

《心理学报》审稿意见与作者回应

题目：自我与他人视角的动态信息加工：一项行为振荡研究

作者：孙楚 耿海燕

第一轮

审稿人 1 意见：

论文通过三个实验探讨了自我和他人视角对面孔信息加工的影响，发现存在两种视角之间的周期性行为振荡现象。论文采用原创的实验范式，报告的结果具有潜在的理论价值，行为振荡的视角也比较新颖，但我对论文的研究设计和数据解读存在一些疑问：

意见 1：论文通过一张俯视角的人偶图像诱发他人视角的加工，但这一操作是否真的能达到所需的效果，缺乏足够的理论和数据论证。比如在实验一的结论中作者提到：“被试会自发地采择人偶的视角，从而具身性地获得人偶视角下的视觉体验”。然而，该实验结果发现“朝向人偶”与“朝向人偶对侧”两者间的差异远低于它们与“朝向自我”间的差异，这很难将“朝向人偶”与“朝向人偶对侧”间的差异直接完全解释为是因为采择了人偶的视角。后续的一系列表述和研究都是基于该未被充分论证的结论，这导致整个论文的逻辑缺乏坚实的基础。建议作者进一步阐明这一实验操作背后的逻辑，并引用相关文献论证其合理性。

回应：感谢审稿专家的意见。关于实验一结果的讨论确实不够充分。对于实验一的结果解释基于面孔加工的方向特异性，个体对于正立面孔的加工要优于其他方向面孔的加工，在实验一中，尽管实验中朝向左侧或朝向右侧的面孔在被试的视角下都非正立面孔，但是由于人偶的存在，朝向人偶与朝向人偶对侧的面孔却表现出了识别正确率上的差异，说明这种差异是由于人偶的引入所导致的，个体自发地采择了人偶的视角，具身性地获得人偶视角下的视觉体验，从而对于人偶视角下的正立面孔产生了加工优势。事实上，这种自发性的视觉观点采择及具身性加工对被试完成任务的影响已经在很多研究中得到了证实，例如在 Surtees 等人(2016)的研究中发现，当被试的对面坐着另一名被试，该被试对于面前桌上数字“6”和“9”的判断反应时和错误率都会高于“5”和“8”（呈现方式如图 a 所示），这是因为被试自发地采择了对面被试的视角，自我视角下的“6”和“9”在对面被试的视角下变为了“9”和“6”（不一致条件），而自我视角下的“5”和“8”在对面被试的视角下仍然是“5”和“8”（一致条件）；又例如在 Ward 等人(2019)的研究中，同样通过在视觉场景中设置人偶图片来引入他人视角，被试的任务是判断旋转了不同角度的字母是“R”还是“Я”，结果发现当字母在他人视角下为正立时，可以促进被试对于字母的判断，这也说明被试通过自发性的视觉观点采择所获得的具身性视觉体验影响了被试的任务表现。除此之外，该研究的结果还表明尽管他人视角下正立字母的加工受到了促进，但仍然远不如自我视角下正立字母的加工速度。这一结果也不难理解，相比于他人视角，自我视角是在生活中更常使用、得到更多训练的视角，因此占据主导地位，在大量视觉观点采择的研究中都可以看到这种自我视角的优势。这也可以解释为什么在实验一中“朝向人偶与朝向人偶对侧两者间的差异远低于它们与朝向自我间的差异”。实际上，实验一的研究除了将实验刺激替换为面孔图片，在实验逻辑上与上述举例的研究是一致的，最终得到的也是相同的结论。根据审稿专家的建议，我们也在实验一的讨论部分进行了内容上的补充。

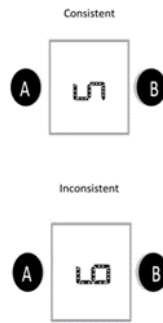


图 a. Surtees 等人(2016)研究中的实验模式图，其中 A、B 代表两名面对面的被试。

Surtees, A., Apperly, I., & Samson, D. (2016). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition*, 150, 43-52.

Ward, E., Ganis, G., & Bach, P. (2019). Spontaneous vicarious perception of the content of Another's visual perspective. *Current Biology*, 29(5), 874-880.

意见 2: 实验二和实验三均对正确率数据做了平移处理，即自我和他人条件下的正确率分别减去各自的整体均值，但是这样处理是否必要？平移处理之后再比较两个条件的正确率，这种比较究竟反映了什么？作者在实验二的讨论中提到“在 0.7~1 秒左右，被试对朝向人偶面孔的识别表现出相对的优势，这体现出自发性视觉观点采择效应的增强”，但从原始正确率来看，个体对朝向自我的加工正确率并未明显减弱，且从未显著低于他人条件，与作者试图论证的“周期性交替优势”相违背。建议作者充分论证平移处理的必要性，并更加仔细地斟酌对结果的表述和解读。

