

《心理学报》审稿意见与作者回应

题目：相对到达时间任务中飞行员对客体特征与运动特征的分离

作者：晏碧华，游旭群

第一轮

审稿人 1 意见：修改后再审。

意见 1：本研究是这样表达研究目的的：“本文拟以运动客体的加工与判断为基本任务，采用 RAT 任务，拟分别采用大小、质地颜色对运动客体进行区分，并考察改变了空间构型的背景（也属非运动特征）对运动特征的影响，而运动特征的设置具有运动方向、大小速率等。通过设置对照组将飞行员与普通被试进行对比，寻求发现两组在运动客体判断及其加工机制上的差异，尤以找到两组在视觉客体特征和视觉空间运动特征上及其相互关系上的处理差异，找到空间能力较高的飞行员在动态空间任务的加工优势，为动态空间能力在飞行员选拔和飞行训练中的运用提供实证研究参考。”我觉得这样的表达逻辑不清楚，作者应该对本研究的目的做一个梳理，并明确表达。

回应：谢谢！确实是这样。已经修改为“本文通过设置 RAT 任务的变式，考察了运动客体大小、颜色的变换、空间构型背景的改变对相对到达时间判断的影响，同时在运动特征上考察了不同运动方向、大小速率对判断的影响。通过设置对照研究将飞行员与普通被试进行对比，寻求发现两组在进行运动客体时间判断时对客体特征和运动特征及其相互关系的处理差异。”

意见 2：引言部分，作者对“视觉客体空间的分离及其神经通路”等五个领域的研究做了介绍。问题 1 中，我已经说了，作者没有清楚的表达其研究目的，因此，我也不能确定这五个部分的研究和本研究主题都有关系？希望作者能够按研究目的，介绍相关的研究。如果一个研究主题和五个领域的研究都有关系，这研究的主题本身就显得过于宽泛了。

回应：综述部分已集中于动态空间能力及 RAT 研究，按照“飞行员空间能力及动态空间能力——RAT 任务特性及研究现状——‘特征分离’研究思路的来源分析、研究必要性——研究目的”进行分析。

意见 3：建议作者专门增加一段来介绍对以往研究的分析，并在此基础上提出自己的研究假设，和研究的创新点。

回应：删除关联不大的内容后已经突出本研究的研究重点，已经清楚说明研究假设和关注点。需要说明的是，本研究为原创研究，现有文献中，RAT 既没有个体差异研究，也没有变换客体特征的研究，因此可参照的文献较少，是现有研究的“平行研究”，非“深入研究”。

意见 4：统计分析中应该指出各实验删除的数据的总的数量。实验一中没有说到总的删除的数据的数量（原文第 10 页）。实验二和实验三没有对实验数据的预处理进行说明。

回应：删除工作是按照删除标准剔除了被试的个别或小部分无效试次（trial），对每个被试的有效测试的数据进行了统计，没有“总的数量”。实验二实验三的预处理同实验一标准，已经添加和说明。

意见 5：结果讨论中，希望能够增加和以往类似研究的结果的对比分析。

回应: 已经添加, 并已删除一些不必要的内容, 删除了讨论部分和综述部分涉及“空间构型”、“表征建构”的相关内容, 这些表述内容应该更适合空间记忆任务的研究而非本篇的相对时间判断研究, 并对研究结果进行集中讨论。

审稿人 2 意见: 我认为研究的结果是比较清晰的, 但是一些问题使得我认为本文至少需要重大修改。

意见 1: 我读完引言后并不理解为什么要做这个研究, 究竟什么是本研究试图解决的问题, 本研究是解决了什么重大理论问题, 有什么突出的新发现? 比如作者提到: “经验变量中, 大小到达效应 (the size-arrival effect, ERS) 说明在进行 TTC 判断时个体倾向于判断大客体先碰上目标, 或认为大的先到达, 即客体大小影响到达时间估计与预期 (DeLucia&Warren, 1994)。……但后期研究采用客体中心刺激也发现了如大小效应和背景依赖性特征等 (Calabro, et al, 2011)。”这样看来你的结果并不是比较新的发现, 也没有说清有什么重大的理论贡献。

回应: 国内还没有动态空间能力的研究 (除本团队前期研究), 国外也没有动态空间能力的飞行员加工优势研究, 本研究发现了飞行员和普通被试在相对到达时间判断中的一系列加工差异, 例如“……飞行员的处理优势还表现不受客体大小、视线方向改变和目标线特征改变的影响……”。这些都是新的发现, 对飞行员选拔训练也有重大现实意义。

