

《心理学报》审稿意见与作者回应

题目：词汇具体性对情绪名词效价加工影响的 ERP 研究

作者：齐正阳 罗文波

第一轮

审稿人 1 意见：

该 ERP 研究采用快速序列视觉呈现范式探究具体性对情绪名词加工的影响规律，脑电结果分析 P1、N170、LPC 三个成分，对应于面孔表情加工三阶段模型提出的不同加工阶段。结果显示，在第二和第三阶段存在情绪效应，并且具体性和情绪在第三阶段交互影响，表明抽象词和具体词的加工差异，抽象词包含更多情感信息。该研究实验目的明确，数据呈现清晰，论文结构完整，是一项很有意义的研究，但也存在以下不足：

意见 1：前言中，作者引用了大量相关文献，调研工作充足，但词汇加工情绪效应部分，以往研究结果存在不一致情况，作者只简单论述，没有深入讨论，也没引出和本研究的相关性，需要整理前言部分逻辑性。

回应：根据专家的意见，前言部分关于词汇加工的情绪效应的 ERP 成分的论述，我们重新修改了表述方式，请见修改稿前言部分(第二和第三段蓝色字)。

意见 2：有关词汇具体性效应，已有研究多采用词汇判定任务，被试判断真/假词。此研究中，作者采用并不常见的快速序列视觉呈现范式，并且控制注意资源，这样做的目的是什么，有什么依据？

回应：感谢专家的审阅。本研究想法的最初来源是，Luo 等人(2010)提出的面孔表情加工的三阶段模型，Zhang 等人(2014)采用同样的双任务 RSVP 范式，探讨了注意资源有限条件下情绪形容词加工的时间进程，研究结果表明面孔表情加工三阶段模型可以推广到情绪词语的加工中去。情绪加工在注意资源有限的条件下，不同类型情绪信息的加工过程可能存在一个大致相同的加工模式，因此，面孔表情加工三阶段模型拓展为情绪加工的三阶段模型。该模型包含了“负性情绪偏向”，“情绪与非情绪的区分”以及“不同情绪的完全分离”三个时间轴上前后相继的阶段，注意资源短缺是情绪加工三阶段模型存在的充分不必要条件。然而，Yi 等人(2015)同样采用双任务 RSVP 范式，发现在注意资源短缺的状况下，情绪名词的加工过程只符合情绪加工三阶段模型中的后两个阶段。情绪形容词和情绪名词加工的差异，我们认为具体性很可能是一个重要的影响因素，因为形容词几乎都属于抽象词，但名词则可以分为抽象词和具体词(如爱情和书桌)。所以，在前人研究的基础上，我们采用双任务 RSVP 范式，探讨注意资源有限条件下，词汇具体性对情绪名词加工过程的影响。其主要目的是为了进一步完善和拓展由我国学者提出，已经在国际上产生重要影响的“表情加工三阶段模型”。

意见 3：作者对行为数据只报告了正确率，没有报告反应时。

回应：感谢您细致地审阅。双任务 RSVP 范式要求被试注意力保持高度集中，任务难度大，实验中既要预防被试疲劳，又要保证任务正确率，以确保分析脑电数据时叠加试次数足够，所以我们选择忽略反应时方面，仅要求被试尽可能准确回答。另外，这个经典的范式要求被试尽可能准确的反应，对反应时不作要求，因此前人采用此范式都只报告正确率，本研究重

点参考的 Zhang 等人(2014)和 Yi 等人(2015)对情绪形容词和情绪名词的研究,任务完成没有时间限制,同样要求被试尽可能回答正确,行为数据分析了正确率,没有分析反应时。因此,兼顾试验任务难度,以及参照前人研究,我们在行为数据分析中选择报告正确率,没有报告反应时。

但根据审稿专家的提示,我们对被试双任务回答正确的试次的反应时进行了分析,结果显示,词汇具体性的主效应 $[F(1, 23) = 1.03, p = 0.32]$,词汇情绪效价的主效应 $[F(2, 46) = 2.16, p = 0.13]$,二者间交互效应 $[F(2, 46) = 0.15, p = 0.83]$ 均不显著。这可能说明被试根据指导语提示,更多关注正确率,没有在意反应时方面。

意见 4: 前言引用的面孔表情三阶段模型中,每个阶段有两个脑电成分指标,本研究中每个阶段只采用了一个脑电成分做指标,请做出合理解释。

回应: 感谢审稿人的提示。首先, Luo 等人(2010)最初提出的面孔表情加工三阶段模型中, P1 和 N1 作为第一阶段指标,表现为负性偏向; N170 和 VPP 作为第二阶段指标,表现为情绪面孔和非情绪面孔的分离; P3 和 N3 作为第三阶段指标,表现为面孔情绪的精细加工。其后续在单试次水平上考察了这 6 个 ERP 成分,结果发现,与其他成分相比, N1 成分的单试次参数受到情绪条件的影响更小,表明 N1 成分可能在单试次水平上包含了较少的面孔表情辨别信息(Zhang et al., 2013)。因此, P1 成分作为面孔表情加工三阶段模型中第一阶段的脑电指标更为可靠。其次,经过 Zhang 等人(2014)和 Yi 等人(2015)探讨注意资源有限条件下情绪词语加工的时间变化过程,表明面孔表情加工三阶段模型能推广到情绪词语的加工中去,而 N170 和 LPC 分别作为词语加工的第二和第三阶段的脑电指标,在此过程中稳定可靠。因此,本研究中对三个阶段都选择了最具代表性的脑电成分作为指标,并与情绪词语加工三阶段的研究形成精准比较。

意见 5: 脑电结果分析中, N170 成分的时间窗在 200 毫秒后,比传统 N170 出现的时间晚,作者是否做出合理解释?