回应: 感谢审稿专家的意见。对数据的平移处理可以将数据的均值调整为 0，从而使得数据的波形在 x 轴附近波动，方便进行傅里叶变换的操作。实际上，审稿专家在之后问题中提到的在行为振荡研究中常用的去除线性趋势(detrend)的处理也是一种相当于去除均值的方法，只是同时还去除了一部分低频信息。而正如在总讨论中提到的，过往的行为振荡研究发现的主要是 4Hz 以上的振荡频率，因此去除线性趋势可以帮助他们排除低频信息的干扰。而在本文中并没有采用去除线性趋势的方式也是由于本文中发现的是低频的行为振荡现象，去除线性趋势就相当于去除了实验效应。因此我们使用平移的方法替代了去除线性趋势的方法。而审稿专家提到“从原始正确率来看，个体对朝向自我的加工正确率并未明显减弱，且从未显著低于他人条件”，这确实是本研究相比于以往行为振荡研究的另一个特殊性。在过往研究中，发生振荡的两种条件往往是完全对等的（例如左视野和右视野，向左的启动与向右的启动等），但在本研究中，发生振荡的双方是自我视角与他人视角，而自我视角下的信息加工相比于他人视角下的信息加工有明显的优势性，因此出现了自我视角下的加工正确率始终高于他人视角下的加工正确率的现象，这也与我们在实验一中得到的结果相一致。尽管我们在实验二和实验三中对自我视角下的面孔图片进行了模糊处理来削弱这种优势性，然而由于个体差异等原因，我们很难通过准确地设置模糊量来使得二者的正确率趋于相等。但这并不影响我们对实验结果的解读。因为傅里叶变换是分别分析每个视角条件的数据结果，对两条曲线各自的平移处理并不会影响傅里叶变换的结果。在对两条数据曲线进行比较时，我们也采用了“相对”这一谨慎的表达方式。举例来说，当两条曲线的均值都为 0，那么朝向人偶的曲线相比于朝向自我的曲线在某一时间段内更高就意味着：在这一时间段内，以朝向人偶面孔加工的平均水平为基准来评判朝向人偶的正确率，要优于以朝向自我面孔加工的平均水平为基准来评判朝向自我的正确率，也就是表现出了相对的优势性。我们也可以用生活中的例子来理解，一名好学生在以前常常拿到 100 分的好成绩，但在最近的几次考试中只拿到 80

分，而另一名差生以前总是在 50 分徘徊，最近却总能考到 70 分，那么我们可以说，差生在最近的这段时间是相对变好了，但好学生的表现相对变差了。尽管差生的成绩从未高过好学生。

意见 3: 本文前言部分的内容组织有待优化：文中对过往研究的综述虽然内容较多，但缺乏侧重，详略分配不够合理，尤其是对关键的论据叙述不够详实。比如，作者的核心逻辑是：加工自我信息与他人信息都依赖镜像神经元，但不明确其是如何同时处理两种信息的，一种可能的假设是神经振荡。那么作者首先需要以充足且具体的证据说明镜像神经元分别是两者的神经基础；其次要论证镜像神经元对两者都有重要意义，否则若是极小部分神经基础共享则理论上未必需要探讨共享神经机制是如何加工信息的；最后要说明大脑为何有必要采取振荡的方式来加工两者的信息而非其他方式（只说振荡行为在很多认知过程中存在并不能说明在所有行为中都存在）。

回应: 非常感谢审稿专家的建议。对于前言中共享表征这一部分的介绍确实不够充分，因此在文章中补充了更多的相关文献来说明共享表征的重要性和普遍性，在各类认知加工过程中都可以发现自我加工与他人加工在神经表征上的重叠，详见前言第三段。这些补充可以进一步说明在共享神经机制的基础上讨论如何区分自我与他人信息的必要性。实际上，对于共享神经表征所带来的如何区分自我与他人这一问题也是研究者们一直都在探讨的热点话题，大量的研究和综述都在讨论哪些脑区可能参与这一区分过程，具体起到怎样的作用，我们在前言第四段进行了进一步的补充，同时在这里列举几篇综述供审稿专家参考 (Decety & Sommerville, 2003; Quesque & Brass, 2019)，而我们提出的行为振荡的假设也是因为前人研究都是在回答当自我信息与他人信息共享神经表征时，大脑如何分辨当前的神经表征是自我信息的表征还是他人信息的表征，然而，这些研究却不能回答，当我们需要同时加工自我信息与他人信息时，如何避免二者共享神经资源所带来的冲突。而行为振荡现象恰恰被认为是一种可以解决不同任务同时加工所带来的时间冲突问题的大脑工作方式。因此我们提出了行为振荡的假设，对于这一假设的提出逻辑我们也在前言第四段进行了完善。

Decety, J., & Sommerville, J. A. (2003). Shared representations between self and other: A social cognitive neuroscience view. *Trends in Cognitive Sciences*, 7(12), 527-533.

Quesque, F., & Brass, M. (2019). The role of the temporoparietal junction in self-other distinction. *Brain topography*, 32(6), 943-955.

