

《心理学报》审稿意见与作者回应

题目：联合评估和单独评估：富有潜力的助推手段

作者：路西 Christopher K. Hsee (奚恺元)

第一轮

审稿人 1 意见：该文系统的介绍了联合评估和单独评估的研究，包括研究实例和可能的机制，并探讨了在该研究基础上可能的助推方向和手段。该文文字流畅，介绍细腻，发人深省，是篇不多见的好的综述文章。我对文章的修改意见不多，只有如下几点。

意见 1：文章花了大量篇幅来介绍在不同评估模式下的偏好反转。偏好 (preference) 大多是通过被试的选择 (choice) 来引申出来的。在联合评估模式下，选择比较直接；但在单独评估模式下，选择只能通过被试对不同选项的判断 (judgment) 来引申，这点在作者给的所有实例中都适用。我的问题是：单独评估模式下的判断真的可以代表被试的偏好吗？最大的问题是人在判断和选择中所用的认知模型可能会不一样，从而导致所表现的行为不同。具体说来，人在判断中最常用的是在加权平均 (weighting-and-adding) 基础上的，补偿性的 (compensatory) 模型；而在选择中，人可用的模型比较多样，既有补偿性的模型，又有启发式 (heuristic) 的，非补偿性的 (noncompensatory) 模型。如果被试在单独评估模式下用的是补偿性模型来做判断，而在联合评估模式下用的是非补偿性模型来做选择，其结果可能会不同，可能会导致偏好反转..... 和广义评估性理论不同的是，这个解释着重于认知系统对外部任务在宏观的信息处理方面上的不同，而广义评估性理论假设信息处理方式没有变化，变化的只是对某些属性的重视度和加权值。

回应：感谢审稿人宝贵的意见，您的问题涉及了多方面，我们尝试一一回应。

首先，您提到，在联合评估 (JE) 中，偏好是通过参与者对两个选项的选择 (choice) 来引申的，在单独评估 (SE) 中，偏好是通过参与者对某个选项的判断 (judgment) 来引申的，两者存在不同。由此，您有所质疑，在 SE 中消费者的判断是否可以代表消费者的偏好。您非常敏锐，这部分我们在介绍时有所疏漏。确实，少量研究中，存在 JE 和 SE 时因变量测量方式不同的问题（正如您指出的），但是，本文中所引用的大部分的研究都注意了这一问题，在 JE 和 SE 时采用了相同的测量方式。例如，Hsee (1996) 关于音乐词典的研究，Hsee (1998) 关于冰淇淋和餐具套装的研究，Kahneman & Ritov (1994) 关于慈善捐款的研究等，无论在 JE 还是 SE 时，消费者的偏好都采用支付意愿 (Willingness to Pay) 来测量；Zikmund-Fisher, Fagerlin, & Ubel (2004) 关于医生选择的研究，无论在 JE 还是 SE 时，消费者的偏好都采用“该选项多大程度是个好的选择 (0-10 点评分)”来测量。除了最终的选择 (choice)，采用支付意愿或者选项好坏程度等“判断” (judgment) 来表征偏好是一种常见的处理方式。只要在实验中确保测量一致，避免偏好测量手段差异带来的偏好不稳定现象即可。

其次，您指出，人们在判断和选择中采用的认知模型不同可能是导致 JE/SE 偏好反转的原因，即人们在 JE 中处于“选择”模式，而在 SE 中处于“判断”模式。正如前面所提到的，要求人们在 JE 中进行“判断”（例如：表达自己的支付意愿、判断选项的好坏程度），并不影响 JE/SE 偏好反转。因此，我们倾向于认为，JE/SE 偏好反转不完全是选择-判断差异的问题。这一部分我们在文中插入了一处简短说明，第 9 页红色字体。