回应: 确如审稿专家所说,传统 P1 和 N170 成分的峰值潜伏期是 100~130ms 和 170ms (Luck, 2005)。然而本研究中 P1 和 N170 成分的峰值潜伏期大概在 180ms 和 250ms,虽然峰值潜伏期延后,但这两个成分的头皮地形图与顶枕 P1 和双侧颞枕 N170 相一致。潜伏期延迟或许是因为 RSVP 范式,这是一个相对困难的实验任务,需要被试在快速呈现的刺激流中先识别 T1 位置数字串的奇偶性,然后在 200ms 后的 T2 位置判断正立名词的情绪类别。我们结果中 P1 和 N170 峰值潜伏期出现的时间与 Luo 等人(2010), Zhang 等人(2014)和 Yi 等人(2015)的研究结果一致。除此之外,也有研究使用双任务 RSVP 范式出现了相同的 ERP 潜伏期延迟(Kessler et al., 2005; Sergent et al., 2005)。

意见 6: 讨论部分,作者从多角度对抽象词加工进行了论述,但具体名词加工只有第二阶段,原因是什么?这结果反映了什么?可以考虑从双编码理论或语境有效性理论方面解释。

回应: 感谢您的提示。双重编码理论和语境有效性理论,对词汇的具体性效应从不同方面提出了解释。双重编码理论认为具体词有高表象性,加工时同时激活表象系统和语义系统,而抽象词的加工只能激活语义系统,因此个体对具体词有加工优势(Paivio, 1986)。语境有效性理论则认为具体词和抽象词都只激活语义系统,但具体词和语境信息的联结更为紧密(Schwanenflugel et al., 1988; Schwanenflugel et al., 1992)。本研究发现,注意资源匮乏条件下,具体名词的加工只符合情绪加工三阶段模型中的第二阶段,我们认为这从侧面反映出抽象词包含更多的情绪信息,因此更符合情绪加工三阶段模型中较稳定和具有普遍性的第二和第三阶段。具体词带有的表象信息,和语境信息的紧密联结,或许某种程度上削弱了其和情感信

息的联结，所以只符合三阶段模型中的第二阶段。

意见 7：建议将实验中采用的词语材料附录在后，方便他人研究参考。

回应：感谢专家的重要提醒，我们已经将实验材料附录在文章后，详见附表 1 和附录 2。

附表 1 实验中选用的情绪名词

	正性词	中性词	负性词
具体词	鲜花 笑脸 首饰	总管 护照 赛场	苍蝇 棺材 老鼠
	绿地 新房 别墅	警卫 狮子 邮票	遗体 病房 太监
抽象词	信誉 美德 名声	廉价 对策 焦点	耻辱 漏洞 创伤
	诚意 深情 乐趣	国情 视野 江湖	隐患 虚荣 困境

附录 2 实验中作为干扰项出现的假词(以倒立方式呈现)

件座 弗据 被别 所野 就出 单岁 典市 茶律 廷寓 汇献 族则 限特

参考文献：

- Kessler, K., Schmitz, F., Gross, J., Hommel, B., Shapiro, K., Schnitzler, A. (2005). Target consolidation under high temporal processing demands as revealed by MEG. *NeuroImage*, 26, 1030–1041.
- Luck, S.J., (2005). *An introduction to the event-related potential technique*. London, UK. The MIT Press.
- Luo, W. B., Feng, W. F., He, W. Q., Wang, N. Y., & Luo, Y. J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, 49(2), 1857–1867.
- Paivio, A. (1986). *Mental representations: A dual coding approach*. Oxford: Oxford Psychology Press.
- Sergent, C., Baillet, S., & Dehaene, S. (2005). Timing of the brain events underlying access to consciousness during the attentional blink. *Nature Neuroscience*, 8(10), 1391–1400.
- Schwanenflugel, P. J., Akin, C., & Luh, W. M. (1992). Context availability and the recall of abstract and concrete words. *Memory & Cognition*, 20(1), 96–104.
- Schwanenflugel, P. J., Harnishfeger, K. K., & Stowe, R. W. (1988). Context availability and lexical decisions for abstract and concrete words. *Journal of Memory and Language*, 27(5), 499–520.
- Yi, S. N., He, W. Q., Zhan, L., Qi, Z. Y., Zhu, C. L., Luo, W. B., & Li, H. (2015). Emotional noun processing: An ERP study with rapid serial visual presentation. *PLoS ONE*, 10(3), e0118924.
- Zhang, D. D., He, W. Q., Wang, T., Luo, W. B., Zhu, X. R., Gu, R. L., ... & Luo, Y. J. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, 9(12), 1897–1903.
- Zhang, D. D., Luo, W. B., & Luo, Y. J. (2013). Single-trial ERP evidence for the three-stage scheme of facial expression processing. *Science China Life Sciences*, 56(9), 835–847.

审稿人 2 意见：

论文采用 ERP 技术，考察了情绪名词的具体性(论文中所用词语为“词汇类型”)和情绪效价如何影响情绪名词效价加工的 ERP 相关物。总的来看，论文在研究目的、额外变量控制、数据分析和讨论等多方面，均存在较大的不足。具体问题如下，主要问题：

意见 1：论文存在的最大的问题是研究目的并不足够清楚。为什么采用注意瞬脱(attentional blink)范式或设置注意资源有限的情境，逻辑并不清楚。这样的情境与日常生活中的语言加

工情境有很大不同,因而降低了研究的生态效度。为了追求研究的理论价值而某种程度地牺牲生态效度,可以理解。遗憾的是,很难看出如此安排的理论价值是什么。如果作者希望考察注意资源如何调节情绪名词加工,那么,作者应该操纵 T1 和 T2 之间的时间间隔。然而,作者并未进行这样的操纵。此外,实验任务要求被试外显地判断名词的情绪效价。不清楚为什么要使用这样的实验任务。这种任务因为强制被试加工词汇的情绪效价,所以必然会引导被试特别注意词汇的情绪内容,而注意瞬脱范式的使用目的似乎是为了减少被试对情绪词汇的注意。两方面的实验安排看起来互相矛盾。不清楚作者这些安排的意图究竟是什么。此外,注意瞬脱范式将工作记忆引入情绪名词加工(因为被试需要等到序列中所有刺激呈现完之后再完成任务),进一步复杂化了情绪名词加工结果的解释。这是本研究中该范式带来的另一个不足。

