探究在线社区中社会规范涌现的动态过程

Shangde Gao[1], Yan Wang[2*], My T. Thai[3]

1 博士候选人，佛罗里达大学设计、建造与规划学院城市与区域规划系及佛罗里达大学建成环境韧性研究所，地址：美国佛罗里达州盖恩斯维尔 Inner Road 1480 号，邮编 32601；电子邮箱：gao.shangde@ufl.edu；ORCID：0000-0003-2218-2872。

2* 助理教授，佛罗里达大学城市与区域规划系及佛罗里达大学建成环境韧性研究所，地址：美国佛罗里达州盖恩斯维尔邮政信箱 115706，邮编 32611（通讯作者）；电子邮箱：yanw@ufl.edu；ORCID：0000-0002-3946-9418。

3 教授，佛罗里达大学计算机与信息科学及工程系及 Warren B. Nelms互联世界研究所，美国佛罗里达州盖恩斯维尔，邮编 32611；电子邮箱：mythai@cise.ufl.edu；ORCID：0000-0003-0503-2012。

摘要：尽管人们已发现社会规范对遏制错误信息的作用，但对于社会规范涌现的模式，例如具有不同特征的在线社区中社会临界转变的模式及其差异，我们仍知之甚少。因此，本研究探讨在线社区中社会临界转变的特征，并考察临界转变特征与在线社区特征之间的相关性。我们以“COVID-19疫苗接种的副作用”为案例主题，首先追踪100个在线社区中临界转变特征的模式。这些社区是根据2020年5月至2021年4月期间Twitter上汇总的通信网络，使用 Louvain 算法检测得出的。随后，我们采用多变量线性回归，探究临界转变特征与社群特征之间的相关性。我们发现，在线社区中的社会临界转变可持续两到四个月，并使接受规范性信念的在线社区成员比例增加50%。回归分析表明，社会临界转变持续时间与社区人数及社会规范的初始接受度呈正相关，而临界转变持续时间与社区成员的度呈负相关。此外，网络模块度和社会规范的初始接受度与社会临界转变的幅度呈负相关，而用户的度和中介中心性则可能与临界转变的幅度呈显著正相关。我们的发现为数字环境中更精准地干预错误信息提供了参考，并为社会规范涌现的时机与机制提供了初步证据。

1 引言

在线平台的广泛发展，推动了公共事件期间各级利益相关者（如政府机构和个人用户）生成的信息传播（Y. Wang et al., 2021）。很大比例的用户生成的在线消息都包含不准确且具有误导性的信息，即错误信息（Del Vicario et al., 2016; Wang et al., 2022）。错误信息的广泛传播从多个方面威胁着人类社会，例如干扰人们就民主、环境和公共卫生问题进行集体决策（West & Bergstrom, 2021）。迫切需要遏制错误信息的传播，并减轻在线错误信息对人类社会造成的负面影响（West & Bergstrom, 2021）。现有研究（例如，D. T. Nguyen et al. (2012)、N. P. Nguyen et al. (2012)、Zhang, Alim, et al. (2015, 2016)、Zhang et al. (2018)、Zhang, Kuhnle, et al. (2016)、Zhang, Zhang, et al. (2015)）倾向于通过以下方式遏制错误信息：（i）辟谣，即在人们接触到错误信息后予以纠正；以及（ii）预先辟谣，即帮助人们

识别虚假或误导性内容（U. K. H. Ecker et al., 2022; Lewandowsky & van der Linden, 2021）。辟谣策略被广泛用于针对特定主题的错误信息采取有针对性的应对措施（U. K. H. Ecker et al., 2022），例如，提供包含事实性阐释的信息（Gao et al., 2021; van der Meer & Jin, 2020; Wang et al., 2022）、事实核查内容（Humprecht, 2020），以及能够促使人们采取健康保护措施的信息（Humprecht, 2020）。当支持错误信息的解释广泛存在时，辟谣策略并非总是有效（Chan et al., 2017）。辟谣信息的效果往往是短期的，之后人们再次接触错误信息便会使其效果消退（Mourali & Drake, 2022）。此外，辟谣策略只能在人们最初接触错误信息之后实施（van der Meer & Jin, 2020），而错误信息的负面后果可能早已出现，并造成显著的社会成本。相反，预辟谣策略有望成为一种有效手段，克服辟谣策略的局限性，并在公众中大规模形成抵御错误信息的能力（van der Linden 等，2020年）。预辟谣策略基于“inoculation”这一社会心理学理论。如果人们事先得到警告，并形成拒绝错误信息的信念，他们就可能对错误信息产生“免疫力”（Lewandowsky 与 van der Linden，2021年）。与辟谣策略相比，预辟谣策略侧重于影响人们对错误信息相关主题的信念，对公众产生长期影响，并减少错误信息造成负面后果的发生（Basol 等，2021年）。在大规模实施时，预辟谣策略会与 社会规范干预 一同开展，旨在形成支持事实证据并拒绝错误信息的社会规范与共识（Dow 等，2022年）。

社会规范干预的基础是人们遵循周围的社会规范（Constantino 等，2022年）。社会规范存在于数字世界和物理世界中（Gao 等，2022年），即社群中共享的信念或可接受的行为；研究表明，它们与人们对错误信息内容的信念显著相关（Andı 和 Akesson，2021年；Gimpel 等，2021年；Lapinski 和 Rimal，2005年）。遵循社会规范可以满足人们避免受到惩罚的愿望，通过与他人协调行动带来益处，并为人们判断特定语境中何种做法得到认可或明智提供简单的启发式判断（Constantino 等，2022年）。基于这一心理现象，研究者已在物理和数字领域实施社会规范干预，以帮助人们形成支持事实证据、拒绝错误信息的信念（Andı 和 Akesson，2021年；Gimpel 等，2021年；Lapinski 和 Rimal，2005年），例如遏制有关气候行动和健康行为的错误信息（Constantino 等，2022年；U. K. Ecker 等，2022年）。具体而言，通过向个人展示描述某一主题相关错误信息的“普遍信念”（即社会规范）的文本，个人往往会调整自己的信念，使之与“普遍信念”一致，并减少对错误信息的依赖（U. K. Ecker 等，2022年）。另一个例子是，向个人展示这样一条信息：“大多数负责任的人在分享文章前都会三思”（一项社会规范），此时他们不太可能分享包含误导性或存在争议内容的社交媒体文章（Andı 和 Akesson，2021年）。

尽管研究已确认社会规范在抑制错误信息方面的作用（Dow et al., 2022; Constantino et al., 2022; U. K. Ecker et al., 2022），但目前鲜有实证证据可为社会规范干预的实施提供指导。仍存在若干知识空白与挑战。首先，近期针对物理世界中对照实验的研究发现，在人工设计的社区中，社会规范涌现往往伴随着一个临界转变过程，即社会临界转变（Berger, 2021; Centola et al., 2018; Ehret et al., 2022）。社会临界转变是指达到“临界点”时，个体社区的微小变化便能引发整个社区对规范性信念的接受度发生突然而非线性的变化（Berger, 2021）。通过预测社会临界转变发生的可能性及其幅度，政策制定者可以调整干预的实施时机和投入力度，从而提升社会规范干预的成效（Andreoni et al., 2021; Ehret et al., 2022）。然而，由于缺乏对在线社区的分析，目前尚不清楚在线社区中是否也存在社会临界转变，以及其临界转变特征（例如持续时间和幅度）是否遵循某些规律。目前，对于如何确定社会规范干预的实施时机和投入力度，以促进规范涌现的临界转变过程，仍缺乏相关知识

。其次，现有研究中的实验已发现一些证据，表明社区特征与规范性信念的传播之间可能存在关联（Hu & Leung, 2017; Savarimuthu & Cranefield, 2011; Sen & Sen, 2010; Yu 等人, 2014）。然而，这些实验通常基于现实世界或虚拟场景中的人工设计社区，因此实验结果未必适用于在线环境中的社区。此外，现有研究尚未揭示社会临界转变过程如何随社区特征而变化。因此，有必要基于现实世界的社区开展实证研究，探讨社区特征与社会临界转变之间的关系，为社会规范干预的设计提供参考。

To 为填补这一研究空白，本研究旨在回答以下研究问题（RQ）： :

• 研究问题 1：在线社区的社会规范涌现过程中是否存在社会临界转变？如果存在，社会临界转变具有哪些特征和模式？• 研究问题 2：社会临界转变的特征是否与个体社区的不同网络特征相关?

