

## 《心理学报》审稿意见与作者回应

题目：知觉负荷对面孔多重社会范畴加工的影响

作者：王冠 周霏 张凯莉 王沛

---

### 第一轮

**审稿人 1 意见：**本文采取侧抑制范式，比较了面孔种族和性别特征的加工特点，揭示了二者的加工过程分别具有自下而上和自上而下性。实验设计严密，结果客观，整体逻辑基本合理。现有以下建议：

**意见 1：**实验 1，实验 2a 的材料筛选，需要提供数据结果。

**回应：**非常感谢审稿专家专业的审稿意见！在实验 1 和实验 2a 中，作为目标刺激，除真名之外总共制作了 72 个假名，由假字和生僻汉字组成。在前侧实验材料包括 72 个假名和 24 个真名（6 个欧洲男姓名和 6 个中国男姓名，6 个欧洲女姓名和 6 个中国女姓名），混合在一起要求被试通过 7 分量表评价真假名的程度。7 为最像真名，1 为最像假名。结果显示，72 个假名的均值为  $M=6.21$ ， $SD=0.78$ ，24 个真名的均值为  $M=1.64$ ， $SD=0.75$ 。配对样本  $t$  检验的结果为  $t(23)=17.47$ ， $p<.001$ ，Cohen's  $d=5.97$ 。这说明真名假名的区别明显，材料操纵有效。同时，男性姓名和女性姓名配对样本  $t$  检验的结果为  $t(11)=.081$ ， $p=.937$ ，Cohen's  $d=.048$  表明男名和女名的真名识别度无差异。欧洲姓名和中国姓名的配对样本  $t$  检验的结果为  $t(11)=.102$ ， $p=.885$ ，Cohen's  $d=.044$ ，表明中欧姓名的真名识别度无差异。对此我们均已在文章中做了补充报告（以黄色标明。其他修改内容亦然。）。

**意见 2：**个别语句联结逻辑跳跃过大或存疑，比如：

“……当名字本身的性别属性不典型的时候，如果面孔性别自动被加工，那么必定对中性名字的性别判断应该会倾向于面孔一致。这样更能进一步证明性别的自上而下的选择性注意加工”；

“……这似乎与 Bruce 和 Young(1986)的面孔信息加工模型相符，认为面孔性别的加工要发生在面孔结构编码之后的注意加工阶段，是自上而下的加工……”。

**回应：**非常感谢审稿专家的非常有价值的建议！我们已经在文章中进行了修改。保证逻辑通畅，并且对面孔信息加工模型与性别加工做了详细的阐述。

**意见 3：**一些观点值得商榷，比如说性别加工是自上而下的加工。性别加工会受到自上而下的注意调节，但不等于就说它是自上而下的加工。

**回应：**非常感谢审稿专家的提醒！的确有些研究者对某些信息加工具有“自上而下的效应”这一观点持保留态度。他们认为，信念、情绪、欲望和动机等状态虽然受到注意和认知资源的影响，但是并不会自上而下地影响知觉过程(O'Callaghan, Kveraga, Shine, Adams, & Bar, 2016)，因此也就不存在自上而下加工效应。本文通过任务指导语，不让被试直接判断性别，而是判断看到的名字性别，从而有效地规避了面孔低水平特征对名字判断的影响。通过侧抑制范式发现高知觉负荷水平下被试对性别的判断不受面孔性别的影响，这一结果虽不能直接证明性别加工是自上而下的，结合低负荷状态下被试对姓名判断受到性别面孔影响这一结论，则基本可以说明注意资源的确影响了性别加工过程，因而不同于种族范畴的那种强制性的加工。也就是说，性别加工是自上而下受到注意调节，而不是一种自上而下的加工。为此，我们在文章中将之改为“性别加工受到自上而下的注意调节”。

**意见 4：**研究结果中发现了一些被试性别效应，但讨论中未做相应分析。

**回应：**非常感谢专家的建议！被试性别效应在实验 2a 和实验 2b 的主效应都很显著。有研究通过元分析发现被试性别的确对于面孔识别的影响非常显著。我们对被试性别的主效应在文中进行了讨论。

10 多年前，Sporer (2001)曾说过：很难讲清楚在多大程度上，跨性别效应是一个稳定的现象。但随着研究的日益加深，加之 fMRI 技术在心理学研究的普及，人们发现女性被试在观看与自己性别相同的面孔时，FFA (fusiform face area) 的激活程度要与男性被试观看男性面孔(Golarai, Liberman, Yoon, & Grill-Spector, 2009; Tahmasebi et al., 2012)。这被称之为女性的自我性别优势效应。也就是说，女性被试对女性面孔的加工具有优势效应。在一些元分析研究中，女性的自我性别优势效应在各种情形下，无论是在有无指导语还是注意/非注意条件下是相当可信的和明显的(Feinman & Entwisle, 1976; Lewin & Herlitz, 2002; Lov n, Herlitz, & Rehnman, 2011; Lov n, Sv rd, Ebner, Herlitz, & Fischer, 2014)，而男性性别优势则并不存在。因此，性别不同的被试在面孔刺激作为干扰刺激会对姓名的性别属性判断产生显著差异，这是因为由于女性被试对面孔加工的自我性别优势效应的存在。因此在实验 2a 中，当姓名和面孔同时呈现时，女性被试的注意资源更多的分配到面孔性别加工过程中，从而造成女性被试对姓名加工的反应时会长于男性被试(Herlitz & Lov n, 2013)。

**意见 5：**文章的创新意义还待强调突出。

**回应：**非常感谢专家的建议！我们文章的创新意义在于首次通过侧抑制范式发现，虽然性别和种族均作为原生社会范畴，但是它们具有不同的加工特征。种族加工是自下而上的强制性

加工，不受知觉负荷影响。性别加工是自上而下受到注意调节的灵活性加工。我们在文中对相关创新性结果进行了强调和突出。

**意见 6:** 文中还有一些错别字。

**回应:** 非常感谢评审专家的建议！我们多次多人对文章进行了校正，对错别字进行了更改。

.....