回应:感谢专家的审阅。本研究想法的最初来源是, Luo 等人(2010)提出的面孔表情加工的三阶段模型。该研究中使用的注意瞬脱范式对三阶段模型的提出十分必要,因为只有在有限注意资源内(短 lag)面孔表情加工呈现出“负性情绪偏向”,“情绪与非情绪的区分”以及“不同情绪的完全分离”三个时间轴上前后相继的阶段。而当注意资源充足时(长 lag),容易产生天花板效应,很难把面孔表情加工的优先级排列出来。面孔表情加工的三阶段模式可能反应了在有限注意资源内情绪加工的时间进程,而且注意瞬脱范式对大脑中情绪加工时间进程的探测十分敏感。

Zhang 等人(2014)采用双任务 RSVP 范式观察到,情绪形容词在有限注意资源内加工的动态变化过程表现出与面孔表情加工类似的三阶段模式,拓宽了面孔表情加工三阶段理论的适用范围。这说明,情绪加工在注意资源有限的条件下,不同类型情绪信息的加工过程可能存在一个大致相同的加工模式。然而, Yi 等人(2015)同样采用双任务 RSVP 范式,发现在注意资源短缺的状况下,情绪名词的加工过程只符合情绪加工三阶段模型中的后两个阶段。我们认为词汇具体性很可能是影响情绪形容词和情绪名词在注意瞬脱效应中产生加工差异的重要因素,因为形容词几乎都属于抽象词,但名词却可以分为抽象词和具体词(如爱情和书桌)。所以,在此基础上,我们采用双任务 RSVP 范式,探讨注意资源有限时,词汇具体性对情绪名词加工过程的影响。

本研究没有操纵 T1 和 T2 之间的时间间隔,主要有两个原因。一是,我们想探讨在有限注意资源内,词汇具体性在注意瞬脱效应中对情绪名词加工过程的影响。本研究重点参考 Luo 等人(2010), Zhang 等人(2014)和 Yi 等人(2015)的先前研究,皆是在有限注意条件下观察到情绪信息加工的时间进程。我们采用相同的实验范式和实验任务,是为了聚焦研究问题,更好的说明词汇具体性的影响。二是,这样也可以节约任务时长,减少被试疲劳感。本研究是 2×3 被试内实验设计,除了要操纵的 6 个实验条件,由于刺激快速呈现时 T2 前后的分心刺激所诱发的脑电活动会影响到 T2 的脑电波形,实验中还要添加 T2 位置空屏条件作为基线,后期通过差异波处理以获得纯粹由 T2 诱发的 ERP 波形。目前正式实验共 720 个试次,总时长一小时,如果再设置长 lag 的条件,实验时长翻倍,双任务 RSVP 范式又是一个相对困难的实验任务,被试极易产生疲劳效应,反而会影响研究结果。

意见 2:前言:“Hinojosa, Méndez-Bártolo 和 Pozo (2010)采用词汇-语义判定任务发现正性词和负性词均比中性词诱发了更大的 LPC 波幅,而 Citron, Weekes 和 Ferstl (2013)采用词汇判定任务发现中性词比正性词和负性词均诱发了更大的 LPC 波幅”。这两项研究(以及论文中引用的其他几项研究)结果模式相反。这种分歧的实验结果是否至少部分地反映了不同类型词语之间多方面的词汇差异,如词频、词语获得年龄?正负性词与中性词之间的比较是项目间(between-item)比较。作者需要讨论这种可能性,而不是简单将这种分歧的结果归因于实验范式或任务因素——“LPC 波幅可以分辨不同情绪效价的词语加工,但在不同实验范式

和实验任务中，详细区分结果各有不同。”

回应：感谢专家的提醒。前言对于词汇加工的情绪效应的 ERP 成分的论述，我们重新修改了表述方式，具体请见修改稿前言部分(第二和第三段蓝色字)。

意见 3：“关于情绪名词具体性的评定，实验前招募 60 名在校大学生，详细解释具体词和抽象词的概念后，采用纸笔测验的方式让其对情绪名词进行评定。”使用的是 5 点量表评定或者其他方法？不清楚。

回应：感谢专家的重要提醒。在修改稿的方法-设计与材料部分，我们补充了情绪名词具体性的评分标准，具体如下：“5 点评分：1 分代表非常抽象，2 分代表一般抽象，3 分代表中性，4 分代表一般具体，5 分代表非常具体。”

意见 4：“同时，具体词和抽象词之间也进行了变量匹配，保证了具体词和抽象词在频率、笔画、唤醒度方面不存在显著性差异。”论文并未报告与此相应进行了何种统计检验以及统计检验结果。

回应：感谢您的专业性意见。结合意见 5，我们重新对实验中所选名词的唤醒度、效价、具体性、整词笔画、首字笔画、尾字笔画、整词频度、首字频度、尾字频度进行了细致分析，采用 2(具体性：具体词、抽象词) × 3(效价：正性、中性、负性)重复测量方差分析发现，选定的名词仅在效价和具体性评分上存在显著性差异， $p_s < 0.001$ 。其他词汇学特征得到了良好控制，六种类型的名词在唤醒度、整词笔画、首字笔画、尾字笔画、整词频度、首字频度、尾字频度上差异均不显著， $p_s > 0.11$ ，具体统计结果请见下表。修改稿中方法-设计与材料部分(蓝色字)做出了相应修改。