本研究以 Twitter 上围绕 COVID-19疫苗副作用的规范为案例。疫苗相关错误信息的传播在大流行期间造成了严重后果（Loomba 等人，2021年）。一项于2020年的调查显示，超过55%的美国成年受访者因相信有关疫苗副作用、政治问题和安全问题的错误信息，而对接种 COVID-19疫苗产生犹豫（Graham 等人，2020年）。接触到有关 COVID-19疫苗的错误信息后，人们可能会犹豫是否接种 COVID-19疫苗，从而加大感染风险（Loomba 等人，2021年）。当前亟需遏制错误信息的传播，并减轻在线错误信息对人类社会造成的负面后果。我们采用 Louvain 算法（Blondel等，2008年），从包含 COVID-19疫苗相关主题的推文中提取 Twitter 账户之间的交流社群。我们采用既有心理学研究对“信念”的定义（Camina 等人，2021年；Durando 等人，2016年；Herzog 等人，2013年；Ritchie 等人，2021年），并关注用户是否认为有关 COVID-19疫苗的“副作用”确实存在，以及是否接受或拒绝接种 COVID-19疫苗。在这一案例中，“支持 COVID-19疫苗接种”是我们所期望的在线社会规范，我们考察各社群中所表达的规范性信念是否出现社会临界转变。我们进一步探究规范涌现的动态如何随社区特征而变化，例如模块度和中介中心性（Winkelmann et al., 2022）。本研究有助于揭示在线环境中社会规范涌现的时间模式和机制。研究发现有助于制定规范干预策略，从而精准缓解错误信息在在线环境中的传播。

2 数据与方法

2.1 概述

如**图 1**所示，本研究首先使用 Twitter 流式 API 收集有关 COVID-19疫苗及相关错误信息的实时推文（Twitter，2022年）。我们依据 Newman（2003）对在线环境中的社群进行定义，即社群由一组顶点构成，这些顶点之间的边密度较高，而与其他群组之间的边密度较低。具体而言，本研究在整个研究期间，从 Twitter 用户之间的“转推”和“提及”网络中检测社群。例如，如果一位 Twitter 用户在整个研究期间转推或提及另一位用户，那么这两位用户之间就存在一条边。在已识别的个体社区中，我们选取人口规模相对较大（即超过 10 位用户）且存续时间较长（即超过 10 天）的社区。针对这些社区，我们追踪遵循规范性信念的社区人口随时间的变化（即追踪规范涌现），并提取社区特征（例如模块度、平均度）。完成准备工作后，我们首先通过观察检测到的个体社区中社会规范涌现的时间趋势是否存在社会临界转变，来回答研究问题 1。

如果存在临界转变，我们将分析社会临界转变特征的模式，本研究考察的特征包括临界转变幅度和持续时间。基于临界转变特征和社区特征，我们回答研究问题 2，并探究社会临界转变与社区特征之间是否存在显著相关性。

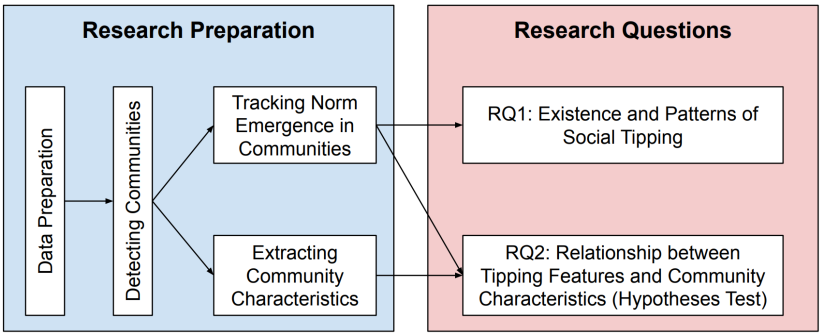


图 1 研究流程

2.2 数据准备

基础数据集通过 Twitter 流式 API 收集，时间跨度为2020年5月1日至2021年4月30日，内容涉及 COVID-19疫苗。具体而言，我们使用 COVID-19疫苗接种相关关键词筛选出与 COVID-19疫苗有关的推文，包括“疫苗”“疫苗接种”等关键词，以及 COVID-19疫苗的品牌名称，例如“辉瑞”。我们基于诸如“提及/回复”消息（即“@username”）和转推消息（即“RT @username”）等通信网络提取在线社区，原因有以下几点。首先，转推/回复行为往往发生在存在关注关系的用户之间，并代表在线用户之间活跃的社会关系（Ozer 等，2016年；B. Wang 等，2021年；Weitzeil 等，2012年）。尤其是，一项关于 COVID-19 转推的研究（B. Wang 等，2021年）指出，超过 50% 的 COVID-19 信息转推是在有关注者/关注关系的用户之间产生的。其次，转推/回复行为能够很好地反映社交媒体用户的社会影响力，因为用户若受到推文内容的影响，往往会转推或回复他人的消息（Evkoski 等，2021年；Yuan 和 Crooks，2018年）。我们可以根据用户之间的互动（例如转推/回复推文），潜在地捕捉某种信念如何在社交媒体用户之间传播（Evkoski 等，2021年）。根据 Skafle 等人（2022年）对新冠疫苗相关错误信息的总结，我们聚焦于所收集推文中有关 COVID-19 疫苗“副作用”的主题，这些推文主要讨论：（a）COVID-19 疫苗是否存在可能严重威胁人类健康的副作用；（b）COVID-19 疫苗是否会导致人死亡；以及（c）COVID-19 疫苗是否未经试验且有毒。我们使用这三个主题的关键词（**表1**），在收集的数据集中识别相关推文。形式为“词语 A + 词语 B”的关键词，表示只有在推文正文中同时找到“词语 A”和“词语 B”，该推文才会被视为与相关主题有关。在线个体社区中的节点是基础数据集中推文的用户。我们仅保留推文提及了基础数据集中其他用户的用户，或被基础数据集中的其他用户提及的用户。新闻机器人账号也会被移除。最终，我们提取出 5,462,900 名不同用户发布的 19,839,188 条推文，这些推文包含与三个错误信息主题相关的关键词（见 **表1**）。我们进一步检测个体社区，并利用该数据集分析规范涌现。

表1 与COVID-19疫苗副作用相关的错误信息关键词

主题	**关键词**

COVID-19疫苗有副作用	“副作用”、“自闭症”、“自闭症患者”、“心理+疾病”, “心理+疾病”、“心理问题”、“心理问题”, “不孕不育”,
COVID-19疫苗可以导致人死亡	"儿童+死亡"、"儿童+已死亡"、"儿童+正在死亡"、"士兵+死亡"、 "士兵+已死亡"、"士兵+正在死亡"、"老人+死亡"、" 老人 +已死亡"、" 老人+正在死亡"、
COVID-19疫苗尚未通过试验，并且有毒	"skip+trail", "毒物", "未测试", "没有接受测试", "未接受测试", "没有接受测试", "没有接受测试", "未接受测试", "未接受测试", "尚未接受测试"