**审稿人 2 意见:** 该研究采用知觉负荷范式考察了对面孔种族和面孔性别的加工的特点。实验设想具有一定的创新性。然而还存在以下问题。

**意见 1:** 实验 1 的刺激和任务的设置有问题，使得结果不太可靠。目标刺激是长度为 2 个字或 3 个字的中国或欧洲男性名，被试的任务是对名字的种族尽快做出判断。首先，中国和欧洲男性名没有经过事前评定，而且都是 2 个字或 3 个字，作者也没有举例。所以有很大可能姓名的种族的操纵有问题。其次，因为刺激是 1,2,4,6 个名字（1 个真名，其余为假名），任务只是让被试对名字的种族尽快做判断。其中，真名假名并没有经过事先评定，假名可能并没有那么假。另外，被试可能会不理解任务。因为并没有让被试从这些名字中搜索出一个看起来像真名的名字，然后再对这一名字的民族做判断。而是只告诉被试对名字的种族做判断。因此，被试会不会对所有刺激中可能是真名的种族的总体做了判断。例如，在图 1 中，第 5 个“赵娜”应该是真名，但第 6 个（看上去是）“樊苜醛”也很像真名啊。实验 2a 和 2b 就对刺激进行了评价然后筛选。所以两个实验的结果差异是不是由实验一的刺激选取问题造成的？

**回应:** 感谢审稿专家的建议！我们在文章实验部分确实没有将指导语描述清楚，从而造成了审稿专家对实验的误解。实际上我们实验 1 的指导语是告诉被试名字矩阵里只出现一个真名（比如，李国勇，卡洛斯，托尼，王刚），要求被试找出真名并对该名字进行种族判断。我们对中国和欧洲男性名字（6 个欧洲男姓名和 6 个中国男姓名，6 个欧洲女姓名和 6 个中国女姓名）做了事先评定，结果是中国名字判断的正确率  $M=99.31\%$ ， $SD=2.532\%$ ，对欧洲名字判断的正确率  $M=98.26\%$ ， $SD=5.483\%$ 。 $t(23)=.827$ ， $p=.417$ ，**Cohen's  $d=0.246$** 。表明中国名和欧洲名区别显著，不存在名字对种族判定可能带来的混淆。其次，我们也对实验材料的真名和假名做了区分度评定。除真名之外总共制作了 72 个假名，由假字和生僻汉字组成。在前侧实验材料包括 72 个假名和 24 个真名（6 个欧洲男姓名和 6 个中国男姓名，6 个欧洲女姓名和 6 个中国女姓名），混合在一起要求被试通过 7 分量表评价真假名的程度。7 为最像真名，1 为最像假名。结果显现，72 个假名的均值为  $M=6.21$ ， $SD=0.78$ ，24 个真名的均值为  $M=1.64$ ， $SD=0.75$ 。配对样本  $t$  检验的结果为  $t(23)=17.47$ ， $p<.001$ ，**Cohen's  $d=5.97$** 。说明真名

假名的区别明显，名字材料操纵有效。相关内容已经补充到了文章中。

**意见 2:** 实验 2a 和 2b 的任务是什么？仍然是让被试判断名字的种族？应该是判断名字的性别，不同于实验一吧。如果是后者，实验 1 和实验 2 的被试的任务就不同，因此很有可能两个实验的结果差别是由于主任务的难度引起的（从表 1 和表 2 来看，实验 2 的对名字的性别判断时间显然长于实验 1 的对名字的种族判断的时间。高负荷时，主任务高难度要比主任务低难度更不易受到 flanker 刺激的影响），而不是对干扰面孔的种族加工或性别加工引起的。

**回应:** 非常感谢专家细致的审阅于宝贵的意见！实验 2a 和 2b 确实是判断名字的性别，我们已经在文中进行了补充说明。后面关于主任务难度的问题，经过我们细致的比对和计算（见下图），发现表 1 和表 2 无论是在一致条件或者不一致条件下的反应时数据并不存在显著差异。因而可以将任务难度的可能性排除掉。

负荷	一致条件( <i>SD</i> ) (种族/性别)	不一致条件( <i>SD</i> ) (种族/性别)
1	635.3(51.59)/ 624.44(72.83)	662.9(62.38)/ 660.68(79.75)
2	738.77(76.07)/ 710.91(75.29)	765.00(85.83)/ 738.93(83.54)
4	974.49(93.30)/ 938.75(119.15)	1004.13(97.24)/ 954.70(132.90)
6	1188.87(106.26)/ 1168.32(109.32)	1207.41(119.51)/ 1151.61(116.64)

评审专家提出的这一建议提醒我们在图表的制作过程中务必要前后保持风格一致，这样不仅有利于专家的审阅，也有利于读者的阅读。再次衷心感谢审稿专家的指导！

**意见 3:** 实验 2a 中“所有名字之前经过评价，真名性别特征明显，假名无意义”的具体结果是怎么样的？比如被试的任务是什么？几点量表？这些名字的取值分别是多少？

**回应:** 感谢审稿专家提出的评审意见！实验 1 和实验 2a 中，作为目标刺激。除真名之外总共制作了 72 个假名，由假字和生僻汉字组成。在前侧实验材料包括 72 个假名和 24 个真名（6 个欧洲男姓名和 6 个中国男姓名，6 个欧洲女姓名和 6 个中国女姓名），混合在一起要求被试通过 7 分量表评价真假名的程度。7 为最像真名，1 为最像假名。结果显示，72 个假名的均值为  $M=6.21$ ， $SD=0.78$ ，24 个真名的均值为  $M=1.64$ ， $SD=0.75$ 。配对样本  $t$  检验的结果为  $t(23)=17.47$ ， $p<.001$ ，Cohen's  $d=5.97$ 。说明真名假名的区别明显，材料操纵有效。对此我们在文中予以了补充说明。

**意见 4:** 为何每个实验中都把被试性别当作一个自变量？有何理论意义？

**回应：**非常感谢专家提出的问题！被试性别在面孔识别过程中存在性别偏向（own-gender bias）。许多研究已经报告过女性在面孔识别任务中的表现要优于男性(De Frias, Nilsson, & Herlitz, 2006; Herlitz, Reuterskiöld, Lovén, Thilers, & Rehnman, 2013; Wahlin et al., 1993), 但是也有一些研究认为, 女性被试只对女性面孔识别具有优势, 男性被试只对男性面孔具有识别优势(Herlitz & Lovén, 2013)。由此可见, 被试性别在面孔加工过程中具有重要影响。因此, 我们将被试性别作为一个自变量进行了考察。而且这也是首次采用侧抑制范式, 考察被试性别不同, 面孔图片如何影响对姓名性别和种族的判断。这对于补充和发展自我性别偏向理论具有重要意义。对此 我们在文中予以相应的补充与说明。