	具体性主效应	效价主效应	具体性×效价的交互效应
唤醒度	$[F(1, 5) = 1.76, p = 0.24]$	$F(2, 10) = 1.10, p = 0.37]$	$F(2, 10) = 0.36, p = 0.70]$
效价	$[F(1, 5) = 2.97, p = 0.15]$	$F(2, 10) = 313.69, p < 0.001]$	$F(2, 10) = 0.48, p = 0.59]$
具体性	$[F(1, 5) = 955.59, p < 0.001]$	$F(2, 10) = 0.03, p = 0.94]$	$F(2, 10) = 0.51, p = 0.53]$
整词笔画	$[F(1, 5) = 0.31, p = 0.60]$	$F(2, 10) = 0.42, p = 0.57]$	$F(2, 10) = 0.28, p = 0.71]$
首字笔画	$[F(1, 5) = 1.98, p = 0.22]$	$F(2, 10) = 0.08, p = 0.86]$	$F(2, 10) = 3.58, p = 0.11]$
尾字笔画	$[F(1, 5) = 1.72, p = 0.25]$	$F(2, 10) = 1.16, p = 0.34]$	$F(2, 10) = 0.96, p = 0.40]$
整词频度(对数)	$[F(1, 5) = 1.33, p = 0.30]$	$F(2, 10) = 0.09, p = 0.85]$	$F(2, 10) = 2.90, p = 0.11]$
首字频度(对数)	$[F(1, 5) = 0.02, p = 0.88]$	$F(2, 10) = 2.44, p = 0.15]$	$F(2, 10) = 2.76, p = 0.14]$
尾字频度(对数)	$[F(1, 5) = 0.08, p = 0.79]$	$F(2, 10) = 1.01, p = 0.40]$	$F(2, 10) = 0.08, p = 0.90]$

意见 5：词频和笔画数数据，应该进行 2(具体、抽象)*3(效价)方差分析，而不是将效价作为一个因素，分别在具体词和抽象词中进行两个独立的单因素方差分析。另外，词频是否进行了对数转换(log-transform)？应该对对数转换后的词频数据进行方差分析。论文还应报告词频数据来源。

回应：感谢您的专业性意见，具体统计结果请见意见 4 的回应。整词频度、首字频度和尾字频度的数据来源于《现代汉语常用词频词典》(刘源 等, 1990)，已在文中追加报告(修改稿的方法-设计与材料部分)，并在对数转换后进行的方差分析。

意见 6：唤醒度、效价和具体性的指标为 7 点或 5 点量表评分？7 点或 5 点量表为顺序量表，而顺序量表评分更合适的统计分析方法是累积混合效应模型，如 CLMM (cumulative link mixed effects models, function clmm, R package ordinal; Christensen, 2019)。建议作者使用 CLMM 而不是方差分析，分析这类顺序数据。

回应：非常感谢审稿人的宝贵意见，根据您的提示，我们认真研读了 ordinal 软件包。确如专家所说，5 点或 7 点评分的李克特量表收集的顺序变量属于离散数据，适合非参数检验，但由于离散变量无法计算平均数，或进行相关、回归等统计分析，因此无法检验有关的研究假设，为了方便解释和有意义分析，心理学实验中也有将其作为连续变量来进行处理的情况。本研究中词语材料的唤醒度和效价的数据来源于汉语情感词系统(CAWS)(王一牛 等, 2008)，该系统中词语的各项参数值是通过平均数计算得来的，因此我们在统计分析中将其视为连续变量处理。为保证与唤醒度和效价的统计方法一致，我们对具体性评分也采用了方差分析方法，并且根据您的意见 4 和 5，重新对词语材料进行了细致分析。十分感谢审稿专家为我们提供了不同的分析思路，使我们深刻认识到实验材料数据处理的重要性。

意见 7：不清楚在具体词与抽象词之间，以及正性、中性和负性词之间，作者是否匹配了其它可能影响词汇加工的词汇变量，如词汇的获得年龄(AoA)、词中所包含的汉字的字频，等等。如果这些潜在的词汇差异未能得到理想的控制，那么，实验所观察到的具体性效应或情绪效价效应，就很难解释。

回应：感谢您的宝贵意见。我们已追加对词语材料的首字字频、尾字字频、首字笔画和尾字笔画的统计分析，具体统计结果请见意见 4 的回应。至此，共匹配了所选名词的唤醒度、效价、具体性、整词笔画、首字笔画、尾字笔画、整词频度对数、首字频度对数、尾字频度对数，共 9 个词汇学特征，保证我们选用的词语材料仅在效价和具体性上有显著差异，其他词汇学变量得到了良好控制。

专家提到的词汇的获得年龄(age of acquisition, AOA)，以及词汇的获得方式(mode of acquisition, MOA)，都是影响词汇加工过程的重要因素。已有研究的国外词语库中(法语、荷兰语、意大利语)已经包括了对 AOA 的评定(Bonin et al., 2018; Della Rosa et al., 2010; Moors et al., 2013)，但目前汉语双字词表中(有词汇具体性的评定)缺少对 AOA 和 MOA 的测量(Yao et al., 2017)。词汇习得年龄效应(AOA 效应)表现为早期习得的词语比晚期习得的词语加工速度更快，反应时更短(Carroll & White, 1973)，这可能来源于词汇累计频率和频率分布模式的影响(Zevin & Seidenberg, 2002)。对 AOA 和 MOA 的评定，大多采用被试主观回忆的方式进行量表评分，记忆可靠性值得商榷。感谢审稿专家的提示，考虑到词汇加工过程的复杂性，未来研究我们将思考从 AOA 和 MOA，及其他影响词汇识别的因素入手进行深入探讨。

意见 8：“具体来看，P1 成分(160-190 ms)选取 Pz, P3/4, POz, PO3/4 电极点分析；N170 成分(230-280 ms)选取 P7/8, PO7/8 电极点分析；LPC 成分(440-600 ms)选取 Cz, C3/4, CPz, CP3/4, Pz, P3/4, 电极点分析。”这种先验的、选取少量电极的方法并不恰当(尤其是在缺乏足够多的、使用类似汉语刺激的研究的情况下，先验的电极选取方法可能漏掉很多有价值的电极数据)。更恰当的方法是将头皮划分为若干个兴趣区，并尽可能选取数量更多的代表性电极(同时也提高数据的利用率，毕竟作者使用了 64 导记录)。

回应：感谢您细致的审阅。我们是根据脑电地形图及已有的相关研究选取的电极位置(Yi et al., 2015; Zhang et al., 2014)，各个脑电成分所选的电极点具有代表性，尤其像 N170 成分，同时与前人研究形成了精准对比。您提示将头皮划分为若干个兴趣区的方法，或许更适合一些对脑电成分或结果有探索性需求的研究。