2.3 社区检测

在检索到的通信网络中，当用户回复其他用户或转发其他用户的推文时，用户之间便形成边。边的权重是一天之内一名用户提及另一名用户的次数。我们使用 Louvain算法（Blondel 等，2008年）从社交网络中检测个体社区。Louvain 算法是一种组合优化算法，旨在最大化检测到的个体社区之间的模块度。该算法首先将每个节点分配到其所属社区，然后依次尝试将每个节点移至其所有邻居社区，以找到模块度增益最大的正值。如果无法获得正向增益，该节点便留在原来的社区中（Blondel 等，2008年）。与其他算法相比，Louvain 算法能够高效地从大规模网络中识别个体社区，例如拥有数百万用户的社交媒体网络。为了更好地揭示大型社区中的社会临界转变，而非小群体（例如成员不足十人的小群体）中的情况，我们从检测到的社区中选出规模最大的 100 个社区进行后续分析。

2.4 对用户针对 *COVID-19*疫苗错误信息所表达的信念进行分类，并追踪社群中的规范涌现

根据用户的推文，我们对特定时期内个人关于 COVID-19疫苗副作用的表达的信念进行分类。我们首先对个人用户推文中表达的信念进行分类。我们使用与 COVID-19疫苗接种相关的 2,000 条推文训练一个长短期记忆（LSTM）模型，并利用该模型估计特定用户发布的推文是否表达了支持或否认有关 COVID-19疫苗接种副作用的错误信息的信念。既有研究表明，LSTM 在文本分类方面表现良好，因为它能够捕捉推文文本中的短语级和句子级特征模式（Zhou et al., 2018）。**图 2**展示了LSTM分类器在训练期间经验证的准确率和损失值；训练结束后，两者分别达到 0.8892 和 0.2292，分类结果的均方根误差为 0.3719。这些指标表明，我们的 LSTM分类器在对个人用户表达的信念进行分类方面表现尚可。

对推文中表达的信念进行分类后，我们根据每位用户当天发布的推文，得出其当天整体表达的信念。具体而言，我们计算一位用户在一天内发布的推文中，否定有关 COVID-19疫苗的错误信息的推文所占的比例。具体来说，如果超过 50% 的推文支持 COVID-19疫苗接种，我们就认为该用户当天接受 COVID-19疫苗接种。如果一位用户某天只发布了一条推文，我们就将该推文中表达的信念视为该用户在特定日期表达的信念。

随后，我们将个体表达的信念汇总到社区层面，并追踪样本社区中的规范涌现。我们将规范性信念定义为“反对关于 COVID-19 疫苗副作用的错误信息”，并通过持有该规范性信念的社区成员比例的时间趋势，追踪社区内规范的涌现。从这些时间趋势中，我们可以识别接受度快速上升的临界点。

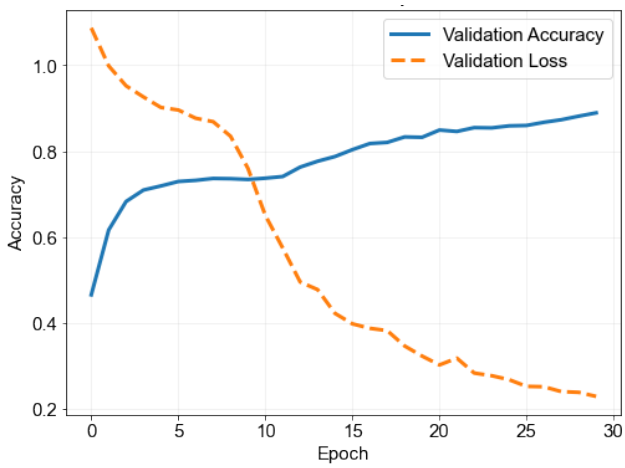


图 2 LSTM分类器在训练过程中的准确率和损失

2.5 社会规范涌现的特征分析

基于样本社区中规范涌现的时间趋势，我们首先观察这些趋势，并检测社群中是否存在社会临界转变（研究问题 1）。我们根据社会临界转变的定义来检测其是否存在：在特定时期，采纳规范的社区成员比例的增长速度相较于此前时期明显更快（Berger, 2021）。我们计算采纳规范性信念的社区成员比例每日增幅，并观察某一时期的增幅是否相较于此前时期明显更快。若是，我们便认为样本社区的规范涌现过程中存在社会临界转变。如果样本社区中确实存在社会临界转变，我们将采用现有研究中的社会临界转变测量方法（Andrighetto & Vriens, 2022），包括社会临界转变的持续时间和幅度（见图 3</style）。持续时间表示社会临界转变存在的时间步数。社会临界转变的幅度则通过临界转变前后采纳规范性信念的社区成员比例变化来衡量。

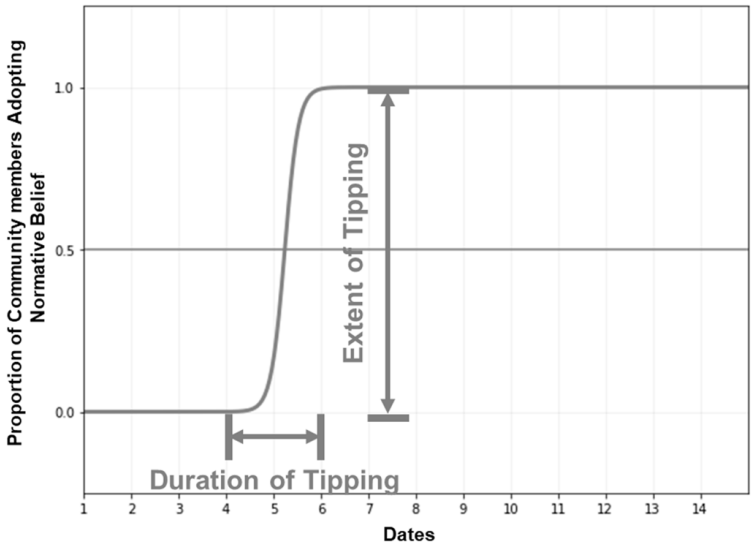


图 3 社会临界转变的持续时间和幅度示意图

2.6 研究社区特征与临界转变特征之间的关系 *res*

在线个体社区的某些特征（**表 2**）可能会通过加快或减慢临界转变过程，影响社会临界转变持续时间。某些社区特征

还可能通过促使个体社区内大范围接受社会规范，影响社会临界转变的幅度，例如在线个体社区的模块度。本研究探讨在线社区的特征（**表 2**）与社会临界转变的持续时间和幅度之间的统计相关性，考察的是接受社会规范的社区成员比例（RQ2）。

表 2 在线社区的特征

特征	参考文献
模块度	(Winkelmann 等人, 2022年)
消息频率	(Centola et al., 2018)
网络规模	(Sabarwal & Higgins, 2021)
社会规范的原始接受水平	(Berger, 2021)
社区成员的度和中介中心性	(Winkelmann et al., 2022)

在检验各项假设时，我们明确规定每个社区特征对社会临界转变持续时间和幅度的影响。具体而言，我们针对**表 2**中的每项社区特征，检验以下假设。

• *H1*：社区的模块度与社会临界转变的持续时间和幅度呈正相关。• *H2*：在线社区成员的平均消息频率与社会临界转变的持续时间和幅度呈正相关。• *H3*：社区规模（即成员数量）与社会临界转变的持续时间和幅度呈负相关。• *H4*：最初接受规范性信念的社区成员比例与社会临界转变的持续时间和幅度呈负相关。• *H5.1*：网络社区的平均度与社会临界转变的持续时间和幅度呈正相关 • *H5.2*：网络社区的平均介数中心性与社会临界转变的持续时间和幅度呈正相关

