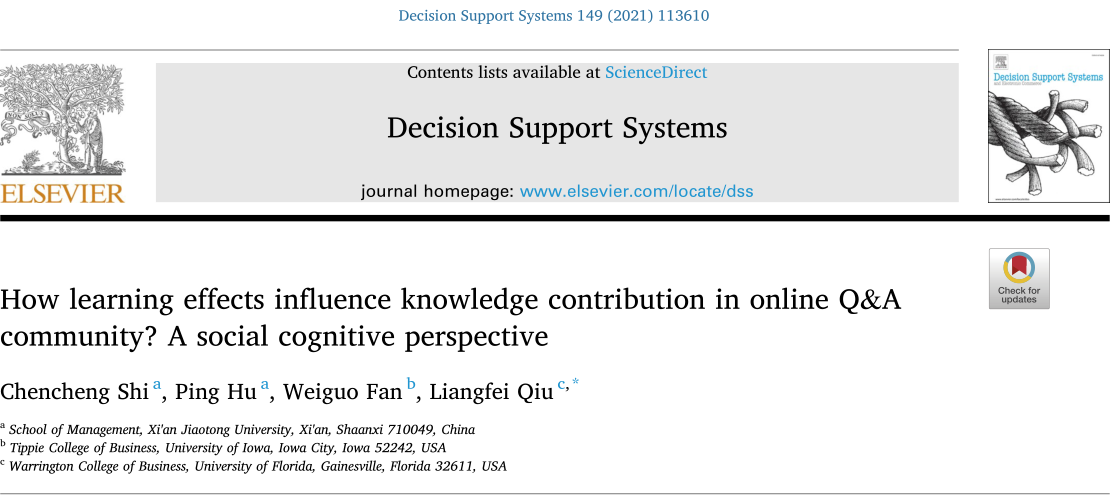
**学习效应如何影响在线问答社区的知识贡献？社会认知角度**



**摘要**：在线问答社区在解决个性化开放式问题方面越来越受欢迎，信息贡献对于在线问答社区的健康发展至关重要。然而，过去的贡献行为和收到的相应社区反馈是否会影响后续的贡献特征尚不清楚。本文运用社会认知理论，考察学习对用户知识贡献行为的影响。具体来说，我们关注两种类型的学习效果：从个人过去的贡献经验中获得的主动学习和从观察他人在问题线索中的表现中获得的替代学习。利用从中国最大的在线问答社区之一收集的数据集，我们发现，获得高度积极反馈的过去用户贡献的长度特征，无论是通过主动学习还是替代学习，都会影响社区中后续贡献的信息性。这些学习效果对于社会地位较高的用户更有效。社会地位高的贡献者的主动学习效果更强。对于高地位贡献者的替代学习，高投票长答案的影响更强，而高投票短答案的影响更弱。本研究对网络知识社区的知识贡献行为有更深入的了解，为建立健康的知识贡献环境提供指导。

**1 引言**

在线问答社区发展迅速，弥补了网络搜索引擎获取个性化信息和知识的不足[1-3]。例如，Quora，最受欢迎的在线问答社区之一，在2018.1有3亿月活跃用户。通过这些在线平台，人们可以提出担忧，分享个人经验，或讨论他人的答案，为整个社区贡献知识。这类社区的可持续性在很大程度上取决于其参与者的高质量和信息量贡献，即用户生成内容(UGC)[4,5]。因此，研究在线问答社区中用户的知识贡献行为一直是一个重要的研究课题。

通过在线社区的知识贡献，虽然没有直接的经济利益，但贡献者可以通过过去的自愿贡献获得积极的声誉、同行的认可和社会地位[6-8]。以往对网络知识贡献的研究表明，获取社会利益是网络社交媒体提供公共物品的主要动机[9,10]。许多在线问答社区，如Quora和知乎，都有“关注”功能的社交机制和“投票”按钮这样的评价机制，每个答案都由平台上的所有用户进行评价。获得较高认可的过去的贡献将吸引其他用户的兴趣，并可能吸引他们跟随贡献者。一个追随者可以收到关于贡献者在社区的最新参与的信息。通过这种方式，贡献者变得更有影响力，享有更高的社会地位。因此，社会对每个贡献的认可的社区反馈可以被视为行为结果，贡献者可以通过这种行为结果了解他们过去贡献的绩效。根据社会认知理论(SCT)[11]中的学习范式，学习要么通过实际做的结果(主动学习)发生，要么通过观察他人的表现(替代学习)发生[12,13]。后一种效应在经济学和信息系统(is)文献中也被称为“观察学习”[14-17]。同样，过去贡献的反馈(来自过去的经验和观察他人)可以帮助用户推断社区的偏好，并执行更多高质量的贡献。

在本文中，我们关注的是信息性维度(informativeness)，即答题卡所包含的信息量[18-20]。以往的研究表明，在线投稿的信息量与其质量或帮助性呈正相关[18,20-22]，一般反映了投稿者在提供方面的努力或勤奋[23,24]。不同于需要其他用户或专家判断的答案质量，每个答案的内容完全由贡献者决定。正是每个贡献者自愿分享的知识构成了这些知识型网站的主要内容。因此，激发信息贡献对于在线知识社区的可持续发展至关重要。

本研究探索了一种新的学习效应线索，可能会影响在线问答社区的信息性贡献。以往关于在线问答社区知识贡献的研究主要集中在用户动机和持续贡献上[3,9,25 - 29]。过去贡献的影响仅限于用户今后是否提供贡献或可能提供多少贡献。虽然最近的研究关注了[30]在线问答健康论坛中同一问题线索中贡献质量的顺序效应，但过去贡献的反馈是否会影响未来知识贡献的内容特征(又称学习效应)尚不明确。因此,我们问:

1. 从过去贡献的反馈中获得的学习效果(包括主动学习和替代学习)是否会影响未来贡献的信息性？

2. 前一个问题中提出的学习效应对不同社会地位的贡献者有不同的影响吗？

SCT认为，那些带来积极结果的行为被视为建模行为，人们可以从中获得象征表示来指导他们未来的行为[11,12]。因此，为了回答研究问题，我们考虑了过去贡献的长度特征和反馈，并提出了一个研究框架，该框架考虑建模的贡献行为源于两个方面：在同一问题线索中，一个人的历史贡献经历和其他人之前的贡献。我们也探讨了相应的贡献者的社会地位的调节效应。此外，我们还特别注意了建模长答案和建模短答案的显著效果。最后，我们讨论了我们的研究如何对现有的网络知识贡献和社会认知理论的文献作出贡献，并为网站管理者提供了几点设计指导。

**2 理论背景**

**2.1 在线问答社区的知识贡献**

知识贡献问题一直被认为是网络社区研究的一个基本问题[9,28,29,31]。研究人员运用社会资本理论[9,32]、自我决定理论[2,29]、社会交换理论[26,31,32]、社会认知理论[33,34]等理论，揭示了影响用户贡献的个体差异和社会机制。他们发现，知识贡献行为一般受贡献时的个人动机状态或社会互动条件的影响。

