blechert and Sheppes, 2015）。另外，在使用执行意图实现外显情绪调节目标的研究，研究者通常需要使用表达特定情绪调节意义的词语来建立情境与反应方式的关联，例如“如果我看到蜘蛛，我将保持平静与放松”（Gallo et al.,2009; Ma et al., 2019）。因此，建立标准化的情绪调节词语库可以从多层面促进情绪调节的未来研究。”

**意见 2：**属性分析的结果与中性比较，各种策略都不一致，说明什么问题？这个属性分析的目的是什么？雷达图是说明什么？讨论部分有讨论属性的差异吗？

**回应：**本研究将情绪调节策略词语的各种属性与中性相比较，是为了说明各种情绪调节策略词语在主要的情绪属性与语言属性上的特点。前人研究表明词语本身的愉悦度、唤醒度等属

性也会对被试造成情绪、记忆和注意等方面的干扰 (Aquino & Arnell, 2007; Smith, 2012); 因此属性分析可以帮助我们更好地了解各种策略相关情绪调节词语的属性。当前, 研究者在制作情绪词语材料库的时候主要考虑效价 (愉悦度), 唤醒度 (激动度), 优势度 (受控-支配度), 动机 (回避-趋近倾向) 与熟悉度这五个情绪维度 (Bradley & Lang, 1999; 王一牛, 周立明, 罗跃嘉, 2008)。此外, 考虑到词语之间的笔画数差异较大, 故将拼写复杂度也作为一个维度纳入考虑。因此, 本研究基于这六个维度构建了雷达图, 从而更直观地呈现各种策略的属性情况; 本研究在讨论部分讨论了各种策略的属性情况, 探讨了不同性别对各种策略的情感偏好。

从属性分析的结果上看, 不同情绪调节策略词语相互之间的属性差异较大, 这反映了不同策略其情绪调节词语具有不同的情绪与语言属性 (比如接纳具有积极情绪属性而表达抑制具有消极情绪属性)。由于本词语库中每个情绪调节词在六个属性维度都具有相应的属性值, 未来涉及多种调节策略对比的研究在选择启动词语的时候, 也可以考虑匹配不同条件下情绪调节词的情绪与语言属性。

**意见 3:** 建议作者考虑做悲伤等负性情绪的调节的验证, 内隐情绪调节用于改善抑郁、焦虑等个体的负面情绪或许更加有意义。

**意见 4:** 建议作者在建立和检验词语库的时候, 不仅用到负性和中性情绪刺激, 也应该考虑正性情绪的调节, 过激的正性情绪也应该属于调节的范畴, 这样才能全面考察内隐情绪词语库在不同情绪效价下的调节效应。

**回应:** 确如审稿专家所说, 内隐情绪调节用于改善抑郁、焦虑等个体的负面情绪的现实意义更为重要。

在本研究团队的前期研究中, 我们已使用情绪调节词语探讨了内隐注意分散对于抑郁人群的情绪调节效果。其结果显示内隐和外显情绪调节都可以降低健康被试的负面情绪状态; 另一方面, 尽管内隐和外显注意分散都可以减少抑郁人群的负面情绪强度, 但仅有使用情绪调节词语的内隐注意分散可以改善抑郁人群的积极情绪状态, 从而初步证明了通过情绪调节词语启动内隐情绪调节干预抑郁的有效性 (李红 & 袁加锦, 2018)。

在本团队的其他相关研究中, 我们也基于健康被试验证了情绪调节词语内隐启动接受, 认知重评或表达抑制调节负面情绪的有效性 (Zhang et al., 2020; Ding et al., 2015; Yuan et al., 2019, 2015)。

然而，上述前期研究更多关注的是对于健康人群的情绪调节，而较少关注通过情绪调节词语启动抑郁，焦虑等情绪障碍人群的情绪调节。尽管我们的前期工作已证实情绪调节词启动内隐注意分散对于抑郁群体的情绪调节效果，但尚未全方位证实其他策略的情绪调节词（如认知重评，接受等）能否成功实现抑郁等情绪障碍人群的情绪调节。

因此，我们在修改稿的局限与未来研究部分，指出了“研究二仅采用了厌恶情绪作为负面情绪的代表进行了验证性研究，未来研究需要以更具现实意义的负面情绪（如焦虑、抑郁）或者积极情绪为对象，进一步验证本词语库启动情绪障碍人群相应情绪调节的有效性。”（p15）。

**意见 5：**内隐情绪调节和自动化情绪调节是一个意思吗？请规范词语，以免读者混淆。

**回应：**非常感谢审稿专家的提醒。两个概念存在着一定的区别。在上一轮修改中所涉及的自动化情绪调节均特指内隐情绪调节。为了避免歧义，在本轮修改中我们已统一规范命名为内隐情绪调节。

.....