意见 4: 实验一在报告了确定样本量的 power 分析，但信息不全，如前人研究的效应量为 0.5，这 0.5 具体是哪一种效应量指标（ d 、 η^2 等）？从计算结果来看，仅需 11 人即可达到 90% 的 power，这意味着参考的效应量是非常巨大的，不太符合一般经验。建议作者澄清计算过程和依据。此外，实验二、三未报告 power 分析，原因为何？如果没有做事前的 power 分析确定样本量，也可以尝试基于当前样本量做 sensitivity 分析，以论证当前的样本量是充足的。

回应: 感谢审稿专家的意见和建议。

首先是对于实验一的答复：在实验一中参考的前人的效应量指标为 η_p^2 ，我们在文中进行了补充。通过 GPower 计算得到的需要被试量较少，一方面是因为参考的效应量较大，另一方面是因为实验一的实验设计较为简单，为单因素三水平的重复测量设计。而在前人研究中的效应量 η_p^2 高达 0.5 是因为面孔加工的方向特异性是一种十分显著和稳定的心理学现象，在大量的研究中都得到了重复。而从本文实验一的结果来看也验证了这一结论：一方面，实验一中面孔方向主效应的效应量 $\eta_p^2=0.909$ ，已经高于参考的效应量指标，说明不同方向间的差异是十分显著的；另一方面，逐一分析单个被试的结果发现，12 名被试中仅有一名被试

的数据模式不完全符合总的的数据模式，说明这一实验效应是十分稳定的。基于以上信息，我们认为通过 GPower 计算得到的需要被试量是合理的，最终得到的实验结果也是可靠的。其次是对于实验二、三的答复：本研究的实验二和实验三为行为振荡研究，并没有采用传统的统计检验方法，因此无法使用传统的 power 分析和 sensitivity 分析，在前人发表的行为振荡研究中也并没有给出被试数量的计算标准，但实验二、三的被试量选择参考了这些研究中的实际被试量，相比较之下本研究的被试量是充足的，以下是前人研究被试量的举例：在一篇研究客体注意的行为振荡现象的文章中，被试数量 $n=14$ (Fiebelkorn et al., 2013)；另一篇研究特征注意的行为振荡现象的文章中，被试数量 $n=15$ (Mo et al., 2019)；在一篇研究启动效应的行为振荡现象的文章中，被试数量 $n=16/18$ (Huang et al., 2015)。而在本研究中，实验二的有效被试量为 26，实验三的有效被试量为 21。

Fiebelkorn, I., Saalman, Y., & Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. *Current Biology*, 23(24), 2553-2558.

Mo, C., Lu, J., Wu, B., Jia, J., Luo, H., & Fang, F. (2019). Competing rhythmic neural representations of orientations during concurrent attention to multiple orientation features. *Nature Communications*, 10(1), 1-10.

Huang, Y., Chen, L., & Luo, H. (2015). Behavioral oscillation in priming: Competing perceptual predictions conveyed in alternating theta-band rhythms. *Journal of Neuroscience*, 35(6), 2830-2837.

意见 5：第 2.3 节应该是“讨论”而非“结论”。

回应：感谢审稿专家的指正，已修改。

意见 6：数据的处理和分析上需要提供更多依据，比如对于实验二/三：①对被试的剔除标准需要有一定依据；②为何没有做常见的去除线性趋势（detrend）处理？

回应：感谢审稿专家的意见。

对于问题①，在实验二、三中，我们对于识别正确率高于 95% 的被试进行了剔除，以避免天花板效应的影响，因为个体差异的原因，个别被试可以轻松地识别实验中快速呈现的面孔图片，从而在不同的 SOA 条件下都达到接近 100% 的识别正确率，最终他们的正确率曲线接近一条直线，而这并不意味着他们不存在行为振荡现象，只是被天花板效应所掩盖，因此对这些被试的数据进行剔除是有必要的。从结果的角度来说，一方面，剔除的被试只占总体被试的一小部分；另一方面，我们也在不剔除被试的情况下进行了同样的统计检验，剔除被试前后的傅里叶变换的统计检验结果是一致的，我们也将这一信息补充到了实验二和实验三的相应位置。实际上，为了避免天花板效应而剔除成绩过高的被试数据的方法在过往的实验研究中也是有迹可循的(Davis et al., 1990; Hickok et al., 2011; Harty et al., 2014)，并且本研究中对于天花板效应的判断标准（95%）比举例中使用的一些标准更为严格。

对于问题②，正如在回答问题 2 时所说的，去除线性趋势的处理的确是行为振荡中常用的处理手段，其作用有两个，一是将曲线均值调整为 0，便于进行傅里叶变换，二是去除曲线中低频信息的干扰。在本研究中，我们探究得到的是低频的行为振荡现象，因此去除线性趋势会削弱实验效应，因此我们使用减去均值的平移处理来代替去除线性趋势的处理。

Davis, A., Haggard, M., & Bell, I. (1990). Magnitude of diotic summation in speech-in-noise tasks: performance region and appropriate baseline. *British journal of audiology*, 24(1), 11-16.

Hickok, G., Costanzo, M., Capasso, R., & Miceli, G. (2011). The role of Broca's area in speech perception: Evidence from aphasia revisited. *Brain and language*, 119(3), 214-220.

Harty, S., Robertson, I. H., Miniussi, C., Sheehy, O. C., Devine, C. A., McCreery, S., & O'Connell, R. G. (2014).