然而，正如SCT所指出的，人类的许多学习都发生在社会环境中[11,12]。在在线问答社区中，知识贡献者不会孤立地或一次性地提供知识。过去的贡献的反馈清楚地显示在个人主页或问题线程页。一个受欢迎的贡献可以给贡献者带来更多的社会认可(通过其他人的投票)，并增加个人在社区中的社会地位(通过获得关注者的数量)。通过这些机制，在选择需要回答的问题时，贡献者不必太关注知识搜寻者的地位或能力，由于之前的研究[9,32]提出的互惠收益的期望，但重点是寻找一个有趣的问题和知识提供本身。这种社会认同的设计为知识贡献者提供了方便学习的机会，用户对贡献的知识和信念可能会根据对在线社区中先前贡献行为及其结果的观察或体验而逐渐演变。即通过实际行动的主动学习或通过观察他人的表现的替代学习，在线问答社区的用户可以了解哪些类型的贡献更有可能被平台上的其他用户认可。这可能会导致他们未来贡献行为的改变，即在线社区的学习效应。据我们所知，网络问答社区中知识贡献行为背后的学习效应尚未被研究。而且，许多关于知识贡献行为的研究都是从一定时期的质或量的角度进行的[2,29,33,35]，而不是直接观察贡献行为的特定内容特征。为了解决这些重要的研究差距，我们应用社会认知理论来探讨个体网络知识贡献行为背后的学习效应。

在本研究中，我们分析了知识贡献的信息量，即答案帖子中所包含的信息量informativeness[18-20]。这一概念是研究在线UGC社区贡献的一个重要特征，也可以用其他类似的词来表达，如信息丰富度information richness [21,636]、全全性comprehensiveness [37]和信息深度information depth[21,38]。在以往的研究中，在线贡献的信息量，如在线评论和答案，通常被视为自变量，与内容质量和帮助性的判断或用户的进一步决策(如消费者的购买决策)有显著相关[18,22,38][19,36]。实际上，这些后续判断或购买行为在一定程度上是贡献者在做出贡献时无法控制的，很大程度上取决于社区内其他用户的偏好。只有内容本身可以完全由贡献者控制。由于我们研究的是具体的知识贡献行为，所以我们将答案的文本内容作为主要研究对象。在在线问答社区的在线知识贡献特征中，以文本长度[18,38,39]等表征的信息量比可读性、客观性等文本特征更为直观。它还反映了贡献者的总体努力[23,24]。因此，我们选择信息性作为在线问答社区知识贡献的研究维度。

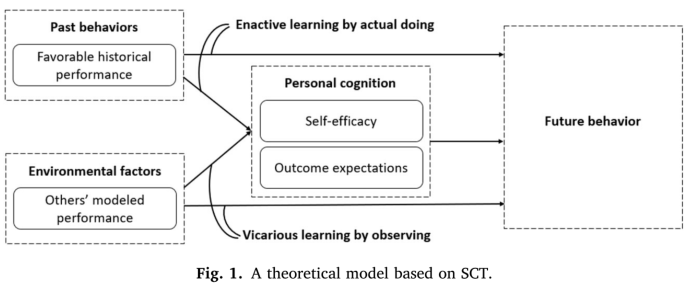
**2.2 社会认知理论和知识贡献**

社会认知理论(SCT)是由社会学习理论[40]发展而来的，它使用一个三元交互模型来解释人类行为，在这个模型中，环境、个人变量(如认知)和行为相互作用[11]。个人对自我效能感和结果预期的认知是影响个人行为的主要决定因素。自我效能指的是“相信一个人有能力组织和执行管理预期情况所需的行动方针”[41]，而结果预期指的是对某些行为产生的结果的预期[42]。环境是指影响个体行为的社会或物理环境。IS研究人员虽然承认SCT的丰富，但在使用SCT讨论利益的影响时，通常侧重于模型内部的单向关系[28,44]。SCT将学习视为一种信息加工活动，其中有关行为和环境的信息可以作为符号表征，为未来行为提供指导[11]。SCT概述了两种学习方法:(1)通过实际行动的结果进行主动学习，(2)通过观察他人的表现进行替代学习[12,13]。在在线问答社区中，用户之间的关系主要是获取有用信息[45]的弱关系，获得社会认可是贡献知识的重要动机[9,27]。因此，贡献者可以通过社区的反馈来了解不同贡献行为的适宜性，并可能进一步调整自己未来的贡献行为，以迎合社区的偏好。

建模是SCT中的一个通用术语，表示学习者在观察一个或多个模型[11]时所发生的心理匹配过程。这个过程为观察者提供了在观察者的环境中进行行为的信息，并帮助他们在头脑中为行为的展示编码规则。人们也可以把自己当成自己的榜样，观察自己的表现，就像“自我塑造”[12]所暗示的那样。Bandura(1986)认为，模型化的行为对观察者的效用要么是通过实际制定行为并观察其有用性，要么是通过间接检查他人行为的结果并获得功能价值来实现。SCT表明，由于预期积极结果[11]，类似于奖励模型行为的行为通常会更频繁地出现。如上所述，这种建模过程也发生在在线问答社区中，我们将获得更高级别认可的贡献(即选票数量)视为建模答案。我们将探索这些答案如何作为社区知识贡献的先决条件。

以往将SCT应用于在线问答社区的研究密切关注认知力量(即自我效能感)和结果预期对后续行为的影响[26,28,33]。然而，在线问答社区中由社区响应[11]促进的先行学习过程仍未得到充分探索，尽管该学习过程已在信息系统研究的其他领域得到检验，如IT培训[13,46,47]和电子学习[44]。因此，我们将社区反馈视为行为结果，作为一种激励(SCT中提出的社会强化)，吸引贡献者执行他们从过去经验中学到的东西(主动学习)或通过观察他人(替代学习)。

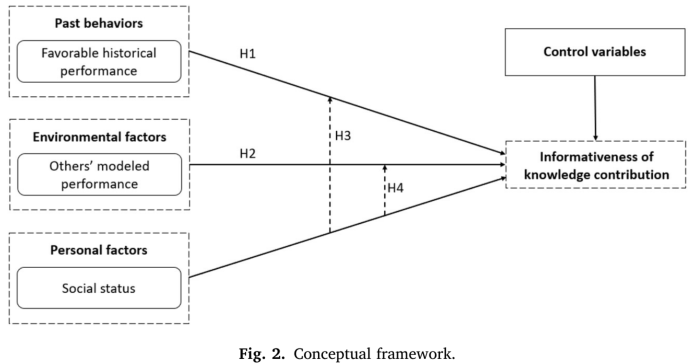
另一方面，在线问答社区的贡献者不一定会根据他们从社会学习中学到的东西来表现。现有行为的结果表明行为的适当性或准确性，并作为动机和信息的来源[12]。SCT表明学习和表现是两个截然不同的过程，并认为一个人是否会完成他所学的内容取决于环境变量，如社会压力和个人影响，如动机和感知需求[12]。这符合SCT在决定个人行为时的三元互反关系。尽管如此，先前与计算机自我效能相关的IS研究也采用了SCT，并揭示了先前表现的后果要么直接影响未来的绩效，要么首先影响一个人对自我效能和结果预期的认知，然后影响未来的绩效[47,48]。在这些研究中，先前表现的结果既包括个人历史行为，也包括通过观察他人表现(环境影响)产生的结果。这些关系如图1所示。



因此，在线问答社区中先前贡献的反馈可能有助于影响未来贡献的环境和个人因素。在SCT中，我们采用了主动和替代学习范式，考察了先前行为对未来知识贡献行为的信息性的影响。

**3 研究模型与假设**

图2为本研究的概念模型。首先，我们探讨了主动学习和替代学习的主要效果。这两种学习效果都源于以回答长度为特征的模式化行为。然后，我们分别研究了贡献者的社会地位与主效应之间的调节效应。此外，随着一个人的社会地位的提高，我们比较了建模(高投票)和未建模(低投票)答案的效果。我们还深入挖掘了先前贡献的长度特征，并特别关注先前贡献中建模的长答案和建模的短答案。