意见 9：“当词语刺激反复出现，具体词包含的表象信息与大脑皮层反应联结更快，在情绪名词加工早期阶段就体现出具体词和抽象词的加工差异。结合 N170 成分的情绪效应，说明注意资源相对缺乏的条件下，情绪名词加工早期阶段受到情绪信息和表象信息的共同影响。”；“早期 N170 成分还受到词汇具体性的调节，说明情绪名词加工的早期阶段受情绪信

息和表象信息共同影响”。上述讨论或结论缺乏数据支持。“具体性影响情绪名词早期加工阶段”的结论要求在早期加工阶段，如 N170 所反映的时间窗口，情绪效价(正性、中性与负性)与具体性(具体、抽象)之间出现显著的交互作用。然而，论文并未观察到这样的交互作用。这意味着，具体性如何影响 N170，并不受情绪效价的影响，因此，正性或负性情绪词与中性词之间，在具体性效应上并无区别(具体性效应并未特异到正性或负性情绪词上)。这样，不能宣称具体性影响情绪名词早期加工阶段。

回应：感谢专家的重要提醒。我们已对上述表达歧义的语句进行修改或删减，详见修改稿中文摘要部分(第 18 页，5~10 行)，讨论部分(第 27 页，17 行~第 28 页，3 行，第 29 页，8~10 行；第 30 页，8~9 行)，结论部分(第 30 页，23~24 行)，英文摘要部分(第 34 页，26~28 行；第 35 页，1~3 行；第 35 页，10~11 行)。

意见 10：因素名称：“具体性”比“词汇类型”更恰当。因为正性、中性和负性词也属于不同的词汇或词语类型。

回应：感谢并接受审稿专家的建议，已将因素名称修改为词汇具体性。

意见 11：半球的主效应没有任何意义，因为它平均了词的具体性和情绪效价的不同水平。所以，不应该报告这样的统计检验结果。

回应：根据专家的要求，修改稿中结果-ERP 结果部分已删除半球主效应的结果报告。

意见 12：“根据 ANOVA 结果显示”--->“ANOVA 结果显示”。

回应：感谢并接受审稿专家的建议，已修改。

参考文献：

- Bonin, P., Môt, A., & Bugaiska, A. (2018). Concreteness norms for 1,659 French words: Relationships with other psycholinguistic variables and word recognition times. *Behavior Research Methods*, 50(6), 2366–2387.
- Carroll, J. B., & White, M. N. (1973). Age-of-acquisition norms for 220 picturable nouns. *Journal of Verbal Learning and Verbal Behavior*, 12(5), 563–576.
- Della Rosa, P. A., Catricalà E., Vigliocco, G., & Cappa, S. F. (2010). Beyond the abstract–concrete dichotomy: Mode of acquisition, concreteness, imageability, familiarity, age of acquisition, context availability, and abstractness norms for a set of 417 Italian words. *Behavior Research Methods*, 42, 1042–1048.
- Liu, Y., Liang, N. Y., Wang, D. J., Zhang, S. Y., Yang, T. Y., Jie, C. Y., & Sun, W (Eds). (1990). *Dictionary of usage frequency of modern Chinese words*. Beijing, China: Yuhang Press.
- [刘源, 梁南元, 王德进, 张社英, 杨铁鹰, 揭春雨, 孙伟 (编). (1990). *现代汉语常用词频词典*. 北京: 宇航出版社.]
- Luo, W. B., Feng, W. F., He, W. Q., Wang, N. Y., & Luo, Y. J. (2010). Three stages of facial expression processing: ERP study with rapid serial visual presentation. *NeuroImage*, 49(2), 1857–1867.
- Moors, A., De Houwer, J., Hermans, D., Wanmaker, S., van Schie, K., Van Harmelen, A. L., ... Brysbaert, M. (2013). Norms of valence, arousal, dominance, and age of acquisition for 4,300 Dutch words. *Behavior Research Methods*, 45, 169–177.
- Wang, Y. N., Zhou, L. M., & Luo, Y. J. (2008). The pilot establishment and evaluation of Chinese affective words system. *Chinese Mental Health Journal*, 22(8), 608–612.
- [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.]
- Yao, Z., Wu, J., Zhang, Y. Y., & Wang, Z. H. (2017). Norms of valence, arousal, concreteness, familiarity, imageability, and context availability for 1,100 Chinese words. *Behavior research methods*, 49(4),

1374–1385.

- Yi, S. N., He, W. Q., Zhan, L., Qi, Z. Y., Zhu, C. L., Luo, W. B., & Li, H. (2015). Emotional noun processing: An ERP study with rapid serial visual presentation. *PLoS ONE*, 10(3), e0118924.
- Zevin, J. D., & Seidenberg, M. S. (2002). Age of acquisition effects in word reading and other tasks. *Journal of Memory and language*, 47(1), 1–29.
- Zhang, D. D., He, W. Q., Wang, T., Luo, W. B., Zhu, X. R., Gu, R. L., ... & Luo, Y. J. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, 9(12), 1897–1903.
-

审稿人 3 意见：

具体性和抽象性可能是影响词汇加工过程的重要因素。该研究采用快速序列视觉呈现范式，采用脑电技术探讨注意资源有限条件下，词汇具体性对情绪名词加工的影响，具有一定的理论意义。文章还存在以下具体问题，希望作者认真加以修改：

意见 1：摘要部分主要阐述了各个成分上的结果，但是缺少总结性的结论，这说明抽象词可能比具体词负载了更多的情绪信息过于简单，信息量不够，建议补充具体的有针对性的内容。

回应：感谢审稿人的提示，我们对摘要部分修改如下：“具体性和抽象性是词汇同一特性的两极，是直接影响词汇加工过程的重要因素。本研究采用快速序列视觉呈现范式，结合脑电技术探讨在有限注意资源内，词汇具体性对情绪名词加工过程的影响。结果发现：名词加工早期和晚期阶段的 ERP 成分受到情绪效价的调节，情绪词比中性词诱发了更大的 N170 和 LPC 波幅；名词加工早期和晚期阶段的 ERP 成分也受到词汇具体性的调节，具体词比抽象词诱发了更大的 N170 和 LPC 波幅；词汇具体性影响情绪名词加工的晚期阶段，LPC 波幅能够分辨出不同情绪效价的抽象词，对于具体词仅能区分出情绪与非情绪，这说明抽象词可能比具体词负载了更多的情绪信息，反映出对情绪信息的精细加工过程。”英文摘要部分也做了相应修改。