对于一般被试, 大小到达效应原初是在 TTC 任务中发现的, 本研究 RAT 任务也发现了这个效应, 这也是一个新的发现。对于一般被试, “客体颜色不影响运动客体的相对时间判断, 客体大小、运动方向、速率大小、视线方向以及背景特征影响判断……”, 这都是 RAT 任务新的发现。

意见 2: 作者提议: “在未来的航线飞行员选拔和训练中, 把极富生态性的动态空间能力测试引入, 任务越具有核心价值 and 代表性, 就越能够在选拔之初解决经过系统训练才能解决的问题, 从而提高选拔效率, 设置综合性核心动态空间任务训练还可以极大提高训练效率。”这个建议是建立在飞行员相对于控制组在 RAT 任务上的优势不是训练产生的, 但是这个假设未经验证。

回应: 飞行员 RAT 优势的产生一方面可能是选拔过程中获得, 另一方面可能是在飞行实践中获得, 但目前民航选拔系统并没有动态空间能力的引入, 我们更相信它是在飞行实践中产生。原文的意思是: 找到相应的飞行员加工优势后, 设置相应的选拔任务和训练任务, 那就可以事半功倍, 不然要用大量的飞行实践去提升动态空间能力不划算。如果有候选者在动态空间能力上有天赋, 更是大大节省了训练成本。从训练上看, 任务越核心, 训练越有效。因此, 原文没有歧义。

意见 3: 统计方法上有问题, 作者能确定使用的是 MANOVA 吗? 还有 $p=.0000$ 是非常奇怪的表达方式。既然一个因素的两个水平产生的主效应显著, 为什么还要 t 检验, 还有数据的报告方式很奇怪, 图上只有各个主效应, 而没有体现出交互效应, 令人看了非常迷惑。

回应: 确定使用的是 MANOVA。需要说明的是, 初稿中数据汇报面面俱到, 应集中汇报有效数据。图上是各水平下较为详细的平均数, 已经可以较直观展现各水平之间的关系, 用数据说明交互作用即可。

意见 4: 作者讲到已知 what 和 where 既分离有交叉的情况, 但是这个研究对我们有什么新的提示吗? 总之虽然提到了很多理论背景, 但是对本文的理论贡献来说很模糊。

回应：已经大量删除该部分内容，综述内容集中于动态空间能力、RAT 及影响因素研究。

审稿人 3 意见：

意见 1：论文前半部分论述了视觉客体空间的分离及其神经通路，指出对运动客体的判断是两条通路的协同作用。后半部分虽然通过实验证实了在相对到达时间任务中飞行员对某些客体特征与运动特征进行了分离加工，抑或进一步说，在相对到达时间任务实验中，飞行员组被试的判断主要受运动特征的影响，而普通被试的判断兼受客体特征（客体大小）与运动特征的影响。后半部分的实验一方面不能有效证实对运动客体的判断是两条通路协同作用的，以及如何协同作用的？另一方面在讨论时也未能在两部分内容之间建立起有机的联系。从总体上来说，这两部分内容感觉是分离的。

回应：当时初稿前半部分论述的是以往研究成果或观点，总结出两条通路的协同作用作为神经基础对运动客体判断的重要性，运动客体的加工有两条通路的功能，本研究则证实了对运动客体运动特征和客体特征的加工存在个体差异，这种个体差异可能体现了两条通路在两类特征上的加工存在个体差异。

修改稿已经对综述内容进行了调整，以运动空间能力为主线而不是以两种特征的分离为主线展开。

意见 2：实验确证在相对到达时间任务中，飞行员的绩效更好。但是，这种更好的绩效既可能与其飞行经验有关，也可能与他们本身就具有良好的空间能力有关（也就是说，他们当初正是因为具有良好的空间能力才被选作飞行员的）。因此，研究结果对飞行员选拔和飞行训练中的实践意义不是很明晰。

回应：一方面，当前民航飞行员选拔系统中空间能力的选拔并没有运动空间能力测试，因此无论是针对飞行员选拔还是飞行员训练都有研究必要性。另一方面，若拿刚入学飞行学员和现役飞行员作比较，虽可以证实飞行员的优势加工是飞行经验的结果，但也会有如年龄、学历等不匹配问题。因此，在动态空间能力研究领域用飞行员和年龄、学历匹配的普通被试作对比可为飞行员选拔和训练带来启发。

意见 3：第 6 页：“RAT 任务研究还较少，有关客体特征与其运动特性的相互关系还没有展开。客体特征包括客体大小、形状、颜色、纹理特征等，而客体运动特征有速度、方向、距离、位置变化等。那么客体特征如何影响运动特征？这正是本研究要关注的。”这句话的表述似乎存在问题，论文未有探讨客体特征如何影响运动特征，探讨的是客体特征与运动特征一起如何影响相对到达时间的判断。

回应：是原文表达错误。修改为“这些客体特征和运动特征的变化是如何影响相对到达时间的判断呢？”

意见 4：第 8 页倒数第 2 行：trail 应为 trial.