在检验假设之前，我们先检查所有纳入考虑的社区特征和社会临界转变特征的统计分布。这样，我们就能判断假设检验的数据是否存在明显偏差。如**图 4**所示，大多数社群的模块度低于 0.1。大多数社群的网络规模小于 200 名用户，社群用户之间的消息频率通常低于每天 10 条。对于社会规范的初始接受度，大多数社群在形成时的接受度低于 40%。但仍有二十多个社群在形成时的初始接受度高于 80%。此外，各社群的平均度和中介中心性往往集中分布在较小的范围内，例如，平均度为 1.8 至 2.0，中介中心性为 0 至 0.12。

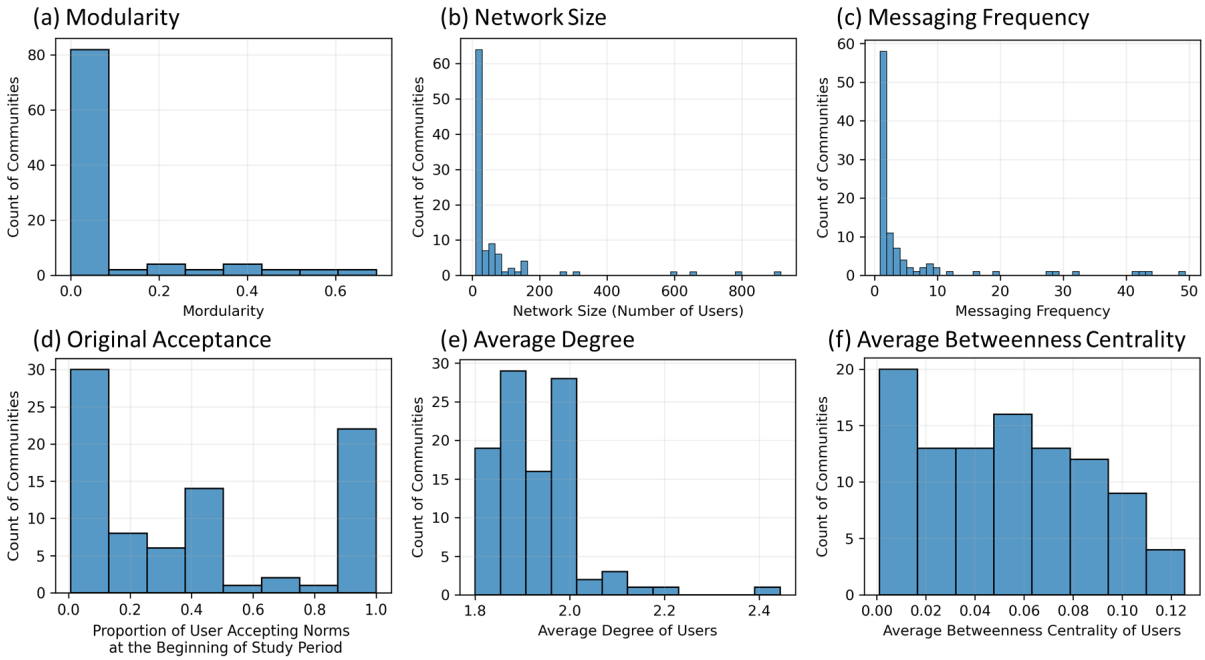


图4 社区特征分布

我们使用多变量线性回归（式 1 和 2）检验上述所有假设。基于已识别的社群，我们根据统计显著性（即 $p-value$）以及社区特征的回归系数为正还是为负，检验提出的假设。例如，为检验有关社会临界转变持续时间的假设 H1，如果变量 $Modularity$ 的$p-value$ 较低，且该变量的系数为正，我们便可以得出结论：在在线社区中，模块度与社会临界转变持续时间呈显著正相关。

$$\text{持续时间}\sim Modularity + Message\ Frequency + Netwrk\ Size + Original\ Acceptance + Average\ Degree + Average\ Betweenness\ Centrality \quad (1)$$

$$Extent\sim Modularity + Message\ Frequency + Netwrk\ Size + Original\ Acceptance + Average\ Degree + Average\ Betweenness\ Centrality \quad (2)$$

3 结果

3.1 样本社区中社会规范涌现的趋势与模式

为回答研究问题 1，我们首先考察社会规范涌现的时间趋势，即样本社群中规范接受度的变化，以确定是否存在“社会临界转变”。具体而言，如果社群中采纳规范性信念（即拒绝错误信息）的人口比例的每日变化远高于此前时期，我们便判定社会临界转变发生在某一特定时段（例如两个具体日期之间）。从规模最大的十个样本社群中社会规范涌现的时间趋势（**图 5**）可以看出，社会临界转变确实存在，而且不同社群的社会临界转变几乎同时发生在2020年12月至2021年4月之间（美国食品药品监督管理局首次批准COVID-19疫苗紧急使用的时期；美国卫生与公众服务部，2022）。尤其是在2020年12月底，采纳规范的人口每日增幅超过10%，远高于此前的每日增幅（通常低于4%）。这些社群发生临界转变后，三个月内，各社群中对COVID-19疫苗接种持有规范性信念的人口比例普遍达到65%。

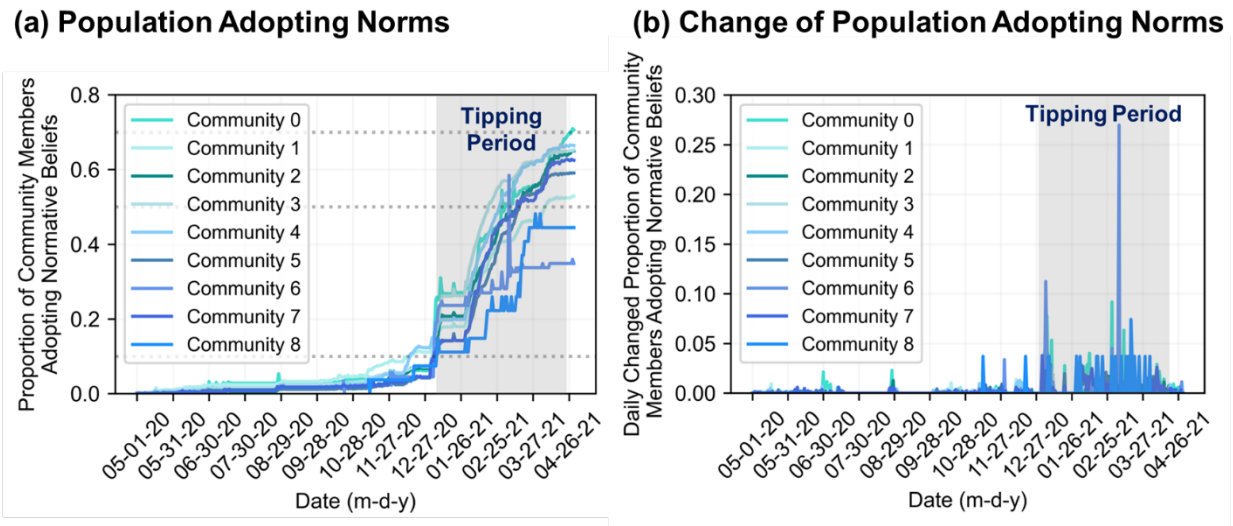


图 5 十个最大的样本社群中社会规范涌现的时间趋势 (a) 和每日变化 (b)

根据所识别出的社会临界转变，我们进一步考察检测到的社区中社会临界转变特征的统计分布，如**图 6**所示。临界转变持续时间和幅度的直方图表明，对于与COVID-19疫苗错误信息相关的规范，在线社区中的社会临界转变往往持续时间较长、程度较强。具体而言，临界转变的平均持续时间为83.26天，中位数为96.5天，95%的样本社区的临界转变持续时间在59至103天之间。17%的社区临界转变持续时间不足一个月，而8%的社区则超过四个月。就临界转变幅度而言，我们发现，86%的样本社区中采纳规范性信念的人口比例增幅超过40%，56%的样本社区中的临界转变幅度甚至超过50%。总体而言，在线社区中围绕拒绝错误信息这一规范的社会临界转变往往持续两到四个月，且超过半数在线社区的临界转变幅度可能超过50%。

