

# 《心理学报》审稿意见与作者回应

题目：无意识信息的刺激表征及其时间特性

作者：罗婷，邱茹依，陈斌，傅世敏

---

## 第一轮

**审稿人 1 意见：**该研究考察了阈上和阈下呈现的 Flanker 与目标在刺激水平和反应水平的冲突效应。阈上 Flanker 与目标的冲突效应与以往研究结果一致，而阈下 Flanker 引发了预想的反应水平的冲突效应，但在刺激水平上却出现了反转的冲突效应。文章存在如下问题供考虑：

**意见 1：**首先文章的研究主题不是很明确，是要考察无意识信息是否存在刺激表征吗？在过去的 100 多年中，已有大量的研究采用各种实验范式证明了无意识信息加工在各级水平上的存在，本研究的主要贡献是什么？这些需要在前言中进一步澄清。

**回应：**的确，有大量研究证据表明无意识信息加工广泛存在，但是由于研究范式和意识测量工具的不同，无意识知觉加工存在许多分歧。一些研究者提出需要明确区分无意识研究中的概念和结论，以免造成混淆(Rothkirch & Hesselmann, 2017)。同时，无意识加工是否广泛存在仍有不小的争论(Hassin, 2013; Hesselmann & Moors, 2015)。近期，Peters, Kentridge, Phillips, 和 Block 等研究者在 Association for the Scientific Study of Consciousness (ASSC)第 20 届年度会议上专题讨论无意识知觉是否存在 (Peters, Kentridge, Phillips, & Block, 2017)。他们认为，探索无意识知觉的问题，需要在知觉概念上、实验范式上进一步努力，在剔除刺激强度和任务反应的基础上考察纯粹的无意识知觉。其中，将反应加工从知觉加工过程分离是考察纯粹无意识知觉一个重要途径。

本研究围绕是否存在刺激知觉表征这一分歧，采用 flanker 干扰范式，通过不同的冲突类型（无冲突、刺激冲突和反应冲突）条件的对比，考察无意识知觉过程。主要的贡献在两点，首先，在 flanker 范式中加入掩蔽技术。一方面，意识水平的目标和无意识水平的干扰项同时出现，促使无意识信息对目标加工产生及时的影响。另一方面，分离不同的冲突类型专门探测无意识信息的知觉加工。其次，由于新方法和范式的使用，发现了无意识知觉加工

的行为学证据。该发现是神经成像和电生理证据的重要补充。

您的意见十分重要，本研究的必要性和创新性已在前言相应地方修改、补充。例如：“综上所述，尽管无意识信息的刺激（语义）表征得到了神经电生理学证据的支持，但仍欠缺行为层面的证据。”(第 1 页)

**意见 2: 1a.** 文中的刺激表征究竟指的是什么？是 flanker 的知觉加工、语义加工（或者包含所有的刺激编码过程）？抑或是目标刺激的知觉加工、语义加工？文章中多次提到“刺激表征”时，有时似乎是指前者，有些似乎是指后者。

**回应：**本文中，刺激表征主要是指信息的知觉加工。我们通过比较任务呈现中存在刺激冲突和无刺激冲突条件的成绩，来推测刺激加工对行为反应的影响，以确认是否存在刺激知觉表征。本研究的 flanker 任务主要涉及到知觉冲突和反应冲突。实验中采用字母作为实验材料，因此假定不存在语义冲突的过程。关于这点，前人的主流观点也是认为，字母 flanker 任务主要涉及刺激知觉冲突和反应冲突，所以本研究中刺激表征是指信息的知觉加工。前言中，引用了许多启动研究，主要因为它们采用了语义材料，所以这些研究里难免会涉及语义加工。因此，在前人的启动研究里，语义加工和刺激加工的概念会混淆。这也是本研究考察纯粹刺激知觉加工的一个原因。

**意见 3:** 文中提到“无意识信息究竟能否影响刺激表征，目前尚存争议(Kiefer & Brendel, 2006; Klinger, Burton, & Pitts, 2000; Lin & Murray, 2014)”，这里似乎是在表述本研究的切入点，那么结合本研究，无意识信息是指阈下呈现的 flanker?刺激表征是指目标的刺激表征？