**3.1 历史贡献的影响**

主动学习指的是通过直接经验进行的学习，在这种学习中，人们通过观察他们的反应结果来形成行为的概念。在在线问答社区中，投票机制的功能使贡献者能够评估社区成员对其过去贡献的社会接受度和可信度。这些越来越多的“投票”或“点赞”可以内化为积极的心理结果，从而提高用户对后续贡献行为的自我效能感[11,29]。SCT假设，成功执行一个行动可以启动高自我效能感，而失败则降低它[11]。此外，具有良好结果的过去贡献也会增加用户执行类似行为的结果期望[11,48]。在IS研究中的一些研究也表明，个人可以根据过去的经验和相关的评价或反馈来调整他们未来的行为[3,49]。过去的贡献获得高的社区评价分数可以加强和作为一个模型行为影响未来的贡献。几项研究发现，长答案通常比短答案更具信息量[21,22,37]。由于我们使用文本长度特征来描述在线问答社区中知识贡献的信息性，因此，我们预计，一个人过去在在线问答社区中建模贡献的长度特征将影响该贡献者未来贡献的信息性。我们因此提出:

H1. 用户贡献的信息性与贡献者过去贡献经验中建模答案的长度呈正相关。

**3.2 其他人在帖子中的贡献的影响**

除了主动学习，学习也可以间接地发生，而不需要学习者公开地执行行为[11,12]。替代学习(例如，通过电子媒体或观察他人的行为来学习)加速了学习，同时消除了人们通过体验可能的负面结果来学习的需求[11-13]。这个概念类似于文献[14,15,50]中的观察学习。根据这些理论，社区中的人们可以通过观察他人的行为和相应的结果来推断社区偏好，推测哪种行为更优越。SCT还表明，观察他人相似的成功可能会增强观察者的自我效能感和自己也能做好的自信 [11,12]。因此，在在线知识贡献的背景下，我们可以预期，在在线问答社区中获得高社区认可的答案可能会被其他贡献者模仿，从而获得积极的结果。

与主动学习部分相似，我们也关注信息性维度。在我们的研究中，为了控制问题的异质性，我们将替代学习的范围限制在先前存在于同一问题线索中的贡献。因此，潜在的贡献者在对一个问题作出贡献之前，可以提前观察现有的答案及其结果，推测社区中可能的读者的偏好。如果用户观察到冗长的答案已经被识别出来，他/她可以了解到这个问题的许多读者更喜欢信息性的答案。为了获得积极的结果，随后的贡献者可能更愿意详细地表达他们的想法。因此，我们提出:

H2. 后续贡献的信息量与已经在同一问题线索中建模的答案的长度呈正相关。

**3.3 社会地位对历史贡献影响的调节作用**

如引言所述，在线问答社区“投票”的评价机制和“关注”功能的社会机制，导致了参与者的社会地位的不同层次。过去流行的贡献可以给贡献者带来许多社会认可和社会效益。研究表明，过去的成功经验增加了用户的自我效能和对该行为结果的预期，进而影响未来的贡献[9,11,26,29,47,48]。在线问答社区中地位越高的用户，过去的贡献经历越好，由于过去的成功经验，可能对自己的知识贡献特征具有更高的自我效能感和期望水平。因此，我们认为社会地位越高的用户会更加看重自己过去的成功贡献。因此，我们提出:

H3. 社会地位越高的贡献者，其过去贡献经历中的建模答案对知识贡献的信息量影响越强。

**3.4 社会地位对他人在帖子中的贡献影响的调节作用**

根据SCT[11,47,48]，用户过去的成功贡献增加了他们对未来贡献的期望。社会地位越高的用户得到的结果越好。也就是说，他们更有可能表现出自我调节的行为，因为他们过去表现出了受欢迎或专业知识。在线产品评论研究人员指出，地位较高的用户倾向于调节自己的行为，表现出的偏差行为较少[51]。在在线问答社区中，同一个问题线索的答案之间存在竞争，即使他们可能在询问个人经验或意见分享的问题上比较弱。一旦决定贡献，以后的贡献者，特别是那些社会地位较高的贡献者，可能不希望在吸引潜在读者的信息方面处于劣势，特别是与之前获得一定程度认可的贡献相比。我们因此有:

H4. 社会地位越高，问答线索模型答案对知识贡献信息量的影响越强。

**4 研究方法**

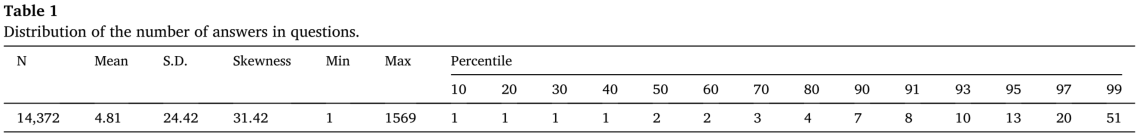
**4.1 研究设置和数据收集**

我们的研究数据来自知乎，这是中国最受欢迎的在线问答社区之一。截至2018年11月底，知乎正式宣布拥有超过2.2亿用户，3000万问题，1.3亿相关答案知乎注册的参与者可以搜索或提问有趣的问题，提供答案，投票或评论他人的答案，并在这个平台上关注或被其他用户关注。与其他在线知识社区不同，知乎回答问题没有时间限制。一旦用户在观察了现有的答案条件后有可能做出贡献，我们鼓励他们随时分享知识。此外，不需要每天提供考勤或答案来获得用户地位提升的分数奖励。在知乎中，所有的奖励或激励机制都只与自己的贡献和其他用户的贡献有关，比如他人的认可和粉丝的增长。这为我们提供了一个相对纯粹的研究环境，从学习的角度来研究知识的贡献。另外，与论坛不同的是，一个特定问题的每个答案都是按照时间顺序或按照一定的排名方法独立排列的。评论和答案是分开的，投票数清楚地显示在问题线索的每个答案下面，这是我们的研究设置中理想的探索替代学习。因此，我们选择知乎作为研究在线问答网站知识贡献可能产生的学习效果的目标体系。此外，知乎还出台规定，保证答案的原创性，并以一定的格式规范引文管理员会“折叠”可能存在剽窃的答案，并警告贡献者修改。

在线问答网站上的一些问题来源于当下的热点事件，可能在短时间内引起众多关注，但在趋势过去后就会保持沉默。这种与时间相关的外生冲动会极大地影响用户的贡献行为[52]。此外，问题主题的属性，如争议性敏感性和专业化，也可能对知识贡献产生影响[53,54]。为了保持话题的相对独立性并消除潜在的周期性影响，我们集中讨论属于一个特定主题的问题子集:“英语学习”。这个主题可能会经历一个相对稳定的增长率，受其他主题的影响较小，潜在地避免了意外的时间内生性问题。

并不是“英语学习”中的所有问题都与技术或学术问题有关。很多问题都是生活导向的，比如询问别人的意见或建议。知乎的用户主要是说中文的，在中国，从小学(根据中国教育部的规定，从2001年开始)到高中，几乎所有的学生都必须学习英语。根据知乎在2020年2月进行的一项网站调查，5518名受访者中有97%的人是1980年后出生的，有89%的人是1990.5后出生的。因此，我们认为知乎的大部分用户可能都有学习英语的经历。即使一个人的英语学习水平不高，他/她仍然可以在英语学习的主题下提供一些问题的答案。换句话说，“英语学习”并不是一个非常特别的话题，只能由有自己特色的用户来贡献。虽然这不是一个热门话题，但它鼓励社区中不断分享技能、意见和建议。

