

《心理学报》审稿意见与作者回应

题目：视觉搜索中的情绪干扰项预习效应

作者：牟兵兵，宛小昂

第一轮

审稿人 1 意见：

意见 1：P5 作者讨论了干扰项预习效应的机制——注意抑制假说，并列举了已有的行为实验、眼动实验和事件相关电位实验来支持这一假说。应简要说明一下这些实验（或代表性实验）的结果，以及这些结果是如何支持注意抑制假说的。

回应：在修改稿中，作者已经增加了对干扰项预习效应的眼动实验和事件相关电位实验的介绍。眼动实验中靶子预习条件下的扫视的潜伏期更长、准确率更低（Caddigan & Lleras, 2010），及事件相关电位实验中靶子预习条件下的 N2pc 成分（与空间选择性注意密切相关的 ERP 成分）更小并且延迟（Shin, Wan, Fabiani, Gratton, & Lleras, 2008），均支持了注意抑制的假说。

意见 2：P6“尤其是在实验 1 中，研究者同时采用了正置和倒置的情绪面孔图片，以检验效应是由面孔的情绪特征引起的，还是知觉特征引起的。”正置面孔和倒置面孔是如何区分情绪特征和知觉特征的？应在“问题提出”部分有所说明。

回应：在修改稿中，作者已在“问题提出部分”增加了一些内容，简要说明了为什么在实验 1A 和 1B 中分别采用了正置面孔和倒置面孔。已有文献表明，正置面孔加工是整体的，而将面孔倒置会干扰对面孔的整体加工，而不影响对面孔的特征加工（Farah, Wilson, Drain, & Tanaka, 1998; Maurer, Grand, & Mondloch, 2002）。因此，将实验 1A 与 1B 的结果进行比较，使作者可以推断出实验 1A 关于正置面孔的结果，是由于对情绪面孔的整体加工导致的，还是由于对面孔的物理特征的加工导致的。

意见 3: P6“选用的是 2 位男性和 2 位女性的面孔照片，其中 1 男 1 女的面孔呈现生气或中性的表情，另外 1 男 1 女呈现高兴或中性的表情。”是 8 张面孔图片，还是 4 张面孔图片？

回应: 在该实验中，作者选用 A 男 A 女的面孔呈现生气或中性的表情（ $2 \times 2 = 4$ 张），另外 B 男 B 女呈现高兴或中性的表情（ $2 \times 2 = 4$ 张）。因此，实验中一共出现 8 张不同的面孔图片。在修改稿中，作者已对这一细节做出了更为具体的说明。

意见 4: 实验 1A 和实验 3A 中的“图片的亮度和对比度保持一致。”图片的亮度和对比度是如何保持一致的？请解释说明具体的匹配过程。

回应: 作者使用了 Matlab 的 SHINE 工具箱来对图片进行处理，使同一实验内的图片之间的亮度（luminance）和对比度（contrast）相互匹配。在修改稿中，作者提供了这一信息并引用了详细说明该工具箱的文献（Willenbockel et al., 2010）。

意见 5: P8“结果和讨论”指出：当靶子为生气面孔时，出现了干扰项预习效应；而其它条件下的干扰项预习效应均不显著。这里的其它条件指的是哪些条件？从表 1 和文中的内容来看，似乎只有“当靶子为中性面孔时”这一种条件。另外，这一表述似乎表明存在交互作用，而前文作者报告并无交互作用，这一表述易使读者感到费解。后文几个实验的结果表述也存在类似现象。如实验 3B 的结果“搜索蜘蛛时，出现了干扰项预习效应；搜索蝴蝶时，干扰项预习效应不显著。”但交互作用不显著。

回应: 当在中性面孔中搜索生气面孔时，干扰项预习条件和靶子预习条件的反应时之间存在显著差异，出现了干扰项预习效应。而对于（1）在中性面孔中搜索生气面孔的错误率、（2）在生气面孔中搜索中性面孔的反应时、（3）在生气面孔中搜索中性面孔的错误率这三个指标，干扰项预习效应和靶子预习效应之间则不存在显著效应。这三者就是作者在原文中提到的“其它”条件。在修改稿中，作者已修改了这个句子使其意义更加明确。此外，作者需要指出的是，尽管文献中的干扰项预习效应往往是以反应时为指标的，但是也需要对错误率进行分析，以证明反应时中出现的效应不是由于速度——准确率权衡而造成的。