意见 2：词汇的具体性和抽象性属于词汇同一特性的两极，作者在关键词中把具体性和抽象性都罗列上了，建议将二者改为一个更具代表性的词。

回应：感谢并接受审稿专家的建议，文中已将数据分析的因素名称由“词汇类型”修改为“词汇具体性”，关键词中也有相应修改。

意见 3：前言中“个体对情绪信息…重要的调节与促进作用”调节作用包含促进作用，这里表述重复，请进一步精炼。

回应：感谢您细致的审阅，已删除促进作用。

意见 4：“词汇情绪性能够增强大脑皮层活动反应，涉及到视觉词汇加工过程各个阶段。”后面在列举各个 ERP 成分的情绪调节效应时最好把相关研究采用的词汇类型(如名词，形容词)加以说明，也便于与本研究结果进行对比。

回应：感谢专家的提醒。前言对于词汇加工的情绪效应的 ERP 成分的论述，我们重新修改了表述方式，具体请见修改稿前言部分(第二和第三段蓝色字)。

意见 5：“早期视觉成分受注意的显著影响，表现为幅值的增强(Hillyard et al., 1998)，由此可

见”这句话可以删除。

回应：感谢并接受审稿专家的建议，已删除。

意见 6：N170 和 EPN 代表的加工阶段和心理意义是不一样的，要详细说明这两个成分在情绪性词语加工中代表的意义。类似地，对 LPC 成分的表述未能充分体现其在情绪性词语加工中的作用，比如正性和负性词汇诱发了大于中性词汇的 LPC 波幅，反映了增强的持续性注意。

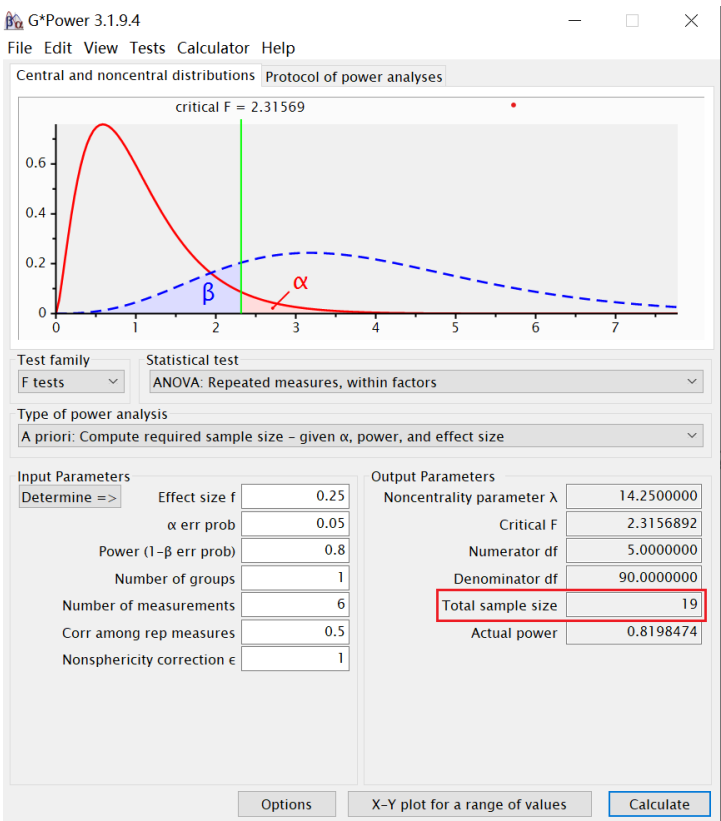
回应：感谢专家的提醒。前言对于词汇加工的情绪效应的 ERP 成分的论述，我们重新修改了表述方式，具体请见修改稿前言部分(第二和第三段蓝色字)。

意见 7：“根据已有研究和理论基础”应提供相应的参考文献。

回应：感谢审稿人的提示，我们已提供相应的参考文献，在修改稿前言部分(第 21 页，11~13 行)。

意见 8：被试量是如何确定的要进行报告。

回应：感谢专家的重要提醒。在修改稿的方法-被试部分，我们报告了被试量的确定方法，具体如下：“本研究为 2(词汇具体性：具体词、抽象词)×3(词汇情绪效价：正性、中性、负性)被试内实验设计，实验前使用 G*Power 软件，依据中等效果量 Effect size $f = 0.25$ ，当 $\alpha = 0.05$ 时需要达到 power $(1 - \alpha) = 0.8$ 所需的被试量约为 19 人。”最终实验共采集 24 人，差异显著效应的偏 Eta 平方均大于 0.13，具有高效果量。



意见 9：每种条件下保留的有效试次数应报告。

回应：感谢审稿人的提示。根据您的意见，在修改稿的方法-脑电数据的记录、处理与分析部分，我们报告了“每名被试每个条件下有效叠加试次数均超过 50 次。”每名被试在每种实

验条件下保留的试次数目具体如下：具体正性词=53±18 个试次，具体中性词=58±22 个试次，具体负性词=61±20 个试次，抽象正性词=60±18 个试次，抽象中性词=51±18 个试次，抽象负性词=63±22 个试次。

意见 10：本研究中并未发现 P1 成分的情绪效应，可以不加讨论。

回应：感谢您的意见。P1 成分结果中，我们仅发现半球的主效应显著，词汇具体性的主效应，词汇情绪效价的主效应，以及二者间的交互效应均不显著。我们认为 P1 成分可能反映了情绪词语加工过程的第一阶段，虽然未发现有意义的情绪效应，但有必要对该结果进行讨论说明。Zhang 等人(2014)发现负性形容词在左半球上比正性形容词和中性形容词诱发了更大的 P1 波幅，Yi 等人(2015)和我们的研究结果却没有发现名词的 P1 情绪效应，这可能是由于刺激词性不同所导致。因为，形容词通常描述个性特征或情绪状态，相较于动词和名词，它和情绪有更直接的关联，也更容易引发自我参照过程(Herbert et al., 2008; Palazova et al., 2011)。

