

## 《心理学报》审稿意见与作者回应

题目：谁该成为体验型产品网络评论的主角，图片还是文字？

作者：杨颖，朱毅

---

### 第一轮

#### 审稿人 1 意见：

《谁该成为网络评论的主角，图片还是文字？》选择了一个新颖的问题，并基于解释水平理论对图片和文字、以及其对怀旧产品的影响进行了探索与解释。文章写作流畅，逻辑较为严密，对已有文献掌握较好，但有一些问题需要予以考虑。

**意见 1：**文章标题为《谁该成为网络评论的主角，图片还是文字？》，但在整个研究设计中并未操作或体现出网络情境。尤其是对于研究中一个主要自变量“购买可行性”，如果在网络情境中，研究一、二设计的以“距离远近”对可行性操作，可能是网络情境中消费者非常敏感的一个因素，此时研究一、二的结果很可能并不适用，很可能购买可行性的作用将大于甚至掩盖评论形式的作用。因此，文章结果可能无法直接推广到网络情境。建议作者修改标题等相应内容，或增加网络情境下的实验数据。

**回应：**谢谢评审专家精准的意见。现已对实验 1 进行了修改，修改后实验 1a 仍用幻灯片呈现评论内容，但实验 1b 则采用网页呈现评论，且评论设置成与现实情况更相近的图、文混杂形式，设置的购买可行性情境也更贴近现实网络消费情境。从实验结果来看，“距离远近”的对被试确实有影响（有主效应），但不会掩盖评论形式的作用（仍存在交互作用），且实验 1b 的结果与实验 1a 一致（详见修改后的实验 1）。受条件所限，实验 1b 的情境仍无法与现实生活中网络消费情境保持完全一致，但实验所展示的这种效应是一种普遍存在的心理规律，因此我们推测实验的结论应该可以推广到网络情境中。

**意见 2：**文章认为“解释水平上的差异”是图片和文字评论、以及购买可行性出现差异的机制，但缺少对图片和文字评论、购买可行性等在解释水平或心理距离上的直接测量，对该机制未进行直接检验。是否可以补充？

**回应：**谢谢评审专家细心的审阅和中肯的意见。在修改后版本中实验 1a、实验 1b 已补充了对解释水平的测量（测量方法见 3.1 部分最后一段），结果与预期一致，接受图片评论的被试的解释水平显著低于接受文字评论的被试，可行性高情境下被试的解释水平显著高于可行性高情境下的被试。

**意见 3：**研究三中“产品类型”，从本质上应该是调节了评论类型、购买可行性对产品评价的作用，但目前的数据分析和讨论，未能体现出这一作用。尤其是数据分析，只能表现出在怀旧和非怀旧产品条件下，均出现了购买可行性和评论形式间的交互作用（即重复了研究二的结果），建议修改数据分析和讨论。

**回应：**谢谢评审专家深刻的剖析和意见。诚如所言，本研究预期怀旧、非怀旧产品类型可以调节评论类型、购买可行性对产品评价的作用。我们已在假设推导部分、实验 3 数据分析部分和讨论部分都对“产品类型”的调节作用进行了进一步的文字说明。但在数据分析方法上，

由于实验 3 没有出现三因素交互作用，不宜进行三因素的简单效应分析，所以修改后的论文中仍是在怀旧和非怀旧产品条件下，分别进行了购买可行性和评论形式的两因素方差分析，然后在后续两因素简单效应分析中阐述了产品类型对购买可行性和评论形式关系的调节作用。这种做法在其他已发表的研究中似乎是被接受的（例如，田阳，王海忠，柳武妹，何浏，黄韞慧.（2014）.品牌承诺能抵御负面信息吗？——自我调节导向的调节作用. *心理学报*. 46(6),864-875）。但如果还有其他更好的统计检验方法，恳请赐教为感！

**意见 4:** 几个实验中对文字和图片评论的操作方式似乎存在问题，为什么要以两种评论夹杂的方式呈现？此时很难说明是否是单纯的文字或图片在发生作用。

**回应:** 谢谢评审专家专业的意见。已对实验 1a 进行了修改，用单纯的文字或图片评论作为实验材料。但因为现实中很多评论是图片、文字兼有，为贴近现实情况，实验 1b 仍采用了两种评论夹杂呈现的方式。从两个实验的结果来看，是否夹杂呈现对被试的影响无差异。

**意见 5:** 各研究在预实验中，对图片与文字评论一致性进行评价，是否均在被试看完所有材料后进行？如果是，这种操作很容易受到被试记忆的影响。

**回应:** 谢谢评审专家细致的审阅和中肯的意见。预实验中呈现完一对图片与文字评论后立即让被试进行一致性评价，然后呈现下一对图片、文字评论，不是让被试看完所有材料后再进行评价。文中已添加说明。

**意见 6:** 文章数据分析中，出现了显著交互作用后，建议使用事后检验对交互作用的具体作用进行分析，而不是直接做一个少一个因素的方差分析。

**回应:** 谢谢评审专家中肯的专业意见。现已改为出现了显著交互作用后，进行简单效应分析，希望能解决这一问题。

**意见 7:** 研究一，图 1，数据似乎有误，与相应文字描述不一致。中文字评论和图片评论的内容具体是什么？

**回应:** 谢谢评审专家细心的阅读和中肯的意见。现已经重新设计和做了实验 1，图 1 已经不是以前的内容，请再次审定。

**意见 8:** “3.3.2”中第三段“外量形式在对网络口碑色彩（正反面口碑）或者口碑传播特征上，目前”，该段文字似乎不应该在这里出现。

**回应:** 谢谢评审专家细致的审阅。已删除该段内容。

**意见 9:** 摘要应说明研究的总结论，而不只是呈现实验结果。

**回应:** 谢谢评审专家细致的审阅和建议。已在摘要中对研究的总结论进行了简要说明。

**意见 10:** 建议在图中标注出现显著差异的部分。

**回应:** 感谢评审专家的建议。已进行标注。

**意见 11:** 实验“结果和分析”部分，除了报告结果外，应简单总结一下该实验得到的结论。

**回应:** 感谢评审专家的建议。已在各个实验的“结果和分析”部分的末尾增加了总结性文字。

**意见 12:** 实验 2 操纵购买行为可行性时, 所陈述的逻辑和最后采取的操纵方式并不一致。

**回应:** 谢谢评审专家细致入微的审阅。已将“预定酒店是否需要用信用卡进行担保”改为“可以直接电话预定房间 vs. 在预定时需要提供指定银行的信用卡进行担保”, 现与后文一致。

---

#### **审稿人 2 意见:**

以往研究关于图片评论和文字评论对消费者的影响的研究结论存在分歧。该研究没有一概而论地说明是图片评论还是文字评论的效果更好, 而是基于解释水平理论, 指出在不同的购买可行性下, 图片和文字评论对消费者产品评价的影响。这种通过独特的理论视角, 找到以往研究存在分歧的原因, 在此基础之上建立理论的做法值得肯定。

**意见 1:** 第一, 研究指出, 图片和文字信息在三点上存在不同之处: 抽象程度, 总结特征的难易程度和个体感知过程。其中, 第二点和第三点论证较为有力, 但第一点缺乏说服力。文字信息一定比图片信息抽象吗? 消费者在文字评论中往往会提及自己的消费过程和体验, 这些具体的信息可以帮助后来的消费者进行决策, 文字评论不一定比图片评论更抽象。请见 Chen & Xie (2008)。

**回应:** 谢谢评审专家深入的剖析和深刻的见解。本研究并不认为所有的图片信息都是具体的, 所有的文字信息都是抽象的; 而是图片和文字信息的抽象或具体程度是相对的、有条件的, 具体来说是指二者在表达相同内容时, 图片信息比文字信息更具体。为确保正式实验中图片评论和文字评论内容的一致性, 本研究还专门做了预实验对此进行控制。同时, 我们在文中也补充了说明 (2.1 部分第二段蓝色字体部分)。

本研究 and Chen & Xie (2008) 的研究也有所不同: 该研究所言的“抽象”和“具体”更接近于口语中的“概括”和“详细”, 是一种客观的表现; 本研究所言的“抽象”和“具体”则是一种心理上的体验, 除了“概括”和“详细”之外, “模糊的”和“清晰的”、“核心的”和“边缘的”、“本质的”和“表面的”等都可以引起心理上的“抽象”和“具体”感。因此, Chen & Xie (2008) 的研究中所说的“抽象”和“具体”跟本研究所言是有差异的。

**意见 2:** 论文并未提供或者描述在实验中采用的图片评论和文字评论, 我们无从得知二者的区别之处。如果实验中提供的文字信息都较为简短或笼统, 那么该研究所探究的则是“抽象的文字信息和具体的图片信息”, 而非“文字评论和图片评论”。

**回应:** 谢谢评审专家细心的审阅和中肯的意见。本研究的初衷不是为了考察所有网络评论中究竟图片评论的影响力大, 还是文字评论的影响力大; 而是为了考察在不同情境里, 人们通过网络评论传达同一信息时, 使用图片评论还是使用文字评论的作用更大。因此本研究所说的“文字评论和图片评论”不是指“所有的文字评论”和“所有的图片评论”, 而是“表达相同内容”的文字、图片评论。基于此, 在本研究的各个实验中, 为使图片和文字内容保持一致, 还专门进行了预实验; 内容不一致的文字和图片评论并不在本研究考察范围之内, 因此未说明二者间的区别。

此外, 文字评论的长短也是按大多消费者发表评论的习惯编写的, 且很多内容也是关于产品具体细节的描写 (如“餐厅的角落陈设了一些老玩意, 包括留声机、老式胶卷相机等, 给餐厅的整体风格增色不少, 显得有文艺气质”, 关于预实验详见 3.1、4.1 和 5.1 部分的内容, 现也已附录中增加了部分图文实验材料, 请评审专家进行鉴定), 因此本研究所用材料

应该不是“抽象的文字信息和具体的图片信息”。

**意见 3:** 研究中的“购买可行性”是否等同于“购买可能性”？以往研究认为消费者在可行性（feasibility）和价值（desirability）做权衡，文中将将可行性作为一个独立的变量，那么此时的可行性是否就等同于可能性，比如在实验中，消费者估计到自己根本不会去车程远的酒店？

**回应:** 谢谢评审专家的深入分析和独到见解。“可行性”和“可能性”确实是两个非常接近且容易混淆的词，它们跟“概率（probability）”应该都有密切的联系（本文 5.2 部分对可行性与概率的关系有所阐述）。本研究所说的“购买可行性”是指“消费者感知到购买到某商品困难还是便捷的程度”（请参见论文 2.2 部分的内容），这意味着本研究中“可行性”是一种主观体验，而“可能性”或许更接近一种客观现实。

第一版论文的实验 1、实验 3 是借鉴其他研究的方法，将“低可行性”情境设置为较远的车程，这在现实中发生的情况可能的确是比较少。因此我们修改了实验 1 的情境，实验 1a 的可行性操纵方式仍采用原来的方法（在购买可行性高的组里，告诉被试“这家餐厅就在您的住处附近，步行只需要约五分钟左右，且顺路就能到达”；在购买可行性低的组里，告诉被试“这间餐厅离您的住处较远（单程约需一个半小时），且需要换乘两趟公交车才能到达”），实验 1b 则设置为与现实中的网络消费情境较为贴近（在购买可行性高的组里，告诉被试“餐厅就在地铁站出口旁边，出站就能看到”；在购买可行性低的组里，告诉被试“餐厅位于某大厦 A 座 307 号店铺，离地铁站出口约需 10 多分钟步行路程，其间要转几个路口”）。