**意见 5：**前两个实验的正确率的结果是怎样的？

**回应：**非常感谢专家提出的问题！实验 1 的平均准确率为 94.9%。实验 2a 的平均准确率为 93.8%。参与这两个实验的被试正确率均超过了 85%。

**意见 6：**表和图内容一样的话呈现一个即可。如, 表 1 和图 1; 表 2 和图 2。

**回应：**非常感谢审稿专家的建议！图可以让读者看到一种明显的趋势, 一目了然。而表可以展现数据的精确性和完整性, 特别是反应时数据本来的要求就很精确。所以, 尽管图和表有些地方内容确实存在重复, 但经过权衡还是保留了图表。

**意见 7：**字号不一致和错字问题。如, 摘要中的“并且”; 全文正文为五号字, 然而 3.1.1 部分为 11 号字; 引言第二段“大理研究表明”。

**回应：**非常感谢审稿专家的建议！我们通过多人多次对文章的文字进行了校正, 并且在排版上严格遵守了期刊的相关要求。

**意见 8：**各个实验的简单效应检验也应该报告 p 值。

**回应：**非常感谢审稿专家的建议！我们在文中的简单效应中补充了 p 值。

**意见 9：**文后文献中要注意姓名缩写之间有空格, 题目首字母小写, 应加上期刊名称。其余期刊名称要注意首字母大写。

**回应：**非常感谢专家的细致阅读！我们再次对文献进行了标准化处理。

**意见 10：**英文摘要最好请人修改一下。

**回应：**非常感谢专家的建议！我们请国外朋友对英文摘要做了修改。

.....  
**审稿人 3 意见：**“知觉负荷对面孔多重社会范畴加工的影响”选题比较有意义，是面孔加工研究的热点问题，但是本文尚存在一些问题。

**意见 1：**摘要中研究背景没有交代清楚。

**回应：**非常感谢专家的建议！我们在摘要中已经对研究背景进行了补充。

**意见 2：**实验 1 没有发现一致性与知觉负荷的交互作用，那是怎么得出“即面孔种族引发的干扰效应会随着负荷水平的提高而有所增强”这个结论的？

**回应：**非常感谢专家的建议！这确实是我们表述不清导致的问题。其实在摘要中，我们想要表达的是种族对面孔的干扰效应在任何一种知觉负荷水平下都存在。种族范畴的加工无论是在高知觉负荷还是低知觉负荷条件下，都会影响被试对姓名种族的判断，造成反应时的延长。对此我们已经在文章摘要中进行了修改。

**意见 3：**前言中关于文献提到最多的是面孔情绪与知觉负荷的那研究，较少提到知觉负荷与性别、种族的研究，作者提到这方面的研究比较少，那么，作者可以适当加一些注意与面孔性别、种族的研究，从而提供更多的研究背景。

**回应：**非常感谢专家的建议！我们在文中补充了下述相关背景：

面孔识别理论(Bruce & Young, 1986)主张大脑加工原生性社会范畴（比如性别，种族）和继发性社会范畴（比如情绪）的神经机制是分离的。Hailing 等人（2017）为了验证这一观点，采用特征选择性注意任务，以事件相关脑电位的 SN（selection negativity）为指标，通过不同任务要求被试注意面孔的性别和情绪。结果发现，性别引发了明显的 SN。这一结果表明：在选择注意过程中，个体对面孔性别的加工更加敏感；同时也表明面孔原发性范畴与继发性范畴具有不同的神经加工机制(Wang, Ip, Fu, & Sun, 2017)。此外，前人关于面孔种族对选择注意的影响主要是考察外种族偏向效应(Ito & Urland, 2005; Park, van Bavel, Hill, Williams, & Thayer, 2016; Wan, Crookes, Reynolds, Irons, & McKone, 2015)，并没有从一般水平上考察种族作为原生社会范畴如何影响选择性注意。比如，有研究让被试检测附加在种族面孔上的字母串中较大的一个字母。发现白人被试在对黑人分心面孔上的字母反应时准确率要低于白人分心面孔。同样地，黑人被试在对白人分心面孔上的字母反应时的准确率要低于黑人分心面孔。而这一外种族偏向可能是由于刻板印象造成的选择注意优势。还有研究尝试通过实验任务比如让被试判断图片中的个体是否喜欢不同类型的蔬菜(Wheeler & Fiske, 2005)或者检测

图片上的点(Macrae, Bodenhausen, Milne, & Ford, 1997)来消除刻板和偏见, 结果表明面孔原生社会范畴的加工(种族)并不是独立的加工系统, 与面孔结构性编码也不是顺序加工的关系(Park et al., 2016)。上述关于种族范畴研究均是考察了个体对外种族注意偏向效应, 迄今为止我们还不清楚种族作为原生性社会范畴在一般意义上的加工特征。所以, 作为两种不同的神经加工机制, 之前的研究主要考察了情绪等继发性范畴在不同知觉负荷下的加工特征, 而对面孔性别或种族等原发性社会范畴在不同知觉负荷下的加工特征考察不足。

**意见 4:** 前言中作者提到“而且在面孔呈现后 45ms-85ms, 中央顶叶部出现了面孔性别判断效应。”这个研究结果的意义是什么, 与本研究之间存在怎样的关系?

**回应:** 非常感谢专家的意见! 这句话的初衷是想说明有研究认为面孔性别的加工并不在面孔的结构性编码之后, 也可能早于(45-85 ms)或不晚于面孔的结构性编码过程(170ms 左右)。目的是想告诉读者对面孔性别这一社会范畴的加工由于研究范式的不同, 其结果也是不统一的。的确这句话放在这里比较突兀, 影响读者理解, 因此我们删除了此句。

**意见 5:** 前言倒数第二段“例如, 当被试观看不同种族和年龄的面孔时(不要求被试对面孔性别和种族做出判断), 不同种族引发的 ERP 波形差异出现在 120ms, 而性别的效应稍晚, 出现在 180ms 左右(Ito & Urland, 2003)”前面提到是不同种族和年龄的面孔, 结果是种族和性别的内容, 请再次确认原文献。紧接上句提到“他们进一步发现最早的种族效应出现在 N100 (潜伏期 120ms 左右)。”如何进一步发现的不清楚。在之后提到结果解释“这可能是由于种族相比性别更加依赖于面孔整体特征的加工、涉及更复杂的结构信息(Olk & Garay-Vado, 2011; Schiltz & Rossion, 2006; Wiese, Kloth, Güllmar, Reichenbach, & Schweinberger, 2012)”, 这里的种族指的是本族和异族面孔更依赖整体加工吗?