我们从知乎收集的数据包含了从2011年1月9日到2017年8月1日产生的37950个问题中的106692个答案。对于每一个问题和答案，我们收集所有的细节，如内容，发布时间，和反馈。对于每个贡献者，我们收集他们的个人信息，如人口统计、参与和反馈信息。我们首先删除没有答案的问题。由于在我们的研究中考虑了贡献者的个人特征，我们排除匿名回答和问题。通过对剩余的69124个答案，隶属于14372个问题的描述性统计(如表1所示)，我们观察到问题中的答案数量服从长尾分布，这与之前的研究[55]一致。只有一个答案的问题占45.3%，超过50个答案的问题占1%。对于答案太少的问题，这个问题可能不会引起太多的注意，或者它可能不是一个开放式的问题，没有明确的答案。对于答案太多的问题，后面的贡献者可能懒得在投稿前阅读之前的所有答案，这就破坏了替代学习效果。因此，我们的样本仅限于包含4至50个答案的问题。此外，如果答案是问题线索中的第一个答案或用户贡献历史中的第一个答案，以使替代和主动的学习有意义，也会被排除在外。为了控制个人历史贡献，对于数据集中的每个贡献者，我们收集他们历史答案的所有信息，包括内容、投票数和发布时间。经过这些处理后，我们最终得到了一个包含20152个答案的语料库，这些答案来自14374个相应贡献者的3569个问题，总共有1824873个历史答案。



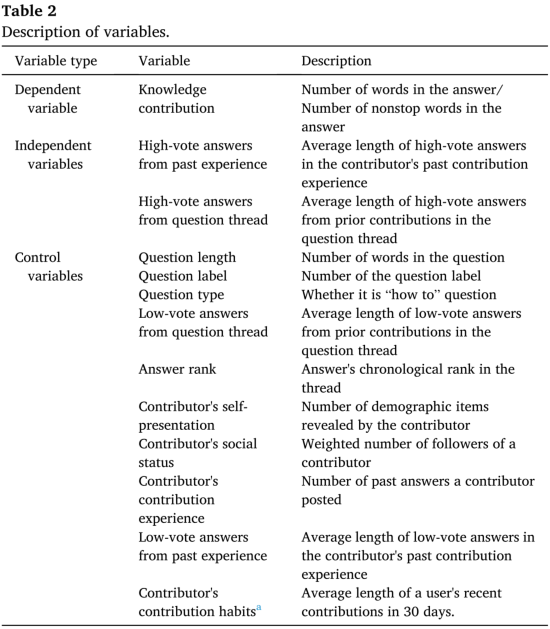
**4.2 先前贡献对网络知识贡献的信息性的影响**

在本节中，我们将探讨一个人的历史贡献和问题线索中的先前贡献和相应结果如何影响后续贡献的信息性。我们计算高票数和低票数答案的平均长度，以比较它们各自的影响。

**4.2.1 变量**

我们用绝对长度和不间断长度来衡量每个贡献的信息量。如上所述，贡献的长度是一个相当直接的维度，表明其信息量和贡献者的总体努力[31,37]。在知乎等在线问答社区中，答案以文字形式为主，因此我们删除了投稿中所有的标点符号、特殊符号、表情符号、图片，并计算剩余文字的长度。由于我们的研究主题并不是特别与图片相关，所以在我们的研究贡献中存在的图片只是作为额外的注释，类似于表情符号的作用。因此，我们不像其他研究那样使用图片的数量作为信息含量的衡量标准[21,656]。此外，我们还使用连续单词的数量来表示度量的稳健性。不间断单词的数量量化了具体信息的水平，即名词、动词和形容词，在每个贡献[23]。表2详细说明了研究中使用的变量。表3显示了这些变量的描述性统计。由于表3所示的偏度，我们对这两个测量值应用对数形式。

对于模型贡献的独立变量，我们需要定义什么是高投票的答案。在知乎，由于每一票都没有具体的时间戳，所以我们无法直接获得后续投稿人提供时的准确票数。知乎的投票功能将投票问题通知给每个投票者的关注者，类似于Twitter上的转发机制。发现75%的转发发生在24小时内，10%的转发发生在一个月内[57]。由于我们只对高得票的答案感兴趣，而对它们在随后的贡献时刻的具体票数不感兴趣，因此我们利用最终票数作为其近似值。根据投票数的分布(如表4所示)，我们将投票数超过5(前20%)和超过10(前12%)的答案分别视为高投票答案，以验证结果的稳健性。然后，我们分别计算一个人在过去贡献经验(主动学习)和问题线索(替代学习)中的高投票答案的平均长度。我们还在控制变量中加入低投票答案的平均长度进行比较。如果一个人的历史贡献或问题线索中没有这类答案，这些变量就为零。由于这些变量的偏态性，我们对它们进行自然对数变换。

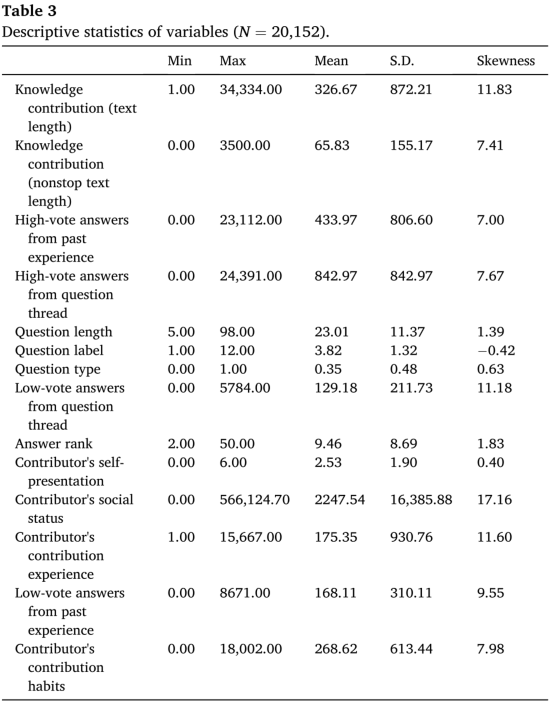


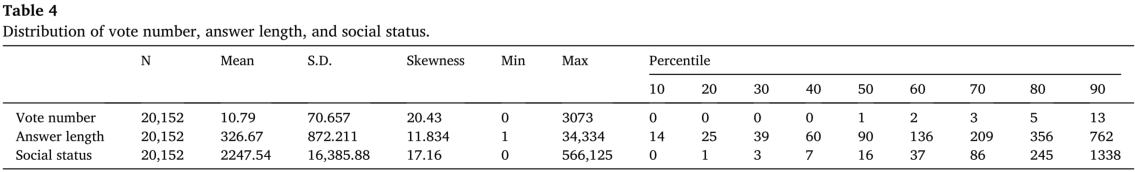
注：虽然我们在回归中包含了高投票和低投票的历史答案，但为了结果的稳健性，我们还额外包含了习惯这个变量。我们已经为这个变量应用了12个变量(详见第4.2.1节的最后一段)。在这里，我们以最近30天内用户的回答的平均长度为例，如表2和表3所示。如果在30天内没有答案，这个变量为零，这占我们样本的15%。

我们控制了可能影响特定知识贡献的各种因素，如表2所示，在三个层次:问题、线索和个人层次。在问题层面，我们控制问题的长度、标号和问题类型等特征。我们特别关注“如何”问题，它是非事实问题，比事实问题具有更多的不确定性，比社会或观点问题[58]更多的信息。从线程级别，我们控制答案的时间排序和线程中低投票答案的平均长度。对于投稿人的个人特征，我们控制了四个因素，包括:(1)自我呈现的完整性;(2)在社区中的社会地位;(3)过去的贡献经历;(4)在贡献者过去的贡献经历中，低票数答案的平均长度。关注者的数量通常用来表示用户在网络社区中的社会地位[5,32,59]。由于知乎并没有显示每一个跟随者的跟随时间，所以我们不得不估计一个人在每次贡献的时候的社会地位。之前在在线问答社区的研究表明，一个人的粉丝数与他过去的贡献所获得的投票总数[55]之间存在正相关关系。因此，我们使贡献时的估计收到的投票、爬取时的总投票和爬取时的总关注者数来衡量他/她给出如下答案时的加权关注者数:

总关注人数\*投稿人估计投票人数/总投票人数

与上述数据处理类似，我们对高度偏态变量(偏态>2)取对数。我们还控制了回答时间假人(每月假人)。





为了结果的稳健性，我们进一步控制了个人习惯性贡献行为在近期贡献和初始贡献中的表现。具体来说，我们分别控制四种类型的习惯行为:(1) 每次贡献之前30天、90天、180天，最近的个人回答的平均长度，(2)每次贡献前，个人最近3、6和9个答案的平均长度，(3) 在每次贡献之前，贡献历史中的最初 30、90 和 180 天的平均答案长度，(4)在每次贡献之前，个人贡献历史中最初的3，6，9个答案的平均长度。对数形式应用于所有这些习惯变量，以避免过度分散。

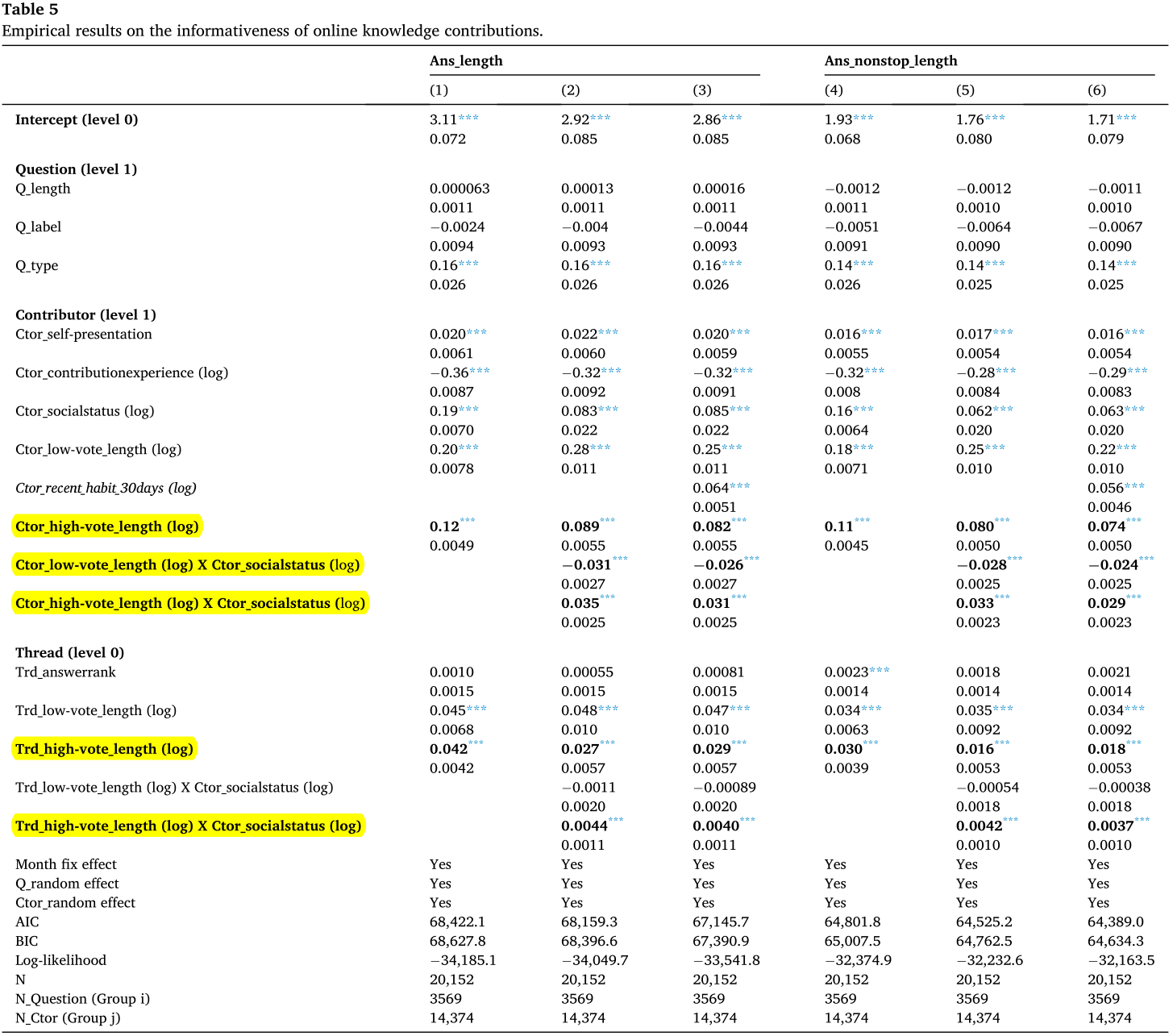
**4.2.2 模型规范**

对于我们的数据结构，答案是按问题或贡献者聚集的。因此，我们使用层次线性模型(HLM)[60,61]。HLM可容纳不平衡的观测值，并可用于将模型方差划分为不同的水平。我们发现我们的数据集不是一个纯粹的嵌套结构，因为观察是在问题和贡献者中聚集的，但问题和贡献者不是彼此嵌套的。在[32]之前的研究基础上，我们进行了交叉分类随机效应的HLM，控制了多个未观察到的异质性来源(同时控制了贡献者随机效应和问题随机效应)。估计方程为:

下标“i”表示Question\_ID，“j”表示Contributor\_ID。“Ctor\_high-vote\_lengthij”表示从个人过去的贡献经验中获得的高票数长的答案的平均长度。“Trd\_high-vote\_lengthij”表示问题线索中高票数答案的平均长度。“Q\_Controlsi”表示问题的特征，如长度、类型、标签数量等。“Trd\_Controlsij”和“Ctor\_Controlsij”表示Answerij的线程特性和贡献者的个人特性，如表2所示。我们还用不同的高投票答案标准(5和10)检验了这些模型，以显示结果的稳健性。对于贡献者的社会地位的调节作用，除了调查它们对导致主动学习和替代学习的模型(高投票)答案的调节作用外，我们还分别探讨了他们对未建模(低投票)答案的调节影响，在一个人的历史贡献和相同的问题线索。这可以全面了解以前贡献的影响。

**4.2.3 实证结果**

表5列出了我们研究模型的回归结果。我们在模型1、2、3中使用因变量ln(Answer\_length)，在模型4、5和6中使用因变量ln (Ans\_nonstop\_length)。



对于主动学习效果，模型1 ~ 6中Ctor\_high-vote\_length的系数显著为正。这验证了H1宣称的在线知识贡献的信息性存在主动学习效应。这些模型中Ctor\_low-vote\_length的显著正系数也表明，无论是否获得较高的社会认知度，个人的贡献信息习惯都会影响未来的贡献。有趣的是，对于社会地位的调节效应，模型2、3、5和6中的Ctor\_high-vote\_length (log) X Ctor\_socialstatus (log)的系数显著为正，而模型2、3、5和6中的Ctor\_low-vote\_length (log) X Ctor\_socialstatus (log)的系数显著为负。这些结果表明，在社区中社会地位较高的用户更容易受到其过去贡献经验的高投票答案的影响，而不容易受到历史低投票答案的影响。这说明，如H3所示，社会地位水平越高，知识贡献的信息性的主动学习效应越强。

对于问题线索中的替代学习效应，与主动学习效应相似，模型1 - 6中Trd\_high-vote\_length的系数均为显著正，验证了H2中提出的在线知识贡献中存在替代学习。这些模型中Trd\_high-vote\_length和Trd\_low- vote\_length的显著正系数都突出了所有先前答案的社会影响(基于回答同一问题的关系)。不同的是，模型2、3、5和6中Trd\_high-vote\_length (log) X Ctor\_socialstatus (log)的正系数表明，从信息的角度来看，社会地位较高的贡献者更倾向于关注问题线索中之前的高投票答案，从而使他们的贡献具有竞争力。这证实了H4。此外，通过对自变量系数的比较，我们发现，个人历史贡献的影响一般大于帖子中先前贡献的影响。

