

《心理学报》审稿意见与作者回应

题目：视空工作记忆负载对真实场景搜索绩效的影响

作者：任衍具 孙琪

第一轮

审稿人 1 意见：本文采用眼动技术考察了不同类型的工作记忆负载对真实场景视觉搜索的影响，结果发现空间记忆负载和客体记忆负载对真实场景视觉搜索有着不同影响。本文研究主题具有一定的创新性（记忆负载对真实场景搜索），扩展了工作记忆负载对视觉搜索的影响研究。实验设计合理、数据处理恰当、结论与数据一致。本文逻辑思路清晰，文字表述准确（包括英文摘要），建议发表。

回应：感谢该审稿专家给出的支持性意见。

审稿人 2 意见：本项研究的选题在视觉搜索领域尚属缺乏探讨的问题，通过实验发现了与采用“传统视觉搜索任务”的研究存在差异的结果，因此值得对这些结果进行更全面的分析和讨论，以为视觉搜索领域的相关研究提供可信的有价值的参考。

现版本的论文欠缺之处在于论文写作中的内容组织和分析论证的逻辑不够简明，“故事”讲得不是很清楚，可能会导致读者理解上的困难。表现在对一些专用术语缺乏解释；对变量的界定和操作方法交代不清；对数据的统计分析报告虽然详尽但与研究结论的论证思路缺乏清楚的对应，论证逻辑不够直接明了；对主要结果的理论解释缺乏依据等方面问题。一些有具体针对性的建议请见审改稿中的批注内容。

意见 1：“工作记忆 (Working Memory, WM) 模型最初是由 Baddeley 等人提出的（参见最近的综述文章: Baddeley, 2012), ...” 这种说法不准确。虽然 Baddeley 和 Hitch 最早系统提出了工作记忆的理论模型，但是工作记忆的概念早在 1960 年就由 Miller 等人提出了（详见：Baddeley, A. (2003). Working memory: Looking back and looking forward. *Nature Reviews Neuroscience*, 4(10), 829-839.）。此外，虽然 Baddeley 的“工作记忆多成分模型”（The multi-component model of working memory）是工作记忆领域最有影响力的模型，但这个模型并不是唯一的工作记忆模型（详见：Shah, P., & Miyake, A. (1999). Models of working memory. *Models of working memory: Mechanisms of active maintenance and executive control*, 1-27.）。

回应：感谢审稿专家澄清概念的来源。已对文字表述做了修改，力求准确。

正文中已改为“Baddeley 及其同事系统提出并发展了工作记忆的多成分模型 (The Multi-Component Model of Working Memory), ...”。

意见 2：搜索目标变化有多种可能性，仅表述为“搜索目标变化”，有可能会使读者难以理解这一变量。这里所说的搜索目标的变化应该是指“不同试次间的搜索目标会变化”，以对应“不同试次的搜索目标相同”这种条件。建议对此变量的操纵方式进行更具体的说明，并阐述清楚此变量的两种取值分别可以解释为目标线索需要保持在工作记忆中和目标线索可直接从长时记忆中提取两种加工过程。

回应：采纳您的建议，对文中表述做了修改。

正文中已改为“...，前额叶（工作记忆的重要脑区，Goldman-Rakic, 1996）受损的猴子很难完成不同试次之间搜索目标不断变化（此时目标线索需要保持在工作记忆中）的搜索任务，但能完成搜索目标不变（此时可直接从长时记忆中直接提取目标线索）的搜索任务...”。

意见 3:“..., 近年来大量运用传统视觉搜索范式进行的研究已经开始探讨真实场景的视觉搜索 ...。”这一句话似乎存在表述上的问题:“运用传统视觉搜索范式”进行的研究一般不会探讨真实场景视觉搜索的现象,而“运用真实场景视觉搜索范式任务”来探讨真实场景视觉搜索现象则更为合理。