**回应:** 非常感谢审稿专家的宝贵意见! 对此, 我们在文中进行了如下更改。

有研究者通过采用 mental oddball 范式, 采用 ERP 技术, 要求被试对戴眼镜的面孔计数, 并在不同实验条件下改变面孔的性别和年龄特征, 结果发现, 无论哪种条件下, 都会出现 N170 效应, 也就是表明性别和年龄知觉加工是类似的(Ito & Urland, 2003)。再此基础上, Ito 和 Urland 等人采用 ERP 技术, 通过系列面孔呈现任务, 每一试次出现 5 张面孔, 其中一张是目标面孔, 要求被试根据任务要求对面孔的种族或性别进行反应, ERP 记录目标面孔出现时的脑电位, 其他非目标图片出现的脑电位作为背景。结果发现对黑人面孔的反应, 无论在何种条件下 120ms 左右都会出现明显的负波, 因此被称之为 N100 效应, 对男人面孔的反应, 无论在何种条件下 180ms 左右都会出现明显的正波, 因此被称之为 P200

效应 (Race and Gender on the Brain: Electrocortical Measures of Attention to the Race and Gender of Multiply Categorizable Individuals)。

由于前人的研究均是考察的白人对外种族的面孔加工，往往看成是一种外群体加工偏向。因此关于种族这一范畴加工是本族和异族面孔加工均依赖整体加工并不确定。这也是本研究拟考察的问题之一。通过对面孔种族与姓名的一致不一致处理，规避了被试本身种族范畴带来的可能影响（本实验中，无论被试是什么种族，均可以分析种族与姓名的一致和不一致带来的差异），进而可以考察种族这一范畴加工在不同知觉水平下的一般特征。

**意见 6:** 前言最后一段预测“具体来说，性别可能更依赖于面部整体加工，而种族信息加工的开始阶段更早，与性别相比加工强制性更高”怎么通过前面的内容推出这个预测不清楚。“据此本文提出假设：种族加工相比性别加工，发生得更早”发生更早的假设将在结果上怎样表现不清楚。总之，最后一段关于研究结果的预期内容不清楚

**回应:** 非常感谢审稿专家提出的疑问！根据前文的论述，我们初步发现，种族和性别范畴的加工时间在不同任务范式下存在不同。但是整体来看，性别加工更倾向于整体结构性加工，发生时间晚于面孔结构性编码阶段，符合功能性面孔识别理论。而种族加工则更早一些（在 150ms 之前，也就是面孔结构性编码之前），也就是前注意阶段，因此种族加工和彭小虎等人（2002）的观点一致。所以，我们提出假设：种族相对于性别加工可能自动化程度或者强制性程度更高。基于此，本实验认为面孔种族的加工发生在前注意阶段，不受知觉资源负荷的影响。面孔性别的加工发生在结构性编码之后，受到知觉资源负荷的影响。

**意见 7:** 每个实验后请添加对该实验的简要讨论，同时实验 1 与实验 2 之间的关系不清楚。

**回应:** 非常感谢评审专家的宝贵意见！根据建议，我们在每个实验结束都对实验进行了讨论。实验 1 是采用侧抑制范式，考察在不同知觉负荷条件下，对姓名的判断如何受到旁边的种族面孔的影响。实验 2 是采用相同的范式，考察在不同知觉负荷条件下，对姓名的判断如何受到旁边的面孔性别的影响。整个实验的目的试图探讨不同的知觉负荷如何影响对面孔多重社会范畴的加工，也就是说这些社会范畴是自动化加工的还是受到注意资源调节。

**意见 8:** 实验 1 中为什么要把被试性别作为统计自变量，前言中没有相关文献提到被试性别在种族加工中的作用，同时后面有一段统计又把该变量去掉的理由也不清楚。

**回应:** 感谢评审专家提出的宝贵建议！被试性别在面孔识别过程中存在性别偏向 (own-gender bias)。许多研究已经报告过女性在面孔识别任务中的表现要优于男性 (De Frias et al., 2006;



Herlitz et al., 2013; Wahlin et al., 1993), 但是也有一些研究认为, 女性被试只对女性面孔识别具有优势, 男性被试只对男性面孔具有识别优势(Herlitz & Lovn, 2013)。由此可见, 被试性别在面孔加工过程中具有重要影响。因此, 我们将被试性别作为一个自变量进行了考察。而且这也是首次采用侧抑制范式, 考察被试性别不同, 面孔图片如何影响对姓名性别和种族的判断。这对于补充和发展自我性别偏向理论具有重要意义。由于被试性别的主效应不显著, 因此我们为了更加简洁明确的考察有显著效应的变量, 就将被试性别变量去掉。我们在文章中也做了相应的补充说明。

**意见 9:** 实验 1 和 2 之间没有做统计分析对比, 怎么说明二者分别表征自上而下和自下而上的内容呢?

**回应:** 非常感谢专家提出的宝贵意见! 我们通过操纵不同的知觉负荷条件, 初步发现种族对姓名判断的影响, 无论在何种知觉负荷条件下都存在。而性别对姓名判断的影响, 只有在低知觉负荷条件下存在, 在高知觉负荷条件下不存在。这说明种族范畴加工是自下而上的强制性加工, 而性别加工受到注意资源的调节, 体现了一种自上而下的灵活性加工特征。

**意见 10:** 为什么实验 1 和 2a 只统计反应时, 而不考虑正确率?

**回应:** 非常审稿专家的宝贵意见! 因为该任务是从反应时的角度来考察不同知觉负荷水平下, 性别和种族对姓名判断影响的特征。首先实验设计比较简洁, 并不要求被试很高的工作记忆能力或者反应能力, 因此被试在这两个试验中的准确率都相当的高, 实验 1 的平均准确率为 94.9%。实验 2a 的平均准确率为 93.8%。其次, 实验 2b 统计正确率的目的并不是验证本文的主要假设, 而是从准确率的角度来证明中性名字操纵是有效的。因此, 实验 1 和实验 2a 并没有对正确率进行深入的统计分析。