回应：已经修改，谢谢！

第二轮

审稿人 1 意见：修后再审。

意见 1：我对文章作者提出的意见，作者基本上都做了较好的回答和补充。有一个问题需要作者进一步补充的。

4、统计分析中应该指出各实验删除的数据的总的数量。实验一中没有说到总的删除的数据的数量（原文第 10 页）。实验二和实验三没有对实验数据的预处理进行说明。

答：删除工作是按照删除标准剔除了被试的个别或小部分无效试次（trial），对每个被试的有效测试的数据进行了统计。没有“总的数量”。实验二实验三的预处理同实验一标准，已经添加和说明。

本系列实验反应时处理标准是：剔除错误反应的反应时，以及三个标准差以外的正确反应时，按照变量维度对正确反应时进行统计。统观反应时，从经验上讲，由于要对运动着的两个客体进行对比，因此判断时间比通常一个客体的知觉判断实验时间要长一些。

上面是我第一次审稿提的意见和作者的回答，和文章中作者的表述。

我的意见是作者需要给出被试的错误反应次数，和三个标准差以外的正确反应时次数，和这些次数在被试在整个实验中的反应次数的比例。因为如果被试的错误反应次数过大，三个标准差以外的正确反应时次数过多都会影响实验数据的可靠性。换句话说，增加这些数据可以更进一步保证文章数据的质量。建议作者加上。

回应：好的，谢谢！已经添加数据说明。

审稿人 2 意见：作者对本文做了重大修改，解决了一部分问题，但是有几个个关键的问题没有得到很好的解决。

意见 1：飞行员组和控制组虽然控制了性别、年龄和文化程度。但是两组之间在其他很多方面（比如智力、以及静态的空间认知能力等）仍然可能存在差异，到达时间判断任务绩效的组间差异究竟是什么引起的仍有疑问。作者应该提供这些方面的证据表明两组之间没有这些差异。更为关键的是两组间的经验和训练差异可能能在很大程度上解释两组间的差异。作者认为即使这种差异来自于训练，也可以节省训练资源。这个也是一个未验证的假设，如果这种 RAT 的绩效是随着飞行员正常的训练程序而自然提高的，并不需要额外的训练的话，节省训练资源的假设就是不成立的。除非有证据表明 RAT 成绩好的人员训练得更快更好。事实上，现在研究思路其实是一个比较组间差异的设计，而不是研究个体差异的设计。

回应：（1）关于两组之间的其他方面差异问题。智力测试中并没有运动空间能力测试，也没有证据表明智力和运动空间能力的关系，“可能有影响”只是猜测，且文化程度在一定程度上可以代表智力水平；静态空间能力和动态空间能力是两个独立因素结构（已有论述，或参考本文所引综述文章），个体运动空间能力的判断绩效和静态空间能力相关不大，因此没有检测。众所周知男女空间能力表现显著不同，因此对性别进行了控制。对年龄进行了控制是基于飞行经验影响的作用。和常人比较飞行员长期接触运动空间，有理由相信飞行实践/经验会对运动空间能力带来影响。至于为什么不认为是选拔带来的优势，已经在第一次审稿意见中回答。

（2）关于“更为关键的是两组间的经验和训练差异可能能在很大程度上解释两组间的差异”。本文的基本设计逻辑就是假定飞行员的飞行实践可以影响 RAT 判断绩效，实验也证明了这种假设（实验目标实现了）。然后接下来的事情（不是本文要解决的）就是用选择合适的动态空间任务去进行集中训练（和选拔），比如尤其是模拟机训练，核心任务的集中训练提升运动空间能力是不是比“随着飞行员正常的训练程序而自然提高”可以更快达到高水平？另外，在模拟机训练中提升能力是不是比在实际飞行中提升运动空间能力更为经济？这是“节省”两重含义。

“除非有证据表明 RAT 成绩好的人员训练得更快更好。”首先，文章仅有的几句文字所说的训练讨论的是指文章发现的运动空间能力的训练和提升，不是说仅仅训练这个 RAT 任