而实验 2 是以酒店为背景材料，购买可行性的操控方式则为“在购买可行性高的分组里，告诉被试‘两间酒店可直接电话预定房间’；在购买可行性低的分组里，告知被试‘两间酒店在预定时需要提供指定银行的信用卡进行担保（即预定时需要报出信用卡卡号，还需要提供卡种、有效期、验证码、持卡人姓名、证件号等相关信息）’”。从结果来看，几种操纵方式无差异，因此虽然低可行性会导致做出消费行为的可能性相应降低，但图片、文字评论的影响差异却不会因此而改变。

---

## 第二轮

### 审稿人 1 意见:

作者较全面地针对审稿意见对文章进行了修改，总体而言，本文在行文逻辑和写作上有了极大的提高，但是我对于部分内容仍有一些疑问:

**意见 1:** 实验三中调节效应的检验方法有误，导致实验 3 的结论不准确。三因素间交互作用不显著，说明产品类型不能调节评论形式与购买可行性的关系。因此，“实验 3 的结果说明，购买怀旧产品不会影响文字形式评论与低购买可行性情境的一致性效应，但会破坏图片评论形式和高购买可行性情境的一致性效应；相反，购买非怀旧产品则不会影响图片形式评论与高购买可行性情境的一致性效应，但会破坏文字形式评论和低购买可行性情境的一致性效应。”这一点结论无法得出，建议作者使用恰当的方法分析调节效应，并修改此部分的统计及结论。

从结果看，产品类型和购买可行性的交互作用，以及产品类型和评论形式的交互作用显著，说明产品类型作为调节变量，分别调节了购买可行性与产品评价的关系，以及评论形式与产品评价的关系。在统计上，应该报告产品类型和购买可行性的交互作用事后检验的结果，

产品类型和评论形式的交互作用事后检验的结果，并依此报告具体调节效应的方向，如购买怀旧产品时，购买可行性上的差异如何，以及购买非怀旧产品时，购买可行性上的差异如何。在图中，也应该报告两个图，图里的 x 轴分别是产品类型和评论形式，以及产品类型和购买可行性。

建议参考：温忠麟，侯杰泰，张雷(2005). 调节效应与中介效应的比较和应用. 心理学报, 37, 268-274.

回应：谢谢专家的意见和指导。事实上，最初的实验 3 结果并不理想，在收到第一次专家意见时，我们就考虑要对实验 2 和实验 3 进行修改。但限于当时条件（临近期末，无法在规定时间内征集到足够的被试）未能实现，因此在收到第二次意见后我们重新设计和实施了实验 2 和实验 3。

或许是因为改变了实验 2 和实验 3 中多处条件（呈现纯图片、文字评论，以网页呈现评论，实验 3 中可行性的操纵方式更接近现实情境，换了一批被试），虽然实验 2 的结果没有发生变化，但实验 3 中却出现了显著的三因素交互作用，两因素交互作用也全部显著。因此，在按照专家意见对统计方法进行修改之外，还对三因素交互作用进行了简单效应分析，并维持了之前的结论（改动部分已用绿色字体标注）。

意见 2：实验 3 中，作者报告：“怀旧产品”条件下的被试在“历史感”4 道题项上的均分为（ $M = 5.53, SD = 0.82$ ）；而“非怀旧产品”条件下的被试在“现代感”4 道题项上的均分为（ $M = 5.04, SD = 1.13$ ），表明实验中所选取的怀旧、非怀旧产品有效的传递了不同的时间距离感，对产品类型的操控有效。”

最好是用一道题测量现代 / 历史感，并对怀旧或非怀旧产品在时间距离上操作的成功与否进行检验。此外，请作者需要报告差异检验的统计结果。另外，研究 1 中增加解释水平的测量，研究 3 最好也能增加。

回应：感谢专家细致的评阅和中肯的意见。已改为通过“这家餐厅整体上让人觉得有历史感”，“这家餐厅的室内装修让人觉得有历史感”，“这家餐厅的评论让人觉得有历史感”三道题项的均分来测历史感，并增加了一道题项“这家餐厅让我感到时间距离久远”来测量时间距离，以及增加了对解释水平的测量。结果表明怀旧/非怀旧产品在历史感、时间距离及解释水平上的统计差异显著。

此外，经讨论，我们还在各个实验的因变量测量和实验 3 的历史感测量中增加了 Cronbach  $\alpha$  系数，以检验各题项的内部一致性信度，结果表明各次测量信度良好。

意见 3：从文章的题目来看，本文的实验目的是为了对比图片和文字评论哪种形式在网络购物中的作用，现实中的评论的确大多是图片和文字兼有，但是按照实验逻辑来讲，本文应该将两种评论形式作为因素的两个水平。因为将二者混合呈现难以避免它们产生的交互作用，而无法准确检验和对比两种评论方式的作用。

回应：感谢专家深刻的分析和见解。已将实验 2 和实验 3 均改为呈现单纯的图片、文字评论。

需要说明的是，本研究认为图片、文字评论本身并不会引起产品评价上的差异，而是它们激发了不同解释水平，进一步跟其他因素产生交互作用，从而影响被试对产品的评价。各个实验的结果也表明，评论形式的主效应是不显著的。我们认为，单纯的还是以某种形式为主的图片、文字评论，它们所激发的解释水平可能在程度上存在差异，但方向却是不变的；即单纯的图片评论和以图片为主的图文混杂评论激发的解释水平可能有所不同，但二者都激

发了相对较低的解释水平，单纯的文字评论和以文字为主的图文混杂评论则相反。实验 1b 结果也印证了这种观点。

此外，我们还查阅了以往一些关于解释水平的研究，发现有的研究也是用这种混杂但以某种因素为主的形式来操控变量，以激发不同的解释水平。如 Fujita, Eyal, Chaiken, Trope 和 Liberman (2008) 将课程的 4 条首要特征、1 条次要特征（或 1 条首要特征、4 条次要特征）混杂，以激发高解释水平（或低解释水平）。又如 Kivetz 和 Tyler (2007) 给被试呈现关于银行的 2 条认同性特征和 2 条实用性特征，并告诉被试 A 银行在认同性特征上更受好评，B 银行在实用性特征上更受好评，以分别激发被试对两个银行的高、低解释水平表征。

经过反复的讨论和权衡后，我们认为，单纯的呈现方式确实更符合通常的实验研究方法，而实验 1b 也已经证实了单纯还是混杂的图片、文字评论对被试产品评价的影响一致；因此这里采纳了专家的建议，使用单纯的图片、文字呈现方式。

#### 参考文献：

- Fujita, K., Eyal, T., Chaiken, S., Trope, Y., & Liberman, N. (2008). Influencing attitudes toward near and distant objects. *Journal of Experimental Social Psychology, 44*, 562–572.
- Kivetz, Y., & Tyler, T. R. (2007). Tomorrow I'll be me: The effect of time perspective on the activation of idealistic versus pragmatic selves. *Organizational Behavior and Human Decision Processes, 102*, 193–211.

**意见 4：**摘要结尾最好是对全文重点结果的概括总结，建议参考已发表的期刊论文进一步进行修改。另外，摘要中说 4 个实验，但具体内容只出现实验 1-3，建议修改。

**回应：**谢谢专家细心的检查。已将“4 个实验”改为“系列实验”，并在具体内容中注明了实验 1a、实验 1b、实验 2 和实验 3。但我们认为摘要中最后部分“本研究表明，当图片、文字评论与购买可行性在解释水平上一致时，影响力最大，且怀旧/非怀旧这一产品类型会调节这种效应。本研究还讨论了如何在实际中运用此结论。”应该算是对全文重点结果的概括总结，不知这种想法是否正确，请指正。

---

#### 审稿人 2 意见：

感谢作者对第一轮评审意见细致阅读和修改。针对评审人提出的对文字信息和图片信息抽象程度的质疑，作者指明抽象程度的比较适用于“表达相同信息时”，这一补充到位，解释合理。纵观全文，还有以下问题与作者商榷。

**意见 1：**本文的一个重要概念是“购买可行性”，购买可行性到底影响了个体决策中的哪些方面？比如，显然购买可行性对信息处理的介入程度有影响。那么本文的效应最合适的解释理论不应是解释水平理论，精细处理可能性模型（Elaboration Likelyhood Model）可能更为适用。（问题（1））

作者关于购买可行性和解释水平的关系有些含糊其辞。在“2.2 购买可行性与评论形式在解释水平上的一致性”这部分中（见下），句 1 中的论点有文献支撑，但不能推出句 2。句 3 支撑的是句 1 中的论点，而不是句 2。

“一般认为，人们在低解释水平下更关注事件的可行性，而在高水平下则更关注事件的价值和意义（句 1），因此可行性和价值分别代表着低解释水平和高解释水平（Liberman & Trope, 1998）（句 2）。基于这种观点，许多研究都发现可行性、价值与心理距离存在交互

作用。例如，在时间距离较远或概率较低的情境下，人们更关注选项的价值；在时间距离较近或概率较高的情境下，人们更关注选项的可行性（Fujita et al., 2008; Wakslak & Trope, 2009）（句3）。”

同样问题也发生在该部分下一段。句4缺乏理论支撑，句6说的是时间距离、社会距离对解释水平的影响，而不是可行性对解释水平的影响。

“但也有一些研究发现，如果将可行性作为一个独立的变量，其高低水平上的差异达到一定程度时，也可代表不同解释水平的心理表征：低可行性代表着相对较高的解释水平，而高可行性则反映了相对较低的解释水平（句4），且高、低可行性与心理距离间也存在交互作用（句5）。如人们在为近期将来做出决策时，会比为远期将来做出决策时对可行性高的选项（如易于操作的软件、难度较低的作业等）偏好更强烈（Eyal, Sagristano, Trope, & Liberman, 2009; Liberman & Trope, 1998）。在为社会距离较近的个体决策时，人们也会比为社会距离较远的个体决策时更偏好可行性高的选项（徐惊蛰，谢晓非，2011）（句6）。”（问题（2））

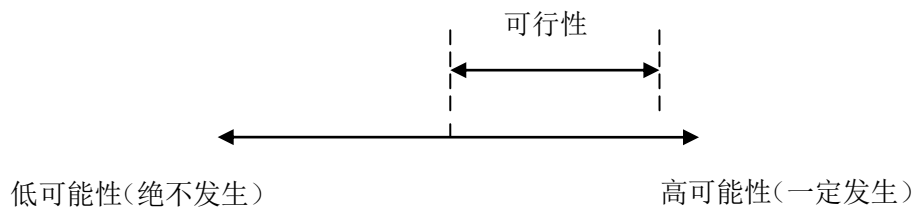
而且，几个实验发现购买可行性的主效应并不一致（1a 购买可行性的主效应不显著，1b 购买可行性的主效应显著，2 购买可行性的主效应不显著），请作者考虑不一致的原因。（问题（3））