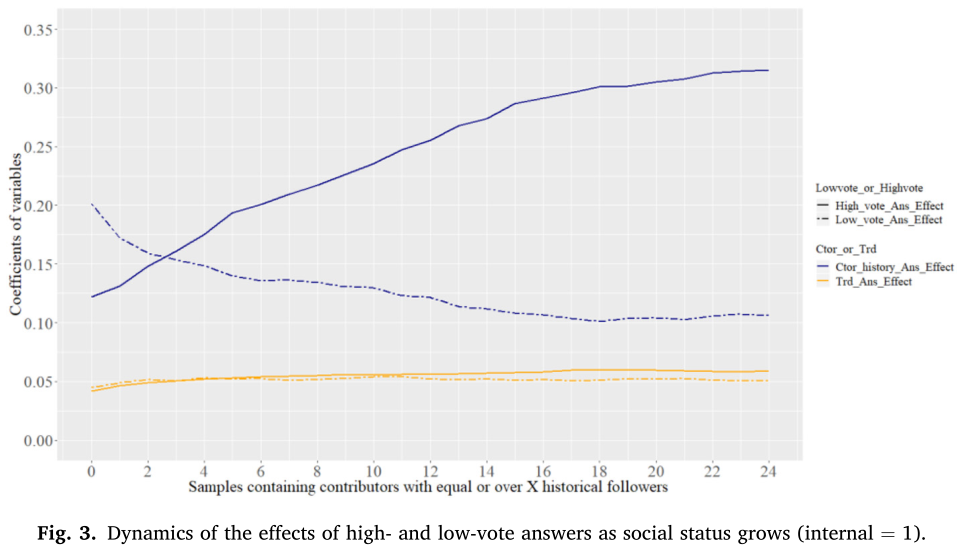
问题类型、贡献者自我呈现水平和贡献经验等变量的控制与知识贡献的信息含量显著相关。在模型1到6中，参与者社会地位的控制变量始终显著为正。这也表明，社会地位较高的贡献者通常提供的信息更丰富。当将个人习惯变量加入回归模型(如模型3和模型6所示)时，Ctor\_high-vote\_length和Ctor\_low-vote\_length的影响以及贡献者社会地位的调节作用均有所降低，但仍然显著。这说明在控制个人贡献习惯后，网络知识贡献存在学习效应(建模效应)。

**5 稳健性检验与其他分析**

**5.1 使用一系列样本比较主效应**

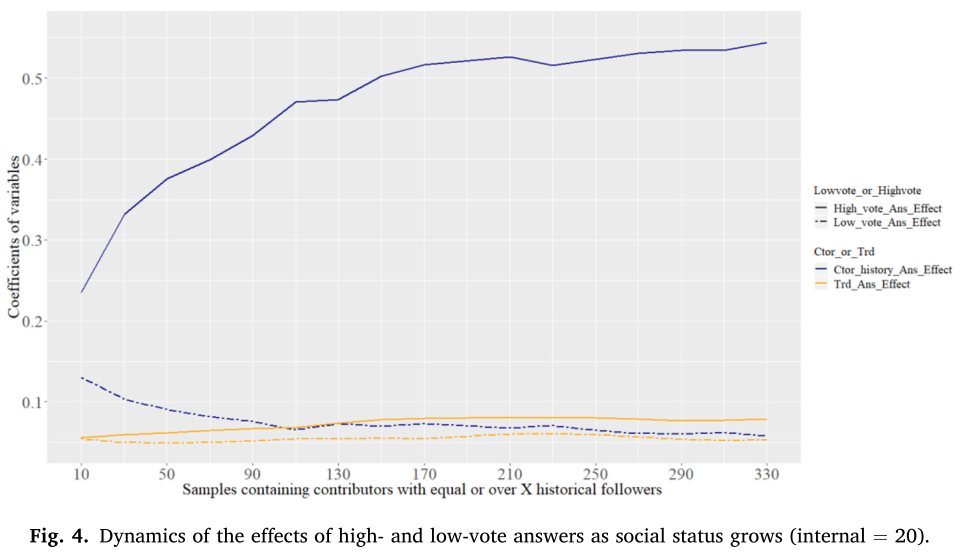
为了更全面地比较未建模贡献和建模贡献的影响，我们进一步采用包含不同社会地位贡献者的样本进行了一系列实证分析。从表4可以看出，在我们的研究样本中，超过10%的贡献者在贡献时没有关注者，30%的贡献者在贡献时收获不到3个关注者。虽然我们的样本中包含了至少一个历史答案的贡献者，但他们中的许多人可能只有很少的高投票答案。

接下来，我们使用回归模型，它只关注表5第1列中建模/未建模答案的直接影响。我们使用的样本包含有相同或超过X个历史追随者的贡献者。我们首先设置X = 0,1,2,3，…，22,23,24，并分别检验Ctor\_low-vote\_lengthij, Ctor\_high-vote\_lengthij, Trd\_low-vote\_lengthij和Trd\_high-vote\_lengthij在这25个回归中的系数。回归分析中四个变量的结果均在0.01水平上显著。结果的动力学如图3所示。蓝线表示历史答案的影响，而橙色表示同一问题线索中答案的影响。实线表示高票数答案的影响，虚线表示低票数答案的影响。



当X在3左右时，图3中的两条蓝色曲线相交，说明当一个贡献者的粉丝数超过3时，其历史贡献中高得票的答案比低得票的答案影响更大。从图3的两条橙色曲线可以看出，当参与者关注人数超过11人时，帖子中高投票数的答案比低投票数的答案更有影响力。

然后，我们将X的区间从1转换为20，设X = 10,30,50，…，290,310,330，并分别进行回归。回归分析中四个变量的系数仍在0.01水平上显著。结果的动力学如图4所示。



与图3类似，图4中的实蓝色曲线显示，随着用户社会地位的增加，用户历史贡献中建模答案的影响也在不断增加。进一步显示出倒U型增长曲线，表明一个人的历史贡献中建模答案(主动学习)的影响随着关注者数量的最初增长而急剧增长，当关注者数量超过110时，这种影响仍在逐渐增加。但增长趋势变得相对稳定。这说明，随着社会地位的提高，在线问答社区的用户会逐渐从过去的高投票贡献中学习，最终在信息量方面形成了比较稳定的写作风格。从图4中蓝色虚线的减小可以看出，随着社会地位的增加，低投票答案对个人历史贡献的影响逐渐减小。

图4中两条橙色曲线相对平坦，实线上升平缓，且始终高于橙色虚线。这说明对于帖子中先前答案的影响，社会地位越高(粉丝超过10人)的参与者对帖子中高投票答案的关注越多，这与我们的假设是一致的。

有趣的是，通过对比图4中两种颜色的曲线，我们发现蓝色曲线普遍高于橙色曲线。这表明，在网络知识贡献的背景下，历史个人行为的影响高于他人的影响。这一发现与以往对网络评级行为的研究一致[62]。当X大于130时，图4中的实橙色曲线超过了蓝色虚线曲线，突出了建模行为和未建模行为之间的明显影响。