**回应：**您的理解是对的。本研究的无意识信息是指 flanker 字母，刺激表征是指目标的刺激表征。尽管阈下 flanker 信息与任务无关，但是该信息的加工会对目标加工产生影响。因此，若这种无意识信息影响了目标的刺激表征，即刺激冲突条件和无冲突条件有差异，那么可以认为无意识信息与意识信息在刺激知觉层面的表征产生了相互作用。因而，我们认为这个结果表明了无意识信息存在刺激表征。

关于这个问题，我们在前言部分对研究的逻辑做了梳理，并对相应内容的表述做了调整，以便读者理解。如“这种刺激冲突情况包含了信息间的刺激干扰但不包含反应干扰，因此刺激干扰效应可以作为信息存在刺激知觉的证据。”（第 2 页）

**意见 4:** 文中提到“主流观点认为无意识信息影响了反应操作加工而非刺激知觉加工(Kiesel,

Kunde, & Hoffmann, 2007; Lin & Murray, 2014)”，这一总结是否存在问题？例如 Kiesel 等人的研究，主要发现是当存在多种刺激-反应的联结规则（stimulus-response (S-R) rules）时，无意识刺激可以同时激活多种刺激-反应联结，即使这一联结不是当前任务要求的。这些结果如何表明“无意识信息影响了反应操作加工而非刺激知觉加工”？

**回应：**这篇研究中，Kiesel, Kunde 和 Hoffmann 等人主要想说明无意识信息激活了反应加工，通过反应加工影响行为。研究里，刺激可以是数字或字母，要求被试判断字母是元音、辅音或者判断数字是大 5、小 5。被试的反应为，例如，元音按左手键，辅音按右手键，大 5 数字按左手键，小 5 数字按右手键。因此，反应方式上存在一致性，如元音和大 5 数字的反应是一致的，而辅音与大 5 数字的反应就不一致。当然，大 5 数字与大 5 数字的反应也是一致的。因为音节和数字是两种不同的刺激加工形式，如果启动效应发生在刺激加工水平，启动和目标属于不同刺激类型时应该不存在启动效应；如果启动效应发生在反应加工水平，那么相同反应的刺激之间会有反应启动。研究结果表明，无论启动刺激是否是当前任务要求的反应，都会影响后续的目标判断。也就是说，不同类型的刺激启动了反应加工，并且这种启动效应在相同刺激类型和不同刺激类型上不存在差异。因此，作者认为无意识启动激活了反应相关的加工而非刺激加工。

**意见 5:**文章的主要意外结果是阈下 flanker-目标刺激冲突效应的反转，作者的解释是“无意识信息与意识信息在刺激表征水平上发生了短暂的竞争，提示无意识刺激表征自动激活并迅速衰退”...“无意识信息与意识信息的知觉整合可能是无意识信息加工通过反应水平影响行为的主要原因。具体来说，当不同意识水平的刺激相同时则整合为相同的刺激表征，此时无意识字母无需额外表征；当它们不同时才会分别激活晚期的反应表征。因此，刺激表征发生冲突时，反应一致的无意识 flanker 促进了目标加工，而反应不一致的无意识 flanker 阻碍了目标加工。”，这一解释主观推测性比较强，如果存在这一机制，阈上结果的 Delta 反应时分布分析是否也应该出现类似的结果？”

**回应：**是的，这一讨论是我们根据现有的实验数据进行的推测。我们的推测与无意识信息加工的一种观点吻合，但需要警惕的是该观点也只是的一家之言。我们采用该观点是为了增强猜测的可信度。总而言之，这只是一个推测，我们的数据表明这种可能性较大。

阈下无意识信息加工的方式可能与阈上意识信息的加工方式不同。首先，阈上意识加工下，flanker 和目标的刺激表征存在很强的竞争关系，这一点同我们的实验（实验 1）及前人的研究结果一致。而阈下无意识加工，实验 1-3 的正确率结果未表现出明显的竞争的关系。

只是通过  $\delta$  反应时分布结果表明, 这种竞争曾短暂存在。由此可见, 无意识信息和意识信息的加工方式有许多的不同, 知觉整合只适用于无意识信息的加工和操作。对此, 我们也只是猜测。但是, 前人的一项研究里考察了阈上 flanker 任务中刺激冲突的  $\delta$  反应时分布 (Davelaar, 2008)。其结果表明, 随着反应时间增加刺激冲突效应并未减小。该结果与阈下无意识 flanker 任务的结果完全不同, 但符合我们对阈上结果的推测。

我们对该部分内容的表述和措辞进行了调整和改动, 如“我们猜测, 无意识信息的知觉整合可能是其加工侧重于反应表征的主要原因。”详细内容见 11 页的第 2 段落。

.....