回应：感谢专家深入的剖析，我们将从下列三个方面来回复意见1。

（1）在回答可行性对个体有哪些影响前，有必要再澄清一下“可行性”与“可能性”之间的关系。其实这在研究构思阶段就曾引起过激烈的讨论，因为有观点认为如果购买可行性低的话，人们根本就不可能选择其进行消费，因此购买可行性对个体的影响差异可能并不是通过解释水平起作用的。但在现实中人们并不是对所有购买可行性低的产品都采取忽略的态度，举例来说，当人们在网上搜索餐厅时，往往会划定一个地理区域（某个行政区或某个商业圈等），对于这个区域外的信息，人们自然会因其购买的可能性太低而选择忽略；但对于区域内的餐厅来说，虽然它们都很可能被选择为用餐的地方，但其购买可行性也存在差异，因为同一个商业圈里也有距离远近之分，也存在是否排队或需要预定等情况。对于市场来说，通过在线评论争取顾客的竞争也正好在这些都具有较高购买可能性，但购买可行性却存在差异的商家之间展开。基于此，本研究中的购买可行性，事实上都属于具有较高购买可能性。换言之，这里的高可行性是跟高可能性情境是类似的，但低可行性的“低”则是一种相对的低，它其实仍位于高可能性的一端，只不过比高可行性的可能性低而已（如下图所示）。相应地，研究中购买可行性对被试的影响也主要是引起了解释水平方面的差异。

关于解释水平理论和 ELM 的关系，以往研究也有所探讨。在解释水平理论看来，无论是高解释水平还是低解释水平都是跟努力加工模式（effortful processing mode，或中枢路径，central route）相联系的，因为人们对不同解释水平的信息都进行了深度加工（Fujita, K., Eyal, T., Chaiken, S., Trope, Y., & Liberman, N., 2008）。在本研究中，我们假设或期望的也是人们会仔细权衡评论所提供的信息（信息介入程度较高），即被试都会采用努力加工模式，这与一般商家所期望的也是一致的。而从本研究中大多实验对可行性情境的设定来看，低可行性情境在现实中也是普遍存在的，人们选择低可行性选项的可能性依然较高，因此被试仍会采取努力加工模式。

综上所述，对于本研究所指的可行性来说，被试的信息介入水平应该都是较高的，可行性主要影响的是被试的解释水平，可能不太适合用 ELM 来进行解释。



(2) 已将句 1 至句 3 部分改为“一般认为，可行性是与低解释水平相关联的，具有低解释水平的特征；价值则是与高解释水平相关联的，具有高解释水平的特征。由于与解释水平相匹配的信息对人们的态度和决策影响更大，而个体采用何种解释水平则主要依据与事物的心理距离远近，基于此，许多研究都发现可行性、价值与心理距离存在交互作用。例如，在时间距离较远情境下，目标的价值对人们的影响更大；在时间距离较近的情境下，达成目标的可行性对人们的影响更大，因此价值和可行性分别代表了高、低解释水平。(Lieberman & Trope, 1998; Fujita et al., 2008) 又如，由于价值和可行性分别代表着高、低解释水平，所以当强调目标的价值时，人们对概率的估计会较低；当强调达成目标的可行性时，人们对概率的估计会较高 (Wakslak & Trope, 2009)。”(文中已标为绿色字体)

同时，需要说明的是，在论文 2.1 部分有这样的陈述：“在心理距离远的情境里，人们会启用高水平解释的心理表征，此时与高水平解释相匹配的信息对人们的态度和决策影响更大；相反，与低水平解释相匹配的信息对人们的态度和决策的影响则更大 (Fujita, Eyal, Chaiken, Trope, & Liberman, 2008)”。因此“在时间距离较远的情境下”实质上是指“在高解释水平情境下”，“目标的价值对人们的影响更大”则意味着“目标的价值是跟高解释水平相匹配的”，即价值代表了高解释水平；反之，“达成目标的可行性则是跟低解释水平相匹配的”，即可行性代表了低解释水平。

同理，句 6 中“如人们在为近期将来做出决策时，会比为远期将来做出决策时对可行性高的选项（如易于操作的软件、难度较低的作业等）偏好更强烈 (Eyal, Sagristano, Trope, & Liberman, 2009; Liberman & Trope, 1998)”，其意为“人们在低解释水平情境下比在高解释水平情境下决策时，更偏好可行性高的选项”；“在为社会距离较近的个体决策时，人们也会比为社会距离较远的个体决策时更偏好可行性高的选项 (徐惊蛰, 谢晓非, 2011)”，其意与前相同。因此，高可行性是跟较低解释水平相匹配的，而低可行性则是跟较高解释水平相匹配的，即“低可行性代表着相对较高的解释水平，而高可行性则反映了相对较低的解释水平” (句 4)。

(3) 本研究中除了实验 1b 之外，其余所有实验结果均显示购买可行性没有主效应，这可能由实验情境、被试偏好差异或其它偶然因素造成的。但本研究的假设中并未考虑变量的主效应，而是把交互作用作为假设的核心和关键；换言之，对于本研究所关注的问题来说，是否产生主效应对我们所要考察的问题来说意义不大。因为，不管可行性是否单独对被试的购买意愿有影响，它都会受到评论形式和产品类型的调节（即存在交互作用），且这种调节作用都是通过因素间的解释水平是否匹配而发生的。基于上述理由，文中没有对各实验主效应的差异进行解释。

**意见 2:** 文章标题不太合适，对服务产品定义不清楚。

**回应:** 谢谢专家细致入微的评阅。为更准确地描述本研究的对象，已将“服务产品”改为“体验型产品”，并在文中增加了关于体验型产品的定义。



**意见 3:** 根据文中关于图片信息和文字信息的理论推导部分，那么文章在概念层面可否从图片评论和文字评论改为图片信息和文字信息？

**建议:** 我们对作者的建议是，深入考虑“购买可行性”这一概念，能否在理论上找到支撑并完善逻辑链条？考虑解释水平理论是否是最合适的理论？如果仍用解释水平理论，我们建议改为研究图片、文字评论在不同时间距离、空间距离、社会距离和概率情境下对消费者产品评价的影响，巧妙地设计一系列实验，将是对该研究领域的重要研究贡献。

以下文献供作者参考: Effects of Social and Temporal Distance on Consumers' Responses to Peer Recommendations; Zhao, Min, and Jinhong Xie; Journal of Marketing Research; Issue: 48 (June); 2011; Pages: 486-496

**回应:** 感谢专家深刻的分析和见解。事实上，本研究可看作是解释水平理论在网络营销领域的一种应用。虽然在理论上本研究所展示的图片、文字跟解释水平的关系是普遍存在的，但由于本研究是通过网络评论来探讨这种关系的，在实验中对图片、文字评论的内容和形式都进行了严格的控制，如果无条件地推广到其他类型的图片、文字信息，可能会出现无法预料的问题。再者，将其他类型的图片、文字信息跟购买可行性、产品类型等因素相结合，也存在许多困难。因此，为谨慎起见，我们没有将“评论”改为“信息”。但专家的意见也对我们有很大启发，在今后的研究中，我们会考虑将其他类型的图片、文字信息作为研究对象。

在解释水平理论研究领域，大多研究确实是将空间距离、时间距离、社会距离作为主要的研究对象，对概率和可行性的研究则相对较少，因此在证据方面也不如其他几个方面丰富，看起来似乎也不如其他方面的理论支撑稳固。但我们认为，从现有的其他研究以及我们之前的研究结果来看，有充分的证据表明可行性的高低确实跟解释水平存在对应关系；而且对可行性的研究也是对解释水平理论研究领域的有益补充，这也正是本研究的亮点之一。此外，本研究是一个更大课题下的子课题，对于图片、文字评论和购买可行性在不同时间距离、社会距离情境下对消费者产品评价的影响，我们已经发表过相关的论文，所以此次转向了相对较“冷门”但又非常具有现实意义的方向。

---

### 第三轮

**审稿人 2 意见:**

作者针对评审意见做出了细致的推敲和修改，理论、实验设计和分析方法等部分都有所加强。在充分肯定作者所做的完善的同时，文章仍存在以下问题：

**意见 1 第一部分:** 作者就理论部分进行了完善。由于文章研究的一些概念在构念并不很清晰（如“购买可行性”），针对第二轮提出的“意见一”的修改并不能令人满意，关键几个概念之间的关系和逻辑链条依然没有梳理得很清晰，特别是购买可行性与解释水平之间的关系。虽然补充了实验测量解释水平，但测量方法比较单一，且对于仅通过改变对体验型产品地理位置的描述，就可以改变解释水平，我们尚存疑虑。

**回应:** 谢谢评审专家深刻的剖析和意见，下面将从三个方面来回应上述意见。

(1) 现已对文中“2.2 购买可行性与评论形式在解释水平上的一致性”部分的内容进行了修改（用紫色字体标注），希望现在对购买可行性的构念更清晰了。可行性、愿望性（desirability，之前译为价值）与解释水平的关系跟图片、文字与解释水平的关系类似，它们分别与低、高解释水平相联系，这在以往研究中已经得到了证实（Liberian & Trope, 1998，摘要原文“In goal-directed activities, desirability of the activity's end state represents a high-level

construal, whereas the feasibility of attaining this end state represents a low-level construal.”)。尽管相对于愿望性来说,可行性是与低解释水平相关联的;但可行性也可分为高可行性和低可行性,当可行性的差异达到一定程度时,所关联的解释水平也会产生差异。这与近年的一篇综述中的观点一致:对真实事物进行模拟的低水平解释表征(具体的)和通过符号来描述的高水平解释表征(抽象的),这两种表征方式本身都是可以在解释水平上再细分的(Trope & Liberman, 2010)。同时,由于人们在对高或低解释水平相关信息进行加工时,会自动地分别使用高、低解释水平的思维模式,所以不同可行性所关联的解释水平特征也会激发相应的解释水平思维模式。总之,可行性与低解释水平相关联,但如果将可行性程度再进行细分,高、低不同的可行性又是分别与低、高解释水平相关联的,并可以激发相同解释水平的思维模式。

(2) 解释水平对信息加工的影响是自动化的,或通俗地讲是“直觉”的,一般来说人们在进行信息加工时无法用言语描述出“解释水平”的差异。因此,跟许多心理测量一样(如内隐记忆的测量),解释水平无法进行直接测量,不能让被试直接报告自己的解释水平是否发生了变化,而只能通过某种任务来间接地测量被试的解释水平是否发生了变化。用分类任务(classification task)对解释水平进行测量的原理在于:由于更概括、更少和上位范畴的类别(superordinate categories,如食物)比更具体、更多和下位范畴的类别(subordinate categories,如面包)更加抽象和具有包容性,因此当个体以一种更抽象的高水平解释的思维加工信息时,会将物品划分为更少和更概括的类别,即分类任务中类别数越少,表明被试解释水平越低;反之,则解释水平越高(Liberman, Sagristano, & Trope, 2002)。

而分类任务也是以往研究中常用的测量解释水平的方法。如 Liberman, Sagristano 和 Trope (2002) 通过实验发现,相比近距离将来的情境,远距离将来情境下的被试将物品划分为更少的类别,这证实了远距离将来激发了高水平解释心理表征这一假设,也表明了分类任务测量解释水平的有效性。Wakslak, Trope, Liberman 和 Alony (2006) 也采用了分类任务来测量解释水平,发现相比事件发生概率较高的情境,事件发生的概率较低情境下的被试将物品划分为更少的类别,这表明后者启动了更高水平解释的心理表征。此外, Lee, Keller 和 Sternthal (2010) 也用分类任务来测量被试的解释水平,发现相比防御聚焦情境下的个体,促进聚焦情境下的个体将物品划分为更少的类别,这表明后者启动了更高水平解释的心理表征。从本研究预实验结果来看,分类任务是可以较好测量被试的解释水平的;且正式实验中的分类任务结果表明,改变购买可行性条件后,被试的解释水平确实发生了变化。