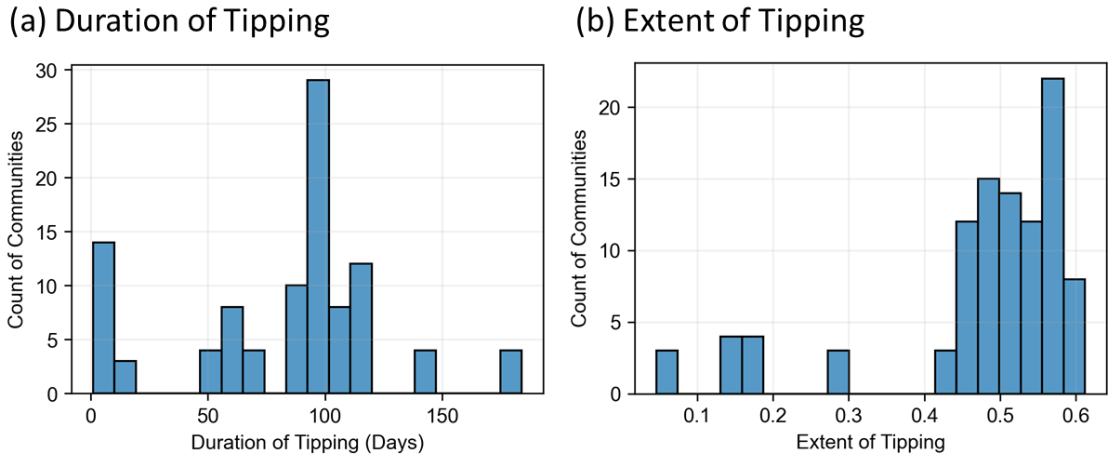


图 6 检测到的社区中社会临界转变特征的分布

3.2 社区特征与临界转变特征之间的关系

在进行回归分析之前，我们首先检查社区特征之间的依赖关系，结果如**图7**所示。具体而言，每对

社区特征的相关性低于0.2。检验结果表明，本研究中考察的社区特征与其他特征相对独立，因此可以纳入多变量线性回归模型。

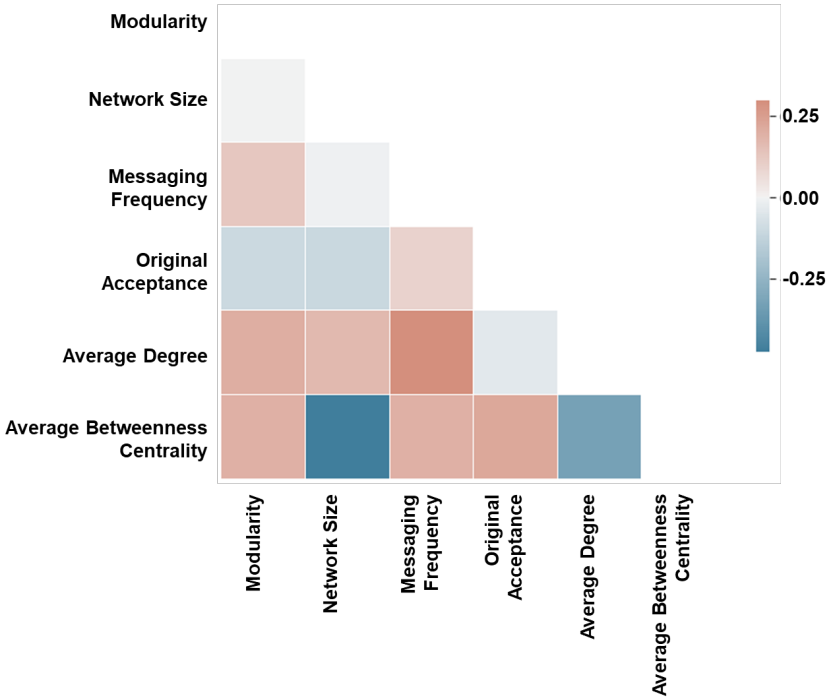


图7 依　　赖性检验：社区特征（颜色越浅表示相关性越低）

社区特征与社会临界转变持续时间之间的关系。回归结果见**表3**，各社区特征的估计系数和95%置信区间（CI）见**图8**。在我们选取的检测到的社区特征中，网络规模（H3）、社会规范的原始接受水平（H4）以及用户的平均度（H5.1）对社会临界转变持续时间有显著的正向影响。具体而言，尽管影响并不显著，但模块度以及社区成员之间的交流频率与社会临界转变持续时间呈负相关（H1、H2）。在线社区中较高的中介中心性与社会临界转变持续时间呈正相关，但这种关系并不显著（H5b）。根据这项回归分析的结果，我们发现，社会临界转变与语境以及社区成员之间的互动密切相关。具体而言，较高的平均度表明，每位社区成员都能与社区内众多同伴交流。原本持有规范性信念的社区成员比例，体现了社区成员对错误信息相关主题的语境素养。我们的结果表明，与互动较少的社群相比，如果个体接触信息并与更多同伴互动，社会规范就更容易传播。此外，原本不否定错误信息的社区成员，如果能与许多原本否定错误信息的同伴进行互动（即初始接受度较高），可能不容易改变自己的信念。另外，大型社群中的规范涌现速度未必会加快，因此大型社群中的临界转变持续时间可能比小型社群更长。

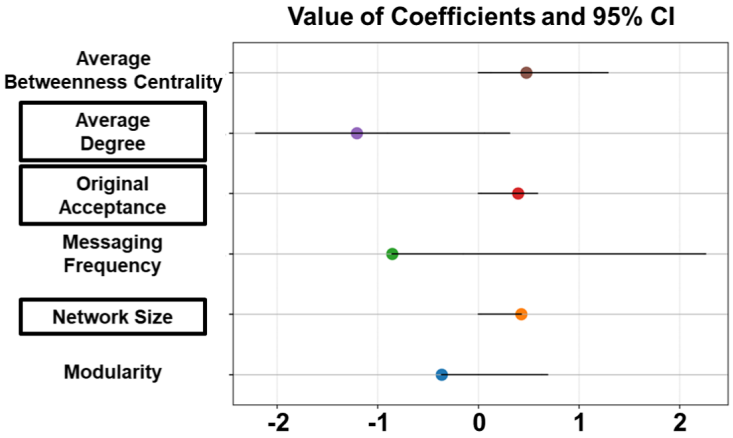


图8 社会临界转变持续时间回归系数的估计值和95%置信区间（显著变量位于红框内）

表3 社会临界转变持续时间的多变量线性回归结果（调整后R^2：0.648）

变量	系数	标准误	t值	P>|t|	[0.025	0.975]
模块度	-0.3648	0.427	-0.855	0.394	-1.207	0.477
网络规模	0.0012	0.001	2.212	0.028*	0	0.002
消息频率	-0.0011	0.002	-0.545	0.587	-0.005	0.003
初始接受水平	0.6879	0.305	2.258	0.025*	0.087	1.289
用户平均度	0.4688	0.08	5.879	< 0.001***	0.311	0.626
用户平均介数中心性	2.3743	2.324	1.022	0.308	-2.212	6.961
显著性水平：0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1						