**意见 11:** 实验中面孔名字图片呈现时间不清楚, 是一直呈现直到被试按键消失还是有呈现时间限制。

**回应:** 非常审稿专家的宝贵建议! 的确, 我们没有明确交代面孔和名字图片呈现时间的问题, 只是在流程图上标注了“按键消失”。在该实验中, 我们的指导语均是要求被试尽可能快速且准确的对刺激做出反应(按键消失)。并且, 我们以平均反应时 3 个标准差以外的数据作为剔除标准。我们在文章补充了详细说明。

**意见 12:** 实验 2 中发现相关交互作用说明了什么，可在该结果后做简要描述。

**回应:** 非常感谢专家的宝贵意见!在实验 2 中，我们在每个小实验的结果部分对相关的交互作用进行了简要的描述。

**意见 13:** 讨论中先把实验主要结果描述一下，然后针对结果逐个展开讨论，内容上会更清晰

**回应:** 非常感谢审稿专家的建议!我们按照评审意见，在讨论部分首先将实验主要结果进行了描述。

**意见 14:** 讨论中提到“且相比于种族判断，认知者更难以对倒置面孔进行性别判断，这说明这些加工过程的发生相对较晚”，怎样得出加工过程发生较晚的结论?

**回应:** 非常感谢专家提出的意见!根据面孔识别理论，个体首先会对面孔的某一特征比如鼻子，嘴巴或者眼睛进行加工，其次会对面孔的各个特征进行结构编码，最后会对面孔进行整体性加工。倒置面孔给认知者结构编码和整体加工带来了难度，所以个体难以对倒置面孔进行性别判断，表明性别加工需要整合各个面部特征，然后进行编码，才能做出最后的判断。因此相比于依靠某种面部特征的种族范畴加工来说，性别范畴加工时间确实发生较晚。

**意见 15:** 结论提到“本研究结果表明，作为重要的社会性信息，面孔加工具有特殊性”作者是如何得出面孔加工具有特殊性的不清楚。

**回应:** 非常感谢专家的意见!面孔加工的特殊性表现为同样作为原生社会范畴的性别和种族，大脑对其加工特征却不尽相同。时至今日，种族作为一种人类进化过程中可能会带来威胁的一种社会范畴信息，其重要性不言而喻，因此种族更可能是一种自动化的强制性加工。然而，在全球倡导男女平等的浪潮愈演愈烈的今天，女性地位稳步提高，能力得到广泛认可，在这一文化观的浸润下，许多国家与地区人们对性别之间的差异敏感性已渐渐降低。因此，虽然性别同样是一种原生社会性范畴，但是并不是大脑优先加工的对象，表现为一种受注意资源调节的自上而下的灵活性加工。这也就是我们所主张的多重社会范畴下面孔加工的特殊性。

**意见 16:** 作者给出的面孔例子带有发型这些外部轮廓，而这些信息对于面孔性别的影响较大，如果没有这些外部信息是否还能得到相似的结果呢?

**回应:** 非常感谢专家的问题!面孔性别的加工确实会受到发型等这些外部轮廓的影响。不过这些外部轮廓作为一种启发式，只是缩短了性别判断的时间，并不会因为发型的存在对性别判断的准确率产生影响(Goshen-Gottstein & Ganel, 2000)。也就是说发型等外轮廓并不会改变性别加工的认知特征。而且本实验并不是要求被试对面孔性别做出判断，而是对名字的性别

做出判断，所以被试在侧抑制任务范式下不会刻意去加工发型等外部线索。此外，如果将发型去掉，那么其生态效度则会大大受到质疑。所以综合考虑，本实验的刺激材料均是完整的面孔信息。并没有进行特殊处理。

#### 小问题

1. 文中字体大小不统一，比如，摘要中“并且被试会倾向于将中性名字的性别判断为与面孔的性别一致”，并且两个字与其他字体大小不一致。

回应：非常感谢评审专家的细致审阅！我们已经通篇进行了文字校正。

2. 文中参考文献引用格式不当，如“Goshen 和 Ganel 的研究发现：面孔性别的加工对枕颞区 N170 没有影响(Goshen-Gottstein & Ganel, 2000)”，请通篇检查修改。

回应：非常感谢评审专家的意见！根据专家的提醒，我们严格按照心理学报的书写规范对文中类似的问题进行了修改。

3. 语句不通顺、句子间缺少连接词、书面语表达不强、标点符号使用不当，如“。而对高兴表情面孔的加工则没有影响。”“所以，在知觉负荷条件下，进行探讨面孔种族加工和面孔性别加工的研究很有必要。”“被试性别的主效应显著， $F(1,24)=4.745$ ， $p=0.039$ ， $\eta^2=0.165$ ，女性被试更倾向于将得出名字性别与面孔性别一致的判断。”“以反应时为因变量，以将被试性别（男性&女性），面孔性别（男性&女性）以及知觉负荷(1,2,4,6)做重复测量方差分析”“说明在本实验中面孔的种族信息的加工一定程度上独立于知觉负荷。表明种族信息早在意志调控之前的前注意阶段便发生了由刺激驱动的，自上而下的加工。因而不受之后意志调控与知觉负荷的影响”“说明面孔性别的加工是受自上而下的意志驱动的，具有灵活性”文中还有很多这样的句子，请仔细检查

回应：非常感谢评审专家的意见！文章的通顺性和意义可达性是一篇质量达标的学术论文基本条件，我们在专家的建议下，通过多人多次对文章进行了挑剔性阅读，对文章中存在的语句不通问题，书面语表达和标点符合等问题进行了细致修改。

4. “进一步的研究表明，面孔种族与性别信息的加工在时间上可能存在一定的差异。”这句话开始的本段内容可以考虑与上段合并

回应：非常感谢评审专家的建议！经过对文章仔细阅读和通篇考量，评审专家的建议确实对于理解文章有很大帮助，因此我们将这段话和上段合并，并且对语序进行了适当的调整，保证了内容通顺和清晰。

5. 结果图中未标注各个条件的标准误，表 1 和图 1 在正文中的位置亦未标注出来

回应：非常感谢评审专家的建议！我们在相应位置补充了各个条件的标准误。对于表 1 和图 1 的位置进行了补充说明。

6. “12 男，14 女， $M=23.4$ ， $SD=2.45$ ”请加上 M 和 SD 相应的单位

回应：非常感谢评审专家对文章细节的建议！我们在相应位置补充了 M 和 SD 的单位。

7. 请添加实验设备屏幕刷新率

回应：非常感谢专家的建议！实验设备的屏幕刷新率均为 60HZ，对此我们在文中进行了补充。

8. 三个实验程序上、刺激材料上有很多重复内容，建议精简，后面的实验只说明与前面实验的不同之处即可

回应：非常感谢评审专家的建议！我们在文章对实验程序和刺激材料的介绍进行了精简，突出了与前面实验的不同之处，从而有助于读者更好的理解三个实验的异同和内在的逻辑关系。