意见 7: 本文的总讨论部分将本研究的发现与过往研究联系得不够紧密，尤其是对于核心发现——个体倾向于以振荡方式处理自我和他人视角的信息，缺乏对自我/他人信息处理的认知和神经机制的过往研究的讨论。将本研究的结果与更多的实证证据对照，将有助于挖掘本研究的价值。

回应: 非常感谢审稿专家的建议。的确在总讨论部分，关于自我-他人区分的前人研究的介绍不够丰富，对加工机制的讨论也不够深入，因此在总讨论中进一步补充了相关内容。

.....

审稿人 2 意见:

本研究通过创设视觉观点采择任务，发现了自我视角和他人视角下的信息加工转换特点。作者认为本研究结果说明在有他人在场的社会场景中，自我视角和他人视角的信息加工是一种具有周期性的交替优势的过程。文章引入的科学问题非常有趣，有助于理解社会互动过程中人的信息加工特点。然而，本研究却存在科学问题与研究方法脱节，以及研究结果不支持假设等较大的问题。

意见 1: 引言部分大篇幅引入社会互动的概念，强调社会互动过程中自我与他人信息交替采择的重要性，似乎作者是想探究社会互动过程中自我与他人视角的动态信息加工，然而作者开发的“视觉观点采择任务”并无社会互动的成分，为了完成实验任务，被试并不需要带入社会互动的场景，仅需判断面孔的朝向。在实验一中，作者发现被试对朝向人偶的正立面孔的加工速度更快，但是这个结果其实可能由其他因素导致。例如，朝向人偶的正立面孔与人偶正好形成一个对称图像，对于对称图像的加工优势也可能造成实验一的结果。在实验二和实验三中，作者发现的结果可能是由于被试在每个试次前的准备状态导致的。被试会在每次试次开始前猜测这次呈现的面孔是什么样的（朝向自我或朝向人偶），这种猜测也可能会随着时间波动，并且导致正确率在不同的时间点上出现波动。所以实验二和实验三的结果可能不是视角波动，而是准备状态的波动导致的。综上，该任务缺乏明显的社会互动操纵，且没有对这些明显的干扰变量进行控制。

回应: 感谢审稿专家的意见。首先对于**第一个问题**，本研究是否是在探讨与社会互动相关的加工过程？本研究中讨论的社会互动是一种更广泛意义上的社会互动，准确地说属于内隐社会认知的范畴。在一个社会场景中，只要有他人在场，我们会自动、内隐地加工他人相关的信息，这可能是长期外显社会互动形成的内隐倾向，是社会适应的结果，也为随时发生的外显互动奠定了基础或提供了指导(Seyfarth & Cheney, 2013; Bargh & Williams, 2006)。本研究重点关注的是内隐发生的视觉观点采择过程，在一个有他人在场的视觉场景中，个体常常会不自觉地、内隐地想象和推测他人视角下的视觉体验，从而处于同时加工自我与他人两种视角下信息的状态，本研究正是在这种同时加工两种视角信息的基础上探究可能存在的行为振荡现象。事实上，很多研究都证实了这种内隐视觉观点采择的存在，例如当被试判断桌面上数字是“6”还是“9”时，会受到桌子对面个体的影响，因为在桌对面个体的视角下数字“6”“9”变为了数字“9”“6”(Zhao et al., 2015; Surtees et al., 2016); 又例如在 Ward 等人(2019)的研究中，同样通过在视觉场景中设置人偶图片来引入他人视角，被试的任务是判断旋转了不同角度的字母是“R”还是“Я”，结果发现当字母在他人视角下为正立时，可以促进被试对于字母的判断。这类研究的特点是要求被试完成的任务是与情境中的他人毫不相关的简单认知任务，而如果实验结果发现被试的任务表现受到了他人视角的影响，而这种影响又恰恰是因为采择他

人视角所产生的,那么就说明了这种自发性视觉观点采择的存在。根据审稿专家的这一问题,我们反思可能是在前言中对于内隐视觉观点采择的铺垫不足,因此在前言第一段补充了相关内容,并在文中多处补充“内隐”“自发”等表达,并补充了“内隐社会认知”作为文章的关键词,使文章的表述更加准确。

其次对于**第二个问题**,实验一中的效应是否可能是其他因素导致的:一方面,正如我们前面提到的,这种自发性的视觉观点采择与具身性加工对被试完成任务的影响已经在许多研究中得到了证实,实验一的研究除了将前述研究中的数字、字母等实验刺激替换为了面孔图片,在实验逻辑上与上述举例的研究是完全相同的,最终得到的也是相同的结论;另一方面,从本研究的刺激设置来看,应该可以排除审稿专家提出的对称性影响的假设。从实验的流程图可以看出,负责引入他人视角的刺激是一张人偶图片的俯视图,而被试需要判断的刺激是正面拍摄的面孔图片,一是前者并不具备面孔这一与后者相匹配的要素,二是两者的图片大小与比例都差异较大,因此很难将其认为是对称刺激;而如果笼统地将方向相对视为一种对称条件,那么朝向人偶对侧的图片反而构成了一种方向一致的条件,这种方向的一致性对识别的促进也有相应的实验证据,例如轮廓整合的相关研究(Field et al., 1993)。因此综合来说,实验一的研究结果不太可能是受到对称性等其他因素的影响。