社群特征与社会临界转变幅度之间的关系。回归结果见**表4**，各项社区特征变量的估计系数和95%置信区间（95%置信区间）见**图9**。在我们选取的检测到的社区特征中，用户之间的平均度（H5a）和中介中心性（H5b）对社会临界转变持续时间有显著的正向影响。同时，社会规范的模块度（H1）和初始接受度（H4）与社会临界转变幅度呈显著负相关。与平均度和中介中心性不同，其他社区特征与社会临界转变幅度之间的关系并不显著。具体而言，网络规模以及社区成员之间的交流频率与社会临界转变幅度之间没有显著关系（H2、H3）。根据回归结果，我们发现社会临界转变幅度也与语境及社区成员之间的互动密切相关。中介中心性和度都与社区成员之间的连接紧密程度有关，而规范性信念的初始接受度则与社区成员对错误信息主题的素养有关。平均度和中介中心性对临界转变幅度的正向影响显著且程度较高，这表明，如果社区成员与其他社区同伴进行大量互动，最终会有更多成员转而接受规范性信念。此外，与临界转变持续时间的回归结果类似，

最初不否认错误信息的社区成员，如果与最初否认错误信息的同伴进行大量互动（即初始接受度较高），或许不容易改变自己表达的信念。

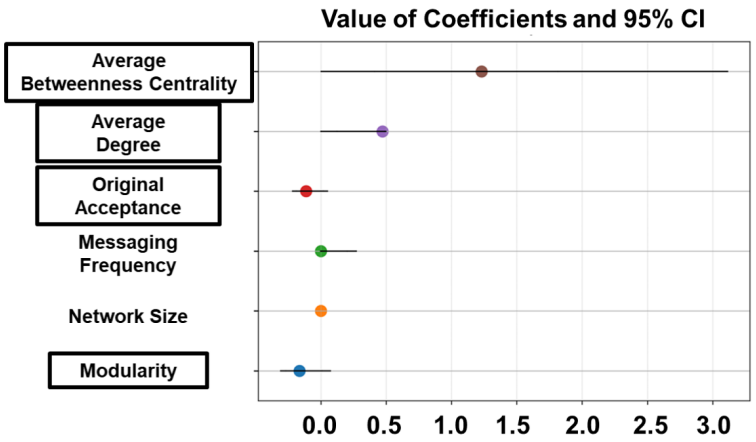


图9 社会临界转变幅度回归系数的估计值及95%置信区间（显著变量以红框标出）

表4 社会临界转变幅度的多变量线性回归结果（调整后R^2：0.972）

变量	系数	标准误	t值	P>|t|	[0.025	0.975]
模块度	-0.1635	0.073	-2.256	0.025*	-0.307	-0.02
网络规模	-0.0001	0.0000897	-1.271	0.205	0	6E-05
消息频率	0.0001	0.000	0.272	0.786	-0.001	0.001
初始接受水平	-0.1127	0.052	-2.176	0.031*	-0.215	-0.01
用户平均度 0.4725		0.014	34.87	< 0.001***	0.446	0.499
用户平均介数中心性	1.2297	0.395	3.113	0.002**	0.45	2.009

显著性水平：0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

4 讨论

社会规范干预有望减缓错误信息的传播，但对于在线环境中社会临界转变是否存在及其模式，以及不同网络特征的社群中临界转变特征有何差异，目前所知仍然有限。本研究探讨规范形成过程中是否存在社会临界转变，并聚焦于驳斥有关 COVID-19 疫苗副作用的错误信息。此外，回归结果表明，临界转变的持续时间与社群规模、平均度以及社区成员对规范性信念的初始接受度关联更大。社会临界转变的幅度（即采纳规范性信念的社区成员数量的增加）与平均度、平均介数中心性、模块度以及社区成员对规范性信念的初始接受度关联更大。

本研究从多个角度拓展了现有知识体系。首先，既有研究更多关注物理世界或人工设计的社区（Berger, 2021；Centola et al., 2018；Ehret等人, 2022），缺乏对在线数字

环境中社会临界转变是否存在及其模式的探索。鉴于数字环境中的社会规范涌现可能不同于其他环境，关于社会临界转变的现有知识未必完全适用于在线社会规范干预。为弥补这一空白，我们利用来自 Twitter 的在线交流数据集开展实证研究，考察 100 个样本社区中的社会规范涌现。在一定程度上，本研究有助于识别在线社区中社会临界转变的持续时间和幅度的统计分布。本研究采用的大型数据集以及具有多种特征的样本社区，使我们能够揭示在线环境中社会临界转变的一般模式。

其次，现有知识体系（例如 Hu & Leung 2017年、Savarimuthu & Cranefield 2011年以及 Sen & Sen 2010年）很少分析社会临界转变模式与在线社区网络特征之间的关系，例如社群的模块度或社区成员的度。为弥补这一空白，我们基于100个样本社区开展的假设检验，有助于识别与临界转变持续时间和幅度显著相关的在线社区特征。我们强调了社会临界转变的特征与模块度、社群规模、平均度、平均介数中心性以及规范性信念的初始接受度之间的显著相关性。我们的研究发现有助于揭示社会临界转变与社区特征之间的一般关系，并为未来设计在线社会规范干预策略提供支持。

本研究仍存在局限，也为我们进一步开展研究留下了机会。首先，仍有一些外部因素会影响个体表达的信念（变化），例如政府政策，而本研究未纳入这些因素。未来的研究将纳入实体社群的因素，以更准确地捕捉社会临界转变与在线社区特征之间的关系。其次，本研究聚焦于 COVID-19 新冠疫苗相关错误信息这一主题，其社区特征和社会临界转变可能遵循不同于其他主题的时间模式。为获得关于在线社区中社会临界转变的更具普适性的研究发现，未来的研究将考察在线交流中的多个主题，例如 COVID-19 的其他预防措施。第三，本研究将每个个体社区视为与其邻近社区相对隔离，但邻近社区中的规范涌现也可能促成个体社区的社会临界转变。未来的研究将考察多社群社交网络情境下的规范涌现和社会临界转变，并探索不同社区中的社会临界转变之间的关系。第四，在本研究期间，我们将社区特征视为相对稳定，但社区特征可能随时间动态变化，并且在不同时间段对规范涌现产生不同程度的影响。未来的研究将捕捉在线社区的动态变化，并考察个体社区网络特征与规范涌现趋势之间的时间交互。

5 结论

探索社会临界转变的模式及其与社区特征之间的关系，对于制定有针对性的社会规范干预以减轻在线错误信息至关重要。我们的研究增进了人们对在线社区中社会临界转变的异质性时间模式及其机制的了解。我们的研究发现可为公共卫生部门、应急响应人员及其他危机管理者提供指导，帮助他们遏制在线错误信息，例如在在线平台上主动传播并认可传递良性规范性信念的信息。通过量身定制的干预策略，危机管理者可以促使在线人群采取适当的预防措施（例如接种 COVID-19疫苗），并减轻无效预防行为造成的不利影响（例如随意拒绝接种疫苗）。借助结合社会规范的反驳式干预，个人有可能对公共卫生宣传活动形成积极态度，并主动抵制和遏制在线错误信息的传播。

