

实用指南：预先揭穿 虚假信息

本文由以下各方联合创作：

剑桥大学

BBC Media Action

Jigsaw

简介

“预先揭穿”是一项日趋重要的技术，旨在通过先发制人的方式建立大众对虚假信息的抵抗力。本指南专为有意抵御虚假信息的从业者制作，帮助他们抗击误导大众和操控人心的信息。文中讲解了预先揭穿的基础知识，目的是将学术研究成果转化为实用方法指南，支持先前并不了解行为心理学的团体和个人部署自己的预先揭穿干预措施。

本文是剑桥大学、Jigsaw (Google) 和 BBC Media Action 的合作成果。剑桥大学社会决策实验室 (Social Decision-Making Lab) 一直处于预先揭穿方法开发工作的前沿，以免疫理论 (Inoculation theory) 为依据，致力于帮助大众建立应对虚假信息和造谣信息的抵抗力。

[Jigsaw](#) 是 Google 的一支团队，与包括剑桥大学在内的全球顶尖大学合作，在各种环境下测试预先揭穿方法，从而了解这种方法的优势和局限性。

[BBC Media Action](#) 是 BBC 的国际发展慈善机构，该机构在积极调整和测试预先揭穿方法的使用，在其工作覆盖的多个国家和地区中将其用作一种[应对信息失序的策略](#)。

本指南的创作人员如下 (按组织分组，各组中的姓名按字母顺序排列)：Mikey Biddlestone、Trisha Harjani、Sander van der Linden、Jon Roozenbeek (剑桥大学)；Alasdair Stuart (BBC Media Action)；Beth Goldberg、Meghan Graham、Mari Iwahara、Bomo Piri、Peter Weigand、Rachel Xu (Jigsaw)。

如您对本指南中的研究有任何问题或疑虑，请联系[剑桥大学社会决策实验室](#)的 [Jon Roozenbeek](#)。

如您需要进一步了解 BBC Media Action 应对信息失序的工作 (包括预先揭穿方法)，或者希望联系 BBC Media Action 咨询其他任何事宜，请发送电子邮件至 [Alasdair Stuart](#)。

引用本文：Harjani, T., Roozenbeek, J., Biddlestone, M., van der Linden, S., Stuart, A., Iwahara, M., Piri, B., Xu, R., Goldberg, B., & Graham, M. (2022). A Practical Guide to Prebunking Misinformation.



目录

第 1 部分：为什么要预先揭穿

1.1 现状	第 5 页
1.2 预先揭穿的作用原理	第 6 页
免疫理论	
可以揭穿哪些类型的虚假信息？	
虚假信息言论	
虚假信息技术	
1.3 形式和技术考虑因素	第 11 页
1.4 预先揭穿的局限性	第 13 页
1.5 未来探索领域	第 15 页

第 2 部分：如何预先揭穿

2.1 应在何时采用这种技术？应由谁实施？	第 19 页
2.2 方法入门	第 21 页
第 1 步：选择主题	
第 2 步：选择受众群体	
第 3 步：设定目标	
第 4 步：选择一种方法：	
“基于问题”与“基于技术”	
第 5 步：选择形式	
第 6 步：设计干预措施	
2.3 衡量成效	第 26 页
2.4 内容素材考虑因素	第 29 页
2.5 注意事项	第 31 页
2.6 预先揭穿核对清单	第 32 页

01: 为什么要预先揭穿

1.1 现状	第 5 页
1.2 预先揭穿的作用原理	第 6 页
免疫理论	
可以揭穿哪些类型的虚假信息？	
虚假信息言论	
虚假信息技术	
1.3 形式和技术考虑因素	第 11 页
1.4 预先揭穿的局限性	第 13 页
1.5 未来探索领域	第 15 页

1.1

现状

网络虚假信息的激增给公共安全和现代民主带来了严重威胁。在现实生活中，它们已经造成了许多严重的后果：在关于新冠的造谣信息猖獗的地区，尽管疫苗供应充足，但新冠致死率要高于相邻地区。¹