回应:这里出现了笔误。现将“传统视觉搜索范式”改为“视觉搜索范式”。因为传统视觉搜索研究和真实场景搜索研究在实验程序上是相似的。

正文中已改为“..., 近年来大量运用视觉搜索范式进行的研究已开始探讨真实场景搜索任务...”。

意见 4:“这些不同点也就使得影响传统视觉搜索和真实场景搜索的因素之间存在着较大的差异。”“传统视觉搜索”这一中文术语与“真实场景视觉搜索”相比,显得含义非常不清。两个术语同时出现时,后者含义明确,而前者容易让不熟悉视觉搜索领域的读者感到困惑:何谓“传统搜索”?根据视觉搜索领域的相关英文文献,当前视觉搜索研究领域中所采用的任务可以分为 traditional visual search 和 visual search in scenes 两大类,后者还可以进一步区分 in artificial scenes 和 in real scenes。traditional visual search 这一术语在英文作者的论文中通常都是作为 paradigm、task 或 experiment 等的限定词。建议论文中用“传统视觉搜索任务”来表述,并对此术语进行解释。

回应:采纳您的建议,对文中表述进行了修改,采用“传统视觉搜索任务”和“真实场景搜索任务”表述。

意见 5:下文中对三种影响因素的介绍没有按照本段中场景情境、目标模版和视觉混乱度的引入顺序来展开,不符合一般论文写作的逻辑习惯。

回应:现已对三种影响因素的引入顺序做了调整。

意见 6:“传统视觉搜索研究发现,目标模板的具体化程度会影响搜索反应时(例如, Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004)。”主要的结论是什么呢?传统视觉搜索任务和真实场景搜索任务中,目标模版的具体化程度的影响有何异同?如何来解释本研究中获得的目标模版具体化程度具有调节作用的结果?

回应:添加了相关的结论,并在讨论部分就本研究中目标模板的具体化程度所起的调节作用进行了解释。

正文中已改为“利用传统视觉搜索任务进行的研究发现,目标模板的具体化程度能够系统地引导搜索,目标模板的具体化程度越高 搜索反应时就越短(例如: Wolfe, Horowitz, Kenner, Hyle, & Vasan, 2004)”。

意见 7:“在传统视觉搜索研究中,搜索画面项目数的多少(集合大小)是影响传统视觉搜索绩效的一个重要因素。而在真实场景的搜索画面中,集合大小通常是难以统计的,由此一些研究者提出采用视觉混乱度来代替传统视觉搜索中的集合大小(Henderson et al., 2009)。视觉混乱度是指由于视觉画面中存在过多的项目使其显示或组织令某些作业绩效下降的一种状态(Rosenholtz et al., 2007)。已有研究发现,视觉混乱度是影响真实场景搜索的一个重要因素,随着视觉混乱度的增加,搜索成绩明显下降(Neider & Zelinsky, 2010, 2011)。在传统视觉搜索研究中,通常使用反应时-搜索集合大小的函数斜率来表示搜索效率,那么真实场景中的搜索效率则可以使用“反应时-视觉混乱度函数的斜率”来表示。”这论断是作者自己提出的观点,还是已有研究者提出的,应该更明确地表述。如果是作者自己的观点,虽然具有

逻辑上的合理性,但在本文中的论证显得不足。应该尝试回答:1、视觉混乱度来代替传统视觉搜索中的集合大小是否合理,是否有更多的文献支持?2、为何不采用本问参考文献之一,即 Wolfe 等(2011)在其研究中采用多边形来计数的方法?