务就可以让飞行员整体训练绩效很好。其次，若有 RAT 任务先天就很好的个体，那在这种能力上就不需要过多训练，已经“更快更好”，这种起点较高的飞行员候选者需要选拔发现。

总之，本文主旨是通过实验研究“发现”差异，今后的后续研究可能会找到对训练绩效和飞行绩效预测力较好的核心动态任务，建议用于选拔与训练。

(3) 本文的研究设计可理解为，在对个体变量进行控制的情况下，设置对照组考察飞行实践对 RAT 任务判断的影响。

本研究设计参照了前人的研究，如以下研究等，以及国内有关飞行员（或模拟机系统训练人员）和对照组（普通被试、飞行学员）的对照研究设计。

Dror IE, Kosslyn SM, Waag WL. (1993). Visual- spatial abilities of pilots. *Journal of Applied Psychology*, 78, 763 – 773

Gordon, H. W., & Leighty, R. (1988). Importance of specialized cognitive function in the selection of military pilots. *Journal of Applied Psychology*, 73, 38–45.

意见 2: 作者认为对比飞行员组比控制组在 RAT 任务上有优势就表明 RAT 任务可以作为飞行员选拔和训练的工具。但是这个优势并不是充分的证据。还需要考虑以下几个方面：1) RAT 任务究竟能够预测飞行员那些任务作业绩效，作者应该提供效标，这样才能说明 RAT 任务绩效是能够预测飞行绩效的选拔工具。2) RAT 的信度怎么样，如果信度不好也是不能用作选拔的，所以最好能提供信度指标。

回应: 如 1 中 (2) 回答，飞行员有加工优势说明飞行实践增强了飞行员的诸如 RAT 这样的运动空间能力，有必要在选拔和训练中引入，文章只是建议，只是阐述了“未来的”训练和选拔引入的重要性。

目前并没有有关 RAT 的飞行绩效效标的研究，如果真要引入选拔和训练，其信度和效度是选拔和训练研究必定重视的(本团队也做航线飞行员选拔系统研究并获教育部科学研究优秀成果奖二等奖)。本文主要是通过实验研究找出差异，至于应用，还需假以时日。

意见 3: 作者说确实是采用了 MANOVA，但是没有明确是对那些因变量做了 MANOVA，也没有报告 MANOVA 的总体结果，只是报告 ANOVA 的结果，从结果看作者只是做了 ANOVA 分析。

回应: 谢谢！很惭愧，已经修改，是重复测量方差分析。

意见 4: 作者认为“在加工运动客体时，对于有良好动态空间能力和飞行实践的飞行员来讲，where 任务的加工具有高自动性，因此能将客体特征和运动特征自动分离。而控制组身上更容易发生视觉客体特征和空间特征的关联。”这个解释不能完全让人信服，也有可能是飞行员和控制组都能很好将客体特征和运动特征关联起来，只是飞行员能更加精确判断距离和速度。事实上这也是可以用实验来检验的，如果能够显示飞行员相对于控制组在客体特征和空间特征的关联上变差了，这样才能支持作者的理论。

回应: 谢谢！虽然文中已经说明是“可能解释”，但是这个推论确实不够谨慎，尤以“自动分离”为甚。

不过，需要说明的是，原意是通过“飞行员在客体条件改变下成绩不受到影响而控制组受到影响”这样的结果推论出“飞行员没有发生视觉客体特征和空间特征的关联而控制组发生了”，貌似推论可以成立。意见提出的“也有可能是飞行员和控制组都能很好将客体特征和运动特征关联起来，只是飞行员能更加精确判断距离和速度”也有歧义，既然能精确判断，说明就没有关联。“如果能够显示飞行员相对于控制组在客体特征和空间特征的关联上变差了，这样才能支持作者的理论。”从实验结果看是飞行员就没有关联，不是变差了。

第三轮

审稿人 1 意见：我对文章提出的意见，作者都进行了很好的修改。我没其他的意见。

审稿人 2 意见：作者已解决了我提出的问题。

意见 1：最后一个建议，关于文章最后两条通路的部分建议再弱化一点，因为基本上是作者作出的推论，没有太直接的证据。然后文章以一系列的关于两条通路的理论问题作为文章的结尾也很奇怪。

回应：已经参照审稿人意见进行了部分修改。因为文章已经说明是可能解释，讨论中有关两条通路的内容也留下少些句子以呼应引言中的研究设计构想来源。最后一段以动态空间能力研究前景作为结尾，以突显该主题研究的重要性。