---

## 第二轮

审稿人 1 意见：

意见 1: 本轮修改论文质量有较明显的提升。但行文规范和逻辑严密精练方面，还可以进一步改善。

回应：非常感谢审稿专家的建议！我们根据主题再次阅读修订了全文，在行文规范和逻辑方面进一步进行了提高性修改。

审稿人 3 意见

意见 1: 摘要第一句“通过两个实验探讨了面孔姓名判断是否会在不同的知觉负荷条件下受到面孔性别与种族信息的影响。”依旧没有清晰表明本研究的目的何在，这里只说明了本研究的实验方法。

回应：非常感谢审稿专家的提醒和建议！本文的主要研究目的是为了考察知觉负荷理论对于面孔多重社会范畴加工机制的解释的可能性。我们已经在摘要中做出了补充。

意见 2: 正文参考文献格式不正确，如“Neumann 等人采用重复启动范式对陌生面孔进行了

考察，也发现了类似的结果(Neumann, Mohamed, & Schweinberger, 2011)，”后面还有很多类似的表达，请更正。有些正文参考文献漏掉年代，如“Yang、Wang、Jin 和 Li 通过操纵视觉短期记忆(visual short-term memory, VSTM)负荷的高低，”请通篇检查其他不当地方。“Hailing 等人(2017)为了验证这一观点，采用特征选择性注意任务，以事件相关脑电位的 SN(selection negativity) 为指标，”这里参考文献引用姓而不是名，请检查其他地方是否存在类似错误。

回应：非常感谢专家的细致审阅！我们已对全文的参考文献的引用格式进行了积极全面的修正。

意见 3：“面孔识别理论(Bruce & Young, 1986)主张，大脑加工原生性社会范畴（比如性别、种族等）和继发性社会范畴（比如情绪）的神经机制是分离的。”请标明这里的原生性和继发性社会范畴英文是如何表达。

回应：非常感谢审稿专家细致和规范的审阅！原生性社会范畴的英文表达为“Primary Social Category”。继发性社会范畴的英文表达为“Posterior Social Categories”。

意见 4：文中有些词使用不当，“外种族”，可以直接用异族来表达；“inattention blindness”指的是非注意盲视。

回应：非常感谢审稿专家对文章细致入微的提醒和建议。我们将全文的“外种族”改为了“异族”，并将“inattention blindness”翻译为“非注意盲视”。

意见 5：“如上所述，我们初步发现，种族和性别范畴的加工时间存在不同：性别加工更倾向于整体结构性加工，发生时间晚于面孔结构性编码阶段，符合功能性面孔识别理论；种族加工则更早一些（在 150ms 之前，也就是面孔结构性编码之前）”这里的文献出处是什么？

回应：非常感谢审稿专家提出的建议。这句话是对上文的相关内容的总结，提出主要的争论就在于 Bruce 和 Young（1986）提出的面孔识别理论与彭小虎等人（2002）的修订模型两者所主张的面孔加工机制不一致。我们在文中进一步引用了相应的主要文献，从时间维度上支持面孔性别和种族加工的特征。

意见 6：“无疑是直接探讨多重原发性社会范畴加工特征的绝佳途径”，绝佳这个词使用过于完美，建议用其他词来代替。

回应：非常感谢审稿专家提出的进一步完善的建议。我们也认为“绝佳”一词的确过于强调完

美，不符合科学精神。鉴于我们想强调的是采用该范式可以很好的解决这一争论，因此我们将该词替换为“有效”。

**意见 7:** “性别偏向 (own-gender bias)”这里准确理解应该是同性别偏向，而不简单是性别偏向。

**回应:** 衷心感谢审稿专家给出的极具严谨性的建议。Own-gender bias 在相应的研究中的的确是代表的同性别偏向。我们已经将文章中的 own-gender bias 翻译为“同性别偏向”。

**意见 8:** “在前侧实验材料包括 72 个假名（由假字和生僻汉字组成）和 24 个真名（6 个欧洲男姓名和 6 个中国男姓名，6 个欧洲女姓名和 6 个中国女姓名）”说明真假名数目存在这么大差异的目的。

**回应:** 衷心感谢专家提出的建议。我们是根据当前实验设计而采取的真假名数量。本实验的知觉负荷水平分为 1,2,4,6 四个条件。由于每个条件只出现一个真名，因此，除去负荷水平为 1 的情况下，其他三个条件下真假名的比值应保证以下等式的要求：

$$\frac{1}{1+1+1+1} = \frac{\text{假名 a}}{\text{真名}} \quad \text{公式 1:}$$

$$\frac{2+4+6}{1+1+1} = \frac{\text{假名 b}}{3 \text{ 真名}/4} \quad \text{公式 2:}$$

$$\text{假名} = \text{假名 a} + \text{假名 b} \quad \text{公式 3.}$$

在本研究中，如果真名的数量确定是 24，那么假名的数量应该保持在 72 左右。这就是真假名数目存在较大差异的原因。

**意见 9:** 结果显示, 72 个假名的均值为  $M=6.21, SD=0.78$ , 24 个真名的均值为  $M=1.64, SD=0.75$ 。”这与前面提到的 7 真名, 1 假名的描述不一致, 统计结果反而是真名比假名分值低。

**回应:** 非常感谢评审专家的细致审阅。由于我们的大意, 将结果汇报反了。正确的结果是 72 个假名的均值为  $M=1.64, SD=0.75$ , 24 个真名的均值为  $M=6.21, SD=0.78$ 。我们已经在文中给予了修改, 并保证了相关数据结果无误。

**意见 10:** 请在所有结果图中标注出标准误。

**回应:** 非常感谢审稿专家的建议。我们在文章的两个图中都标注了标准误。

**意见 11:** “结果表明无论知觉负荷高低, 被试对种族的判断总是会受到种族加工的影响。这说明了种族信息作为无关任务信息仍然影响到了选择性注意。”这句话理解起来比较困难, 建议在适当位置补充关键词, 改成“结果表明无论知觉负荷高低, 被试对姓名种族的判断总是会受到面孔种族加工的影响。这说明了面孔种族信息作为无关任务信息仍然影响到了选择性注意。”其他实验及总讨论中也存在同样的问题, 建议都做修改。