最后对于**第三个问题**,实验二和实验三的结果是否可能是被试准备状态的波动所导致的。首先,如果像审稿专家提出的,被试会在每个试次开始前猜测这次呈现的面孔是什么样的,一方面,被试在每个试次中的准备状态可能是不同的,在一部分试次中可能猜测是出现朝向自我的面孔,另一部分试次中可能猜测是出现朝向人偶的面孔;另一方面,不同被试在任务中的准备状态也可能是不同的,一部分被试可能更多地猜测出现朝向自我的面孔,另一部分被试可能更多地猜测出现朝向人偶的面孔。这就导致即使准备状态也存在行为振荡现象,不同被试和不同试次间行为振荡的初始相位也是不同的,当这些试次的数据相叠加,效应就会相互抵消,最终结果中不会表现出振荡波。其次,即使假设所有被试在所有试次中都表现出相同的准备状态,但是在分析数据时,每个试次中起始点的定义是人偶图片呈现的时刻而不是每个试次开始的时刻(动态掩蔽刺激的呈现)。在这里,我们的假设是人偶图片的呈现会使得被试在这一时刻启动他人视角下的信息加工,并以他人视角下的信息加工为初始状态开始表现出行为振荡。这一操作是行为振荡范式的关键点,在前人研究行为振荡的范式中,也是通过线索的呈现来启动被试行为振荡中的某一个状态,从而能够以此为初始点将不同试次的数据整合到同一个时间轴上而又不会使效应相互抵消。而正如在实验二和实验三的程序部分所介绍的,从试次开始到人偶图片呈现我们设置了 0.5-1s 的随机间隔时间来排除试次开始的时间点对振荡的影响。即使每个被试在试次开始的一瞬间都保持同一种预期状态并在随后表现出行为振荡现象,但由于随机间隔的插入,当试次以人偶出现的时间为锚定点进行叠加时,试次开始的时间点其实是错位的,这种错位使得准备状态的效应被相互抵消,最终保留下来的还是自我视角和他人视角信息加工的行为振荡模式。

Bargh, J. A., & Williams, E. L. (2006). The automaticity of social life. *Current directions in psychological science*, 15(1), 1-4.

Zhao, X., Cusimano, C. J., & Malle, B. F. (2015). In Search of Triggering Conditions for Spontaneous Visual Perspective Taking. In *CogSci*.

Surtees, A., Apperly, I., & Samson, D. (2016). I've got your number: Spontaneous perspective-taking in an interactive task. *Cognition*, 150, 43-52.

Ward, E., Ganis, G., & Bach, P. (2019). Spontaneous vicarious perception of the content of Another's visual perspective. *Current Biology*, 29(5), 874-880.e4.

Field, D. J., Hayes, A., & Hess, R. F. (1993). Contour integration by the human visual system: evidence for a local

“association field”. *Vision research*, 33(2), 173-193.

意见 2: “视觉观点采择任务”中自我与他人界定不明确。该任务中, “人偶”被研究者假设为他人, 被试需要采用观点采择的办法, 具身认知判断人偶对面面孔的朝向。那么, 在该任务中, 自我是通过什么体现的? 按照作者的观点, 自我并不出现在视觉观点采择任务中, 自我就是在电脑面前完成按键任务的被试, 这样看来, 作者为自我和他人设定了不同的场景, 他人的场景是社会互动, 而自我的场景是完成实验任务, 处在不同场景中的自我与他人, 如何比较?

回应: 感谢审稿专家的问题。正如我们在回答问题 1 时所讨论的, 本研究中的社会互动体现在被试自发地、内隐采择了视觉场景中他人的视角, 而这种自发性视觉观点采择的效应是通过面孔方向识别的任务探测出来的。因此在这一逻辑下, 自我(被试)并不是没有出现在视觉观点采择任务中, 而恰恰是视觉观点采择的主体, 只不过这种视觉观点采择过程是内隐发生的。单纯的他人场景也并非是社会互动, 我们不是要探究人偶图片与面孔图片之间可能存在的社会互动, 而是关注自我(被试)在这种有他人(人偶)在场的情境中与他人的社会互动(视觉观点采择)所产生的影响(对面孔方向判断的影响)。而自我与他人的比较体现在自我与他人视角下面孔朝向的不同。在实验所设置的这样一个俯视场景中, 被试直接看到的面孔刺激被认为是自我视角下的信息, 而场景中的人偶视角下的面孔刺激被认为是他人视角下的信息, 通过想象可以理解两种信息仅存在方向上的不同, 因此具有可比性。的确, 在这种实验范式中自我是真实的, 而他人是虚拟的。但这是一种方便在实验室中研究社会互动的简化情景, 优势是可操控性好, 并且可以有效地控制实验中的混淆变量, 大量研究视觉观点采择或其他社会互动过程的文章都采用了这种实验形式(例如: Santiesteban et al., 2015; Todd et al., 2017; Ward et al., 2019)。当然也有许多研究将自我与他人同时设置在虚拟环境或真实环境, 前者的代表是采用虚拟现实技术的相关研究, 后者的代表则是同时邀请两名被试或采用一名假被试的相关研究。这些研究范式各有优劣, 但都被广泛用于探究自我与他人间的交互过程。

Santiesteban, I., Banissy, M. J., Catmur, C., & Bird, G. (2015). Functional lateralization of temporoparietal junction-imitation inhibition, visual perspective-taking and theory of mind. *European Journal of Neuroscience*, 42(8), 2527-2533.