回应: 1.该问题在原稿中未表述清楚,已有研究者(Henderson et al., 2009)提出并验证了利用视觉混乱度代替集合大小的有效性,现已对表述做了修改。2. 关于真实场景画面的集合大小,不同的研究者给出了不同的方法,本文采用的是 Rosenholtz 等(2007)提出的视觉混乱度作为集合大小的替代量。Wolfe 等(2011)采用多边形计数的方法源自 LabelMe 的数据库(该数据库是由 MIT 的 Antonio Torralba 领导的实验室开发的),LabelMe 数据库的大部分图片拍自国外,而且里面会出现人物(有研究表明画面中的人物对观察者的注意有特殊的吸引力)。我们权衡利弊,采用了视觉混乱度代替集合大小的方法。当然在以后的研究中也可采用 Wolfe 等(2011)的方法进行相互印证。

正文中已改为“搜索画面项目数的多少(搜索集合大小)是影响传统视觉搜索任务绩效的一个重要因素。针对真实场景搜索画面中的集合大小难以统计这一事实,一些研究者提出采用视觉混乱度来代替传统视觉搜索任务中的集合大小,视觉混乱度是指由于视觉画面中存在过多的项目使其显示或组织令某些作业绩效下降的一种状态 (Rosenholtz et al., 2007), 随后有研究者验证了这种替代方法的有效性 (Henderson et al., 2009)。并有研究发现,随着场景画面视觉混乱度的增加,搜索绩效明显下降 (Neider & Zelinsky, 2010, 2011)”。

意见 8:“同时也需注意到,在真实场景中,虽然特定的搜索目标一般与特定类别的场景相联系,如订书机与书房相联系,但是不同的书房在空间结构和组成物品种类、数量和特征等方面上是存在差别的,进而视觉混乱度也就存在差异,也就导致搜索效率存在差异。据此我们认为客体工作记忆是否对真实场景搜索效率产生影响可能与搜索目标是否变化有关。”这一段的逻辑显得非常不清晰,到底是要进一步介绍关于视觉混乱度的影响作用呢,还是搜索目标变化与否的影响作用,还是场景情境的影响作用。并且,如果是要介绍搜索目标变化的影响,和本文前面关于前额叶受损猴子一段应该有着密切的联系。为何分开写?如果是笼统的场景情境的影响作用,又显得和后面的实验操作联系不大。

回应: 删除该段内容。对文中相关表述进行调整。

意见 9:“我们采用视空工作记忆任务和场景搜索任务相结合的双任务范式设计了如下实验:操纵搜索目标物体是否变化(不变和改变)、和视空工作记忆负载(无负载、客体负载和空间负载)以及目标模板表征的精细程度(物体的提示形式:文字和图片);通过测量搜索场景图片的混乱度(具体方法参见 Rosenholtz et al., 2007)对搜索画面的混乱程度进行分组(高和低),...”行文至此,应该明确提出研究的论题,以利于读者依照论题来理解研究的实验设计。而不是到讨论部分的第一段再来明确论题。

回应: 采纳您的建议,对文中表述进行修改,并调整了相关的内容。

正文中已改为“综上所述,本研究尝试采用视空工作记忆任务和真实场景搜索任务相结合的双任务范式,考察视空工作记忆负载对真实场景搜索任务绩效的影响机制,同时探讨试次间搜索目标是否变化、目标模板的具体化程度以及搜索场景画面的视觉混乱度所起的调节作用。具体而言:①视空工作记忆负载(包括空间负载和客体负载)对真实场景搜索任务成绩的影响与对传统视觉搜索任务成绩的影响是否相同,目标模板的具体化程度、试次间搜索目标是否变化以及搜索场景画面的视觉混乱度是否会调节这些影响;②视空工作记忆负载(包括空间负载和客体负载)对真实场景搜索效率的影响与传统搜索效率的影响是否相同;目标模板的具体化程度、试次间搜索目标是否变化以及搜索场景画面的视觉混乱度是否会调节这些影响”。

意见 10: “...目标模板...”建议用“目标模版的具体化程度”作为影响变量的表述。

回应: 对该变量名称进行了修改。

意见 11: “...以此来考察视空工作记忆对真实场景搜索成绩和效率的影响,同时考察目标模板表征的精细程度与搜索画面的视觉混乱度所起的调节作用。”建议此段与后面实验设计部分结合在一起写,逻辑会更清楚。