意见 11：在抽象词与情绪信息的联结部分，作者提到了 fMRI 研究的结果，这对于解释 LPC 成分的交互作用关系并不大。后面的内容也偏离要讨论的交互作用，建议删除。

回应：感谢审稿人的意见。本研究发现词汇具体性影响情绪名词加工的晚期阶段，LPC 波幅能够分辨出不同情绪效价的抽象词，对于具体词仅能区分出情绪与非情绪，说明抽象词可能比具体词负载了更多的情绪信息。本研究结果与情感具身认知理论突出情绪与抽象概念间联结的特点一致，该理论近年来得到了众多研究的支持(Borghi et al., 2017)，对抽象概念表征的研究多数采用 fMRI 技术，讨论部分最后对情感具身认知理论的探讨，希望对未来词汇具体性研究提供更多技术方法的思路。

参考文献：

- Borghi, A. M., Binkofski, F., Castelfranchi, C., Cimatti, F., Scorolli, C., & Tummolini, L. (2017). The challenge of abstract concepts. *Psychological Bulletin*, 143(3), 263–292.
- Herbert, C., Junghofer, M., & Kissler, J. (2008). Event related potentials to emotional adjectives during reading. *Psychophysiology*, 45(3), 487–498.
- Palazova, M., Mantwill, K., Sommer, W., & Schacht, A. (2011). Are effects of emotion in single words non-lexical? Evidence from event-related brain potentials. *Neuropsychologia*, 49(9), 2766–2775.
- Yi, S. N., He, W. Q., Zhan, L., Qi, Z. Y., Zhu, C. L., Luo, W. B., & Li, H. (2015). Emotional noun processing: An ERP study with rapid serial visual presentation. *PLoS ONE*, 10(3), e0118924.
- Zhang, D. D., He, W. Q., Wang, T., Luo, W. B., Zhu, X. R., Gu, R. L., ... & Luo, Y. J. (2014). Three stages of emotional word processing: An ERP study with rapid serial visual presentation. *Social Cognitive and Affective Neuroscience*, 9(12), 1897–1903.
-

审稿人 4 意见：

本论文采用快速序列视觉呈现范式结合事件相关电位技术探讨了词汇的情绪效价和具体性对词汇加工的调节作用。发现词汇的情绪效价在N170和LPC上的主效应，词汇具体性在N170上的主效应，以及LPC上的交互效应。本研究讨论的问题具有一定的理论意义和新颖性，但是存在以下问题。

意见1：论文的标题没有反映出所研究的科学问题，更像一个实验报告的题目。建议修改。

回应：感谢您的建议。根据您的意见我们拟将题目改为《词汇具体性对情绪名词效价加工影响的 ERP 研究》。

意见 2：被试量比较小（N=24），这个被试量是如何确定的？有何根据？此外，被试存在较大的性别差异，这对结果会有怎样的影响？为何不能避免这样的问题？

回应：感谢审稿人的意见。在修改稿的方法-被试部分，我们报告了被试量的确定方法，具体如下：“本研究为 2(词汇具体性：具体词、抽象词)×3(词汇情绪效价：正性、中性、负性)被试内实验设计，实验前使用 G*Power 软件，依据中等效果量 Effect size $f = 0.25$ ，当 $\alpha = 0.05$ 时需要达到 power $(1 - \alpha) = 0.8$ 所需的被试量约为 19 人。”最终实验共采集 24 人，差异显著效应的偏 Eta 平方均大于 0.13，具有高效果量。

本研究中女性被试占比较大，为检验性别差异对实验结果的影响，我们将性别因素暂定为被试间因素，重新对正确率和三个脑电成分进行 2(词汇具体性：具体词、抽象词)×3(词汇情绪效价：正性、中性、负性)重复测量方差分析。行为和脑电结果均表明，性别和词汇具体性间的交互作用不显著[$F_s < 1.19, p_s > 0.29$]，性别和词汇情绪效价间的交互作用不显著[$F_s < 1.90, p_s > 0.17$]，三者间的交互作用不显著[$F_s < 0.82, p_s > 0.43$]，性别因素并未影响实验结果。实验当初采用了随机招募被试的方式，没有严格控制性别比例，非常感谢您的提醒，未来研究过程中我们会更加重视被试性别平衡的问题。

意见 3：表 1 只报告了具体词和抽象词之间的比较结果，为何没有呈现不同情绪水平之间的比较结果？文中有提到“六种类型的名词在以上变量上差异均不显著”这句话如何理解呢？比如负性词汇和中性词汇之间唤醒度差异不显著这一点不太容易做到，希望作者更加具体的说明材料的筛选。

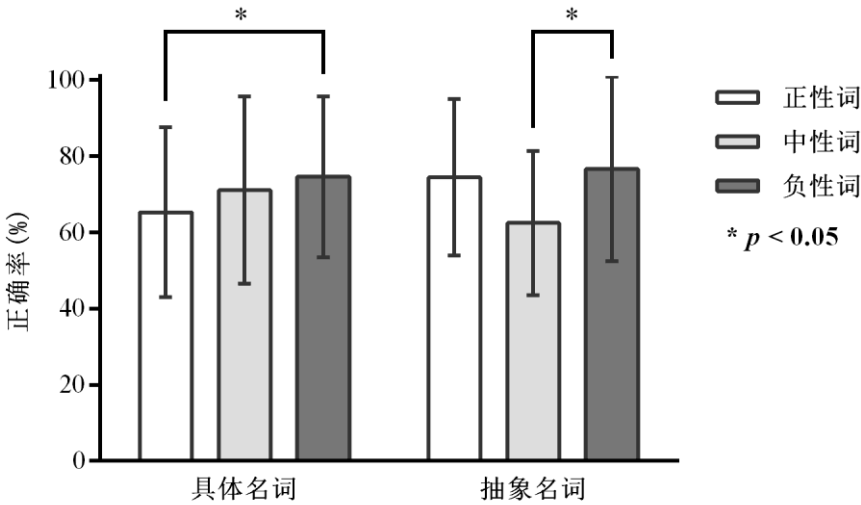
回应：感谢审稿人的提示。我们重新对实验中所选名词的唤醒度、效价、具体性、整词笔画、首字笔画、尾字笔画、整词频度、首字频度、尾字频度进行了细致分析，采用 2(具体性：具体词、抽象词)×3(效价：正性、中性、负性)重复测量方差分析发现，选定的名词仅在效价和具体性评分上存在显著性差异， $p_s < 0.001$ 。其他词汇学特征得到了良好控制，六种类型的名词在唤醒度、整词笔画、首字笔画、尾字笔画、整词频度、首字频度、尾字频度上差异均不显著， $p_s > 0.11$ ，具体统计结果请见下表。修改稿中方法-设计与材料部分对材料的筛选和分析进行了更详细和准确的阐述，请您查阅。