(3) 通过改变地理位置来操控可行性或解释水平这种方法在以往研究中也运用得比较普遍。如 Todorov, Goren 和 Trope (2007) 分别用“餐厅就在当地” vs. “餐厅离当地 30 分钟路程”来操控高、低可行性。徐惊蛰和谢晓非(2011)用“超市就在住处附近,步行距离 5 分钟,且顺路就能到” vs. “超市距离住处较远,单程 1 小时左右,需乘地铁再换乘公交方可到达”来操控高、低可行性。Nunes 和 Park (2003) 分别用“当地商场” vs. “需要 15 分钟车程的另一商场”来操控高、低可行性。Lu, Xie 和 Xu (2013) 用“餐厅距离非常近,步行只需五分钟”, vs. “餐厅位置距离较远,需要两个小时才能到达”来操控高、低可行性。Baskin, Wakslak, Trope, 和 Novemsky (2014) 也用“餐厅需要 5 分钟路程” vs. “餐厅需要一个小时路程”来操控高、低可行性。且在本研究中不仅仅是改变了地理位置,也改变了到达目的地的“方便程度”,如“步行 5 分钟,顺路就能到达”中的“步行”、“顺路”,“乘车一个半小时,换乘两次公交车”中的“乘车”、“换乘”,它们在方便程度上是有差异的;“地铁站口,出站就能看到”与“某大厦 A 座 307 号店铺,离地铁站出口约需 10 多分钟步

行路程，其间要转几个路口”也不只是位置的差别，二者更大的差异在于路程的复杂程度，即到达目的地的“方便程度”。操控检验的结果也证实了不同可行性条件下的被试在方便程度感知上的差异。因此，我们认为本研究实验中通过改变对地理位置的描述，改变了被试的购买可行性，从而改变了他们的解释水平。

**意见 1 第二部分：**其次，在对第一位评审人第二轮意见三的回复中，作者说“图片、文字评论激发了不同解释水平”，这似乎与文章的理论并不一致。

**回应：**感谢专家细致的评阅。不知道这里的不一致是否指论文中说的是文字、图片分别“具有高、低解释水平的特点”（现已改为：文字、图片信息分别是与高、低解释水平相关联的），回复中说的是文字、图片分别“激发了不同的解释水平”。如果是的话，可以从以下角度对此进行解释。

基于解释水平理论的观点，解释水平的差异可以体现在信息媒介上（如文字、图片），也可以体现在事物或对象的不同属性上（如可行性与价值，主要特征与次要特征等），还可以体现在个体的思维模式上（如抽象思维模式，具体思维模式）等不同方面。而且人们对高、低解释水平相关联的信息会自动地采用相同解释水平的思维模式，如 Lee, Keller, & Sternthal(2010)研究发现，相比接收防御聚焦信息的个体，接收促进聚焦信息（高解释水平的事物属性）的个体将物品划分为更少的类别，这表明后者启动了更高水平解释的心理表征（高解释水平的思维模式）。对于文字、图片而言，Amit, Algom, & Trope（2009）的研究也已证实，文字和图片分别具有高、低解释水平的特点，并与不同心理距离相关联；并且，相比文字形式而言，被试在分类任务中将图片形式呈现的物体划分为更多的类别，表明文字和图片（不同解释水平特征的信息）分别激发了高、低解释水平的思维模式。

在本研究中，购买可行性是实现购买行为这一事件的一种属性，不同购买可行性与不同的解释水平相关联，因此激发了个体不同解释水平的思维模式；而文字、图片评论则是与不同解释水平特征相关联的信息媒介，它们也会激发个体不同的解释水平思维模式；当可行性和评论形式激发的解释水平一致时，个体的认知加工最流畅，因此觉得产品更有吸引力，而二者激发的解释水平不一致时，则会对个体的认知加工形成干扰，降低其对产品的评价（在意见 2 的回复中，对“解释水平一致性效应”的机制有更详细的探讨）。

**参考文献：**

- Baskin, E., Wakslak, C. J., Trope, Y., & Novemsky, N. (2014). Why feasibility matters more to gift receivers than to givers: a construal-level approach to gift giving. *Journal of Consumer Research*, 41(1), 169-182.
- Lee, A. Y., Keller, P. A., & Sternthal, B. (2010). Value from regulatory construal fit: The persuasive impact of fit between consumer goals and message concreteness. *Journal of Consumer Research*, 36(5), 735-747.
- Liberman, N., Sagristano, M., & Trope, Y. (2002). The effect of temporal distance on level of construal. *Journal of Experimental Social Psychology*, 38, 523-535.
- Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology*, 75, 5-18.
- Lu, J., Xie, X., & Xu, J. (2013). Desirability or feasibility self-other decision-making differences. *Personality and Social Psychology Bulletin*, 39(2), 144-155.

- Todorov, A., Goren, A., & Trope, Y. (2007). Probability as a psychological distance: Construal and preference. *Journal of Experimental Social Psychology, 43*, 473–482.
- Nunes, J. C., & Park, C. W. (2003). Incommensurate resources: Not just more of the same. *Journal of Marketing Research, 40*, 26–38
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review, 117*(2), 440–463.
- Wakslak, C., Trope, Y., Liberman, N., & Alony, R. (2006). Seeing the forest when entry is unlikely: Probability and the mental representation of events. *Journal of Experimental Psychology: General, 135*, 641–653.
- 徐惊蛰, 谢晓非.(2011). 解释水平视角下的自己—他人决策差异. *心理学报, 43*(1), 11–20.

**意见 2:** 作者认为信息与解释水平的匹配是“文字评论和图片评论谁更有效”的关键机制,这一理论视角体现了作者的深刻洞察。但对文献的引用和梳理尚有不足,使得理论推导比较薄弱,缺乏支撑。分析部分没有直接的机制分析,理论贡献上有所欠缺。

**回应:** 谢谢专家的意见和指导。我们对论文的 2 文献综述部分再次进行了修改和完善(用紫色字体标注),希望能解决这方面的问题。

对于心理距离和解释水平的工作机制,目前似乎并没有一个统一的说法,即使是理论的创立者,也是用不同的方式来解释其机制。如,在一篇综述中, Trope 和 Liberman (2010) 曾表达“当心理距离与信息媒介(文字、图片)在解释水平上一致时,认知加工最有效”,“心理距离的增加减少了对低解释水平信息的运用”,“增加行为结果的心理距离,会增加中心特征的重要性”,“当心理距离增加时,愿望性(desirability)会更受重视”等等,这些说法各不相同,但都用来对心理距离和解释水平的工作机制进行解释。

在心理距离与解释水平关系的研究中, Stroop 范式(经典的 Stroop 效应:当个体看到字体是红色的“蓝”字时报告字的颜色,其反应时比字体是蓝色的“蓝”字要长;或看到蓝色字体的“红”字时报告字的颜色,反应时比看到红色字体的“红”字要长。出现这种现象是因为人们会自动地对字的意义进行加工,但当要求其加工字体的颜色时,就必须抑制这种自动加工,去加工字体的颜色,对意义的加工干扰了对字体的加工,因此增加了认知加工的难度,所以反应更慢。)是常用的研究方法。如汽车、马车(信息的内容)因时间久远不同分别属于低、高解释水平的事物,图片、文字(信息的媒介)因抽象程度不同也分别属于低、高解释水平;当个体在对文字“汽车”进行认知加工时,信息的内容和信息的媒介分别要求或激发其用不同的解释水平思维模式对信息进行加工,两种不同的思维模式相互干扰,所以反应时比对文字“马车”的反应时长。同理,对图片“马车”的反应时比对图片“汽车”的反应时长(Amit, Algom, & Trope, 2009)。类比 Stroop 效应的机制,我们倾向于认为“解释水平一致性效应”的原理是:如果不同的事物激发了个体相同的解释水平思维模式,个体的认知加工就会更容易(或认知流畅性高);如果不同的事物激发了个体不同的解释水平,个体的认知加工就会因受到干扰变得困难(或认知流畅性低)。而正是这种认知加工的难易(或流畅性)影响了人们对事物的态度和决策。

尽管目前关于“解释水平一致性效应”的具体机制还没有更多的直接证据,但这种现象或规律的存在却是得到了众多研究的证实。而本研究的目的也不是直接考察“解释水平一致性效应”的工作机制,而主要是对这种效应的应用,要回答的问题是“图片、文字评论中谁的影响力更大”,因此在文献综述中没有将探索效应的机制作为一个重点来专门阐述,但我

们在讨论部分“6.1 图片、文字网络评论对消费者的影响”对这个机制进行了说明（详细请参见论文 6.1 紫色字体部分）。

参考文献：

Amit, E., Algom, D., & Trope, Y. (2009). Distance-dependent processing of pictures and words. *Journal of Experimental Psychology: General*, 138, 400–415.

Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440-463.

**意见 3：** 研究对实验进行了严谨的设计，并针对意见进行了实验的补充。但实验的方法比较单一，既是网络购物环境，可否考虑增加现场实验，提升研究的外部效度。

**回应：** 感谢专家细致的评阅和中肯的意见。如果进行现场实验，目前存在以下困难：（1）尽管很多网络评论平台都有将评论进行分类的功能，但无法像实验中那样对评论的呈现进行控制，因为当前网站还不具备相应功能来区分出哪些评论是以图片为主，哪些是以文字为主；如果在网络评论平台上开发相应的评论分类呈现功能，所需的时间、人力、财力成本都非常高，显得很经济；（2）不易对被试进行区分，因为无法分清哪些消费者是受实验设置的评论所影响前来消费，哪些是因为其他因素前来。因此，虽然我们非常赞成并且也希望能进行现场实验研究，但目前开展现场实验的条件却尚不成熟。

---

**编委专家意见：**

文章运用解释性水平理论探讨网络评论中文字和图片评论在何种条件下对消费者产品评价影响更大，视角比较新颖，有一定的新意，论文的结论对企业营销也有一定的指导意义。论文在如下方面需要修改：

**意见 1：** 论文多处有文字错误。

**回应：** 谢谢专家细心的检查。经过修改和反复校对，我们发现和纠正了一些文字错误，希望这一问题已经得到解决。

**意见 2：** 目前论文文字接近 1 万 9 千字，建议大幅度压缩，压缩到 1 万 2 千字左右。

**回应：** 谢谢专家的意见。我们已经对论文进行了压缩，目前字数是 1 万 5 千多字，但如果压到 1 万 2 千字左右的话，可能会失去一些重要的信息，不知能否请编委酌情考虑下此情况。

**意见 3：** 将 desirability 翻译成“价值”似乎不妥，因为价值是一种关于成本与收益的主观比较，desirability 反映的是欲求程度或水平，两者不宜画等号。

**回应：** 谢谢专家的意见和指导。目前对 desirability 的翻译主要有“价值”和“愿望性”两种，从内涵来说，翻译成“愿望性”更贴近其本意；但目前翻译为“价值”的文献更多一些，本文选择将其翻译为“价值”主要是出于避免术语滥用方面的考虑。但为了便于理解，我们接纳了专家的意见，将“价值”改为了“愿望性”。