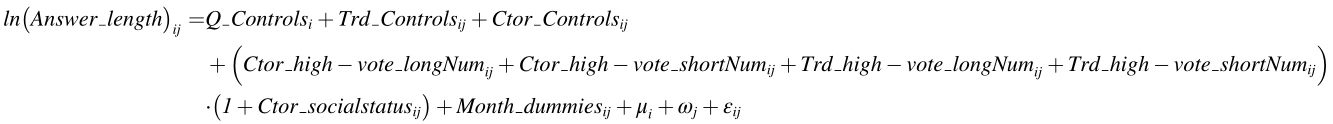
**5.2 对建模贡献的进一步探索**

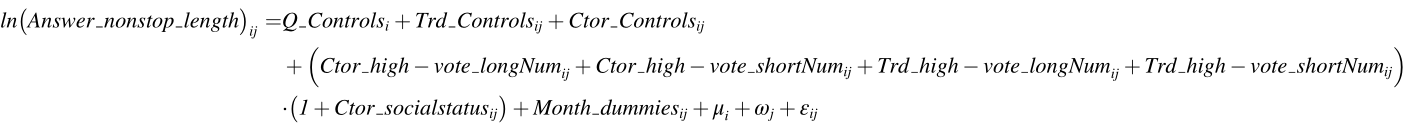
在这些在线问答社区的高投票贡献中，真正翔实的贡献可能会因为他们详细和深刻的知识贡献而获得无数的认可(赞、赞)。然而，并不是每一个高投票贡献都是真正的信息[50]。与其他社交媒体一样，在线问答社区也会抑制吸引眼球的内容[14]，这些内容往往质量较差，但却可以通过感官唤起或刺激获得巨大的人气。这些引人注目的贡献的出现可能会削弱内容贡献者对严格贡献的热情，导致社区中贡献的认真氛围降低。

从这个观点来看，长答案和短答案对随后的知识贡献有不同的影响，并可能被贡献者的社会地位不同的调节。除了像以前的研究通常做的那样计算平均值之外，更详细地描述先前的建模行为将是有趣的[63,64]。因此，我们分开关注两种建模行为的影响：建模的“one-linear”(高投票短答案)和建模的“real stuff”(高投票长答案)，并分别调查它们对知识贡献的影响。基于心理学文献中的社会影响理论(SIT)[65,66]，我们采用高投票答案的数量作为每种效应程度的代理。在营销和信息系统领域采用SIT的研究证明，具有一定影响力的信息源数量对社会影响水平有积极的影响[67-69]。

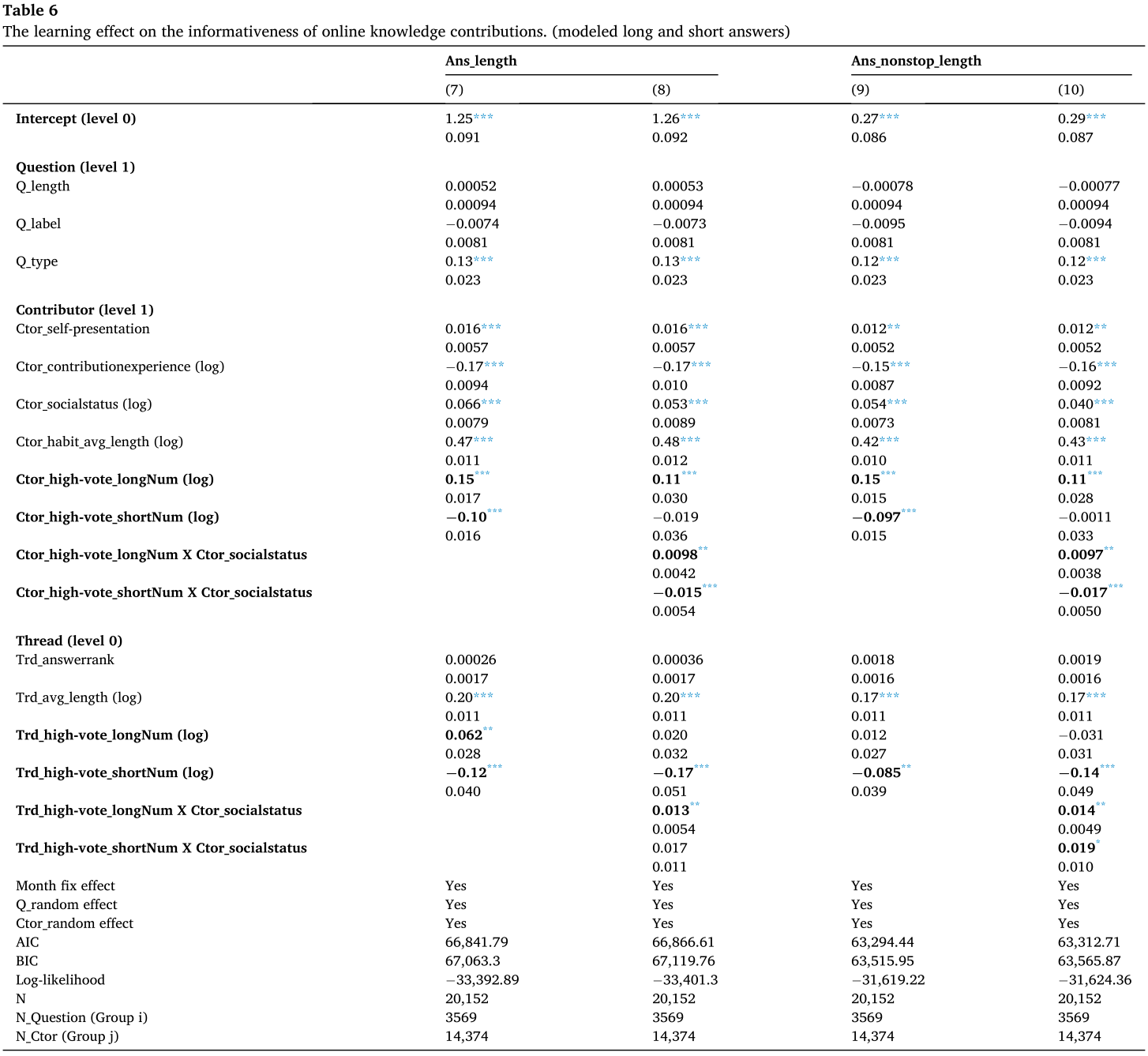
根据表4所示的答案长度分布，我们选择长度前20%(超过356个单词)的答案作为长答案，最后20%(少于25个单词)的答案作为短答案。高票答案(5和10)的标准与前一节相同。基于这些分类和SIT，我们计算贡献者过去的贡献经验和问题线索中每种类型的答案的数量，以建立新的独立变量。由于它们的偏态性，我们对这些变量进行ln(1 +X)的自然对数变换。

也采用HLM，公式如下:





“Ctor\_high-vote\_longNumij”和“Ctor\_high-vote\_shortNumij”表示从一个人过去的贡献经验中获得的高投票长的和短答案的主动学习效果。“Trd\_high-vote\_longNumij”和“Trd\_high-vote\_shortNumij”表明在问题线索中高投票长答案和短答案的替代学习效果。因为我们特别注意本节的高投票答案，不同于方程1和2的控制变量。，我们进一步控制贡献者的历史贡献和同一问题线索中先前答案的平均长度。结果如表6所示。



与4.2.3节的结果类似，模型7和9中Ctor\_high-vote\_longNum、Ctor\_high-vote\_shortNum Trd\_high-vote\_longNum, Trd\_high-vote\_shortNum的系数结果，模型7和9也验证了主动学习和替代学习对在线知识贡献的信息性的影响。modeled的长答案和modeled的短答案对未来贡献的信息量有相反的影响。无论是从历史经验还是先前的贡献，这些长答案都将带来更多的信息贡献，反之亦然。不同的是，在模型7 ~ 10中，Ctor\_high-vote\_longNum系数高于Ctor\_high-vote\_shortNum，而Trd\_high-vote\_longNum系数低于Trd\_high-vote\_shortNum。这表明，在一个人的历史贡献中建模的长答案通常比建模的短答案有更高的影响，而建模的短答案在问题线索中的影响比建模的长答案更大。