	具体性主效应	效价主效应	具体性×效价的交互效应
唤醒度	[$F(1, 5) = 1.76, p = 0.24$]	$F(2, 10) = 1.10, p = 0.37$	$F(2, 10) = 0.36, p = 0.70$
效价	[$F(1, 5) = 2.97, p = 0.15$]	$F(2, 10) = 313.69, p < 0.001$	$F(2, 10) = 0.48, p = 0.59$
具体性	$F(1, 5) = 955.59, p < 0.001$	$F(2, 10) = 0.03, p = 0.94$	$F(2, 10) = 0.51, p = 0.53$
整词笔画	[$F(1, 5) = 0.31, p = 0.60$]	$F(2, 10) = 0.42, p = 0.57$	$F(2, 10) = 0.28, p = 0.71$
首字笔画	[$F(1, 5) = 1.98, p = 0.22$]	$F(2, 10) = 0.08, p = 0.86$	$F(2, 10) = 3.58, p = 0.11$
尾字笔画	[$F(1, 5) = 1.72, p = 0.25$]	$F(2, 10) = 1.16, p = 0.34$	$F(2, 10) = 0.96, p = 0.40$
整词频度(对数)	[$F(1, 5) = 1.33, p = 0.30$]	$F(2, 10) = 0.09, p = 0.85$	$F(2, 10) = 2.90, p = 0.11$
首字频度(对数)	[$F(1, 5) = 0.02, p = 0.88$]	$F(2, 10) = 2.44, p = 0.15$	$F(2, 10) = 2.76, p = 0.14$
尾字频度(对数)	[$F(1, 5) = 0.08, p = 0.79$]	$F(2, 10) = 1.01, p = 0.40$	$F(2, 10) = 0.08, p = 0.90$

意见 4：行为结果中准确率比较低（60%~70%），此外准确率的标准差小到离奇，请作者检查下是否标准差在作图是没有乘以 100%。并解释这个准确率的意义。

回应：感谢专家细致审稿。确如您言，我们在行为结果报告和作图中，标准差没有乘以 100%，在修改稿中已改正，行为结果图也重新修正(如下)，非常感谢您的提醒。我们的实验采用双

任务 RSVP 范式，同时用了导致经典注意瞬脱效应的时间间隔，准确率比较低很好地说明了本实验能够重复之前实验所产生的注意瞬脱效应。



意见 5：由于行为结果准确率低，而 ERP 分析剔除了错误反应的 Trial，请具体报告每种条件下剩余多少 Trial 纳入 ERP 分析。并解释分析这个重复次数是否足够。

回应：感谢审稿人的提示。在修改稿的方法-脑电数据的记录、处理与分析部分，我们报告了“每名被试每个条件下有效叠加试次数均超过 50 次。”每名被试在每种实验条件下保留的试次数目具体如下：具体正性词=53±18 个试次，具体中性词=58±22 个试次，具体负性词=61±20 个试次，抽象正性词=60±18 个试次，抽象中性词=51±18 个试次，抽象负性词=63±22 个试次。由此可见所有条件下最终脑电分析时有效叠加次数均超过 50 次，这个重复次数在 ERP 分析能够接受的范围。

意见 6：讨论部分存在两个问题：首先，止步于对结果的讨论而缺乏对研究发现的意义的深入讨论，其次，应该充分讨论本研究存在的缺陷。

回应：感谢审稿人的意见。本研究发现词汇具体性影响情绪名词加工的晚期阶段，LPC 波幅能够分辨出不同情绪效价的抽象词，对于具体词仅能区分出情绪与非情绪，说明抽象词可能比具体词负载了更多的情绪信息。这结果与情感具身认知理论突出情绪与抽象概念间联结的特点一致。因此在讨论最后部分对情感具身认知理论的探讨，以及引用些抽象概念表征的研究多数采用的 fMRI 研究，是希望对未来词汇具体性研究提供更多技术方法的思路。

词汇的获得年龄(age of acquisition, AOA)，以及词汇的获得方式(mode of acquisition, MOA)，都是影响词汇加工过程的重要因素。感谢审稿专家的提示，考虑到词汇加工过程的复杂性，未来研究我们将思考从 AOA 和 MOA，及其他影响词汇识别的因素入手进行深入探讨。

第二轮

审稿人 1 意见：

作者已对我提出的意见进行了很好的修改和回复，推荐发表。

回应：诚挚感谢专家老师对稿件的宝贵建议。

.....

审稿人 3 意见：

对摘要的修改与正文更加符合：“词汇类型”修改为“词汇具体性”也更加明确；前言中“个体对情绪信息…重要的调节与促进作用”调节作用去掉了促进作用，更加简练；对于词汇加工的情绪效应的 ERP 成分的论述表述更为细致，补充了部分数据，论文修改基本达到《心理学报》的发表标准，文章个别字句尚需审核，建议修后发表。

回应：诚挚感谢专家老师对稿件的宝贵建议，我们再次通读全文进行自查，把一些语句修改为更加准确、严谨的表述。