最为核心的是，您认为，选择-判断差异的解释中，最重要的是假设了 JE/SE 时，个体的信

息处理方式存在差异；但是广义可评估模型，假设信息处理方式没有变化，变化的只是对属性的权重。信息处理方式的差异，确实是一个重要的解释，我们在这一点上非常认同您的观点。不过，我们觉得，广义可评估性模型并没有假设信息处理方式没有变化，比如，JE 比 SE 涉及了更多的比较加工(comparative mindset)，因为 JE 提供了另一选项作为参照点。当然，正如您提到的，广义可评估模型所关注的重点并不在信息处理方式的差异上。因此，除了广义可评估理论，我们也介绍了一些其他理论的解释，例如双系统理论，这一解释重点关注了 JE/SE 信息处理方式的不同，目前学者倾向于认为单独评估更多采用系统 1，而联合评估更多采用系统 2。您所提到的补偿性模型和非补偿性模型可能是另一种值得未来关注的解释方式。

意见 2: 在作者介绍的实验中，实验设计是组间的还是组内的？在我看来，好像单独评估大多用的是组间。如果真是这样的话，那结果可能会受到被试间差异的影响。用组内实验来显示评估条件对偏好的影响会让人更信服。

回应: 您的观察非常正确，大部分研究中，单独评估采用的是组间设计而不是组内设计。有少量的研究在单独评估中采取了组内设计，例如 Bazerman, Loewenstein, & White (1992), Yang, Hsee, & Zheng (2011)，单独评估的参与者是顺序观看了两个选项。我们在文中(第 13 页)增加了相应的说明，用红色字体标出，对此问题进行了澄清。

意见 3: 在应用广义评估性理论时，有对属性的“评估性”单独测量吗？还是评估性只是一个理论假设？

回应: 属性的可评估性进行过单独测量。例如：在 Hsee(1996)研究二中，询问了参与者对于选项的属性是否有所了解(e.g. “do you have any idea with Do you have any idea how good a GPA of 3.0 from UIC is?”)。又比如：在 Hsee, Zhang, Wang, & Zhang (2013)，询问了参与者对救助了 0 只、100 只、200 只北极熊是否容易评估(easy to evaluate)还是难以评估(hard to evaluate)，结果表明参与者对于 0 这个数量感到更容易评估，因为它可以被加工为一种类别变化，对于 100、200 这种数量变化感到更难以评估。我们在文中(第 11 页)也加入了一个简短的说明。

意见 4: 我很喜欢作者关于评估模式的助推可能性的讨论。我主观的结论是联合模式可以助推更理性的决策，而单独模式可以助推更多的正面感受。无论我的结论正确与否，希望作者对到底是联合模式好还是单独模式好有一个小的总结，以帮助读者在应用该想法时有个更清晰的认识。

回应: 您的推论非常准确，当决策有客观标准的情况下，联合评估模式有助于更加理性的选择。但是当决策的目标是个体获得更好的体验时，为了减少预测-体验偏差，单独评估的模式更适合。感谢您的建议，我们在助推部分的结尾处(第 26 页)新增了一个小的总结，以红色字体标出，敬请指正。

.....

审稿人 2 意见:

意见 1: 该文总结了联合评估与单独评估的研究成果，并拟从助推应用的角度，梳理整合了在日常决策时的规律，有理有据，可以发表。但整体感觉文章结构比较松散，也不知道为何作者总是用一小段一小段的陈述方式，分析也略显表层化，归纳不够，总结不够。比如：在 P14 上有一段：“单独评估和联合评估可以影响函数的非线性程度（如图 3 所示），在单独评估的模式下，函数的非线性程度较高；在联合评估的模式下，函数的非线性程度较低(Hsee &

Zhang, 2010; Hsee et al., 2013)。”表述, 知其然不知其所以然。这一类的情況还比较多。另外, 最后从助推的角度考虑应用的可能性时, 什么样的规律应该联合评估, 反之又适合单独评估呢? 应该不仅仅是这样单个单个的陈述, 而最好有更高一层次的提炼和总结。

回应: 感谢您的建议, 这篇综述确实采用了现象-解释-应用的顺序组织, 在应用部分也是按照现象的不同领域进行了分类, 有结构松散的问题。根据您的建议, 我们进行了一些修改。

首先: P14 关于函数的非线性程度的部分, 我们进行了补充说明, 以红色字体标出, 对原因进行了说明。另外, 在助推部分, 我们增加了一个小的总结(第 26 页), 对联合评估和单独评估适用的情景进行了归纳, 敬请指正。