**审稿人 2 意见：**

**意见 1：**第 2 页第 24 行“Gross 的情绪调节模型(James J Gross, 2013)”前面没有介绍这个模型；而且这个模型有五个维度，为什么只选取了三个？前文对情绪调节理论介绍过少。

**回应：**非常感谢审稿专家的意见。我们已在引言第一段丰富了情绪调节及其模型的介绍；根据 Gross 的情绪调节过程模型，情绪调节过程包括五个阶段，分别是：情境选择、情境修正、注意分配、认知改变和反应调节。情绪调节过程的前两个阶段（情境选择和情境修正）均特指旨在改变情境的情绪调节，该两种方法不属于通过认知方法实现情绪调节的范畴，因此也没有涉及具体的认知策略。所以，前两种策略不在我们通过认知方法调节情绪的范围之内，故只选取了后三个。

在修改稿中，我们已经在文章开篇对情绪调节过程模型以及为何关注后三个阶段的调节策略予以了清晰说明，具体如下（p3）：

“根据 Gross 等人(2013)提出的情绪调节模型（见图 1），情绪调节过程包括五个阶段，分别是情境选择（Situation selection）、情境修正（Situation modification）、注意分配（Attentional deployment）、认知改变（Cognitive change）和反应调整（Response modulation）。每一个

情绪调节阶段都有对应的情绪调节策略。其中，情境选择，情境修正，注意分配和认知改变属于先行关注策略（Antecedent-focused strategies），反应调整属于反应关注策略

（Response-focused strategies）。另一方面，情境选择与情境修正均是通过改变情境从而改变情绪的策略，因此不属于认知策略，即改变自身的注意、认知方式与行为实现情绪调节的策略范畴。”

意见 2：第 4 页第 4 行，请说明取样方式？是否排除了有心理疾病或精神疾病的被试，所有参与实验的被试都计入统计了吗。样本数量对编制问卷来说，似乎过少。

意见 3：被试量过少。

回应：非常感谢审稿专家的意见。研究二取样方式为发布招募信息进行招募，经过状态焦虑量表(STAI-state)(Spielberger, 1970)、特质焦虑量表(STAI-trait) (Spielberger, 1970)、贝克抑郁量表(Beck depression rating scale)(Beck & Beamesderfer, 1974)的筛选，均未发现有焦虑症、抑郁症患者。

本研究的研究 1 部分为建立情绪调节词语库。具体而言，本研究要求被试在实验室情境下评价备选情绪调节词语的情绪属性（效价，唤醒，优势度，动机），语言属性（熟悉度，拼写复杂度）和情绪调节属性（5 种调节策略），从而制作出符合情绪调节目标的情绪调节词语库。因此，本研究所使用的并不是传统意义上的问卷法，而是问卷法与实验法的结合。

关于样本量的问题，我们在开展本研究之前参照过王一牛等人（2008）编制“汉语情感词系统”（CAWS; N=124），徐舒靖等人(2008)编制“情绪障碍研究用汉语情绪词分类系统（N=54）；白露等人编制中国情绪图片系统（CAPS, N=46）以及 Lang, Bradley, & Cuthbert（1997）制作国际情绪图片系统（IAPS，未明确说明 N；只说 subjects were college students taking an Introductory Psychology class and participating as part of a course requirement）所使用的样本量。我们使用了略高于类似情绪数据库制作的样本量标准，来确定本研究的样本量（N=128）。

为了回应审稿人关于样本量的意见，我们根据研究 1 的设计通过 gpower 软件计算了事后实际获得的统计力度。根据当前设计，每个被试（N=128）需要评价 5 种情绪调节策略相关度（分心，重评，抑制，宣泄，接受）和 6 个情绪维度（效价，唤醒，优势度，动机，熟悉度，复杂度）共 30 个维度，结合  $p < 0.05$  作为显著性阈限和实际获得的最低效应量  $f = 0.25$ ，事后得到的统计力度值是  $\text{power}(1 - \beta \text{err prob}) = 0.999$ ；在进一步将性别因素（考虑到后续分析包含了性别差异）纳入计算以后， $\text{power}(1 - \beta \text{err prob})$  依然=0.999。这些结果提示基于