回应: 采纳您的建议,删除该部分内容,与实验设计部分结合来写。

意见 12: “2.4.1 实验设计.....”按照本人对作者实验设计思路的理解,本研究的逻辑主线是对比三种视空工作记忆负荷情况下完成真实场景中的视觉搜索任务的绩效及搜索效率,而搜索目标是否在试次间变化、目标模版表征的具体化程度、搜索场景图片的混乱度三个自变量在实验设计中是作为调节变量来分析的。作者没有把诸多同时在操纵的自变量之间的这种关系清楚交代,增加了读者理解其实验设计的难度。

回应: 采纳您的建议,先交代一下自变量之间的关系,然后再说总的实验设计,以利于读者理解。

正文中已改为“本研究的主要目的是为了对比三种类型的工作记忆负载条件下观察者完成真实场景搜索任务的绩效,同时探讨试次间搜索目标是否变化、目标模板的具体化程度以及搜索场景画面的视觉混乱度三个变量的调节作用。因此该实验采用 3 (工作记忆负载类型: 无负载、客体负载和空间负载) \times 2 (目标模板的具体化程度: 图片和文字) \times 2 (搜索场景画面的视觉混乱度: 高和低) \times 2 (试次间搜索目标是否变化: 不变和变化) 的混合实验设计”。

意见 13: “实验中有四种搜索类型, ①图片提示搜索目标-搜索场景混乱度高; ②文字提示搜索目标-搜索场景混乱度高; ③图片提示搜索目标-搜索场景混乱度低; ④文字提示搜索目标-搜索场景混乱度低。在一次实验中, 对于同一张搜索场景, 当用文字提示搜索目标时, 不可再用图片进行提示, 即每位被试在一次实验中对同一图片搜索一次。对于这四种搜索类型采用被试间平衡。”如前文所述, 目标提示形式和搜索场景混乱度是两个被试内变量, 由两个变量 2 \times 2 组合出来的四种搜索类型如何进行被试间平衡, 240 张图片又如何在四种搜索类型中分配?

回应: 原稿未表述清楚, 实际上每个被试在正式实验中共搜索 120 张图片 (240 张包括原件和复件), 视觉混乱度高低各 60 张。平衡方法说明如下: 现用 P 代表图片提示, W 代表文字提示, H01-H60 代表视觉混乱度高的 60 张图片, L01-L60 代表视觉混乱度低的 60 张图片。例如, 参加者 1 完成的提示形式与材料的组合为 PH01-PH30, WH31-WH60, PL01-PL30, WL31-WL60; 参加者 2 完成的提示形式与材料的组合为 WH01-WH30, PH31-PH60, WL01-WL30, PL31-PL60。这样做的目的是为了使每一张搜索图片的目标都有两种可能的提示形式。因此, 至少需要两个被试才能完成提示形式与材料组合的平衡。我们的总被试量为 16 人 (2 的倍数)。为方便读者理解, 我们对文中表述做了进一步的修改。

正文中已改为“...每种实验条件中有四种搜索类型, ...。对于同一张搜索场景画面, 当用文字提示搜索目标时, 不可再用图片进行提示, 即每位被试在所完成的实验中对同一图片搜索一次, 正式实验中每位被试共搜索 120 张 (视觉混乱度高低各 60 张) 场景图片, 这 120 张场景搜索图片在实验中呈现的顺序是随机的, 而图片提示和文字提示与材料的组合则采用被试间平衡; 具体来说, 用 P 代表图片提示, W 代表文字提示, H01-H60 代表视觉混乱度高的 60 张图片, L01-L60 代表视觉混乱度低的 60 张图片; 参加者 1 完成的提示形式与材料的组合为 PH01-PH30, WH31-WH60, PL01-PL30, WL31-WL60; 参加者 2 完成的提示形式与材料的组