**意见 4：** 用 feasibility 这样一个心理变量来解释另一个心理变量“产品评价”也是值得推敲的，在具体操作中，实际上是用“方便程度”来反映 feasibility，方便程度似乎比 feasibility 有更多客观的成分，或许从企业营销的角度也更容易把握。

**回应：**感谢专家深刻的分析和见解。根据解释水平理论，个体采用何种解释水平对事物进行心理表征，比起事物本身来说，对个体偏好的影响更大。feasibility 是虽然是一种主观感受，但却是通过解释水平来影响个体偏好（即“产品评价”）的；而“方便程度”虽然显得更客观，但也是通过解释水平来影响个体的“产品评价”。因此，不管是用 feasibility 还是“方便程度”，在解释“产品评价”的差异上，二者并不会有很大差异。

此外，除了“方便程度”，feasibility 还可以有多种其它表现形式。如去驾校学习的目标是考取驾照，练车场的远近、是否有车接送属于实现目的“方便程度”，是实现考取驾照这一目标的 feasibility 的一种反映；而教练的水平、车辆的好坏则不属于“方便程度”，但也反映了实现考取驾照这一目标的难易程度，是 feasibility 的另一方面。本研究之所以采用“方便程度”来反映 feasibility，是因为在服务产品消费中，这是最为常见的 feasibility 表现形式。但如果将 feasibility 直接换成“方便程度”的话，就缩小了 feasibility 的范畴，因此我们认为不宜用“方便程度”代替 feasibility。

---

### 审稿人 3 意见：

文章《谁该成为体验型产品网络评论的主角，图片还是文字？》的选题具有重要的现实意义。作为一篇学术论文，本文在理论框架，尤其是变量界定方面，仍存在问题，具体的评审意见及几点建议请见下面。祝研究顺利！

随着互联网的普及，网络评论已经对消费者的行为和决策产生了重要影响。故此，文章《谁该成为体验型产品网络评论的主角，图片还是文字？》的选题具有重要的现实意义，其发现的结论一方面可以帮助消费者了解自身的信息处理模式，另一方面也可能为企业的营销实践提供积极的指导意见。作为一篇学术论文，本文写作思路清晰、文献梳理认真，且已经进行了两轮的修改，文章的大部分细节问题都已得到改进。但在理论框架，尤其是变量界定方面，仍存在以下问题，希望作者加以解释说明，并提供相应的实证依据。

**意见 1：**关于核心变量的界定，作者在原文中提到“本研究认为文字评论具有抽象、核心、本质和去背景化的特点，与高解释水平的心理表征相关联；而图片通常更加具体、特殊以及背景化，反映了低解释水平的心理表征。”尽管作者引用了一些相关文章，但是否有实证研究证实了这一点推论？举个例子，相对于一张图片来说，10 个字的文字评论可能比较抽象，但如果是 100 字的文字描述呢？文字描述随着字数增加，可以提升细节的刻画，会使得更加具象。同样对于图片来说，不同像素和角度的图片也会体现不同的抽象水平。所以很难单纯将文字和图片进行对比。

网络评论形式（图片 vs 文字）是本文的一个核心变量，但这个变量的界定并不清晰。从理论上说，我们可以将文字评论分为抽象描述（比如笼统地说餐厅陈列精致）vs 具象描述（比如具体地描述餐厅的布置、摆设等等），在实验操纵时能够控制两段文字的字数和语言风格（rational vs. emotional），只改变抽象还是具象描述。这样的变量界定是清晰的，并且实验控制很干净。但是目前文章粗略地将文字和图片进行对比，作为同一个变量的两个维度，包含了太多的其他因素，比如文字的字数，图片的质量和角度，这些都会影响结果。

**回应：**谢谢评审专家深入的剖析和深刻的见解。我们非常认同评审专家的“抽象”和“具体”可以再细分的观点，根据解释水平理论的观点，对真实事物进行模拟的低水平解释表征（具体的，如图片）和通过符号来描述的高水平解释表征（抽象的，如文字），这两种表征方式

本身都是可以在解释水平上再细分的 (Trope & Liberman, 2010)。因此图片可以有相对具体的图片 (如注重写实) 和相对抽象的图片 (如注重意境), 文字亦然。

但对于网络评论而言, 大都是将同一商家的评论集中在一起, 或根据产品的某一共同属性将评论信息集中在一起; 所以人们在浏览网络评论时看到的评论内容都是对同一商家, 或同一产品属性的描述, 即评论的对象或内容具有一致性。在正式实验前, 我们先选出了一些餐厅和酒店网络评论最常见的一些产品属性, 然后根据这些产品属性寻找图片评论和编写相应的文字评论。通过预实验确定了**内容一致的文字和图片评论**作为实验材料 (关于预实验详见论文 3.1、4.1 和 5.1 部分的内容)。因此本文并不是笼统粗略的将所有文字评论和所有图片评论进行对比, 而是研究表达同一内容的文字和图片形式评论, 而对于内容相同的图片和文字评论来说, 文字评论的抽象水平显然更高, 图片则更具体。举例来说, 就算用一大段文字来说明餐厅的桌子是红色的 (如可以用朱红、粉红、暗红、大红等词汇来形容, 还可以举例“像某某东西一样红”, 也可以跟其他事物对比“比某某事物红色深一些, 比某某事物红色浅一些”, 等等), 但每一种红色其实都是一个红色的子范畴 (“抽象”的红色), 都可根据其颜色的深浅再进行细分, 描述得更具体仅仅意味着这个子范畴更小而已 (“抽象”程度更低)。但如果是看到一张红色桌子的图片, 那么不管这张桌子是属于哪类红色, 人们头脑中都只有一种红色, 就是图片上所呈现的那种红色 (即 “具体” 的红色)。

而图片和文字在解释水平上的差异在以往的研究中也得到过证实, 如对物品的识别任务中, 相比图片形式, 被试对文字形式呈现的心理距离较远的同一物品 (如马车) 反应时较短; 而相比文字形式, 被试对图片形式呈现的心理距离较近的同一物品 (如汽车) 反应时较短 (Amit, Algom, & Trope, 2009)。在时间距离、社会距离远的情境下, 人们更偏好以文字形式发送信息, 在时间距离、社会距离近的情境下, 人们更偏好以图片形式发送信息 (Amit, Wakslak, & Trope, 2013)。这些研究都证实了在表现同样或相似内容时, 相比图片, 文字是与高水平解释关联的表征; 而相比文字, 图片是与低水平解释相关联的表征 (我们已在文章中对这些文献进行了补充说明, 并对文献综述部分进行了梳理和完善, 具体请参见论文 2.1 紫色字体部分)。综上, 我们认为, 尽管文字和图片的解释水平都可以再细分, 但对于本研究来说却不是必须的。

此外, 在所有实验中, 对文字的字数、字体, 图片的大小等因素也专门进行了处理, 尽量消除了额外因素的影响。

#### 参考文献:

- Amit, E., Algom, D., & Trope, Y. (2009). Distance-dependent processing of pictures and words. *Journal of Experimental Psychology: General*, 138, 400-415.
- Amit, E., Wakslak, C., & Trope, Y. (2013). The use of visual and verbal means of communication across psychological distance. *Personality & social psychology bulletin*, 39(1), 43-56.
- Trope, Y., & Liberman, N. (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440-463.

**意见 2:** 关于文章的理论框架存在一个核心问题, 即两个自变量 (购买可行性和评论形式) 都分别会影响解释水平, 它们之间并不是正交的关系。同样地, 怀旧产品 vs 非怀旧产品也会影响解释水平, 这使得实验结果是三个变量的叠加效果。从严格意义上来说, 这并不是调节作用, 而是一个简单的匹配效应。以实验 1 为例, 购买可行性低 (其实就是较远的空间距离, 本身会激发 high construal level), 因为文字也是更为抽象的, 所以当告知被试餐厅较

远时会喜欢文字评论，这其实就是 *construal level* 的一个匹配。如果本文的理论框架就是基于解释水平的匹配与否，那其实就是将解释水平应用到网络评论的情景，本身从理论贡献上来讲有所欠缺。

**回应：**感谢专家深刻的分析和见解。解释水平理论是一个比较成熟的理论，本研究并没有突破解释水平理论的基本框架，但仍有一定的理论贡献和价值。首先，本研究主要目的在于探究网络评论对消费者的影响机制，并如何提高网络评论在营销中效用；所要考察的首要问题是“图片和文字评论在什么情况下影响力更大”，而这在以往研究中一直存在争议。本研究则假设并证实了“文字和图片只有在跟其他因素解释水平一致时，各自的影响才最大”，这为深层次理解图片、文字的影响机制提供了新的思路。

其次，以往研究大多认为可行性与低解释水平相关联，愿望性 (*desirability*, 之前译为价值) 与高解释水平相关联，忽略了可行性本身也与不同程度的解释水平相关联，并对人们的评价和决策产生影响 (这与贵专家在意见 1 中的观点也有共通之处，如文字也可分为具体的文字和抽象的文字；论文“2.2 购买可行性与评论形式在解释水平上的一致性”部分对此进行了专门讨论)；本研究则恰恰证明了不能简单地把可行性划只归于低解释水平，不同程度的可行性也能激发不同程度的解释水平，并对个体产生影响。这是通过实证研究回应了 Trope 和 Liberman (2010) 所提出的理论观点“即对真实事物进行模拟的低水平解释表征 (具体的) 和通过符号来描述的高水平解释表征 (抽象的)，这两种表征方式本身都是可以在解释水平上再细分的”。从这个意义上而言，本研究结论进一步拓了解释水平的研究，对解释水平理论的发展也是一项积极尝试。

第三，目前关于网络口碑的研究大多忽略了人们对口碑的内部心理加工过程和机制，本研究将解释水平引入网络口碑研究，既为理解网络口碑影响机制提供了一种新的理论框架，也推进了网络口碑研究纵深方向的发展。在本修改稿中，我们也从以上三个方面对论文讨论部分进行了一些修改和补充，以期明确本研究的理论贡献。

**意见 3：**除了作者提出的假设之外 (*match story*)，是否有可能存在一种 *compensation story*，比如在购买可行性低的情境下，本身被试是一种高解释水平的思维模式，此时被试有可能更需要一些具象的信息来帮助他们做决策，如果这个解释成立，图片评论应该在购买可行性低的情境下更有效果。

**回应：**感谢专家深入的见解。这确实是一个非常有趣和有价值的观点，但从现有关于解释水平理论的研究来看，我们还没有发现跟匹配效应“不匹配”的研究结果。但这并不排除专家提出的“补偿”理论，或许在某些特殊的情况下，可能会出现跟匹配效应正好相反的补偿效应。例如，一般来说解释水平的匹配效应对人们的影响是自动化的，不被人们所意识的；但如果一个人了解解释水平理论的原理，为了让自己的决策更合理，可能会有意识地搜寻不同解释水平的信息进行补偿。再比如，人们对高、低解释水平的信息会自动地采用相同解释水平的思维模式，如果通过训练来改变这种自动加工的习惯，可能会出现一些不一样的结果。这在以后的研究中确实是非常值得探讨的。

**意见 4：**实验 1 中对于购买可行性的操控存在一些问题。其实作者操纵的是餐厅的空间距离，*spatial distance* 本身就是心理距离的一个维度，会直接影响解释水平和思维模式。从这个角度来说，其实本文提到的购买可行性完全可以换成其他变量，比如心理距离。



参考文献: Trope, Yaacov, & Liberman, Nira (2010). Construal-level theory of psychological distance. *Psychological Review*, 117(2), 440-463.