**回应:** 非常感谢审稿专家的积极建议。我们已经根据专家的建议, 同时结合实际情况, 为了进一步提升文章的易读性, 对全文进行了补充修改。

**意见 12:** “该结论验证了彭小虎等人(2002)的理论主张的种族等直接视觉处理的加工可能与结构性编码一起发生或者发生在比结构性编码更早的前注意阶段。这说明了种族作为原生性社会范畴, 具有一种自下而上的强制性加工的特征。”前后字体大小不一致。

**回应:** 非常感谢专家细致入微的审阅。我们已经将全文文字按规范格式进行了修订。

**意见 13:** “所有名字之前经过评价, 真名性别特征明显, 假名无意义”, 请举例说明真假名的特点。

**回应:** 非常感谢审稿专家的建议。在该实验(2a)中, 真名不仅看起来更像真名, 而且具有明显的性别特征, 比如女性真名“于晓丽, 江秋萍”, 男性真名“王刚, 张腾飞”。假名均无意义, 比如“沃缶, 三未, 芝士始”。我们在文中也相应做了举例补充说明。

**意见 14:** “负荷水平为 4 时不一致条件反应时长于一致条件, 但差异不显著,  $SE_{一致}=16.81$ ,  $SE_{不一致}=17.49$ ,  $p=0.065$ 。在负荷水平为 6 时, 不一致条件的平均反应时短于一致条件,  $SE_{一致}=21.81$ ,  $SE_{不一致}=20.29$ ,  $p=0.077$ , 也没有达到显著水平(表 2)”。差异不显著的结果没有意义, 仅是数字上存在大小差别。

**回应:** 非常感谢专家提出的建议。我们已经在文中补充了该实验简单效应下的各个  $F$ 、 $SE$ 、 $p$  和  $\eta^2$  的结果。让专家和读者一目了然。

“知觉负荷与一致性的交互作用显著(图 3),  $F(3,96)=31.195$ ,  $p=0.035$ ,  $\eta_p^2=0.88$ , 简单效应分析表明, 在负荷水平为 1,2,4, 不一致条件的平均反应时均长于一致条件, 其中负荷水平为 1 ( $SE_{一致}=13.01$ ,  $SE_{不一致}=13.19$ ,  $F(1,31)=30.576$ ,  $p=0.001$ ,  $\eta_p^2=0.613$ ) 和 2 ( $SE_{一致}=14.81$ ,  $SE_{不一致}=15.29$ ,  $F(1,31)=12.171$ ,  $p=0.002$ ,  $\eta_p^2=0.329$ ) 都达到了显著水平。负荷水平为 4 时不一致条件反应时长于一致条件, 但差异不显著, ( $SE_{一致}=16.81$ ,  $SE_{不一致}=17.49$ ,  $F(1,31)=1.592$ ,

$p=0.205$ ,  $\eta_p^2=0.063$ )。在负荷水平为6时,不一致条件的平均反应时短于一致条件,  $SE_{一致}=21.81$ ,  $SE_{不一致}=20.29$ ,  $F(1,31)=0.136$ ,  $p=0.713$ ,  $\eta_p^2=0.006$ , 也没有达到显著水平(表 2)。不同负荷水平下反应时的差异, 表明性别范畴加工受到了自上而下的注意效应的调节。”

**意见 15:** “当名字本身的性别属性不典型的时候, 当注意负荷较低时, 对中性名字的性别判断倾向于与面孔一致。当注意负荷较高时, 被试无法对面孔性别进行加工, 所以对中性名字的判断与面孔性别无关。”请标明参考文献出处。

**回应:** 非常感谢审稿专家的建议。我们已经在正文中标出了参考文献。本段话目的是为了引出进一步验证性别加工的自上而下受注意调节的特征的必要性, 因此在文中我们也进一步予以了突出和完善。

**意见 16:** “筛选出了 12 个评分在 4.5 到 5.5 之间的名字 ( $M=5.15$ ,  $SD=0.38$ ), 所有名字与 5 做单一样本 T 检验均无显著差异”, 请标明这个差异不显著对本实验的意义。

**回应:** 非常感谢审稿专家细致的建议。我们在文中补充了这个结果对本实验的意义。“这一结果说明该实验选取的名字不具有明显的性别特征”。

**意见 17:** “表明当注意资源耗竭时, 被试对中性名字的判断属于随机猜测水平。”这个实验结果前面标明 4 和 6 不存在差异, 怎么能说 6 是资源耗尽条件?

**回应:** 非常感谢审稿专家对本文提出的这条建议。前文表明的 4 和 6 不存在差异“指的是高知觉负荷主效应下的事后检验的结果。“表明当注意资源耗竭时, 被试对中性名字的判断属于随机猜测水平。”这句指的是知觉负荷和面孔性别的交互作用, 在高知觉负荷条件(4, 6)下, 当知觉负荷为 4 时, 由于女性面孔加工优势的存在, 女性面孔下被试对名字的准确率反应较男性面孔有一定的提高。当进一步把知觉负荷提高到 6 时, 被试的注意资源几乎完全耗竭, 女性面孔加工优势也随之消失, 无法对背景图片的性别性别加工, 因此对中性名字的判断属于随机猜测水平。由于我们的表述不当, 引起了理解上的困难, 我们已经在文中进行了修改。

**意见 18:** 建议表 3 中的 M 与 SD 用  $M\pm SD$  表示

**回应:** 感谢审稿专家的建议。的确使用  $M\pm SD$  的表述方式更显规范, 我们已经将表 3 进行了修改。



**意见 19:** “与种族相比，性别更加依赖于面孔特征整体加工，影响性别判断的面部线索主要有眼睛、下颌线轮廓、鼻子、眉毛以及下巴”，请标明文献出处

**回应:** 非常感谢审稿专家细致负责的评审意见。我们确实在陈述这一观点时忘记引用相关文献。我们已经在文章中补充了相关文献。

---

### 第三轮

**审稿人 3 意见** 部分需要作者在文中更正的内容，作者只在回复意见中提到有修改，但文中尚未修改请作者认真对待需要在文章中更正的内容，另外还有一些小问题需要修改。

**意见 1:** 面孔识别理论(Bruce & Young, 1986)主张，大脑加工原生性社会范畴（比如性别、种族等）和继发性社会范畴（比如情绪）的神经机制是分离的。”请标明这里的原生性和继发性社会范畴英文是如何表达。

**回应:** 非常感谢审稿专家细致和规范的审阅！原生性社会范畴的英文表达为“Primary Social Category”。继发性社会范畴的英文表达为“Posterior Social Categories”。我们已经将英文表达在正文中补全完整。