合为 WH01-WH30, PH31-PH60, WL01-WL30, PL31-PL60, 以此类推。...”。

意见 14: “...由于本研究的目的是为了考察视空工作记忆任务对真实场景搜索绩效的影响, 我们关注的是不同实验条件下与搜索任务有关的因变量结果, 而非记忆任务的成绩, 因此下面的结果分析是针对与搜索任务有关的因变量进行的...” 记忆任务的成绩如果不是比较高 (一般 0.85 以上的正确率可以接受), 就可能存在被试在两个任务上认知资源分配的权衡, 所以应该报告记忆任务的成绩, 剔除记忆任务成绩明显低的被试数据。但在本研究确实无须进行进一步的统计分析。

回应: 在修改稿中报告了不同条件下视空工作记忆任务的正确率。审稿人认为记忆任务的正确率一般在 0.85 以上可以接受。我们查阅 Woodman, Vogel, & Luck (2001) 的研究结果, 发现单客体和单空间工作记忆任务的正确率大约分别在 0.85 和 0.70 的水平。实际上我们也用另外两组被试分别测试了单客体和单空间工作记忆任务的正确率, 分别为 0.81 ± 0.09 , 0.67 ± 0.02 , 与 Woodman 等人的单工作记忆任务的正确率数据基本一致。Woodman 等人双任务条件下客体和空间工作记忆任务的正确率大约在 0.75 和 0.65 的水平, 我们在修改稿中补充的对应数据大约为 0.67 和 0.59。由于本研究是关于工作记忆负载对真实场景视觉搜索的影响机制, 不能够将本研究的数据与 Woodman 等人对应条件下的数据直接做比较。基于以上分析, 我们认为修改稿中表 2 的正确率数据是可以接受的。

修改稿正文中增加了表 2 的正确率数据。

意见 15: “...不同条件下结果的描述统计如表 2a 至表 2f 所示, ...” 由于研究设计中自变量和因变量都比较多, 导致描述统计结果颇占篇幅, 且不易解读。建议结合推断统计部分有选择性地报告重要的或直接支撑论点的描述统计指标。描述统计表如果确有必要, 建议作为论文的附录。

回应: 采纳您的建议, 结合推断统计选择性报告重要结果(增加了统计图 2、3 和 4)。同时将描述统计结果放到附录中。

正文中已改为“...不同实验条件下搜索任务的描述统计结果 见附表 1a 至附表 1f 所示...”。

意见 16: 推断统计部分分析的思路与研究论题的两个方面 (本文讨论部分的第一段概括的两个问题) 之间的逻辑关系不够对应。根据《心理学报》的审稿指南: “数据分析的顺序要符合研究逻辑, 即与文献综述和讨论部分的表述顺序 (思路) 相一致”。

回应: 采纳您的建议, 对文中小标题的层次进行调整, 与研究论题相呼应。

意见 17: “在研究中作为调节变量的两个自变量之间存在交互作用, 如何解释呢? 这一分析与研究的论题是否有密切的关系? 做四因素分析是否有必要? (下同)

回应: 采纳您的建议, 删减不必要的统计分析。这里说一下修改稿中本研究结果的分析思路: 主要分析工作记忆负载类型的主效应及与调节变量的交互作用和简单效应, 选择有新意的调节变量之间的交互作用做分析, 突出要研究的主题。一般来说, 如果变量之间的交互作用显著了, 就不再对主效应做更细致的分析, 而关注简单效应。在讨论部分对主要的结果和有新意的结果做进一步的讨论分析。

意见 18: “...搜索画面视觉混乱度高时的搜索效率高于视觉混乱度低时的, $F(1, 90) = 544.216$, $p < .001$, $\eta_p^2 = .858$...” 这是个有趣的结果, 但如何从理论上进行解释?

