

# How to Critique

---

译自：Humphreys, Macartan. *How to Critique*. <https://macartan.github.io/teaching/how-to-critique>

---

Here are some pointers on the things to look for when discussing or reviewing a paper.

本文是关于讨论或述评论文时的一些注意事项的提示。

## Discussants 讨论参与者的注意事项

Generally discussants have 10 – 15 minutes to give comments on a paper, sometimes less. With that much time you can make **3 good comments**. You should not use this time to say everything you liked or did not like about a paper and you should not get lost in the weeds. If you describe errors you have to get to the so what. The fact that there is an error is not in itself of interest. You should select your comments so that:

一般来说，讨论者有10-15分钟的时间对一篇论文进行点评，有时时间会更短。在这样的时间（限制）下，你可以提出**3点好的点评**。你不应该用这段时间（不分轻重地）谈你对一篇论文喜欢或不喜欢的所有点，也不应该迷失在杂乱无章（的评论）中。如果你描述了（文章犯的）错误，你就必须讲清楚这个错误的后果是什么（so what）。「文章中有错误」这个事实本身并不令人感兴趣。你应该有选择地点评，以：

1. they **open up** a conversation **敞开**对话空间；
2. they speak to the **major issues** the paper addresses 与文章尝试解决的**主要议题**进行对话；
3. they provide pointers to **how to do better**. 对「**如何做得更好**」进行提示。

Remember as a discussant it is **not about you**, it is about making the paper better and helping people understand its strengths and limitations. Mostly it's about the speaker. If you think the paper is great you do not have to drum up a critique, but you should still try to help people see why it is great. Having slides helps organize your presentation and helps people follow. A single slide with three bullets on the three big points is enough. If you have a laundry list of smaller points, share it with the speaker afterwards.

请记住，作为讨论者，这（讨论的目的）**不是关于你**，而是（为了）让论文变得更好，并帮助人们了解其优点和局限。这主要是关于演讲者（论文作者）的。如果你认为论文写得很好，倒也不必强行批评，但你仍应努力帮助人们了解它好在哪里。准备幻灯片有助于组织

起你的展示，帮助人们跟上（你的节奏）。一张幻灯片上有三大要点就足够了。如果你有一长串较小的要点，可以事后再与演讲者分享。

**Same language, different perspective:** The really useful critiques often come from taking a really fresh perspective on a piece of work. This requires stepping back and not becoming beholden to the author's spinning of their findings. Often useful to figure out what this is a case of? What is the general class of phenomena this speaks to? If you had lots of resources how would you address the question? If you could set it up as an experiment how would you do it? If you really had to take a policy action based on this work, which elements would give you pause? But as you take different perspectives you should try to speak the same language otherwise you can end up talking to yourself and influencing no one.

**同样的语言，不同的视角。**真正有用的批评往往来自于对一篇作品的全新视角。这需要「后退一步」，不要受制于作者对其研究结果的转述。通常情况下，弄清楚下列问题是很有用的：这是一个什么案例？它所涉及的现象的一般类别是什么？如果你有很多资源，你将如何解决这个问题？如果你能把它设计成一个实验，你会怎么做？如果你真的要在这项研究的基础上采取政策行动，哪些因素会让你暂停？但是，当你采取不同的视角时，你应该试着用同一套语言（体系）说话，否则你最终会自说自话，而影响不了任何人。

# Reviewing 述评的注意事项

For a formal review or referee report you have space to go into much more depth. A standard approach is to divide these reviews into three parts.

对于正式的述评或审阅报告，你有更多空间去深入研究。一个标准的方法是将述评分为三个部分。

- The first part can be a single paragraph — it **summarizes** the key contribution of the paper as you see it, gives an overall assessment, and points to the key issues, concerns, or strengths. **Don't forget the strengths.** Try to articulate succinctly what you know now that you didn't know before you read the piece. Often a quick summary can draw attention to strong features you were not conscious of, or makes you realize that what you were impressed by is not so impressive after all. 第一部分可以是一个单独的段落——**总结**你所看到的论文的主要贡献，给出总体评价，并指出关键的议题、关切或优点。**不要忘记（介绍文章的）优点。**试着简明扼要地阐述你现在知道、而你在读这篇文章之前并不知道的东西。通常情况下，快速的总结可以让你注意到你（之前）没有意识到的显著特征，或者让你意识到你（之前）印象深刻的东西其实并不是那么令人印象深刻。
- The second part discusses **3 – 6 major features** of the paper; the checklist below lists features that could be

useful to think through when selecting themes. Try to organize by theme (measurement, explanation etc.). 第二部分讨论论文的**3-6个主要特点**；下面的备忘清单列出了在选择主题时可供思考的特点。尽量按主题组织（比如测量、解释等）。