**意见 2:** 无疑是直接探讨多重原发性社会范畴加工特征的绝佳途径”，绝佳这个词使用过于完美，建议用其他词来代替。

**回应:** 非常感谢审稿专家提出的进一步完善的建议。我们也认为“绝佳”一词的确过于强调完美，不符合科学精神。鉴于我们想强调的是采用该范式可以很好的解决这一争论，因此我们将该词替换为“有效”。

**意见 3:** 结果显现，72 个假名的均值为  $M=6.21, SD=0.78$ ，24 个真名的均值为  $M=1.64, SD=0.75$ 。”这与前面提到的 7 真名，1 假名的描述不一致，统计结果反而是真名比假名分值低。

**回应:** 非常感谢评审专家的细致审阅。由于我们的大意，将结果汇报反了。正确的结果是 72 个假名的均值为  $M=1.64, SD=0.75$ ，24 个真名的均值为  $M=6.21, SD=0.78$ 。我们已经在文中给予了修改，并保证了相关数据结果无误。我们已在文中进行了纠正。

**意见 4:** 为了考察知觉负荷理论对于面孔多重社会范畴加工机制的解释的可能性, 通过两个实验探讨了姓名判断是否会在不同的知觉负荷条件下受到面孔性别与种族信息的影响。”这句话缺少主语。

**回应:** 感谢专家非常负责的建议, 我们已经在文中改为: “为了考察知觉负荷理论对于面孔多重社会范畴加工机制的解释的可能性, 我们采用两个反应时实验探讨了姓名判断是否会在不同的知觉负荷条件下受到面孔性别与种族信息的影响”。

**意见 5:** 如上所述, 我们初步发现, 种族和性别范畴的加工时间存在不同: 性别加工更倾向于整体结构性加工, 发生时间晚于面孔结构性编码阶段, 符合功能性面孔识别理论; 种族加工则更早一些 (在 150ms 之前, 面孔结构性编码之前) (Ito & Urland, 2003), ”前面是”我们初步发现”, 后面又引用”Ito”, 这里的结果是你们发现的, 还是引用文献的结果, 表述不清楚。

**回应:** 非常感谢专家对文章提出的建议。的确是我们没有表述清楚。这句话实际是对前文综述的一个概括性表述, 并不是我们的发现, 而是我们接下来通过研究试图要解决的问题。Ito 和 Urland (2003)的研究在前文已经有了表述, 放在该句中反而起到了影响读者理解的作用, 因此我们将此引用文献去掉, 并且重新组织了语言。

**意见 6:** 实验采用 2 (被试性别: 男/女)  $\times$  2 (名字与面孔种族信息的一致性: 一致 vs. 不一致)  $\times$  4 (知觉负荷水平: 1, 2, 4, 6) 被试内设计” 每个被试只有一个性别, 怎样进行被试内设计

**回应:** 感谢专家的严格审阅和积极的建议。我们的确出现了严重的笔误, 实际研究是一个重复测量两个因素的三因素混合设计。我们在文中给予了纠正。

**意见 7:** 文中表格的表头上标注的是“平均数 $\pm$ 标准差”, 但表中的内容格式与此不对应

**回应:** 非常感谢专家对文章规范性方面的要求。我们已经在表头和表内的数据格式进行了统一。

**意见 8:** 当被试加工性别面孔时, 任务知觉负荷较低时, 面孔性别可以得到加工, ”前后都是面孔性别, 让读者难以理解

**回应:** 感谢专家的建议。我们重新组织了这句话。改为: 任务的知觉负荷较低时, 被试对面

孔的性别信息可以进行加工。

---

#### 第四轮：编委复审意见

英文摘要质量实在太差，远远达不到发表水平，必须进行仔细修改，且有条件的话找英语为母语者润色。 仅以摘要的第一段为例：”Generally, load perceptual theory claimed stimulus of task-unrelated will not be processed along with the increasing of load perception. But there is still a debate on face processing. The most prominent showed it is uncertain that face processing especially race and gender under the influence of perceiving load.” (1)load perceptual theory: 专业术语不对，应为 perceptual load theory; 而且 theory 一次拼写错误； (2) will be: 应该为现在式, is; (3) stimulus of task-unrelated: 应为 task-unrelated (task-irrelevant) stimulus. 可能还需要考虑是否用 stimuli. (4) the increasing of: the increasement of (5) load perception: perceptual load. 应该还有更多问题，但我想这已经足以说明修改英文摘要的必要性了。

回应：感谢编委的积极建议。由于我们不小心，摘要写好了没有保存成功，提交的还是以前写的摘要，我们已经将最新摘要上传，并进一步修改完善。

---

#### 第五轮：主编终审意见

需要修改的问题：

1. P23 中，“配对样本 t 检验的结果为  $t(1,23)=17.47, p<.001$ ”，请作者检查 t 检验的自由度的写法是否正确？你此处写的自由度的格式到底是 t 检验还是 F 检验？

回应：谢谢主编老师对文章仔细的审查。我们误将相关的 t 检验的报告格式写成了 F 检验的报告格式。应该是： $t(23)=17.47, p<.001, Cohen's d=5.97$ 。同时我们也在正文中进行了修改。

2. 作者稿件中给出的  $\eta^2$  还是偏  $\eta^2$ ，如果是偏  $\eta^2$ ，则应写出偏  $\eta^2$  或  $\eta_p^2$ ；如果不是偏  $\eta^2$ ，作者是否应考虑回答你是怎样计算  $\eta^2$  的？

回应：谢谢主编老师提出的宝贵意见。稿件中的  $\eta^2$  指的是 partial  $\eta^2$ ，我们已在文中进行了修改。

3. 英文摘要还润色、提升。

回应：谢谢主编老师对论文提出进一步完善的意见。我们请具有国外留学背景的老师对英文摘要进行了润色和修改。