回应: 谢谢评审专家细心的审阅和中肯的意见。实验 1a 用餐厅的路程来操纵购买可行性是借鉴了以往研究的做法, 事实上不少其他研究对可行性的操控都涉及到路程。如 Todorov, Goren 和 Trope (2007) 分别用“餐厅就在当地” vs. “餐厅离当地 30 分钟路程”来操控高、低可行性。徐惊蛰和谢晓非(2011)用“超市就在住处附近, 步行距离 5 分钟, 且顺路就能到” vs. “超市距离住处较远, 单程 1 小时左右, 需乘地铁再换乘公交方可到达”来操控高、低可行性。Nunes 和 Park (2003) 分别用“当地商场” vs. “需要 15 分钟车程的另一商场”来操控高、低可行性。Lu, Xie 和 Xu (2013) 用“餐厅距离非常近, 步行只需五分钟”, vs. “餐厅位置距离较远, 需要两个小时才能到达”来操控高、低可行性。Baskin, Wakslak, Trope 和 Novemsky (2014) 也用“餐厅需要 5 分钟路程” vs. “餐厅需要一个小时路程”来操控高、低可行性。

其次, 从以往关于空间距离与解释水平关系的研究来看, 大多对不同空间距离的设置差异巨大。如 Fujita, Henderson, Eng, Trope 和 Liberman (2006) 用“距离这里 3 英里以外” vs. “距离这里 3000 英里以外”来操控近和远的空间距离。Henderson, Fujita, Trope 和 Liberman (2006) 针对纽约的被试, 分别用在“在纽约曼哈顿” vs. “在意大利佛罗伦萨”来操控近和远的空间距离。Stephan, Liberman 和 Trope (2010) 分别用“在离被试不远的一栋建筑物内” vs. “在另一个城市”来操控近和远的空间距离。Amit, Wakslak 和 Trope (2013) 针对纽约的被试, 分别用“位于东海岸的某城市” vs. “位于西海岸的某城市”来操控近和远的空间距离。而本研究的设置中(“步行 5 分钟, 顺路就能到达” vs. “乘车一个半小时, 换乘两次公交车”, 以及“地铁站口, 出站就能看到” vs. “某大厦 A 座 307 号店铺, 离地铁站出口约需 10 多分钟步行路程, 其间要转几个路口”)空间距离的差异相对较小, 可能对被试的解释水平影响并不大。

事实上我们也注意到空间距离和购买可行性可能会混淆, 因此在实验 1b 和实验 3 中, 除了使实验情境更接近现实情况外, 也希望通过将餐厅限定在同一步行街内, 让被试在心理上将不同餐厅划定在同一区域内, 并在情境设置中极大地缩小了不同餐厅的空间距离(“地铁站口, 出站就能看到” vs. “某大厦 A 座 307 号店铺, 离地铁站出口约需 10 多分钟步行路程, 其间要转几个路口”)。经过这样的调整后, 仍得到了跟实验 1a 相同的结果。此外, 实验 2 采用了“是否需要信用卡担保”来操控购买可行性, 这样一来就完全排除了空间距离的影响, 得到的结果仍跟其他实验一致。

综上所述, 我们认为, 虽然不同空间距离很可能会引起被试解释水平的变化, 但本研究中对空间距离的设置差异相对较小, 引起结果差异的可能主要还是由到达目的地的“麻烦程度”造成的。如“步行 5 分钟, 顺路就能到达”中的“步行”、“顺路”, “乘车一个半小时, 换乘两次公交车”中的“乘车”、“换乘”, 它们在方便程度上是有差异的; “地铁站口, 出站就能看到”与“某大厦 A 座 307 号店铺, 离地铁站出口约需 10 多分钟步行路程, 其间要转几个路口”也不只是地理位置的差别, 因此不宜将本研究中的“购买可行性”等同于“心理距离”。

参考文献:

Amit, E., Wakslak, C., & Trope, Y. (2013). The use of visual and verbal means of communication across psychological distance. *Personality & social psychology bulletin*, 39(1), 43-56.5.

- Baskin, E., Wakslak, C. J., Trope, Y., & Novemsky, N. (2014). Why feasibility matters more to gift receivers than to givers: a construal-level approach to gift giving. *Journal of Consumer Research*, *41*(1), 169-182.
- Fujita, K., Henderson, M., Eng, J., Trope, Y., & Liberman, N. (2006). Spatial distance and mental construal of social events. *Psychological Science*, *17*, 278-282.
- Henderson, M. D., Fujita, K., Trope, Y., & Liberman, N. (2006). Transcending the “here”: The effect of spatial distance on social judgment. *Journal of Personality and Social Psychology*, *91*, 845-856.
- Lu, J., Xie, X., & Xu, J. (2013). Desirability or feasibility self-other decision-making differences. *Personality and Social Psychology Bulletin*, *39*(2), 144-155.
- Nunes, J. C., & Park, C. W. (2003). Incommensurate resources: Not just more of the same. *Journal of Marketing Research*, *40*, 26-38
- Stephan, E., Liberman, N., & Trope, Y. (2010). Politeness and social distance: A construal level perspective. *Journal of Personality and Social Psychology*, *98*, 268-280
- Todorov, A., Goren, A., & Trope, Y. (2007). Probability as a psychological distance: Construal and preferences. *Journal of Experimental Social Psychology*, *43*, 473-482.
- 徐惊蛰, 谢晓非. (2011). 解释水平视角下的自己—他人决策差异. *心理学报*, *43*(1), 11-20.

**意见 5:** 目前实验刺激物的选择并不是很符合现实。真实的网络评论多数是文字+图片，很少会有单纯只有图片的评论。即便作者在其中一个实验用了图片和文字混合的方式，但也是5条评论中四条纯图片，一条纯文字。而真实的网络评论是每一条评论都会有文字描述，然后配上图片。比较贴近现实的一种操纵方法是给被试呈现一条评论，这条评论可以是以文字描述为主，配上一张简单的图片（文字为主型评论），或是只是简单的一句话文字描述，配上多张细节的图片（图片为主型评论）

**回应:** 谢谢评审专家专业的意见。在以往的研究中，有不少研究确是用混合的方式来呈现实验材料（如 Fujita, Eyal, Chaiken, Trope, & Liberman, 2008; Kim, Zhang, & Li, 2008），在本研究最初的设计里，所有实验也都是用的混合方式呈现图片、文字评论。但在有专家提出，这样做很难说明是否是单纯的文字或图片在发生作用。在经过讨论后，我们也认为混合的形式可能存在一些问题，例如，假设图片评论对被试没有任何作用，但文字评论条数的多少会对被试造成影响，那么混合呈现的结果差异就不是由文字为主或图片为主所造成的了，而是单纯由文字评论的条数影响了结果。基于此，我们采取了一种折中的方式，先在实验 1a 中采用纯文字、图片评论证明评论形式与购买可行性交互作用的存在，然后实验 2b 中采用混合的形式来呈现评论；两个实验得到的结果一致，所以在实验 2、实验 3 中继续采用了纯文字、图片评论的呈现方式。

#### 参考文献:

- Fujita, K., Eyal, T., Chaiken, S., Trope, Y., & Liberman, N. (2008). Influencing attitudes toward near and distant objects. *Journal of Experimental Social Psychology*, *44*, 562-572.
- Kim, K., Zhang, M., & Li, X. (2008). Effects of temporal and social distance on consumer evaluations. *Journal of Consumer Research*, *35*, 706-713.

**意见 6:** 实验中的分类任务是放在 DV 测量之后进行。在此之前，被试已经接受了购买可行性和评论形式（图片 vs 文字）的操纵，所以最后的分类任务结果其实受到这两个变量的叠加影响，很难得出结论认为看图片评论比看文字评论更能激发低解释水平。

**回应:** 谢谢评审专家细致的审阅和中肯的意见。首先，分类任务似乎只能在 DV 测量之后，如果放在看评论之前的话，是无法测到两种评论的影响效果的；而且购买可行性情境的设置与呈现评论二者的顺序也不宜调换，因为人们一般都是先有消费的需求再看评论，先看评论再产生消费需求的情况较少；再加上分类任务耗时较长，如果放在呈现评论之后，设定购买可行性情境之前，被试可能会遗忘评论内容。

再者，看图片和看文字的被试都各有一半分别接受了高、低购买可行性的操控，如果有叠加效应的话，应该可以抵消。具体来说，图片组和文字组结果中购买可行性带来的变异应该是相等的，对 *t* 检验结果应该没有影响。

---

## 第四轮

### 审稿人3意见:

经过上一轮的修改，本文的理论逻辑和研究框架更加清晰。但对于文章中涉及的一些核心变量，我仍然存在以下疑问。

**意见1:** 关于评论形式：作者关注的是文字/图片这两种评论形式对产品评价的影响。然而对于文字VS 图片，两者除了解释水平的差异外，还存在其他许多维度的差异，比如唤起水平。相对于文字，彩色图片色彩鲜明，imagery vividness 程度更高，可以提升被试的唤起水平，也更加吸引被试的眼球。实验当中是否有测量唤起水平、感知的吸引程度？如果要排除这一 confounding factor，作者需要比较文字和黑白图片的评论，由于两种都是黑白像素，不涉及色彩和生动性方面的影响，会使得实验结果更加干净。

**回应:** 非常感谢评审专家深刻的剖析和见解。我们将从下面三个方面来回应意见1。

(1) 我们也主张图片比文字更能唤起人们的情绪反应，这在讨论的6.1部分对此有简略阐述：“图片是一种形象表征，更容易被人们视为审美的对象，会引发个体带有情绪和主观体验的审美反应”。但同时我们也认为：“情绪和体验都是依赖并随情境而变化的，且欢乐、忧伤等情绪体验可以拉近个体所感知的心理距离 (Van Boven, Kane, McGraw, & Dale, 2010)，因此情绪和作为诱因的图片都是与低解释水平心理表征相关联。”因此，如果imagery vividness对不同购买可行性条件下被试的影响有差异的话，一定程度上也会通过解释水平来表达这种差异。(2) 如果色彩和生动性对被试有显著影响的话，那么图片评论（或以图片为主的评论）都应该比文字评论（或以文字为主的评论）更有吸引力，但本文所有实验结果均表明评论形式（图片/文字）对产品评价并没有主效应，即图片并不总是比文字更有吸引力。因此，即使图片评论“可以提升被试的唤起水平，也更加吸引被试的眼球”，但在本研究中对被试的产品评价的影响与文字相比差异却是不显著的。(3) 在现实中，使用黑白图片作为评论是非常罕见的，如果以此作为实验材料，可能会降低研究的生态效度。

### 参考文献:

Van Boven, L., Kane, J., McGraw, A.P., & Dale, J. (2010). Feeling close: Emotional intensity reduces perceived psychological distance. *Journal of Personality and Social Psychology*, 98(6), 872–885.