- The third part is for “**smaller issues**” where you can bullet point things from ambiguities, to estimation issues, to pointers to other work. 第三部分是「**小问题**」，你可以分点列出一些东西，例如模糊性、估计问题，以及对其他研究的提示。

Other things:

其他事项：

- It's useful to authors when you can **point to literature** they have not read, if relevant. 如果你能指出他们没有读过的相关文献，这对作者是很有用的。
- It's useful to authors to know **what to cut**: reviews tend to worry about length but still ask for more. 对作者来说，知道**应该削减什么**是很有用的：审稿意见往往担心长度，但同时却要求作者（写）更多。
- Your **tone** should be such that you would not feel embarrassed if someday your review gets into the public domain by mistake. 你的**语气**应该是这样的：假如有一天你的审稿意见被不小心传到了公共领域，你也不会感到尴尬。

- You should feel free to ask for **extra material** such as replication data or analysis plans. Sometimes reviewing can go quicker if you can access data. 你大可以自由地要求（作者）提供**额外材料**，如数据复本或分析计划。有时，如果你能获得数据，审稿会更快。
- Don't ask the authors to ask and answer a **different question**; respond to the paper you have been sent. 不要要求作者提出和回答**另一个问题**；要针对发给你的那篇论文做回应。
- Be **generous**: share references if they are missing but don't assume that researchers intentionally ignored the work of others (or your work!); raise ethical issues if you see them but don't assume researchers acted without ethical concern; ask for multiple comparisons corrections but don't assume deliberately misleading reporting. 要**宽宏大量**：如果（原文）缺少（你认为重要的）参考文献，要与大家分享，但不要认为是研究者故意忽视了别人的研究（或你的研究！）；如果你看到（学术）伦理问题，要提出来，但不要认为研究者的行为（是故意）不符合（学术）伦理要求；可以要求进行多重比较校正（以减少结果误报），但不要认为（作者）是故意（作出）误导性的报告。
- **Pronouns**. For anonymous review it's usually safe to use pronouns "you" or "they" even if single authorship has been indicated. **代词使用**。对于匿名评审，使用「你们」

或「他们」作代词通常更安全，即使（文章）已经表明只有一位作者。

## The Checklist 备忘清单

Here is my list of what to look out for as I read a paper:

以下是我列出的阅读一篇论文时的注意事项清单：

### Theory 理论

- Is the theory internally consistent? 理论是否内在一致？
- Is it consistent with past literature and findings? 是否与过去的文献和发现一致？
- Is it novel or surprising? 是否新颖或令人惊讶？
- Are elements that are excluded or simplified plausibly unimportant for the outcomes? 被排除或简化的元素是否真的对结果（变量）不重要？
- Is the theory general or specific? Are there more general theories on which this theory could draw or contribute? 该理论是一般的还是具体的？是否有更多的一般理论可供这个理论吸收或作出贡献？

## From Theory to Hypotheses 从理论到假说

- Is the theory really needed to generate the hypotheses? 该理论是否对于产生假说来说真的必要吗?
- Does the theory generate more hypotheses than considered? 该理论产生的假说是否比（我们）考虑的多?
- Are the hypotheses really implied by the theory? Or are there ambiguities arising from say non-monotonicities or multiple equilibria? 这些假说真的是由理论所推导出的吗？还是存如非单调性或多重均衡等问题所导致的模糊性？
- Does the theory specify mechanisms? 该理论是否明确指出了机制？
- Does the theory suggest heterogeneous effects? 该理论是否暗示了多样化的效应？

## Hypotheses 假说

- Are the hypotheses complex? (eg in fact 2 or 3 hypotheses bundled together) 假说是否太过复杂？（例如事实上是2或3个假说捆绑在一起）
- Are the hypotheses falsifiable? 假说是可证伪的吗？