Todd, A. R., Cameron, C. D., & Simpson, A. J. (2017). Dissociating processes underlying level-1 visual perspective taking in adults. *Cognition*, 159, 97-101.

Ward, E., Ganis, G., & Bach, P. (2019). Spontaneous vicarious perception of the content of Another's visual perspective. *Current Biology*, 29(5), 874-880.

意见 3: 正像作者所说, 正立面孔的识别容易产生天花板效应, 因此在许多面孔识别任务中, 反应时是一个重要的行为指标。尤其是在本研究中, SOA 是通过严格的时间来界定的, 因此将面孔判定的反应时来作为区分信息采择对象的指标是非常合理的, 不清楚为何作者使用了正确率作为行为指标。

回应: 感谢审稿专家的问题和建议。在传统的面孔识别的实验范式中, 反应时确实是更常使用的指标, 并且使用反应时作为因变量相比于正确率确实具有优势, 例如可以避免天花板效应。然而在行为振荡的实验范式中, 正确率是更受青睐的因变量指标, 在许多行为振荡研究中都采用了正确率作为实验的因变量(例如: Landau & Fries, 2012; Fiebelkorn et al., 2013; Mo et al., 2019)。这是因为行为振荡范式需要更高的时间分辨率, 而这这就要求刺激呈现的时间越短越好。因为更短的呈现时间意味着其跨越的振荡波上的相位区间越短, 最终将不同条件汇

总形成的振荡曲线就越精确，如果刺激呈现的时间过长，可能其本身的呈现时间就已经跨越了多个周期的振荡，这样被试对这一刺激的反应就可能综合了不同振荡状态下的反应，就无法准确地探测行为振荡现象。而过往使用反应时作为因变量的面孔识别实验，面孔呈现时间较长，甚至直到被试完成反应才消失，在这样的设计下，被试基本不会做出错误的判断，因此反应时是更优越的因变量指标，但这样的设计是无法满足行为振荡的研究的。在我们的实验中，面孔仅呈现 66.67ms，并且立即被掩蔽刺激所取代，在这样的设置下，被试识别面孔具有较高的难度，因此错误率较高，正确率就成为了更加合适的因变量指标。

Landau, A., & Fries, P. (2012). Attention samples stimuli rhythmically. *Current Biology*, 22(11), 1000-1004.

Fiebelkorn, I., Saalmann, Y., & Kastner, S. (2013). Rhythmic sampling within and between objects despite sustained attention at a cued location. *Current Biology*, 23(24), 2553-2558.

Mo, C., Lu, J., Wu, B., Jia, J., Luo, H., & Fang, F. (2019). Competing rhythmic neural representations of orientations during concurrent attention to multiple orientation features. *Nature Communications*, 10(1), 1-10.

意见 4: 图 4C 中，朝向自我面孔的正确率在 SOA 的不同阶段几乎处于相同水平，并未随着他人的正确率出现波动，这与作者的假设并不符合。

回应: 感谢审稿专家的问题。实验二中的这一结果的确是本研究不完美的地方。实验二中对朝向自我面孔的正确率曲线进行的傅里叶变换的结果也没有达到统计显著的水平。结合实验二与实验三的结果，我们推测这可能是受到了被试注意力分配的影响。通过将所有试次合并并进行统计我们发现，被试对于呈现时间位于整个人偶-面孔 SOA 范围的中部的面孔刺激的识别正确率更高，对于实验二来说，人偶-面孔 SOA 的范围是 0.2-1.2s，被试正确率的高峰集中在 0.8-0.9s 附近，对于实验三来说，人偶-面孔 SOA 的范围是 0.2-2.2s，被试正确率的高峰集中在 1.1-1.3s 附近，这可能是因为被试在习得了面孔刺激呈现时间的大致范围后，对面孔呈现的时间点存在一定预期，这一预期即是面孔呈现的平均时间点。因此在这一预期时间附近，被试的注意水平更高，对刺激的判断也就更准确。而对于实验二来说，0.8-0.9s 恰恰是自我视角的信息加工相对减弱的阶段，我们猜测这种注意水平的提高掩盖了一部分信息加工减弱的效应，从而使得实验二中朝向自我面孔的正确率曲线趋于平缓。同理，在实验三中，1-1.3s 恰恰是他人视角的信息加工相对减弱的阶段，而这种预期导致的注意力水平的提高又使得实验三中朝向人偶面孔的正确率曲线趋于平缓。但是尽管可能存在这样的影响，一方面，实验二中我们通过两条曲线不同位点间的配对 t 检验，从另一角度说明了两条曲线间的相对优势变化；另一方面，实验三中的两条曲线的傅里叶变换的统计检验在 1Hz 处都达到了显著水平，在实验二的基础上进一步验证了我们的假设。因此综合以上的内容，我们认为研究得到的结论是可靠的。