参考文献

Andı, S., & Akesson, J. (2021). 助推远离虚假新闻：来自一项社会规范实验的证据。《数字新闻学》，9(1), 106–125. https://doi.org/10.1080/21670811.2020.1847674 Andreoni, J., Nikiforakis, N., & Siegenthaler, S. (2021). 预测社会临界转变与规范变迁：对照实验的发现。《美国国家科学院院刊》，118(16), e2014893118. https://doi.org/10.1073/pnas.2014893118 Andrighetto, G., Castelfranchi, C., Mayor, E., McBreen, J., Lopez-Sanchez, M., & Parsons, S. (2013). （社会）规范动态 [Application/pdf]。36页。https://doi.org/10.4230/DFU.VOL4.12111.135 Andrighetto, G., & Vriens, E. (2022). 社会规范变迁研究议程。《皇家学会哲学汇刊A辑：数学、物理与工程科学》，380(2227), 20200411. https://doi.org/10.1098/rsta.2020.0411 Antonakaki, D., Fragopoulou, P., & Ioannidis, S. (2021). Twitter研究综述：数据模型、图结构、情感分析与攻击。《专家系统及其应用》，164, 114006. https://doi.org/10.1016/j.eswa.2020.114006 Basol, M., Roozenbeek, J., Berriche, M., Uenal, F., McClanahan, W. P., & Linden, S. van der. (2021). 迈向心理群体免疫：针对新冠肺炎错误信息的两种预先辟谣干预的跨文化证据。《大数据与社会》，8(1), 205395172110138. https://doi.org/10.1177/20539517211013868 Berger, J. (2021). 社会临界转变干预可促进可持续消费规范在实际场景中的传播或消退：一项准实验干预研究的证据。《可持续性》，13(6), 3529. https://doi.org/10.3390/su13063529 Blondel, V. D., Guillaume, J.-L., Lambiotte, R., & Lefebvre, E. (2008). 大型网络中的社区快速展开算法。《统计力学杂志：理论与实验》，2008(10), P10008. https://doi.org/10.1088/1742-5468/2008/10/P10008 Camina, E., Bernacer, J., & Guell, F. (2021). 心理科学实证研究中的信念操作化。《科学基础》，26(2), 325–340. https://doi.org/10.1007/s10699-020-09722-9 Centola, D., Becker, J., Brackbill, D., & Baronchelli, A. (2018). 社会惯例中临界点的实验证据。《科学》，360(6393), 1116–1119. https://doi.org/10.1126/science.aas8827 Chan, M. S., Jones, C. R., Hall Jamieson, K., & Albarracín, D. (2017). 辟谣：反驳错误信息的信息之心理效力荟萃分析。《心理科学》，28(11), 1531–1546. https://doi.org/10.1177/0956797617714579 Constantino, S. M., Sparkman, G., Kraft-Todd, G. T., Bicchieri, C., Centola, D., Shell-Duncan, B., Vogt, S., & Weber, E. U. (2022). 扩大变革规模：利用社会规范促进气候行动的批判性综述与实用指南。《公共利益心理科学》，23(2), 50–97. https://doi.org/10.1177/15291006221105279 Del Vicario, M., Bessi, A., Zollo, F., Petroni, F., Scala, A., Caldarelli, G., Stanley, H. E., & Quattrociocchi, W. (2016). 错误信息的在线传播。《美国国家科学院院刊》，113(3), 554–559. https://doi.org/10.1073/pnas.1517441113 Dow, B. J., Wang, C. S., Whitson, J. A., & Deng, Y. (2022). 缓解和管理新冠肺炎阴谋论信念。《英国医学杂志：领导力》，leader-2022-000600. https://doi.org/10.1136/leader-2022- 000600 Durando, P., Alicino, C., Dini, G., Barberis, I., Bagnasco, A. M., Iudici, R., Zanini, M., Martini, M., Toletone, A., Paganino, C., Massa, E., Orsi, A., & Sasso, L. (2016). 意大利地区医务人员季节性流感疫苗接种依从性的决定因素：一项横断面研究的结果。《英国医学杂志开放版》，6(5), e010779. https://doi.org/10.1136/bmjopen-2015-010779 Ecker, U. K. H., Lewandowsky, S., Cook, J., Schmid, P., Fazio, L. K., Brashier, N., Kendeou, P., Vraga, E. K., & Amazeen, M. A. (2022). 错误信息信念的心理驱动因素及其

纠正抵抗。《自然综述：心理学》，1(1), 13–29。https://doi.org/10.1038/s44159-021-00006-y Ecker, U. K., Sanderson, J. A., McIlhiney, P., Rowsell, J. J., Quekett, H. L., Brown, G. D., & Lewandowsky, S. (2022)。结合反驳与社会规范能够增强信念改变。《实验心理学季刊》，174702182211117。https://doi.org/10.1177/17470218221111750 Ehret, S., Constantino, S. M., Weber, E. U., Efferson, C., & Vogt, S. (2022)。群体认同可能削弱干预后的社会临界转变。《自然·人类行为》。https://doi.org/10.1038/s41562-022-01440-5 Evkoski, B., Mozetič, I., Ljubešić, N., & Kralj Novak, P. (2021)。转推网络中的社群演化。《公共科学图书馆·综合》，16(9), e0256175。https://doi.org/10.1371/journal.pone.0256175 Gao, S., Wang, Y., & Platt, L. S. (2021)。运用系统动力学模拟美国卫生机构在 Twitter 上的信息传播以及用户接触疫苗相关错误信息的情况。载于第18届国际危机响应与管理信息系统会议（ISCRAM）论文集。美国弗吉尼亚州布莱克斯堡。Gao, S., Wang, Y., & Webster, G. D. (2022)。基于 Twitter 和物理世界中的描述性社会规范，对表达的态度变化进行因果建模：COVID-19疫苗接种案例研究。《网络心理学、行为与社交网络》，cyber.2022.0153。https://doi.org/10.1089/cyber.2022.0153 Gimpel, H., Heger, S., Olenberger, C., & Utz, L. (2021)。社交媒体上的社会规范对打击假新闻的有效性。《管理信息系统期刊》，38(1), 196–221。https://doi.org/10.1080/07421222.2021.1870389 Han, X., Cao, S., Shen, Z., Zhang, B., Wang, W.-X., Cressman, R., & Stanley, H. E. (2017)。社交网络中社群与多样性的涌现。《美国国家科学院院刊》，114(11), 2887–2891。https://doi.org/10.1073/pnas.1608164114 Herzog, R., Álvarez-Pasquin, M. J., Díaz, C., Del Barrio, J. L., Estrada, J. M., & Gil, Á. (2013)。医务人员的疫苗接种意向是否与其知识、信念和态度有关？一项系统综述。《BMC公共卫生》，13(1), 154。https://doi.org/10.1186/1471-2458-13-154 美国卫生与公众服务部。（2022）。COVID-19疫苗。https://www.hhs.gov/coronavirus/covid-19-vaccines/index.html Hu, S., & Leung, H. (2017)。社群网络下多智能体系统中的局部规范现象。《第16届自主智能体与多智能体系统会议论文集》，1570–1572。https://dl.acm.org/doi/10.5555/3091125.3091365 Humprecht, E. (2020)。他们如何辟谣“假新闻”？事实核查透明度的跨国比较。《数字新闻学》，8(3), 310–327。https://doi.org/10.1080/21670811.2019.1691031 Iacopini, I., Petri, G., Baronchelli, A., & Barrat, A. (2022)。群体互动调节社会惯例中的临界规模动力学。《通讯·物理》，5(1), 64。https://doi.org/10.1038/s42005-022-00845-y Lapinski, M. K., & Rimal, R. N. (2005)。阐释社会规范。《传播理论》，15(2), 127–147。https://doi.org/10.1111/j.1468-2885.2005.tb00329.x Lewandowsky, S., & van der Linden, S. (2021)。通过 inoculation 和预先辟谣应对错误信息与假新闻。《欧洲社会心理学评论》，32(2), 348–384。https://doi.org/10.1080/10463283.2021.1876983 Loomba, S., de Figueiredo, A., Piatek, S. J., de Graaf, K., & Larson, H. J. (2021)。衡量 COVID-19疫苗错误信息对英国和美国接种意向的影响。《自然·人类行为》，5(3), 337–348。https://doi.org/10.1038/s41562-021-01056-1 Luo, H., Zhao, L., & Zhang, Z. (2020)。基于社会互动的因素对家庭废弃物相关行为的影响。《废弃物管理》，118, 270–280。https://doi.org/10.1016/j.wasman.2020.08.046