关于评审专家提出“交互作用不显著时成对比较仍然显著”的问题，对于实验 1A 确实如此。ANOVA 的结果表明，预习条件和靶子类型之间的交互作用在统计上不显著。但是，参照已有文献并从干扰项预习效应的操作定义出发，作者进行了事先计划的成对比较（planned pairwise comparison），对每一对靶子预习条件与干扰项预习条件的结果进行比较，计算干扰项预习效应并判断是否显著。作者认为，事先计划的检验与事后检验（post-hoc tests）不同，

不以 ANOVA 中的交互作用是否显著为转移。但是，作者也同意的是，考虑到 ANOVA 中的交互作用并不显著，这样的事先计划的成对比较结果尽管显著，但其说服力不如 ANOVA 交互作用也同时显著时强。

关于评审专家提出蝴蝶蜘蛛实验（原稿中的实验 3B，修改稿中的实验 2B），情况是这样的：当提交时，作者只收集并报告了 15 位被试的数据。而对于这个样本，预习条件与靶子类型之间的交互作用不显著。而事先计划的成对比较则揭示，搜索蜘蛛时出现了显著的干扰项预习效应，搜索蝴蝶时干扰项预习效应在统计上没有达到显著（尽管两个条件下的反应时相差 41 毫秒）。在作者接到两位评审专家的修改意见之后，作者根据第二位评审专家的第 7 条意见对这个实验又补收了 5 个被试的数据，这样就一共收集了 20 个被试的数据。对这 20 个被试的数据进行分析发现，搜索蜘蛛时仍然有显著的干扰项预习效应，同时搜索蝴蝶时的干扰项预习效应（58 毫秒）也达到了显著水平。作者在修改稿中报告了 20 个被试的实验结果。因此，对于蝴蝶蜘蛛实验，就不存在实验 1A 中的那种令人费解的情况了。

意见 6：请在统计结果中补充效应量(partial eta squared)或 Cohen's d。

回应：在修改稿中，作者已经对 ANOVA 中的显著效应报告了效应量(partial eta squared)，对 t 检验的显著效应报告了 Cohen's d。

意见 7：总讨论显得较为单薄，不够深入。(a)“Aston-Jones 和 Cohen（2005）提出，注意有探索（Exploration）与固着（Exploitation）两种模式”。情绪效价对这两种注意模式的不同影响（如积极情绪扩展注意，促进探索）与情绪干扰项预习效应及其不对称性有何关系？

回应：在修改稿中，作者更加深入讨论了注意效价和情绪效价的交互作用。本研究中实验之间比较一致的结果是，如果当前搜索的靶子为负性的、传递威胁性信息的，则会出现稳定的干扰项预习效应。已有文献表明，负性的、传递威胁性信息的刺激，如生气的脸或蜘蛛、蛇的图片，往往比中性、不传递威胁性信息的刺激更迅速地吸引人们的注意（Lipp et al., 2004; Öhman et al., 2001）。而本研究的结果说明，尽管这些刺激本身是吸引注意的，但如果它们与之前失败的搜索经验（搜索过、但找不到靶子）联系起来，那么人们在当前的搜索则将注意偏离这种信息，使注意进入 Aston-Jones 和 Cohen（2005）提到的“探索模式”而移向之前没有选择过的信息。另一方面，在实验之间，本研究关于正性的或中性的情绪信息的结果不尽一致。在中性面孔中搜索高兴面孔或是搜索中性面孔时，没有引起显著的干扰项预习效应；在与恐惧有关的动物中搜索与恐惧无关的靶子，则出现了显著的干扰项预习效应。也就是说，