**审稿人 2 意见:** 该研究利用侧抑制任务来探究无意识刺激表征, 并且使用  $\delta$  反应时分布方法探索了无意识刺激表征的时间进程, 具有一定的创新性。但还存在一些小的问题。如下:

**意见 1: Flanker 翻译为侧抑制更合适**

**回应:** 感谢审稿人的建议, 本稿中 Flanker 的中文翻译统一为“侧抑制”, 见第 2、3 页(第 1 段尾句、第 2 段首句、第 3 段首句) 及下文相应的部分。

**意见 2: 做意识测试时为什么将中央字母改为“A”, 与正式实验不一样?**

**回应:** 实验 1B 的无意识测验中, 中央字母为“A”, 被试报告 flanker 的内容。这种实验操作独立地考察了注意条件下 flanker 字母的意识水平。为了避免这种无意识测量与真实实验情景的差异, 实验 2 的无意识测量采用了实验情景一致的刺激内容和按键反应。因此, 我们可以从多个角度考量 flanker 刺激的无意识觉察水平。

**意见 3: 实验 1 的结果中提到“刺激冲突条件与无冲突条件的正确率差异不显著 ( $t_{12} = 2.11$ ,  $p > 0.05$ )”。这里的  $p$  值具体值是多少? 是大于 0.1 还是边缘显著?**

**回应:** 实验 1 中, 此处的  $p$  值为 0.157。该值远高于边缘显著水平, 可以认为是统计结果不显著。为了更清楚地说明这个结果, 本稿中该处统计值显示为  $p > 0.1$ 。(见第 4 页)

**意见 4: “无意识的刺激表征可能在建立联结的过程中衰退, 从而使无意识下刺激冲突的趋势与意识下有所不同”, 这里的“意识下”有所歧义。**

**回应:** 本稿中, 此处的表述改为“由于无意识信息消退较快(Greenwald et al., 1996; Salti et

al., 2015), 无意识的刺激表征可能在规则转换的过程中快速衰退, 从而与意识知觉中刺激表征之间的冲突有所不同。”(见第 6 页)

**意见 5:** “反应时分析还剔除了错误试次及错误后试次。”这句话中的“错误后试次”表述不清楚。

**回应:** 该句的表述修改为“反应时分析不仅剔除了错误试次还剔除了错误反应之后的第一个试次”。(见第 8 页)

**意见 6:** “本研究结论的推广仍需要考虑以下问题。研究中的无意识 flanker 信息处于非注意加工, 而且 flanker 效应受空间距离的影响”, 无意识信息并不能够说明它一定处于非注意加工, 这里表述不够严谨, 另外, 既然 flanker 效应受空间距离影响, 作者应该在方法处表明该研究所采用的空间距离大小。

**回应:** 同意审稿专家意见。本研究中被试的任务是判断中央字母, 此时 flanker 字母与任务无关。尽管如此, 被试也有可能分配注意资源到 flanker 信息上。因此, 本稿相应位置修改为“研究中的无意识 flanker 信息与任务加工无关”(见第 11 页)。本研究的另一个局限是 flanker 效应与空间距离的关系, 也就是说任务无关的信息得到无意识加工的过程是否会因为注意加工范围的变化而不同。在审稿人的建议下, 我们补充了本研究材料的视角大小, 以便后续研究参照。研究的方法部分(第 3 页), 增加内容为“Flanker 任务里, 五个字母水平排列, 构成  $3.5 \times 1.4^\circ$  视角(见图 1)”。