媒体对话：COVID-19期间公众佩戴口罩的在线随机研究。《医学互联网研究杂志》，24(3)，e34831。https://doi.org/10.2196/34831 Newman, M. E. J.（2003）。复杂网络的结构与功能。《SIAM评论》，45(2)，167–256。https://doi.org/10.1137/S003614450342480 Nguyen, D. T., Nguyen, N. P., & Thai, M. T.（2012）。在线社交网络中的错误信息来源：应怀疑谁？MILCOM 2012 - 2012年IEEE军事通信会议，1–6。https://doi.org/10.1109/MILCOM.2012.6415780 Nguyen, N. P., Yan, G., Thai, M. T., & Eidenbenz, S.（2012）。在线社交网络中的错误信息传播遏制。第三届ACM网络科学年度会议论文集——WebSci '12，213–222。https://doi.org/10.1145/2380718.2380746 Ozer, M., Kim, N., & Davulcu, H.（2016）。使用非负矩阵分解方法检测政治Twitter网络中的社区。2016年IEEE/ACM社交网络分析与挖掘进展国际会议（ASONAM），81–88。https://doi.org/10.1109/ASONAM.2016.7752217 Ritchie, D., Van den Broucke, S., & Van Hal, G.（2021）。将健康信念模型和计划行为理论应用于乳腺摄影筛查：系统综述与荟萃分析。《公共卫生护理》，38(3)，482–492。https://doi.org/10.1111/phn.12842 Sabarwal, T., & Higgins, J.（2021）。网络中传染的控制与传播。SSRN电子期刊。https://doi.org/10.2139/ssrn.3831245 Savarimuthu, B. T. R., & Cranefield, S.（2011）。规范的形成、传播与涌现：多智能体系统中规范仿真模型综述。《多智能体与网格系统》，7(1)，21–54。https://doi.org/10.3233/MGS-2011-0167 Sen, O., & Sen, S.（2010）。社交网络拓扑与选项对规范涌现的影响。载于J. Padget、A. Artikis、W. Vasconcelos、K. Stathis、V. T. da Silva、E. Matson与A. Polleres（编），《智能体系统中的协调、组织、制度与规范（第五卷）》（第6069卷，第211–222页）。Springer Berlin Heidelberg。https://doi.org/10.1007/978-3-642-14962-7_14 Shakya, H. B., Christakis, N. A., & Fowler, J. H.（2014）。社交网络社区与健康行为之间的关联：对印度农村厕所拥有情况的观察性社会中心网络研究。《美国公共卫生杂志》，104(5)，930–937。https://doi.org/10.2105/AJPH.2013.301811 Skafle, I., Nordahl-Hansen, A., Quintana, D. S., Wynn, R., & Gabarron, E.（2022）。社交媒体上关于COVID-19疫苗的错误信息：快速综述。《医学互联网研究杂志》，24(8)，e37367。https://doi.org/10.2196/37367 Twitter。（2022）。实时推送推文。https://developer.twitter.com/en/docs/tutorials/stream- tweets-in-real-time van der Linden, S., Roozenbeek, J., & Compton, J.（2020）。抵御关于COVID-19的假新闻。《心理学前沿》，11，566790。https://doi.org/10.3389/fpsyg.2020.566790 van der Meer, T. G. L. A., & Jin, Y.（2020）。为公共卫生危机中的错误信息干预寻求公式：纠正性信息类型与来源的影响。《健康传播》，35(5)，560–575。https://doi.org/10.1080/10410236.2019.1573295 Varsha, K., & Patil, K. K.（2020）。社交网络中社区检测算法概述

.

*2020*年国际创新计算技术会议 *(ICICT)*，121–126。https://doi.org/10.1109/ICICT48043.2020.9112563 Wang, Y., Gao, S., & Gao, W. (2022). 事实信息与错误信息之间动态关系的研究：关于 COVID-19 期间预防措施相关推文的实证研究。《突发事件与危机管理杂志》，*30*(4), 427–439。https://doi.org/10.1111/1468- 5973.12385

Wang, Y., Hao, H., & Platt, L. S. (2021). 在 Twitter 上考察政府机构与利益相关者在 COVID-19 早期阶段的风险与危机沟通。《人类行为中的计算机》，*114*，106568。https://doi.org/10.1016/j.chb.2020.106568 Wang, B., Liu, B., & Zhang, Q. (2021). COVID-19 疫情期间 Twitter 的使用情况及危机转发动态实证研究。《自然灾害》，107(3), 2319–2336。https://doi.org/10.1007/s11069- 020-04497-5Weitzel, L., Quaresma, P., & de Oliveira, J. P. M. (2012). 衡量 Twitter 微博客中的节点重要性。第二届网络智能、挖掘与语义国际会议论文集 - *WIMS '12*, *1*。https://doi.org/10.1145/2254129.2254145 West, J. D., & Bergstrom, C. T. (2021). 科学领域内及有关科学的错误信息。《美国国家科学院院刊》，*118*(15), e1912444117。https://doi.org/10.1073/pnas.1912444117 Winkelmann, R., Donges, J. F., Smith, E. K., Milkoreit, M., Eder, C., Heitzig, J., Katsanidou, A., Wiedermann, M., Wunderling, N., & Lenton, T. M. (2022). 迈向气候行动的社会临界转变过程：概念框架。《生态经济学》，*192*，107242。https://doi.org/10.1016/j.ecolecon.2021.107242 Yu, C., Zhang, M., & Ren, F. (2014). 网络化多智能体系统中社会规范涌现的集体学习。《*IEEE*控制论汇刊》，*44*(12), 2342–2355。https://doi.org/10.1109/TCYB.2014.2306919Yuan, X., & Crooks, A. T. (2018). 研究 Twitter 上的在线疫苗接种讨论与社群。第九届社交媒体与社会国际会议论文集，197–206。https://doi.org/10.1145/3217804.3217912 Zhang, H., Alim, M. A., Li, X., Thai, M. T., & Nguyen, H. T. (2016). 在线社交网络中的错误信息：以有限预算检测所有错误信息。《*ACM*信息系统汇刊》，*34*(3), 1–24。https://doi.org/10.1145/2885494 Zhang, H., Alim, M. A., Thai, M. T., & Nguyen, H. T. (2015). 在在线社交网络中部署监测点以及时检测错误信息。*2015*年*IEEE*国际通信会议 *(ICC)*，1152–1157。https://doi.org/10.1109/ICC.2015.7248478 Zhang, H., Kuhnle, A., Smith, J. D., & Thai, M. T. (2018). 不确定条件下的抗争：遏制错误信息并推动真相传播。*2018*年*IEEE/ACM*社交网络分析与挖掘进展国际会议 *(ASONAM)*，266–273。https://doi.org/10.1109/ASONAM.2018.8508402 Zhang, H., Kuhnle, A., Zhang, H., & Thai, M. T. (2016). 在为时已晚之前检测在线社交网络中的错误信息。*2016*年*IEEE/ACM*社交网络分析与挖掘进展国际会议 *(ASONAM)*，541–548。https://doi.org/10.1109/ASONAM.2016.7752288 Zhang, H., Zhang, H., Li, X., & Thai, M. T. (2015). 在有效提高社交网络认知度的同时限制错误信息传播。载于《计算社会网络》，*9197*，35– 47。施普林格国际出版。https://doi.org/10.1007/978-3-319-21786-4_4 Zhou, P., Xu, J., Qi, Z., Bao, H., Chen, Z., & Xu, B. (2018). 利用分层选择性注意力进行关系抽取的远程监督。《神经网络》，*108*，240–247。https://doi.org/10.1016/j.neunet.2018.08.016