中性的、不传递威胁性信息的刺激，不如负性的、传递威胁性信息的刺激那样能强烈地吸引注意，而与不成功的搜索经验联系起来时，也不一定能够引导注意在当前搜索中的分配。

意见 8：“情绪干扰项（预习）效应中可能存在一种不对称性”，作者简单解释了可能的原因，这一原因是否有相关研究证据的支持，对情绪效价与注意抑制的关系有何启示？

回应：关于情绪干扰项预习效应中存在的不对称性，本研究只是提出这样一种可能。如上面一条回复所说，本实验中关于负性靶子的结果比较稳定，在实验间也比较一致；但是关于中性的、不具有威胁性的靶子，实验之间的结果不完全一致。一种可能的解释是，情绪干扰项预习效应中存在不对称性；但是也有其它的可能性。例如，中性和负性刺激除了在效价上有差别外，在唤醒程度上也有差别。在修改稿中，作者对这个问题作出了更为明确的讨论。相关研究证据的支持，来自 Lamy 等（2008）对跳出的情绪启动效应（Emotional Priming of Pop-out）的研究。他们发现当靶子为生气或高兴面孔时有跳出的情绪启动效应，而靶子为中性面孔时则没有该效应。但是同样地，他们也不能排除这种差异是由于效价的影响，还是唤醒程度的影响造成的。作者认为需要说明的是，尽管本研究中的实验无法区分效价和唤醒水平的影响，作者在其它实验中已致力于区分它们。在其中一个实验中，作者选取了唤醒水平相等的正性和负性情绪图片，也就是控制了唤醒水平，然后检验效价的作用。

意见 9：“Blanchette(2010)等人强调在情绪与认知的研究中应区分被试的“整体情绪”(integral affect)和“附带情绪”(incidental affect)。本研究中的情绪属于哪一种？

回应：在修改稿中，作者在总讨论部分对这一问题进行了讨论。作者认为，本研究中的情绪属于整体情绪，即被试完成实验任务的过程中，任务本身给被试带来的情绪体验；而附带情绪则指先由与实验任务无关的材料引发被试的情绪，在本研究中没有涉及。

意见 10：已有研究发现年龄或认知老化对干扰项预习效应有影响（Wan et al., 2011, 2014），这对情绪干扰项预习效应有何启示？

回应：对于 Wan et al. (2014)关于青年和老年被试对面孔示意图进行搜索，出现干扰项目预习效应的研究，由于和情绪干扰项预习效应直接相关，作者在修改稿的问题提出部分对之进行了说明。但是，老化对干扰项预习效应的影响，不在本论文的考虑范围之内，因此本论文没有详细地进行讨论。但在这里对此进行说明：Wan 等（2011）比较了青年人和老年人在颜色搜索的干扰项预习效应，实验任务是让被试在搜索画面中找到颜色与另外两个刺激不同

的靶子，再对靶子的形状特征作出反应。结果表明，老年人在任务中的反应时更长，干扰项预习效应也成比例地增长，说明老年人可能会加强自上而下对注意的引导作用（经验对当前注意分配的影响），来弥补老化过程中的自下而上加工的减缓。而对面孔示意图进行情绪搜索时，老年人对正、倒置的负性面孔进行搜索时出现了干扰项预习效应；青年人则只对正置的负性面孔进行搜索时出现了干扰项预习效应，对倒置的负性面孔则不存在这种效应（Wan et al., 2014）。这说明青年人在对面孔示意图时进行搜索时候进行的是整体加工，而老年人则很有可能是依赖图片之间的特征差别来进行搜索的，因此对倒置面孔也出现干扰项预习效应。作者认为，这些研究说明测量干扰项预习效应也可以成为对注意抑制能力的一种测量。

审稿人 2 意见：

意见 1： 研究使用不同类型的图片通过一系列的实验研究情绪干扰项预习效应，其实只是分别采用了不同的实验材料来考察情绪干扰项预习效应。实验之间缺乏一定的连接性和逻辑性，使得研究整体上较为单一。