对于社会地位对主动学习的调节作用，模型8和模型10中Ctor\_high-vote\_longNum X Ctor\_socialstatus和Ctor\_high-vote\_shortNum X Ctor\_socialstatus的系数均显著，且与主动学习效应方向一致。这说明社会地位越高的用户在未来进行供给时，其成功的历史经验对其影响越大。历史建模行为(包括建模的长答案和短答案)对社会地位较高的贡献者的影响更大。

有趣的是，对于替代学习的调节作用，模型8和10中的Trd\_high-vote\_longNum X Ctor\_socialstatus和Trd\_high-vote\_shortNum X Ctor\_socialstatus的系数是积极的，尽管模型8中的变量Trd\_high-vote\_shortNum X Ctor\_socialstatus 并不显著。即贡献者的社会地位对Trd\_high-vote\_shortNum的调节作用与Trd\_high-vote\_shortNum的直接影响相反。这表明，虽然后续用户的所有贡献都受到之前在线程中建模的简短答案的影响，但社会地位较高的用户受这些信息较少的答案的影响较小。相反，他们更多地受到问题线索中先前高票数的长答案的影响。这表明，一旦社会地位较高的用户决定提供答案，他们倾向于规范自己的行为，并关注更多信息性贡献。

**6 讨论与意义**

受同一问题或同一贡献者在线知识贡献的不同信息水平的激励，我们研究了过去贡献的长度特征和相应的反馈(收到的社会激励)是否可以激励贡献者在随后的贡献中贡献更多或更少的内容。我们发现，无论是在过去的经验中，还是在同一个问题线索中，未来贡献的信息量都受到过去贡献和相应反馈的显著影响对于社会地位较高的用户，在四个决定因素中，从过去的成功经验中获得的主动学习效果是最具影响力的因素。有趣的是，建模答案的影响对社会地位较高的用户更有效，而社会地位较低的用户更倾向于重视未建模的答案。我们进一步关注两种特定类型的建模答案：高投票长答案和高投票短答案，并验证它们对未来信息贡献的相反影响。在同一问题中，社会地位较高的参与者更看重高票数的长答案，而较少受到高票数的短答案的影响。

**6.1 理论贡献**

首先，我们借鉴社会认知理论中的学习范式，对过去贡献绩效的内容特征影响下的在线问答社区贡献行为提供了新的理解。SCT认为，个体可以通过实际行为(主动学习)或观察他人(替代学习)来学习，并强调那些获得积极结果的行为将被视为模式化行为，并影响未来的行为。这种学习模式已被应用于诸如IT培训[13,46,47]和e-learning[44]等环境中。同样，在线问答社区中过去的知识贡献产生的社会激励也可以主动这种学习范式。我们的研究扩展了这一理论视角的适用性，并验证了主动学习效应和替代学习效应对在线问答社区知识贡献的信息量的影响。此外，当更详细地探究过去的贡献时，我们的实证结果也表明了一些不同的发现。这包括:(1)在个人历史贡献中，高票长答案的影响通常高于高票短答案，而高票短答案在问题线索中的影响要强于高票长答案。(2)社会地位高的参与者在作出知识贡献时，对问题线索中高票长答案的考虑更多，对高票短答案的影响更小。总之，我们认为学习效应是一个有助于理解有激励或反馈机制的网络社区贡献行为的理论视角。

其次，我们的研究也有助于网络知识贡献的文献。之前的研究已经确定了过去的贡献对未来的知识贡献行为的影响[3,30]或社区反馈如何影响用户的持续在线知识贡献[29]。我们通过结合过去的贡献和相应的反馈来补充这些研究，并调查它们对未来贡献的信息性的复合影响。过去在在线问答社区的贡献将被考虑在一个人的历史贡献和之前的贡献在同一问题线索。我们的研究结果表明，控制个人贡献习惯后，网络知识贡献的学习效果显著。贡献者，尤其是那些社会地位较高的人，会受到他们过去成功贡献的影响，而高投票长答案的影响更大。问题帖中的高投票答案，特别是高投票简短答案，也会影响后续帖子的信息量，社会地位较高的用户对帖子中的高投票长的答案关注更多，而对高投票简短答案关注更少。此外，我们还说明了建模和未建模答案对不同社会地位的贡献者的不同影响。个人在帖子中的历史贡献通常比其他人的贡献更有影响力，这一发现也与在线评级行为的研究一致[62]。这些结果通过摆脱特定贡献行为和相关反馈之间的关系，提供了对在线知识贡献的更全面的理解，并提出了激发信息性贡献的建议。

对于只针对中国的研究概率性，如选题所述，知乎在线问答社区的用户主要是说中文的年轻人，而英语几乎是中国所有年轻人的必修课程。这确保了这个主题的广泛潜在用户。由于知乎在学术上不是一个学习英语的网站，所以知乎的英语学习话题是一个比较随意的平台，通过提问来分享技巧、观点和建议。这个话题也与热点事件无关。因此，本研究可以推广到中国以外的其他环境中，在与外语学习等即时热点无关的主题中进行随意的个人分享。

**6.2 实践意义**

我们的研究结果强调了过去贡献的价值，以及对未来贡献的信息性所收到的反馈，并指导了在线社区的管理和设计，以激发信息性贡献(更多的“真实”答案)。这可能包括:(1)通过在投稿界面显示提示来引导投稿者的行为，比如“关注过去的表现，信息丰富的答案更容易获得更多的选票”;(2)提供更多的正面反馈机会，获得信息丰富的答案;当一个读者在一个很长的答案中停留了很长时间，网站可以通过提示来提醒他们投票，比如“贡献并不容易，不要嫉妒你的赞赏。”这可能会使知识型知识贡献者得到更多的认可。(3)以“优秀答题者”等称号奖励答题者的勤奋和贡献。

另一方面，如果在线知识网站决定采用一种鼓励多样化贡献风格的发展策略，类似的鼓励设计方法也可以应用于其他类型的贡献，如在社区的某些部分或主题中有趣的“one-linear”答案。这可能会吸引喜欢这类在线贡献的用户。

**6.3 局限性及展望**

我们的研究并非没有局限性。首先，尽管我们提供了选择一个相当稳定的数据主题以防止意外的外源性冲动的理论依据，但在其他类型的主题(如热点事件)中检验学习效应的存在将是有趣的。像知乎这样的在线问答社区，短时间内就可以有几十万个热点问题的答案。学习对内容生成的影响可能在这个在线贡献的即时背景下表现得不同。在更广泛的背景下，未来的研究可以提供关于在线社区中的贡献行为、IT工件和学习效果的更深刻的视角。

第二，本研究中被解释变量的维度局限于知识贡献的信息量。文本分析还有其他几个角度，如客观性和情绪性。除了基于文本的在线社区，学习效应也可能影响YouTube或Instagram等其他类型的UGC社区的个人风格。

第三，我们的研究重点是在控制低投票答案和个人贡献习惯的影响后，高投票答案的学习效果。这两种控制方式也会对未来的贡献产生重大影响。虽然我们已经探讨了高投票和低投票回答对不同社会地位贡献者的显著影响，但更深入地挖掘这些贡献行为的潜在机制将是有趣的。其他理论如在线评论中的策略行为[51]和社会认知早期研究[11]也可以为个体知识贡献行为提供更全面的理解。

第四个限制与研究设计有关。虽然我们控制了综合的个人和环境因素来证明对知识贡献的主动和替代学习效应，但这样的设置可能不足以验证理论模型的因果关系。未来的研究可以采用随机实验来加强因果识别。