意见 5: 建议在引言部分多多铺垫行为振荡的逻辑与假设，有助于读者的理解，而不是根据研究结果部分的内容来揣测。

回应: 非常感谢审稿专家的建议。对于之前没有接触过行为振荡相关研究的读者来说，首次阅读本文确实不容易理解其中的实验逻辑，因此，我们在实验二的引言部分增加了对行为振荡的实验逻辑的介绍。

第二轮

审稿人 1 意见: 作者对我在第一次审稿中提出的问题做出了较好的回应。我仅剩的疑虑在于

研究结论的表述方式。如前次审稿意见所述，研究 2 和 3 发现的震荡效应是基于平移后的正确率，而在原始正确率上，自我条件从未显著低于他人条件。作者选择使用“相对”二字来表述结果，但“相对”既可以是相对于各个条件自己的均值基线，也可以是相对于另一个条件，这种说法依然容易引起读者误解。建议作者做两方面的补充修改：

意见 1：使用更精准的语言表述结果，如点明“相对”的参照对象是什么。

回应：感谢审稿专家的建议。仅使用“相对”这一表述确实不够准确，因此我们在文章中将原本使用“相对”的表述替换为了更准确的表达。

意见 2：作者在回应中为这种不对称性提出了解释：“在过往研究中，发生振荡的两种条件往往是完全对等的（例如左视野和右视野，向左的启动与向右的启动等），但在本研究中，发生振荡的双方是自我视角与他人视角，而自我视角下的信息加工相比于他人视角下的信息加工有明显的优势性，因此出现了自我视角下的加工正确率始终高于他人视角下的加工正确率的现象，这也与我们在实验一中得到的结果相一致。”这一解释很合理，建议加入正文的讨论部分。

回应：感谢审稿专家的建议。我们将这部分内容添加到了实验二的讨论中。

.....

审稿人 2 意见：

作者认为只要有他人在场，就会出现观点采择，这样的思考太绝对了。他人在场的确会影响个人的行为表现，但是是否会有观点采择，却并不一定。作者如果要证实自己的观点，可能需要添加实验条件。

意见 1：无人偶时，不同朝向面孔的正确率，假设面孔判定任务时，三个方向无差异，由此证明是人偶的出现导致左右面孔判定出现差异。

回应：感谢审稿专家的建议。被试对于朝左面孔和朝右面孔的识别成绩的确有可能存在差异，不过在原本的实验一中，我们对人偶的呈现位置在试次间进行了平衡，在一半的试次中，人偶出现在面孔左侧，此时朝向左侧的面孔为朝向人偶的面孔，而朝向右侧的面孔为朝向人偶对侧的面孔；在另一半试次中，人偶出现在面孔右侧，此时朝向右侧的面孔为朝向人偶的面孔，而朝向左侧的面孔为朝向人偶对侧的面孔。因此通过这样的设计，我们认为不需要无人偶的控制条件也可以排除朝左面孔和朝右面孔之间的识别差异对实验结果带来的影响。

意见 2：人偶背向面孔图片。如果此时左右面孔判定无差异，证明即使他人在场，并且只有当位置恰当，才会出现观点采择。如果左右面孔判定有差异，证明并非是观点采择导致差异，而是人偶与面孔的其他特征影响了判断。

回应：感谢审稿专家的建议。为了进一步确定实验中得到的效应的确是视觉观点采择产生的，排除是人偶的其他特征影响了判断，一个非常有效的手段就是加入您提出的人偶背向面孔的对照条件。因此我们重新进行了实验一，在保留实验一中原有三个条件的基础上又添加人偶背向面孔的三个条件。实验结果发现，只有在人偶面向面孔时，被试对朝向人偶面孔的识别才会优于朝向人偶对侧面孔的识别，而当人偶背对面孔时，二者没有显著差异。这一结果进一步说明了实验中表现出的对朝向人偶面孔的加工优势是视觉观点采择的结果。其次，比较人偶面向面孔时对朝向人偶面孔的识别正确率与人偶背对面孔时对朝向人偶面孔的识别正确率可以发现前者显著高于后者，从另一角度说明了视觉观点采择对面孔识别的促进作用。我们将重做的实验一替换掉了原来的实验一，因此对实验一中的内容进行了重写，并对

英文摘要中的相关部分进行了更新。

第三轮

审稿人 2 意见：

感谢作者的工作。在修改稿中，作者修改了实验一，设置了背对面孔的条件，并且发现这一条件下，被试对面向他人的面孔和背向他人的面孔的加工没有差异。这个修改应该可以在一定程度上证明实验结果是由观点采择造成的。对于一审意见的回复，我对图片中这个问题还是有点疑惑。我感觉我对行为震荡的理解和作者的理解有很大的偏差。我对于行为震荡的理解是：行为震荡是以每个试次下的 SOA 为单位的，每个试次下，面孔呈现之前，被试会在不同的 SOA 下出现自我和他人视角的震荡性转换。如果在面孔刺激呈现时，被试采用的视角和面孔的朝向正好相对，被试的正确率就会高一些，反之，正确率就会低一些。例如，对于一个面向娃娃的图片，在前一个 SOA 时，被试采用自我视角，所以正确率低一些；但是在下一个 SOA 时，被试震荡为娃娃视角，所以正确率高一些。但是，作者的意思好像是：行为震荡是以试次为单位的，被试在不同的试次会出现自我和他人视角的震荡性转化（有的试次是自我视角，有的试次是他人视角）。例如，对于一个面向他人的图片，在前一个试次下，被试采用自我视角，所以正确率低一些，但是在后一个试次下，被试震荡为他人视角，所以正确率高一些。如果作者对行为震荡的理解的确是这样的话，就引发了两个问题。第一个问题是，实验二和实验三中，使用 SOA 作为数据分析的时间单位（自变量）可能是不合适的，试次可能是更合适的时间单位？第二个问题是，受行为震荡影响，每个被试在每个试次下的视角也是不一致的，这会不会造成作者所说的抵消现象呢？