---

## 第二轮

**审稿人 1 意见:**

**意见 1:** 这篇文章通过冲突效应重点考察阈下呈现的 flanker 是否存在刺激水平和反应水平的信息加工, 可是修改稿中多处表述还是让人费解, 例如前言中的第二段, “无意识信息究竟能否影响刺激表征”, 这会让人理解为无意识信息能否影响目标的刺激表征, 而文章的重点似乎应该是“无意识信息是否存在刺激表征”, 而不是目标的刺激表征; 前言中 Klinger 等人 (2000) 的研究只是说明了当无意识启动的信息与目标任务无关时, 不会产生启动效应。作者在文中的表述“发现无意识启动内容的语义信息并不总是影响目标加工”是不恰当的, 没有产生启动效应, 并不能认为没有影响目标加工, 只能说没有影响目标的任务相关的加工。。

“因此无意识信息的反应加工而非刺激加工影响了认知加工”这样的句子更是让人费解。。希望文章在表述上进一步提高。

**回应：**为了方便读者准确把握研究的问题，我们修改了文中相应部分的表述，如第 1 页第 2 段相关内容修改为“其中，无意识信息究竟是否存在刺激表征，”；如第 11 页第 1 段“由于刺激冲突条件和无冲突条件下 flanker 和目标的反应方式都相同，因此，这两个条件之间的差异至少与它们在刺激水平的加工有关，支持了无意识信息的刺激表征。”。

我们对 Klinger 等人 (2000)研究结果的意义做了新的限定和补充，“例如，Klinger 等人 (2000)采用启动-掩蔽范式研究语义材料的无意识加工，发现并非总存在无意识启动内容的语义信息影响目标成绩。”见第 1 页第 2 段。

除此之外，我们也对文章中表述可能有歧义的有关内容进行了修改，见红色标记。

**意见 2：**对于 flanker-目标刺激冲突效应的反转，主要表现在正确率上，可是如果仔细观察数据，阈上（实验 1A）和阈下(实验 1B)的错误率结果的 pattern 是非常相似的，只是阈上无冲突和刺激冲突条件的差异没有达到显著（95.3% vs. 97.1%， $p=.157?$ ），实验 1A 只有 13 个被试，也许样本量增加一些，差异就显著了。那么，各个实验的样本量是如何确定的？实验 1A 和 1B 分别有 13 名被试，实验 2 只有 11 名被试，实验 3 稍多一些，有 21 名被试，样本选择有什么依据吗？

**回应：**3、研究中的样本量是综合德国杜塞尔多夫大学编写的 G\*Power 工具包 (<http://gpower.hhu.de/>) 的预算样本量和前人研究的样本量确定的。在该工具包下，我们采用 F 检验的统计效应  $f$  为 0.8（较高）， $\alpha$  为 0.05 等参数。本研究为 3 水平的单因素重复测量实验设计，该工具给出的参考样本量为 6。但前人研究中的实际样本量远比预估值大。如，Davelaar (2008)的研究考察阈上刺激、反应水平冲突量的 delta 分布，采用了 9 人的样本量。此外，也有其它研究考察刺激、反应水平冲突特征的研究采用了 20 多人的样本量，如 Wendt, Heldmann, Munte, 和 Kluwe 等人(2007)采集了 23 名被试的数据，Verbruggen, Notebaert, Liefoghe, & Vandierendonck 等人(2006)的实验被试量为 26 人。鉴于此，我们在实验 1 和实验 2 中采用了小样本量分别为 13 和 11。为了避免样本量过小的批评，我们在研究三中增加样本量至 21。当然，研究中实验间样本量的差异过大，会引起诸多的质疑。我们选取实验 3 中最早参加实验的 13 名被试的数据，重新分析实验 3 的数据。由于一名被试的正确率在 2 个标准差以外，因而未纳入统计。12 名被试的反应时、正确率结果与 19 名（两名被试的正确率数据在 2 个标准差外而剔除）被试的数据统计结果一致。综上，本研究中的样本量虽然