回应： 如前所述，为了使实验之间的关系更为密切、主题更为明确、统一，作者从以下三方面作出修改：（1）作者删除了原稿中的实验 2（即在生气和高兴面孔中进行视觉搜索），使修改稿中只包括 2 个实验，而且这两个分别使用不同类型的情绪刺激检验干扰项预习效应（实验 1 中面孔图片，实验 2 中动物图片）。（2）在实验 2A 的结果部分，作者对实验 1A 和实验 2A 的数据合并进行了分析，2（实验材料，人脸或动物图片） \times 2 靶子类型（具有威胁性的或不具有威胁性的） \times 2 预习条件（干扰项预习或靶子预习）的混合设计（mixed-design）ANOVA。阐述情绪材料类型对干扰项预习效应的影响。（3）作者还在问题提出和总讨论部分加强了对各实验的综合分析，使实验之间的连续性和逻辑性更加明确。

意见 2： 研究创新点之一是考察过去对情绪性刺激进行不成功的视觉搜索对当前的注意的抑制，揭示了情绪与注意之间的交互作用。这在以往的研究中得到了证实，本研究是综合了不同的情绪材料进行重复。此外，研究考虑了不同效价的情绪刺激在与注意交互作用时的“不对称性”。无论情绪性的刺激材料还是非情绪的刺激材料，视觉搜索的不对称性研究也已经较为成熟。

回应： 在修改稿中，作者修改了“问题提出”和“总讨论”部分，以便使读者更加明确地看到本研究并非仅仅是综合了不同的情绪材料而对已有的研究进行重复。Wan 等（2014）使用面

孔示意图 (schematic face)，初步揭示了之前对情绪材料的不成功的搜索引起当前的注意抑制。但是值得注意的是，该研究同时对青年人和老年人进行了测试，青年人对正置的面孔出现了干扰项预习效应而对倒置的面孔示意图没有出现干扰项预习效应；而老年人对正、倒置的面孔示意图都出现了干扰项预习效应，说明他们对正置面孔表现的干扰项预习效应可能不一定是基于面孔示意图所表现出的情绪特征。此外，在已有的视觉搜索文献中，面孔示意图缺乏生态效度、与真实情绪面孔之间存在重要差异等问题，也得到了讨论。因此，使用情绪面孔的真实图片而检验干扰项预习效应也是非常必要的。已有文献表明，生气的脸（实验 1A）和蜘蛛/蛇图片（实验 2A）都属于具有威胁性的信息，对它们的视觉搜索会更快，即威胁优先效应 Frischen, Eastwood, & Smilek, 2008 ; Lipp et al., 2004; Öhman et al., 2001)。即使是同样具有威胁性，不同的情绪材料之间也存在着重要差异并受到关注 (Fox, Griggs, & Mouchlianitis, 2007)。正如以上这些非常具有影响力的论文彼此之间并非简单的重复和验证，本研究的各个实验之间也并不是简单的重复。

此外，评审专家提到“不同效价的情绪刺激在与注意交互作用时的不对称性”研究已经较为成熟，这个观点作者是同意的，关于视觉搜索的不对称性的研究是比较多的，尤其是作者也在论文中引用了很多关于情绪搜索不对称性的研究。但是，已有文献大量证明的是当前刺激的特征（无论是情绪性的还是非情绪性的）对当前视觉搜索的不对称影响，而本研究的结果初步证明的是之前不成功的搜索经验中的情绪效价与注意交互作用时的不对称性，是强调了自上而下加工（经验的影响）和自下而上加工（刺激的情绪效价）的交互作用中的不对称性。根据作者对相关领域的理解，这应是比较新的结论。如果评审专家能够推荐其它提出了这方面证据的文献，作者也非常愿意在正文中加以引用，与作者的结果相互佐证。

同时，作者也考虑到需要加强本研究与文献之间的相互联系，明确本研究在这个领域中的作用。因此在修改稿的总讨论部分，作者简要讨论了情绪干扰项效应与情绪启动效应 (affective priming effect) 和跳出的情绪启动效应(emotional priming of pop-out)的区别。简而言之，干扰项预习效应的特点在于，已有研究表明它是一种更为纯粹的注意效应，可以完全排除反应抑制的可能性。作者决不希望对本研究的创新意义夸大其辞，但希望这些修改能够使读者更加客观、清楚地看到本研究的意义。