# Evidence I: Design 实证（一）：研究设计

- External validity: is the population examined representative of the larger population of interest? 外部效度：研究对象群体是否代表了更大的相关群体？
- External validity: Are the conditions under which they are examined consistent with the conditions of interest? 外部效度：研究所处的条件是否与（理论）感兴趣的条件一致？
- Measure validity: Do the measures capture the objects specified by the theory? 测量效度：测量标准是否准确体现了理论所规定的对象？
- Consistency: Is the empirical model used consistent with the theory? 一致性：使用的实证模型是否与理论一致？
- Mechanisms: Are mechanisms tested? How are they identified? 机制：是否检验了机制？机制是如何被识别的？
- Replicability: Has the study been done in a way that it can be replicated? 可复制性：该研究是否可以重复进行？
- Interpretation: Do the results admit rival interpretations? 解释：（研究）结果是否给对立的解释留下了空间？

## Evidence II: Analysis and Testing 实证（二）：分析与验证

- Identification: are there concerns with reverse causality?  
因果识别：是否存在反向因果的隐忧？
- Identification: are there concerns of omitted variable bias?  
因果识别：是否存在遗漏变量偏差问题？
- Identification: does the model control for pre treatment variables only? Does it control or does it match?  
因果识别：模型是否只控制了干预前的变量？它是控制还是匹配？
- Identification: Are poorly identified claims flagged as such?  
因果识别：那些识别性差的论断是否被标记出来？
- Robustness: Are results robust to changes in the model, to subsetting the data, to changing the period of measurement or of analysis, to the addition or exclusion of plausible controls?  
稳健性：（检验）结果对模型的变化、对数据的子集、对改变测量或分析的时间段、对增加或排除合理的控制变量是否稳健？
- Standard errors: does the calculation of test statistics make use of the design? Do standard errors take account of plausibly clustering structures/differences in levels?  
标准误：检验统计的计算是否利用了研究设计？标准误是否考虑到了合理的聚类结构/水平差异？
- Presentation: Are the results presented in an intelligible way? Eg using fitted values or graphs? How can this be

improved? 呈现方式：结果是否以一种可理解的方式呈现（例如使用拟合值或图表）？（呈现方式）如何改进？

- Interpretation: Can no evidence of effect be interpreted as evidence of only weak effects? 解释：没有效果的证据是否可以解释为只有微弱效果的证据？

## Evidence III: Other sources of bias 实证（三）：偏差的其他来源

- Fishing: were hypotheses generated prior to testing? Was any training data separated from test data? 「钓鱼执法」：假说是否产生在检验之前？训练数据是否与检验数据分开？
- Measurement error: is error from sampling, case selection, or missing data plausibly correlated with outcomes? 测量误差：取样、案例选择或数据缺失的误差是否与结果有可能的关联？
- Spillovers / Contamination: Is it plausible that outcomes in control units were altered because of the treatment received by the treated? 溢出效应/污染：控制组样本的结果是否因（干预组）被试者接受的干预而发生改变？
- Compliance: Did the treated really get treatment? Did the controls really not? 依从性：接受干预的人是否真的得到了干预？对照组是否真的没有（被干预）？
- Hawthorne effects: Are subjects modifying behavior

simply because they know they are under study? 霍桑效应：被试者是否仅仅是因为知道自己正在被研究而改变行为？

- **Measurement: Is treatment the only systematic difference between treatment and control or are there differences in how items were measured?** 测量：干预是干预组和对照组之间唯一的系统性差异，还是说在样本的测量上也存在差异？
- **Implications of Bias: Are any sources of bias likely to work for or against the hypothesis tested?** 偏见的影响：是否有任何偏见来源可能对所验证的假说起支持或反对的作用？

## Explanation 解释

- **Does the evidence support the particular causal account given?** 证据是否支持所给出的特定的因果解释？
- **Are mechanisms examined? Can they be?** 是否对机制进行了检查？它们可以（被检查）吗？
- **Are there observable implications we might expect to see associated with different possible mechanisms?** 是否有我们可能预料到的、与可能存在的另一种机制有关的、可观察的影响？

## Policy Implications 政策意义

- Do the policy implications really follow from the results? 政策意义是否真的来自于研究结果?
- If implemented would the policy changes have effects other than those specified by the research? 如果实施, 政策变化是否会产生不同于研究指出的其他效果?
- Have the policy claims been tested directly? 政策主张是否得到了直接检验?
- Is the author overselling or underselling the findings? 作者是否对研究结果夸大或贬低了?