回应：感谢审稿专家的问题。本研究首先通过实验一验证这一实验范式可以有效地探测内隐视觉观点采择所产生的效应，随后实验二和实验三是在实验一的基础上探测这一过程中可能存在的不同视角间（自我 vs. 他人）的行为振荡现象。审稿专家对于行为振荡的理解是非常正确的，它是指被试的行为表现随着时间进程而产生节律性的波动，因此被试会在不同的 SOA 下具有不同的表现，而不是在不同试次下有不同的表现。因此最理想的实验形式应该是在一个试次中的多个时间节点探测被试的行为表现，通过行为表现的波动来验证行为振荡，如下图 x 中 A 所示。然而实际上这种形式是无法实现的，一方面，被试在一个试次中只能进行一次反应（在我们的实验中为对面孔方向的判断），反应完成后被试的任务就结束了，无法提供多次探测的机会；另一方面，即使通过一定的实验设计使得被试可以在一个试次中进行多次反应，时间精度也无法达到行为振荡的研究要求（例如在本文实验三中相邻 SOA 条件相差为 66.67ms，我们无法要求被试每 66.67ms 就做出一次反应），并且反应本身也可能破坏振荡过程。因此，研究者们使用另一种巧妙的方法来研究行为振荡过程，即将不同的 SOA 条件分配到不同的试次中去，这时每个试次中只要求被试进行一次反应，但不同的试次探测的是不同的 SOA 条件下被试的表现。如果我们通过一定的实验操控来保证每个试次中行为振荡的初始相位是相同的（在本研究中是将人偶的出现时间点作为试次的初始点，这一时刻他人视角被引入到实验情景中），那么我们就可以通过试次间结果的叠加来得到一条被试的行为表现随 SOA 变化的曲线，进而探究行为振荡现象，如下图 x 中 B 所示。而本研究正是采用这种经典的行为振荡的研究方法，这也就解释了为什么汇总不同试次的实验数据能够得到随 SOA 变化的振荡曲线。

可能是原文中在解释行为振荡研究范式的原理时对举例的表述方式容易使读者产生误解，因此我们对原文中的实验原理介绍部分进行了修改。

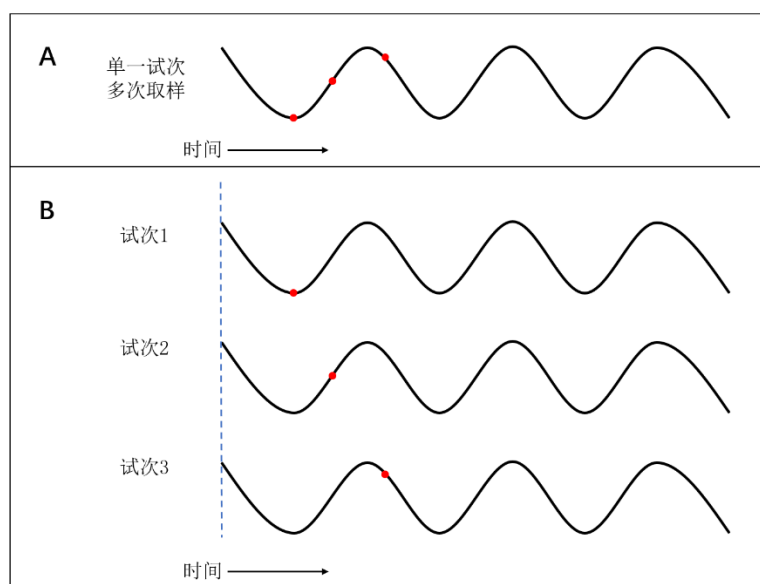


图 x. 行为振荡研究方法示意图。图中黑色曲线代表可能存在的行为振荡曲线，红点代表探测的时间点。

第四轮

审稿人 2 意见：建议接受

编委意见：建议发表

主编意见：

根据审稿人及编委的意见，结合稿件修改情况，建议修改后发表。此外，建议：（1）将“总讨论”按理论逻辑加上小标题，以帮助作者了解研究的具体理论或机制。（2）增加“结论”部分。在文章的最后一段中，“总结一下”过于口语化。此外，文中还有一些标点符号的使用不够规范，如该用顿号的地方用了逗号、该有标点的地方却省略了。

回应：已在总讨论中加入小标题并将最后一段改为结论部分并删掉过于口语化的表达；对文章中标点的使用进行了检查和修改