意见 3：被试为国内大学生群体，但刺激材料选自日本与高加索人情绪面孔库、日本与高加索人中性面孔情绪库及卡罗林斯卡定向情绪面孔库，种族差异能排除么？为何不从国内的情绪库里面选取？此外，研究中的情绪性动物图片选自互联网，那么图片的情绪评定怎样控

制？为何不从国内的刺激材料库中选取？

回应：首先，种族差异不在文中这个研究的考虑范围之内。这个研究没有采用国内的情绪库，因此确实无法排除种族差异。实验 1 中采用的是 Elkman 的日本与高加索人面孔库，主要是因为该面孔库在情绪面孔视觉搜索的研究中应用非常广泛。在修改稿中已经删去的生气/高兴面孔搜索实验 2 采用了卡罗林斯卡定向情绪面孔库，主要是因为 Lamy 等（2008）研究“跳出的情绪启动效应”(Emotional Priming of Pop-out)时用的是该面孔库，因为该效应与情绪干扰项预习效应是最为接近的，因此作者在实验 2 中也采用了它。不过，在作者的其它实验中，作者研究了种族搜索是否能引发这种试次间效应，即实验任务为在画面内找到一张与另外两张面孔种族不同的面孔作为靶子，然后对它的其它特征（左边还是右边有白色横杠）进行反应。

其次，本研究使用的情绪性动物图片的确是选自互联网。该研究主要参照的文献是情绪动物视觉搜索文献（如 Lipp, 2006; Lipp et al., 2000; Öhman et al., 2001），这些研究中的情绪动物图片最初也选用于互联网，经过处理后在一系列研究中使用。因本研究需要用到 4 种动物的图片，每种动物 8 张，且要求是动物的全身照而且背景是草地、森林等自然背景，因此作者未能找到合适的情绪刺激材料库能提供这全部 32 张图片。选取这些动物主要基于马和猫在已有文献中被定义为与恐惧无关的动物，而蜘蛛和蛇被定义为与恐惧有关的动物（如 Lipp et al., 2004）。在该实验中，视觉搜索任务是一种动物类别搜索，即找到与其它两张图片动物类别不同的靶子，区分这两类图片是从所表现的内容出发。因此，在本论文中，作者不对每张图片的情绪效价进行区分。

但是，作者也考虑到评审专家可能希望明确知道这些图片的情绪效价。因此，在修改本论文的过程中，作者请 14 位被试（包括男 10 人、女 4 人）对这些图片的情绪效价和唤醒水平在 9 点量表上进行了评价，效价评分高则更为积极，唤醒水平分数高则唤醒水平更高。被试平均年龄为 19.9 ± 1.0 岁，没有参加过本研究中的任何实验，但是与本研究各个实验的被试来自同一被试库。他们对四类动物的评价分别是：马图片，情绪效价 6.0 ± 0.8 ；唤醒水平 4.9 ± 1.0 ；猫图片，情绪效价 $6.5 \pm 1.$ ，唤醒水平 5.2 ± 0.9 ；蜘蛛图片，情绪效价 2.3 ± 0.9 ，唤醒水平： 5.8 ± 2.0 ；蛇图片：情绪效价： 2.4 ± 0.9 ，唤醒水平 7.1 ± 1.7 。作者对四类图片的情绪效价进行了多重成对比较并对结果进行了 Bonferroni 校正，发现对蜘蛛和蛇图片的效价评分显著低于马和猫的图片，all $t_s > 7.96$, corrected $p < 0.001$ ；但是蜘蛛和蛇之间、马和猫之间的效价评分差异不显著，both $t_s < 1.97$, corrected $p > 0.42$ 。就唤醒水平来说，蛇图片的唤醒水平评分要显著高于马和猫图片，both $t_s > 3.29$, corrected $p < 0.05$ ；但是蜘蛛图片的唤醒水平评

分与马、猫图片之间的差异则不显著, $\text{both } ts < 1.17$ 。由此可见, 如果一定要对该实验中的 4 类动物图片进行情绪评定上的划分, 马和猫图片是中性、唤醒程度中等图片; 蜘蛛图片是负性、唤醒水平中等的图片; 蛇图片是负性且高唤醒水平的图片。但是, 作者认为还是从动物类型对本实验中的图片进行分类, 更符合该实验的目的。因此, 这些对图片的情绪评定结果, 作者没有在论文正文中进行报告。