回应: 删除这里的主效应, 将该结果放在与目标提示形式的交互作用中, 在讨论部分从理论上对之进行解释, 具体解释参见文中表述。

意见 19：“...综合正确率和搜索反应时两个指标发现，与无负载条件相比，空间负载会降低真实场景搜索的成绩，而客体负载的影响较小（仅表现在正确率指标上），...”这一结果应该是与传统视觉搜索任务中所获得的结论最有差异的地方，建议进行深入的分析 and 论证。

回应：后面部分对该结果的原因做了较深入的分析（见下一段标蓝色的部分，详见下一个问题的回应部分）。

意见 20：“所不同的是：该研究发现客体负载对真实场景搜索成绩的影响较小，究其原因可能是由于真实场景的复杂性造成的。...因而不问之前是否有记忆负载当个体一看到真实场景，客体工作记忆便都处于满载状态，也就使得客体工作记忆不能对多余特征进行存储，使得返回抑制的作用减弱，表现为有无记忆客体负载对搜索的反应时没有多少影响。”这一论断太牵强，搜索任务中场景特征数量大会导致客体工作处于满载状态，这难以让人信服，逻辑上不通，也缺乏实验证据。

回应：原稿表述不清，现已对文字表述做了进一步的修改，力求清晰准确。

正文中已改为“...究其原因可能是由于真实场景搜索画面比传统视觉搜索画面所包含的信息更为丰富和复杂所致。近期利用传统视觉搜索任务进行的研究表明，客体工作记忆在传统视觉搜索任务中的一个主要作用是对先前搜索过的干扰子产生返回抑制，而且该作用的产生受制于客体工作记忆的容量，一旦搜索集合大小达到客体工作记忆的容量后，客体工作记忆的负载对搜索成绩的影响就很小了(Emrich, Al-Aidroos, Pratt, & Ferber, 2010)。经典研究发现，客体工作记忆的容量约为 4 个项目（如，Luck & Vogel, 1997），而这个数量与真实场景中物体的数量相比是微乎其微的，因此，不管之前是否有客体工作记忆负载，个体在真实场景中搜索一段时间后，由于返回抑制的原因，客体工作记忆便处于满载状态，这使得个体在后续搜索过程中的返回抑制能力减弱，表现为有无客体工作记忆负载对真实场景搜索的成绩没有多少影响”。

意见 21：“...这说明不是由于注意资源分配造成的，而是实验的要求造成的”这一结果是否会影响其他结论的可靠性？

回应：本研究的数据还无法直接回答您提出的这一问题，由于目前这方面的研究还比较少，该研究中的参数设置是参考了利用传统视觉搜索任务进行的研究，我们认为对该问题的回答还需要将来的同类研究做进一步的印证。

正文中已改为“由此可见，该结果似乎不是由于注意资源分配造成的，而是当前实验的要求所造成的，未来研究可通过延长搜索时间的上限来对这一解释做进一步的检验。”

意见 22：“...本研究中仅仅探究了目标模板、视觉混乱度对“视空工作记忆影响真实场景搜索”的调节作用，将来的研究...”调节变量不是还有“搜索目标是否变化”吗？

回应：已对文字表述进行修改。

正文中已改为“...本研究仅探究了试次间搜索目标是否变化、目标模板的具体化程度和搜索画面的视觉混乱度对“视空工作记忆影响真实场景搜索绩效”的调节作用，将来的研究...”。

意见 23：“...将来的研究可以考察执行工作记忆负载真实场景搜索绩效的影响。”次句话似乎不通顺。

回应：已对文字表述进行修改。

正文中已改为“...将来的研究可以考察执行工作记忆负载对真实场景搜索绩效的影响。”。

第二轮

审稿人 2 意见：最后还有一处文字表述细节的修改建议已在审改稿中标出。

意见 1：再吹毛求疵一下，建议把“目标模板的具体化程度、试次间搜索目标是否变化以及搜索场景画面的视觉混乱度”三个变量的表述顺序改为与本段第一句中“试次间搜索目标是否变化、目标模板的具体化程度以及搜索场景画面的视觉混乱度”一致的顺序。

回应：感谢审稿专家细致的修改意见，现已对三个变量的表述顺序进行了修改，在修改稿中已用蓝色标注（见正文第四页批注）。