**意见2:** 关于购买可行性 (feasibility): 根据作者的推论, 可行性跟低解释水平相关, 而图片相对文字来说又具有低解释水平的特点, 所以当购买可行性高时 (对应低解释水平), 图片形式的评论更有效。如果这一推论成立, 那么 desirability 也可以提出这样的假设, 即当购买愿望性低时 (对应低解释水平), 图片形式评论更有效, 但当购买愿望性高时 (对于高解释水平), 文字形式评论更有效。为什么本文只是关注了购买可行性这一变量? 对于任何一个效应, 都可能存在许多调节变量, 一篇文章当然很难涵盖所有的变量, 但在文中所选择的应该是最为重要和最具有解释力的变量。在这一情境中, 可行性和愿望性两者不相上下, 为什么仅关注可行性?

**回应:** 非常感谢评审专家深刻的洞察。我们在最初构思本研究时, 对于是否在研究中将愿望性 (desirability) 也作为一个独立变量进行研究也引发过一些讨论。本研究中之所以将可行性 (feasibility) 作为一个单独的变量来进行研究, 是作者因为在阅读文献时“无意中”发现可行性与心理距离存在交互作用: 如人们在为近期将来做出决策时, 会比远期将来对可行性高的选项偏好更强烈 (Eyal, Sagristano, Trope, & Liberman, 2009; Liberman, & Trope, 1998); 同等愿望性条件下, 个体自己决策比为他人决策更偏好可行性高的选择 (徐惊蛰, 谢晓非, 2011); 在高概率的情境里比在低概率情境里, 人们对不同可行性选项购买意愿的差别也更大 (Todorov, Goren, & Trope, 2007)。所以推论出如果将可行性作为一个独立的变量, 其高低水平上的差异达到一定程度时, 也是跟不同程度的解释水平相联系的。而在很多研究中愿望性 (desirability) 与心理距离则没有发现交互作用: 如选项价值和决策对象 (社会距离) 对选项吸引力的主效应不显著 (徐惊蛰, 谢晓非, 2011); 选项价值和发生概率对购买意愿的交互作用不显著 (Todorov, Goren, & Trope, 2007)。所以, 出于谨慎起见, 我们没有做出“当购买愿望性低时, 对应低解释水平, 图片形式评论更有效; 当购买愿望性高时, 对应高解释水平, 文字形式评论更有效”这一推论。

另一方面, 从生活经验来讲, 愿望性 (desirability) 对消费者的影响可能远比可行性 (feasibility) 大, 例如, 在一个物美价廉的餐厅 (高愿望性) 跟一个味道一般、价格偏贵的餐厅 (低愿望性) 中进行二选一, 如果两个餐厅的购买可行性存在差异的话, 人们可能会因为贪图便利 (高可行性) 而选择后者 (低愿望性); 但如果二者在可行性上不存在差异, 其差异只在于呈现的形式不同的话, 人们可能不会因为看的图片就选择味道一般、价格偏贵的餐厅。或许, 这也是以往一些研究中 desirability 跟心理距离没有交互作用的一个重要原因吧。

#### 参考文献:

Eyal, T., Sagristano, M. D., Trope, Y., & Liberman, N. (2000). When values matter: Expressing values in behavioral intentions for the near vs. distant future. *Journal of Experimental Social Psychology, 45*, 35–43.

Liberman, N., & Trope, Y. (1998). The role of feasibility and desirability considerations in near and distant future decisions: A test of temporal construal theory. *Journal of Personality and Social Psychology, 75*, 5–18.

Todorov, A., Goren, A., & Trope, Y. (2007). Probability as a psychological distance: Construal and preference. *Journal of Experimental Social Psychology, 43*, 473–482.

徐惊蛰, 谢晓非. (2011). 解释水平视角下的自己—他人决策差异. *心理学报, 43*(1), 11–20.

**意见3:** 解释水平作为一个庞大的理论体系, 除了影响可行性 VS 愿望性之外, 还会影响其

他许多方面，比如对于secondary feature versus primary feature 的关注。以往研究已经发现高解释水平的消费者更偏好primary feature，决策时更容易受primary feature 影响，而低解释水平的消费者则更容易受到secondary feature 影响。这一理论也可以用来解释本文的研究发现。图片和文字形式的评论在传递这两种features 方面也存在不同。

回应：非常感谢评审专家中肯的意见。其实我们与贵专家的意见是不谋而合的。在文献综述部分（“2.1 图片和文字在解释水平上的差异”部分第二段）有这样的陈述：“模拟真实对象的图片会将对象细枝末节的特征都呈现出来，让接收者很难分清主次；而文字却更能抓住对象的核心特征和中心思想，省略了琐碎的细节（Bar, 2004）。即相比图片，评论者对产品的文字说明可以更清晰地让接收者了解其想表达中心思想，因此文字比图片更核心和去背景化”，从这段陈述可知，我们是主张图片和文字分别跟secondary feature、primary feature相联系的。不过我们把这一观点作为理论基础和假设提出的依据，放在了文献综述部分，而没有用它来对结果进行解释。

参考文献：

Bar, M. (2004). Visual objects in context. *Nature Reviews Neuroscience*, 5, 617–629.

意见4：实验中操纵了文字VS 图片评论形式，那么评论的效价呢？所有评论都是正面的，还是中性的？如果是负面评论，文字VS 图片评论形式还有这样的结果吗？

回应：非常感谢评审专家深刻的见解。本研究所有实验的评论内容都是正面的，在“3.1 实验材料与预实验”部分对此有说明（为了防止评论正负性对结果的干扰，所有评论均为正面评论）。之所以这样做，首先，是考虑到人们对负面信息有“偏好”，往往会对负面信息投入更多的关注（Ahluwallia, Burnkrant, & Unnava, 2000；金立印，2007）。在现实生活中，一条负面评论的影响力往往比十条正面评论的影响力还大，在实际案例中也可以发现，许多商家的好评率都是百分之九十几，但左右消费者是否购买产品的却常常是数量上仅占百分之几的负面评论。由于这种作用效果的不对等性，在进行实验控制时存在一些困难，因此我们暂时未将其纳入本研究中。其次，用图片形式来呈现负面评论的时候有一些困难。我们在前测中进行过这样的尝试，发现寻找负面的图片评论是非常困难的，且图片形式的负面评论和文字形式的负面评论很难让被试感知到内容一致。最后，因为本研究主要是想考察评论呈现形式和购买可行性两个变量对产品评价的影响，而评论正负性不是本研究要考察的主要变量，所以将评论的效价均控制为正面评论。

但这确实是本研究的一个不足之处，我们在“6.5 研究局限与未来的研究方向”部分也对此进行了说明（“本研究只选取了正面的图文评论为实验材料，并未探究评论的正负性是否会影响研究结果。但根据以往的研究，评论的正负性是影响评论效果的重要因素（Ahluwallia, Burnkrant, & Unnava, 2000），因此未来的研究可以探究正负性评论是否会影响评论形式、购买可行性间的解释水平一致性效应”），并准备在以后的研究中进一步考虑这一因素。

参考文献：

Ahluwallia, R., Burnkrant, R. E., & Unnava, H. R. (2000). Consumer response to negative publicity:

The moderating role of commitment. *Journal of Marketing Research*, 37(2), 203–214.

金立印. (2008). 消费者对网络评论信息的可信性评价—网站主导者、专业性及产品类别的影响. *营销科学学报*, 4(4), 23-44.

**意见5:** 本文并没有对中介效应进行检验? 提出的假设背后的机制和原理还有待进一步的实证支持。

**回应:** 非常感谢评审专家的建议和指导。遵照评审专家的意见, 我们尝试对实验结果进行了中介效应分析。

首先, 根据本研究的理论假设, 在可行性高和低的情境下, 解释水平中介了评论形式对产品评价的影响, 因此将每个实验总的的数据根据购买可行性的高低分为两组, 然后采用 Baron 和 Kenny(1986)的中介分析程序分别分析两组数据的中介效应。其中评论形式为预测变量, 分析中处理为虚拟变量, 解释水平为假设的中介变量, 被试的产品评价为结果变量。具体步骤为: 第一步, 采用预测变量预测结果变量, 进行回归分析; 第二步, 采用预测变量预测中介变量。第三步, 同时采用预测变量和中介变量预测结果变量, 进行多元回归分析。分析结果见表 1。

**表 1 解释水平的中介效应检验**

<b>实验 1a</b>						
		因变量	自变量	B	SE B	$\beta$
购买可行性低	第一步	产品评价	评论形式	-1.14***	0.20	-0.53
	第二步	解释水平	评论形式	1.03***	0.25	0.42
	第三步	产品评价	评论形式	-0.46**	0.13	-0.22
			解释水平	-0.66***	0.05	-0.76
中介效应占总效应比例: 59.63%						
购买可行性高	第一步	产品评价	评论形式	1.40***	0.22	0.56
	第二步	解释水平	评论形式	1.50***	0.28	0.50
	第三步	产品评价	评论形式	0.44**	0.15	0.18
			解释水平	0.64***	0.05	0.76
中介效应占总效应比例: 68.57%						
<b>实验 1b</b>						
		因变量	自变量	B	SE B	$\beta$
购买可行性低	第一步	产品评价	评论形式	-1.10***	0.21	-0.54
	第二步	解释水平	评论形式	1.03***	0.23	0.49
	第三步	产品评价	评论形式	-0.30*	0.13	-0.15
			解释水平	-0.77***	0.06	-0.81
中介效应占总效应比例: 72.10%						
购买可行性高	第一步	产品评价	评论形式	1.09***	0.18	0.58
	第二步	解释水平	评论形式	0.88*	0.40	0.26
	第三步	产品评价	评论形式	0.75***	0.11	0.40
			解释水平	0.39***	0.03	0.71
中介效应占总效应比例: 31.48%						
<b>实验 2</b>						
		因变量	自变量	B	SE B	$\beta$

购买可行性低	第一步	产品评价	评论形式	-0.67**	0.25	-0.33	
	第二步	解释水平	评论形式	1.09***	0.29	0.44	
	第三步	产品评价	评论形式	-0.12	0.23	-0.06	
			解释水平	-0.51***	0.09	-0.61	
中介效应占总效应比例：82.97%							
购买可行性高	第一步	产品评价	评论形式	0.75*	0.32	0.29	
	第二步	解释水平	评论形式	2.06***	0.45	0.50	
	第三步	产品评价	评论形式	0.06	0.32	0.02	
			解释水平	0.34***	0.08	0.54	
中介效应占总效应比例：93.39%							
<b>实验 3</b>							
		因变量	自变量	B	SE B	$\beta$	
怀旧产品	购买可行性低	第一步	产品评价	评论形式	-1.49***	0.22	-0.68
		第二步	解释水平	评论形式	1.20***	0.21	0.61
		第三步	产品评价	评论形式	-0.38**	0.12	-0.17
				解释水平	-0.93***	0.06	-0.83
中介效应占总效应比例：74.90%							
非怀旧产品	购买可行性高	第一步	产品评价	评论形式	0.86**	0.29	0.37
		第二步	解释水平	评论形式	2.10***	0.36	0.60
		第三步	产品评价	评论形式	0.03	0.31	0.01
				解释水平	0.40***	0.09	0.59
中介效应占总效应比例：97.67%							

从表 1 可以看出，实验 1a、实验 1b 的结果表现为部分中介效应，实验 2 的结果表现为完全中介效应，实验 3 的非怀旧产品情境下得到了完全中介效应，怀旧产品情境下得到了部分中介效应。