数量不一，但不会影响统计结论。我们在总讨论中增加了样本量的局限性。

我们与您有相似的疑问，如果增加实验 1A 的样本量， $p$  值是否能得到显著水平。于是，我们将实验 1A 的样本量和效应量输入 G\*Power 工具包中做检测。结果显示，我们接受当前无冲突条件和刺激冲突条件差异显著时犯 II 型错误的概率为 35%（图 1 上）。当然， $p$  值显著性会随着样本量的增加（图 1 上）。当样本量增加到 30 时，我们接受备择假设犯 II 型错误的概率会下降到 5%，但这也意味着极大的样本代价。与之相比，实验 1B 在样本量为 10 时，接受备择假设犯 II 型错误的概率只有 1%（图 1 下）。因此，实验 1A 的差异未达显著水平可能并不能归于样本量。

综上，我们认为 1) 研究中的样本量大体符合统计检验的标准；2) 实验 1A 中无冲突条件和刺激冲突条件的正确率不存在显著差异。

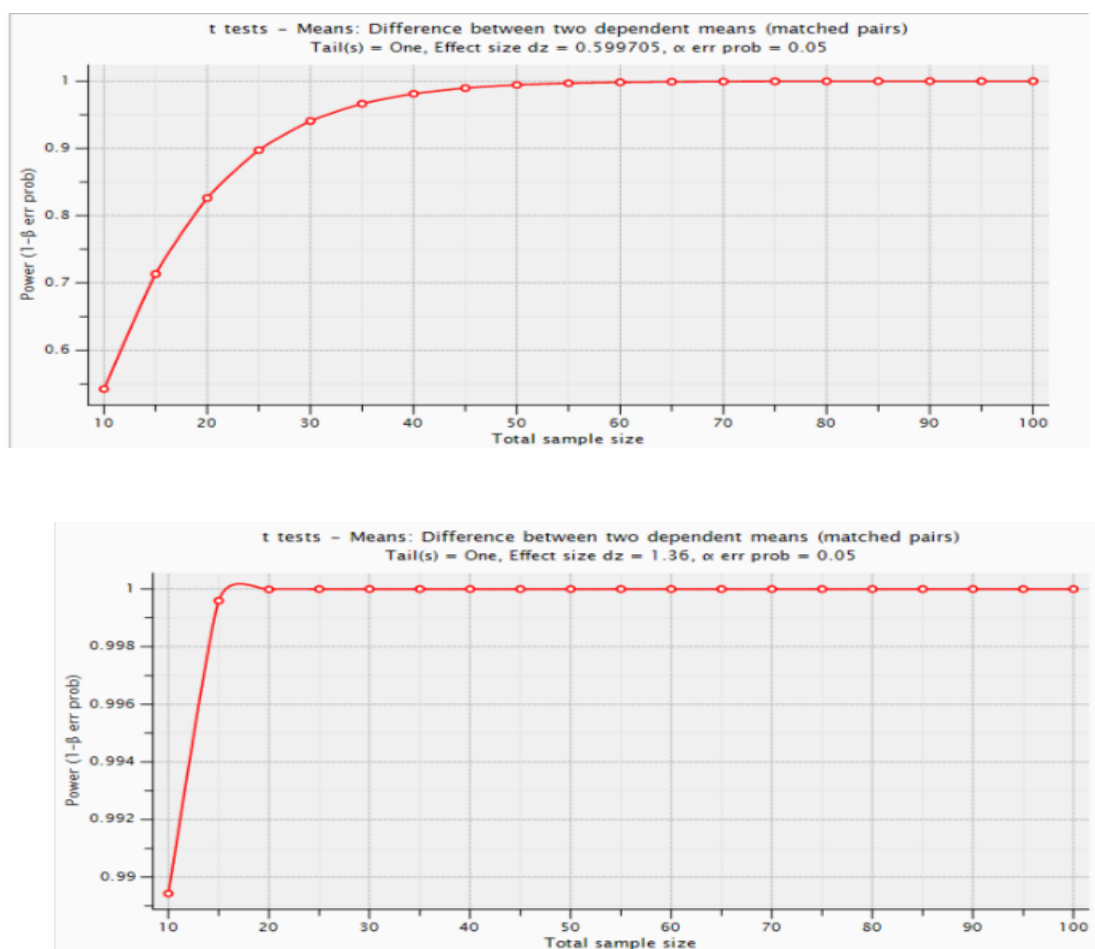


图 1. 无冲突条件和刺激冲突条件正确率差异的统计效应与样本量关系的曲线图。上图表示实验 1A 中变化趋势，当样本增加至 30 人时，统计效应  $(1-\beta)$  才达 95%。下图表示实验 1B 中变化函数，当样本量为 13 时，统计效应  $(1-\beta)$  已达 99%。

意见 3: Flanker 的意识测试的程序写的不够详尽, 是要被试报告 Flanker 字母是什么? 还是提供几个备选项让被试选择? 为什么实验 1B 的随机猜测概率是 25%, 而实验 2 的随机猜测概率就变成了 50%?