意见 4: 方法部分的实验设计部分开始应该介绍实验设计类型, 否则无法清楚地了解自变量和因变量, 也就不知道研究主要操控了什么变量, 通过什么指标来考察情绪干扰项预习效应。

回应: 在修改稿中, 作者在各个实验的方法部分更加明确地说明了本实验是被试内设计, 两个自变量是预习条件和靶子类型; 并在实验 1A 的方法部分明确说明了本研究各个实验中均使用反应时和错误率作为因变量。

意见 5: 被试数量的选取较少, 特别是有的实验只有 15 人, 这可能存在一定的随机性, 代表性差。

回应: 以往关于干扰项预习的文献表明, 干扰项预习效应是一种十分稳定 (robust) 的试次间效应, 单个实验的被试人数均为十余人, 如 11 人 (Goolsby et al., 2005, 实验 1.1)、12 人 (Ariga & Kawahara, 2004; Caddigan & Lleras, 2010)、18 人 (Shin et al., 2008) 等。因此, 作者采用的样本数与已有文献的范围一致。但是, 作者也理解第二位评审专家的考虑, 并在修改论文前对专家提到的只有 15 个被试的蝴蝶蜘蛛实验补收了实验数据, 修改稿中该实验有被试 20 人。增加了这些被试之后, 结果模式基本没有变化。唯一明显的变化是, 搜索蝴蝶时, 样本量为 15 人时干扰项预习效应 (41 毫秒) 在统计上不显著; 样本量增加到 20 人时干扰项预习效应 (41 毫秒) 在统计上显著了。这是符合统计学基本原理的, 当样本量增加, p 值会减小。

意见 6: 统计结果的报告中应该报告统计检验力的大小。

回应: 在修改稿中, 作者已经对 ANOVA 中显著的效应报告了效应量 (partial eta squared), 对 t 检验报告了 Cohen's d 。

意见 7: 文中提到的“事先计划的成对比较”中的“事先计划”是什么意思?

回应: 原文中提到的事先计划的成对比较为 planned pairwise comparison, 是与统计中的

post-hoc tests（事后检验）相对应的常用统计方法。在修改稿中，作者在第一次提到该词时提供了英文原文，使读者更加明确作者使用的统计方法。根据已有文献及干扰项预习效应的操作定义，作者对每种靶子/干扰项组合的靶子干扰项预习和靶子预习条件下的反应时和错误率进行了比较。

意见 8：总讨论部分较为单薄，并没有针对研究得出的结果逐一进行解释和讨论。文章的讨论应该结合结果的每一个有差异的数据点进行讨论，因此本文的讨论不具有实质性内容。

回应：根据第一位评审专家的部分修改意见，作者在修改稿中对总讨论进行了扩展和深化。本论文是一篇多实验论文，在每个实验的讨论部分，已经对该实验中“有差异的数据点”进行了讨论。在总讨论中，主要是对实验之间的结果进行分析比较，并联系文献讨论本研究的意义。

第二轮

意见 1：2.2 结果与讨论部分“除此之外的干扰项预习效应(在平均反应时或错误率上)均不显著， p 值均大于 0.26。”作者在对第一轮修改意见的回复中解释了“除此之外的干扰项预习效应”。但“除此之外”，仍给人感觉是其他实验条件，而不是同一实验条件在反应时和错误率上的干扰项预习效应，作者可对反应时和错误率的结果分开表述，这样可能就不会产生误解了。

回应：在修改稿件，作者将该句修改为：“对于（1）在中性面孔中搜索生气面孔的错误率、（2）在生气面孔中搜索中性面孔的反应时、（3）在生气面孔中搜索中性面孔的错误率这三个指标，干扰项预习和靶子预习这两个条件之间则不存在显著差异， p 值均大于 0.26。”

意见 2：4.3 结果与讨论部分，“即具有无威胁性的靶子的反应(1356 ms, 6.3%)要比对有威胁性的靶子的反应(1397 ms, 4.7%)更快但是错误率也更高。”的结果是否意味着本研究存在如作者对第一次修改意见（第 6 条意见）的回复中提及的那种“速度——正确率权衡现象”？