根据大多数研究的惯例，我们对完全中介效应和部分中介效应进行了区分，但对于部分中介效应，有以下两点情况需要说明。（1）完全中介和部分中介概念是有问题的。当总效应小，样本也小的时候容易得到完全中介效应，即使间接效应占总效应的比重可能不到七成。而且完全中介效应排除了将来探索其他中介的可能性，因此应当放弃完全中介的概念，将所有中介都看作是部分中介（Preacher & Hayes, 2008；温忠麟，叶宝娟，2014）。（2）本研究中大多数结果为部分中介效应实验的总效应都相对较大，因此在加入中介变量后即使回归系数变化很大，也仍难以降低到不显著的水平，但中介效应占总效应的比重都较高（除实验 1b 的购买可行性高情境下外）。由此可见，本研究中解释水平确实起着中介变量的作用。因此，我们认为本研究的结果是支持理论假设的，但在今后的研究中确实需要对理论假设所涉及的机制和原理进行更深入的研究。

另外，由于编委专家和编辑部多次向作者提出压缩篇幅的要求，如编委专家曾提到“目前论文文字接近 1 万 9 千字，建议大幅度压缩，压缩到 1 万 2 千字左右”。虽经过压缩，目

前论文的字数仍超过 1 万 5 千字，超过了编委的要求。考虑到篇幅的要求，同时也考虑到文章考察的核心问题为“图片、文字评论在不同购买可行性情境下对消费者产品评价的影响差异，及产品类型（怀旧/非怀旧）在其中的调节作用”，我们没有在论文中报告中介效应检验的结果。

再次感谢专家所提出的宝贵意见！

#### 参考文献：

- 温忠麟, 叶宝娟. (2014). 中介效应分析: 方法和模型发展. *心理科学进展*, 22(5), 731–745.
- Baron, R. M., & Kenny, D. A. (1986). The moderator-mediator variable distinction in social psychological research: Conceptual, strategic, and statistical considerations. *Journal of Personality and Social Psychology*, 51, 1173–1182.
- Preacher, K. J., & Hayes, A. F. (2008). Asymptotic and resampling strategies for assessing and comparing indirect effects in multiple mediator models. *Behavior Research Methods*, 40, 879–891.

---

## 第五轮

#### 审稿人意见：

经过几轮的修改，本篇论文的逻辑更加严谨。许多问题都已在之前几轮的修改中得到了较好的解释，从中可以看出作者在整个文章评审过程中付出的努力。希望以下几点建议能对文章的改进有所帮助。

**意见 1: 解释水平的测量:** 各个实验都是在最后才测量解释水平。以实验 1 为例，在此之前，被试已经接受了购买可行性和图片 VS 文字评论形式的操纵，因此很难说最后分类任务中表现的差异是来源于哪个变量。关于解释水平的结果也写的不够清晰，是否存在购买可行性和评论形式的交互作用？还是只有两个变量的简单主效应？更合理的方法是用独立的前测来单独验证购买可行性与解释水平的对应，以及评论形式与解释水平的对应。

**回应:** 谢谢评审专家深刻的剖析和意见。我们已经将论文中解释水平测量部分的  $t$  检验改为了方差分析，结果表明所有实验中评论形式和购买可行性对解释水平的主效应均显著，二者交互作用均不显著（论文中 3.2.3、3.3.3、4.4、5.4 橙色字体部分）。

此外，为了稳妥起见，我们还遵照评审专家的意见进行了两个补充测验，分别单独验证了购买可行性与解释水平的对应关系，以及评论形式与解释水平的对应关系。由于实验 1b 设置的情境更具有代表性和推广性（贴近现实网络消费情境，且实验 3 采用了同样的可行性情境设置），因此参照实验 1b 的实验设置进行了补充测验。

(1) 补充测验 1: 非参加正式实验的 49 名大学生参加了验证购买可行性与解释水平对应关系的补充测验。首先，用与实验 1b 一致的方式对购买可行性进行操控（要求被试设想以下情景：自己正乘坐地铁去某步行街，并准备在那里就餐。在该步行街有一间餐厅，它环境卫生、价格公道、食物美味。告诉购买可行性高组的被试“餐厅就在地铁站出口旁边，出站就能看到”；告诉购买可行性低组的被试“餐厅位于某大厦 A 座 307 号店铺，离地铁站出口约需 10 多分钟步行路程，其间要转几个路口”），随后要求被试对餐厅进行评价，并进行了购买可行性的操控检验，及用分类任务测量了被试的解释水平（以上测量方式均与实验 1a 一致）。结果表明，接受“购买可行性高”操控的被试评估自己去餐厅的方便程度( $M=5.58$ ,  $SD=1.47$ )显著高于“购买可行性低”的被试( $M=2.76$ ,  $SD=0.88$ ),  $t(47)=8.19$ ,  $p<0.001$ ,  $d=2.33$ ,



说明对购买可行性的操控有效。另外，购买可行性低条件下的被试划分的类别数( $M = 3.88$ ,  $SD = 0.93$ )显著少于购买可行性高的被试( $M = 6.58$ ,  $SD = 1.72$ ),  $t(47)=6.90$ ,  $p < 0.001$ ,  $d = 1.95$ , 表明相比购买可行性高的被试, 购买可行性低的被试与更高解释水平的心理表征相关联。证实了低、高购买可行性与高、低解释水平间的对应关系。

(2) 补充测验 2: 非参加正式实验的 56 名大学生参加了验证评论形式与解释水平对应关系的补充测验。首先, 用与实验 1b 一致的方式对评论形式进行操控(请被试设想有这样一间餐厅, 它环境卫生、价格公道、食物美味。然后通过网页向被试呈现餐厅的评论, 每一页只呈现一条评论, 点击“下一条”按钮观看下一条评论。网页界面的设计效仿国内某知名点评网站, 但为了控制其他因素的干扰, 网站上所有评论者均为匿名, 餐厅名称以“某某店”代替, 网站名称等其余信息均做空白处理。“文字”组的被试看的 5 条评论中包含 4 条文字评论和 1 张图片评论(以文字评论为主); “图片”组的被试看的 5 条评论中包含 4 张图片评论和 1 条文字评论(以图片评论为主)。随后要求被试对餐厅进行评价, 并用分类任务测量了被试的解释水平(以上测量方式均与实验 1a 一致)。结果表明, 接收文字评论的被试划分的类别数( $M = 4.12$ ,  $SD = 1.45$ )显著少于图片评论的被试( $M = 7.03$ ,  $SD = 1.61$ ),  $t(54)=7.09$ ,  $p < 0.001$ ,  $d = 1.90$ , 表明相比图片评论, 文字评论是与更高解释水平的心理表征相关联的。证实了文字、图片形式评论分别与高、低解释水平间的对应关系。

总体而言, 两个补充测验均验证了低购买可行性、文字评论在高解释水平上匹配, 而高购买可行性、图片评论则在低解释水平上匹配。此外, 由于补充测验的结果跟正式实验方差分析的结果一致, 以及考虑到编委专家和编辑部压缩篇幅的要求, 我们没有将补充测验加入正文中。

再次感谢专家所提出的宝贵意见!

**意见 2: 中介检验:** 虽然基于篇幅所限无法在文中报告逐步回归的中介检验法, 但建议作者可以考虑用目前更为新兴的 **bootstrapping** 方式对中介效用进行分析, 并且用此方法的结果汇报内容较为简洁, 不会对篇幅控制有影响。当然, 本文重心并不是中介效用, 不放入中介检验部分的内容也是可以理解的。

**回应:** 感谢评审专家的建议。在上一轮修改中, 我们尝试过多种统计方法, 也尝试了 **bootstrapping** 方式对中介效用进行分析, 但结果与逐步回归法检验的结果差异不大, 因此在回复中只报告了最保守的逐步回归法的结果。但为了突出文中的重点和缩减篇幅, 我们打算在正文中加入中介效应检验结果, 在此感谢专家的理解!

**意见 3: 服务/产品类别:** 文章所使用的情境都是关于体验型服务, 其结论是否同样适用于信用型服务 (**credence service**) 呢?

**回应:** 感谢评审专家深入的见解。这对本研究结果的现实推广来说是一个非常具有意义的挑战, 我们尝试从以下几个方面进行回答。

首先, 由于信用型服务产品在使用后也很难对产品进行评价 (Darby & Karni, 1973; Harrison-Walker, 2001), 因此消费者根据自己体验所做的评论对他人的参考价值可能不如体验型服务产品。而且在现实消费情境中, 相对体验型产品而言, 搜索型产品和信用型服务产品消费者使用后发表的网络评论, 以及评论的网络平台也都相对较少。但人们可能会更看重专家或权威的评论 (Galetzka, Verhoeven, & Pruyn, 2006), 就这点来说, 如果将本研究的结果运用于信用型服务产品的话, 可能需要将消费者评论换成专家评论才会更有效。

其次，信用型服务产品的核心特征一般很难用图片来描述，例如医疗、教育、电气维修等服务，人们很难用图片来展示医生的医疗水平，教师的教学水平，以及维修工人的技术水平，因此用图片来对信用型服务产品的核心特征进行评论难度较高。

第三，相比体验型产品，现实中提供同类信用型服务产品的商家相对较少，服务水平和价格等因素也相近的商家就更少了，因此购买可行性对信用型服务产品的影响可能较小，甚至可能被服务的水平或价格等其他因素的影响所湮没。

尽管如此，我们仍认为解释水平理论所揭示的是一种基本的认知规律，它的影响具有很大的普适性，解释水平的匹配效应对信用型服务产品依然适用，但通过解释水平来影响产品评价的因素可能不是本研究所探讨的这些因素。综上所述，我们认为本研究所揭示的解释水平匹配效应对信用型服务产品同样适用，但适用的情境，起作用的变量可能跟本研究不同，这还有待进一步的探究。

#### 参考文献：

Darby, M. R., & Karni, E. (1973). Free competition and the optimal amount of fraud. *Journal of Law and Economics*, 16, 67-86.

Galetzka, M., Verhoeven, J. W., & Pruyn, A. T. H. (2006). Service validity and service reliability of search, experience and credence services: A scenario study. *International Journal of Service Industry Management*, 17(3), 271-283.

Harrison-Walker, L. J. (2001). The measurement of word-of-mouth communication and an investigation of service quality and customer commitment as potential antecedents. *Journal of service research*, 4(1), 60-75.

意见 4：图片上的标识不太清晰，白色和灰色柱状图各代表什么需要清楚表明。另外，图片的纵轴和横轴也最好有清晰的标识，方便读者了解相关信息。

回应：非常感谢评审专家的细心检查，非常抱歉！这应该是 WORD 软件的一个 bug 造成的，导致所有文本框中的文字消失，现已纠正了该错误。再次对我们的疏忽致歉！

---

## 第六轮

#### 编委专家意见：

意见 1：英文摘要太长；

回应：谢谢编委专家的意见。遵编委意见，修改稿中已将英文摘要缩减为 535 字。

意见 2：文字可以进一步精炼，建议压缩到 1 万 2 千字以内。

回应：谢谢专家的建议。遵编委意见，经过多轮修改，现已将修改稿中论文正文部分缩减到 12078 字。由于内容较多，如果再减少字数，恐怕会影响论文的完整性。未能完全达到要求，请编委专家见谅！

意见 3：个别地方仍有文字错误或语意含混处，建议认真核对。

回应：谢谢编委专家细心的检查。经仔细核查，已改正多处错误。此外，因缩减字数、修改文字地方较多，此次没有用颜色标记改动之处，请编委专家见谅！

主编终审意见：可以发稿。