回应: 实验 1 的意识测验是让被试选择具体的字母, 由于有 4 个备选字母, 因此有 4 个反应项, 所以随机猜测水平为 25%。具体修改见第 4 页“除了: 1) 要求被试判断 flanker 位置的字母是哪一个字母, 2) 并对无意识字母的选择做信心评估。”, 及第 5 页“由于该任务的字母-反应备选项为 4 个, 因此随机猜测水平为 25%。”。

实验 2 中的意识测验是让被试根据任务规则对 Flanker 字母反应。由于实验中只有两个反应选择项, 所以随机猜测水平为 50%。具体见第 7 页“1) 根据实验程序的反应规则判断 flanker 字母, ”, 及“该结果与随机概率(二选一反应的猜测水平为 50%)相比差异不显著 ( $t_{10} = -0.87, p > 0.05$ )”。

意见 4: 我倒觉得实验 3 的结果更为有趣, 实验 3 使用 3 对 (6 个) 字母, 实验难度增加, 结果发现刺激冲突条件相对无冲突条件反应加快 (592ms vs. 601ms), 而阈上的结果刚好相反 (523ms vs. 496ms), 反应时的数据似乎更能说明问题? 但文章中似乎重点关注了正确率的数据。

回应: 本研究的因变量选择有 2 个方面的考虑。首先从理论角度, “是由于无意识信息影响了决策过程中的证据累积, 因而正确率相比于反应时能更敏感地标识无意识加工(Vlassova et al., 2014)”。其次, 从数据驱动的角度来讲, 正确率比反应时的结果更稳定。本研究的三个实验中, 正确率的结果极其一致。尽管反应时的结果也表现出类似的趋势, 但只在实验三中发现了显性的统计差异。实验三中反应时的显性结果可能与难度有关, 这个结果确实更清楚、且更有意义。其意义远非如此, delta 反应时分布也是证明无意识刺激表征的重要支撑。但另一方面, 这似乎也暗示了反应时反映无意识加工的限制性, 如稳定性较差, 依赖特定的实验条件。

.....

审稿人 2 意见:

意见 1: 作者没有很好地回答“做意识测试时为什么将中央字母改为“A”, 与正式实验不一样”这个问题。意识测试实验和正式实验中央字母不一致的设计, 导致意识测试实验中得到的结



果并不能很好地说明在正式实验中被试对 flanker 字母的意识水平。建议作者重新补做实验 1 B，确保意识测试中的中央字母与正式实验中一致。其他问题作者均已经进行了很好地回答和修改。

**回应：** 审稿人指出，我们在无意识测试与无意识侧抑制任务中呈现不同的实验刺激，因而无意识测试的结果不能保证无意识侧抑制任务中的意识水平。因此，审稿人建议我们补做实验 1B。我们接受审稿人的意见。新的实验 1B 与原来的实验 1B 相同，除 1) 无意识测试中采用侧抑制任务中相同的刺激呈现及 2) 补充无意识测试的信心评估。本实验收集了 13 名被试的数据。统计结果显示，无意识信息的加工方式与本研究中其它实验的发现一致。重要的是，无意识测试的结果表明，被试不能知觉到阈下信息的内容。具体内容见方法（第 4 页）、结果（第 5 页）及图表部分。

需要声明的是，由于投稿系统的一些原因，遗漏了审稿人 1 的评审意见。在此，我们仍保留了一审中的改动，方便审稿人 1 审阅和评价。

---

### 第三轮

**审稿人 1 意见：** 没有其他意见。

**审稿人 2 意见：** 作者已经针对审稿意见补充了相应的实验。建议发表。