当前样本量的统计力度是充足的。

事实上，从我们事后所获得的主要统计指标来看，进入情绪调节词语库的词语具有较高的重测信度（ $r=0.703$ ， $p<0.001$ ）和评分者一致性系数（ $w=0.276$ ， $p<0.001$ ），这表明参加评定的 128 名被试具有较为一致的评价标准，提示我们所获得情绪词语库具有较高的可信度。

但无论如何，我们必须承认一个情绪调节材料库的可靠性，有效性，必然需要未来研究在更多样本量的基础上予以检验和完善。因此，我们在修改稿的讨论部分对样本量偏小的局限和未来研究方向进行了讨论和补充，具体如下（p15）：

“本研究也存在一些局限，需要未来研究进一步解决。首先，研究一被试量偏少（ $N=128$ ），情绪调节词语数据库的有效性需要在未来研究中，通过样本量的积累进行进一步验证和完善。第二，研究二仅采用了厌恶情绪作为负面情绪的代表在健康人群中进行了验证性研究，未来研究需要以更具现实意义的负面情绪（如焦虑、抑郁）或者积极情绪为对象，进一步验证本词语库启动情绪障碍人群相应情绪调节的有效性。第三，有研究表明，内隐情绪调节策略对被试生理反应的调节更为有效(Williams et al., 2009; Yuan et al., 2015)，本词库中内隐情绪调节策略对生理反应的调节效果还未得到验证，需在今后进一步完善。”

**意见 4：**第 4 页第 16 行，前文没有提过中性词，可在第 9 行补充；没有说明中性词的作用。

**回应：**非常感谢审稿专家的意见。已在相关位置提及和补充了中性词的作用，具体如下（p6）：“从中选取与五种情绪调节策略相关的双字词语和四字成语以及与情绪调节策略无关的中性词。纳入中性词的目的主要有两点：一、建立可以与情绪调节词语相对照的，标准化的中性词语材料；二、在情绪调节词汇间起缓冲作用，从而避免被试因连续地加工情绪调节词汇，猜测到实验的目的。”

**意见 5：**第 4 页第 21 行，为什么选这六个属性，没有说明。

**回应：**非常感谢审稿专家的意见。选择六个属性的原因已在 2.1.1 实验材料部分进行了说明，具体如下（p7）：

“六个维度的选取是基于已有理论和中文语言的实际情况。例如，英文情绪词库（ANEW）包含了愉悦度、唤醒度、优势度三个指标（Bradley & Lang, 1999）；汉语情感词系统（CAWS）



包括了愉悦度、唤醒度、优势度、趋向度、熟悉度五个指标（王一牛，周立明，罗跃嘉，2008）。另外，本研究在五个指标基础上，考虑到某些四字词的笔画数差异较大，故将其拼写复杂度也作为一个维度。”

**意见 6：**统计，为什么不用因素分析？

**回应：**由于本研究并非严格意义上的问卷编制，待评价的 5 个情绪调节维度和反映词语信息的六个属性维度均是根据研究目的和已有情绪理论（情绪维度理论）确定。所以用因素分析的方法提取主成分与本研究目的没有直接关系。

**意见 7：**第 9 页第 2 行，没有给出三组统计的 F 值和 p 值。

**回应：**负性材料和中性材料的 F 检验结果已经补充。需要指出的是，尽管在评价过程中我们纳入强度分布很广的负性图片（N=120），为了控制刺激的情绪强度对结果的影响，我们纳入研究 2 的负性刺激仅包括一组中等强度的负性情绪图片（N=15）。在本轮修改稿中，我们对这一点进行了明晰。

具体修改如下（p12，paragraph 1）：

“根据评定结果随机选取了 15 张中等强度的负性情绪图片（唤醒度为  $6.16 \pm 0.61$ ，效价为  $2.55 \pm 0.38$ ）作为负性情绪诱发材料，15 张中性图片（唤醒度为  $2.14 \pm 0.21$ ，效价为  $4.92 \pm 0.42$ ）作为中性对照材料。所选取的负性图片相比中性图片在唤醒度（ $F(1,28) = 583.08, p < 0.001$ ）和效价（ $F(1,28) = 265.25, p < 0.001$ ）评分上差异都显著”。