回应：在对第一轮修改意见进行回复时，作者提到，在分析干扰项预习效应时，需要比较干扰项预习和靶子预习两个条件中的反应时是否有差异，也要比较这两个条件下的反应错误率。这是因为，干扰项预习效应可能会受到速度——正确率权衡现象的影响。在以往的研究中作者曾发现，A 组被试在反应时上的干扰项预习效应比 B 组被试更小，即 A 组在干扰项预习

和靶子预习这两个条件中的反应时之差变小了,但是同时他们在这两个条件中的错误率之差变大,即在错误率上的干扰项预习效应变大了。因此,这种情况下作者只能得出结论说 A 组被试的干扰项预习效应就受到了速度——正确率权衡现象的影响,而不能说 A 组被试的干扰项预习效应比 B 组被试小。也正是因为如此,在分析干扰项预习效应时,作者需要同时分析反应时上的和错误率上的效应。

但是,评审专家在第二轮修改中这条意见中提到的结果,作者认为并不说明本研究的干扰项预习效应受到了速度-正确率权衡现象的影响,而是说明本研究中靶子类型(具有威胁性的,或不具有威胁性的)的作用受到了反应速度——准确率现象的影响。评审专家提到的这一结果来自于作者将实验 1A 和实验 2A 的结果进行合并,并进行了 2 (实验材料:人脸或动物图片) \times 2 (预习条件:干扰项预习或靶子预习) \times 2 (靶子类型:威胁性的或不具有威胁性的)的混合设计 ANOVA 分析时。该分析的结果表明,靶子类型在反应时上的主效应显著, $F(1, 33) = 10.58, p < 0.01, \eta_p^2 = 0.14$, 在错误率上也显著, $F(1, 33) = 5.04, p < 0.05, \eta_p^2 = 0.13$ 。因此,这个结果并不说明本研究中的干扰项预习效应受到了速度——正确率权衡现象的影响,而是说明被试对不同类型的靶子(具有威胁性的 vs.不具有威胁性的)的反应可能受到了速度——正确率权衡的影响。在修改稿中,作者已明确指出这一点。

意见 3: 用“很难把注意从它上面移走(Most, Chun, Widders, & Zald, 2005)”来解释“对威胁性靶子的反应要比对不具威胁性的靶子的反应慢”的实验结果,看起来较合理,也符合已有研究的结果。但文章作者又如何解释实验 1B (倒置情绪面孔图片引起的干扰项预习效应)中发现的“靶子类型的主效应在反应时上显著, $F(1, 16) = 5.82, p < 0.05, \eta_p^2 = 0.27$ ”这一结果呢?

回应: 作者非常感谢评审专家提出这个问题。作者重新审视并思考了作者的结果及讨论。首先,在实验 1A 中,被试对中性面孔靶子的反应比对生气面孔靶子的反应更快;而在实验 1B 中,被试对倒置中性面孔靶子的反应比对倒置生气面孔靶子的反应更快。因此,综合实验 1A 和实验 1B 的结果进行考虑,被试在实验 1A 中对中性面孔和生气面孔的反应时之差,很有可能不是由于面孔的情绪特征造成的。其次,如评审专家在上一条意见中所说,当作者把实验 1A 和实验 2A 的数据合并进行分析时,靶子类型对反应的影响受到了反应速度——准确率权衡现象的影响。因此综合这两条进行考虑,本研究的结果并没有说明“对威胁性靶子的反应要比对不具威胁性的靶子的反应慢”。因此,在修改稿中,作者从总讨论部分移除了这部分讨论。在上一轮的修改稿中,作者曾经在对这个结果的讨论中提到,被试在情绪干扰

项预习效应的实验中，需要通过靶子的情绪性找到它，并对它的非情绪性特征进行反应，因此被试的干扰项预习效应能够反映其情绪灵活性（**affective flexibility**）。作者现在仍然持这一观点，但是也认识到需要进一步的研究工作进行验证，在目前这篇论文中就进行论述未免为时过早，因此作者移除了这部分讨论。