**意见 8：**第 9 页第 9 行，“情绪调节启动句子”句子是不是也要先评定，控制情绪效价、唤醒度？

**回应：**非常感谢审稿专家的建设性意见。尽管在词语库制作的过程中我们测试了每个词语材料的属性，但词语的情绪属性的确并不等于句子的情绪属性。因此，我们非常赞同审稿专家的意见，并认为应该在以后类似的研究中测试和匹配句子的情绪属性。

在本研究之中，我们的确没有事前单独测试每个情绪条件下所使用句子的效价和唤醒。为了回应审稿专家的意见，我们另外邀请了 2 名（1 男 1 女，年龄：34/29）对实验目的不知情的评分者对研究 2 所涉及的情绪调节启动句（6 种策略（分心，重评，抑制，宣泄，接受，中性）\* 每个条件 15 个句子）进行了情绪效价与唤醒度的评定。两名被试的评分一致性程度较高（ $w = 0.360, p < 0.001$ ），表明评分结果较为可信。

随后将两名评分者对每个句子的效价与唤醒度评分分别进行平均，最后纳入多变量方差分析。结果显示情绪调节策略在效价上的主效应不显著（ $F(5, 84) = 2.25, p > 0.05$ ）但在唤醒度上具有显著的调节策略主效应（ $F(5, 84) = 6.11, p < 0.001$ ）。对唤醒度的调节策略主效应进行事后多重比较分析发现，注意分散、表情宣泄、接受、重评与抑制 5 种策略之间的唤醒度差异不显著（ $p_{\min} = 0.17$ ），但五种情绪调节句子的唤醒度均高于中性句子（all  $p < 0.001$ ）。

这些结果提示在研究 2 中，6 个实验条件所使用情绪调节启动句不存在情绪效价的显著差异；尽管各调节策略的启动句之间不存在唤醒度的显著差异，但调节策略启动句的唤醒度显著高于中性句子。

在修改稿中，我们在研究 2 方法部分增加了对启动句情绪属性检验结果的简要介绍，具体如下（p12, paragraph 2）：

“本研究邀请了两名对实验目的不知情的评分者对 6 种情绪调节启动句进行了效价与唤醒度的检验，结果显示 6 种情绪调节启动句在效价评定上不存在显著差异（ $F_{5, 84} = 2.25, p > 0.05$ ；但在唤醒度上存在显著的调节策略主效应（ $F_{5, 84} = 6.11, p < 0.001$ ）。对唤醒度上的调节策略主效应进行事后多重比较检验，发现注意分散、表情宣泄、接受、重评与抑制 5 种策略之间的唤醒度差异不显著（ $p_{\min} = 0.17$ ），但五种情绪调节句子的唤醒度均显著高于对照条件句子的唤醒度（ $p_s < 0.001$ ）。”

**意见 9：**第 9 页第 18 行，为什么评定困难程度没有说明；被试的情绪状态需要再测一次吗，怎么保证任务二之前被试的情绪状态是一致的？

**回应：**非常感谢审稿专家深思熟虑的意见。

首先，经过仔细检查，我们发现在研究 2 的方法部分（3.1.3.实验设计；第 13 页）上一稿对句子组织任务的困难程度已经进行了说明，但的确遗漏了对需要付出努力程度评分标准的说明。在本轮修改中，我们对困难程度和努力程度的评分标准进行了如下清晰地说明：

“句子整理任务完成后，被试对句子整理任务的困难程度（1=一点不难，9=非常困难）和需要付出努力的程度（1=非常低，9=非常高）进行评分。”

在修改稿中，我们对这一部分用蓝色字体进行了突出显示。

第二，针对审稿专家提出的如何确保在第二个任务开始前被试情绪状态的一致性这个问题，我们目前的实验设计的确没有在任务 2 开始之前再次测量情绪状态。没有再次测量情绪状态的主要考虑是两点：1），在研究 2 整个实验程序开始之前，我们匹配了 6 组被试的情绪相关特质（如特质焦虑等）、状态（状态焦虑等）和情绪调节倾向；2），我们所使用的

句子整理任务的难度较低，无需被试大量付出主观努力，而针对句子整理难度和主观努力程度评分的分析也发现，6组被试无论是在句子整理难度评分还是主观努力程度评分上均无显著差异（见正文 **3.2.1 操作性检验**），因此排除了不同条件下句子整理难度不同而造成任务2开始前被试情绪状态显著不同的可能。另外，由于不同情绪调节条件下的启动句之间不存在情绪属性的显著差异（见正文 **3.1.2 实验材料，p12**），所以也不太可能因为启动句子情绪属性的不同造成任务2开始前不同条件的被试情绪状态的不对等。

概言之，我们非常感谢审稿专家的建设性意见，我们也赞同应该确保被试在开始任务2之前情绪状态的同质性；但基于上述三个方面的证据，我们认为在本研究中，可能影响任务2开始前被试情绪状态的主要因素得到了较好的控制，因此不太可能造成任务2开始前不同条件下情绪状态的不对等。

**意见 10：** 3.1.4 统计分析部分，为什么不采用两因素方差分析，比较六组被试对中性 and 负性图片的差异？

**回应：** 本研究在设计上采用了 2（刺激效价：负性，中性）\*6（启动策略类型：分心，重评，抑制，宣泄，接受和对照）实验设计的方案。我们纳入中性刺激的核心原因是为了检验所采用的负性图片诱发负面情绪的有效性。其结果也证明负性图片相比中性图片能显著诱发被试的情绪体验（效价更负与唤醒度升高，见 3.2.1 操作有效性检验）。因此，在情绪诱发的有效性得到验证的基础上，为了数据分析流程的简洁我们便将因变量指标锁定于负性图片的情绪效价与唤醒度评分，进行基于情绪调节策略（6 水平）的单因素方差分析。

事实上，与审稿人的意见相一致，我们曾对评分数据进行 2（刺激效价：负性/中性）\*6（启动策略类型）的混合方差分析。其结果无论在效价评分还是唤醒度评分上，均显示出显著的策略\*刺激效价交互作用（效价：  $F_{(5, 190)}=35.75, p<0.001$ ；唤醒度：  $F_{(5, 190)}=3.83, p=0.003$ ）。随后的简单效应分析拆分效价因素均在负性刺激条件下发现显著的启动策略主效应（效价：  $F_{(5, 190)}=4.09, p=0.001$ ；唤醒度：  $F_{(5, 190)}=4.94, p<0.001$ ），但未在中性条件下发现显著的启动策略主效应（ $F_{(5, 190)}<1.50, p>0.20$ ）。随后在负性刺激条件下，进一步对效价与唤醒度各自的主效应进行多重比较便得到了与直接基于负性刺激分析一样的结论。

因此，出于数据分析流程简洁和结果交代清晰的角度，本研究采用了先行进行操作有效性检验证明负性图片的情绪诱发有效性，随后基于负性图片进行单因素方差分析的方法。

### 第三轮

**审稿专家 1:** 无进一步修改意见。

**回应:** 谢谢!

**审稿专家 2:**

经过两轮的意见，作者较好的回答了问题，并针对相关问题对文稿进行了修改和完善。本研究旨在建立标准化的情绪调节词语库的建立，情绪调节词语库的建立对促进本土化的情绪及情绪调节方面研究具有重要的意义和贡献力，望作者能在后期能从词语库的数量上、词语的多样性上（比如过度积极情绪调节，悲伤相关的负面情绪词语等）、词语有效性的验证上对该库有更多更加全面的检验和完善。

**回应:** 非常感谢审稿专家对本研究意义的认可和对我们的修改工作的肯定!

诚如上一轮审稿意见回复所述：虽然本研究以厌恶情绪为代表初步验证了情绪调节词语库的材料启动相应情绪调节策略的有效性，但一个标准化情绪调节词语库在实际应用中的有效性，必然需要未来研究通过样本的积累、使用情境的多样化来多方面验证和进一步完善。另一方面，目前该词语库中注意分散相关的四字成语数量较少，也需要未来研究的进一步挖掘来丰富和完善。在修改稿中（p16-17，蓝色字体），我们对上述情况予以了清晰说明。

再次感谢审稿专家的中肯建议，我们必将在后续研究中对该词语库进行多维度、多角度的检验、补充和完善，为促进我国情绪调节领域的发展不懈努力!

### 第四轮

**主编意见:** 本文经过修改后改进明显，发表前需对一些细节进行润色修改。

1、自检报告中作者报告已经撰写了英文摘要，但是在文件中没有看到，是遗漏了吗?

**回应:** 非常感谢主编的提醒，英文摘要已随本文件一同上传。

2、一些统计的效果量没有给出，目前只给出了方差分析部分，其他统计检验比如 t-test 等均未给出，需要加上。

**回应:** 非常感谢主编的意见，在修改稿中已补全统计量，请审阅。

3、一些细节错误需要全文检查，例如图 7 的 A 和 B 部分没有标出；参考文献中不少页码不完整等问题。

**回应:** 非常感谢主编的意见，类似问题已自查和修正，请审阅。