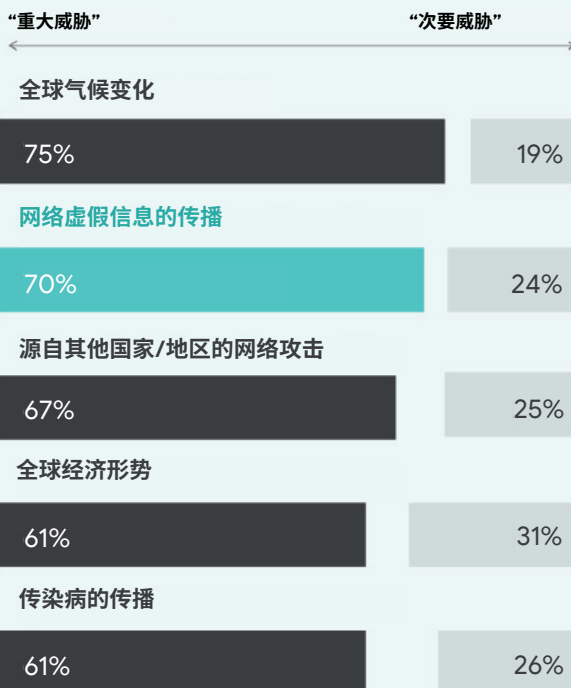
与此同时，在皮尤研究中心 (Pew Research Center) 于 2022 年面向 19 个国家/地区发起的一次投票中，他们发现有 70% 的受访者将虚假信息评为其国家/地区的主要威胁，仅次于作为全球威胁的气候变化。²

打击虚假信息任重而道远。人们已经设计出许多干预措施来尽可能减少虚假信息和造谣信息的传播和使用，³包括但不限于信息打假、及时提醒、自动贴标签和提高信息素养。⁴但仍有各种困难导致这些措施无法大规模取得成功，包括实践中的困难和概念上的困难。一种比较突出的方法通常称为“信息打假”，针对的是已经传播开来的虚假信息。在人们看到虚假信息之后再为其展示纠正信息（比如事实核查）可能有效，但这往往非常耗时、成本高昂，也很难达到必要的部署速度。虚假信息可能会非常顽固，因为一旦接触到虚假信息，人们往往会不断将这类信息用作依据，即便在信息被纠正后也不改变。⁵除此之外，事实核查也很难凑效，因为其参与度历来不高：对 Facebook 上 5 万篇信息打假帖子的调查发现，在接触过虚假信息的受众群体中，仅有少数人与事实核查帖子互动。⁶

因此，研究人员已经在探索提前防范虚假信息的方法，目的是避免此类信息在人们的头脑中形成固有认知。先发制人的方法关注人们接触虚假信息之前的阶段，通常称为“预先揭穿”。预先揭穿干预措施分为许多不同的类型，但通常都以免疫理论为依据。预先揭穿宣传信息在人们遇到虚假信息之前提供警告和反驳论点，从而帮助大众筑起对虚假信息的心理防线。应该注意，免疫法通常在个人接触到虚假信息之前最为有效，但如果某人在接触过虚假信息后并未完全相信，那么仍然有可能实现免疫。（请参见“[2.2: 定义受众群体](#)”中的进一步讨论）。

人们认知中的主要全球威胁

数据来源：皮尤研究中心



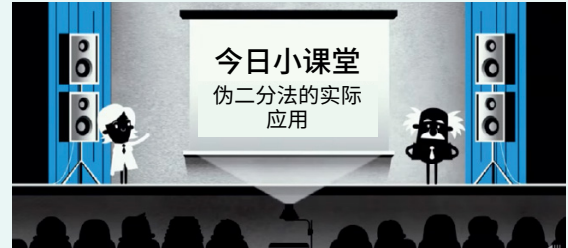
1.2

预先揭穿的作用原理

预先揭穿关注大众通常被网上的内容操纵、误导的方式，而不会针锋相对地质疑虚假信息传言，也不会指手画脚地告诉人们应该相信些什么。

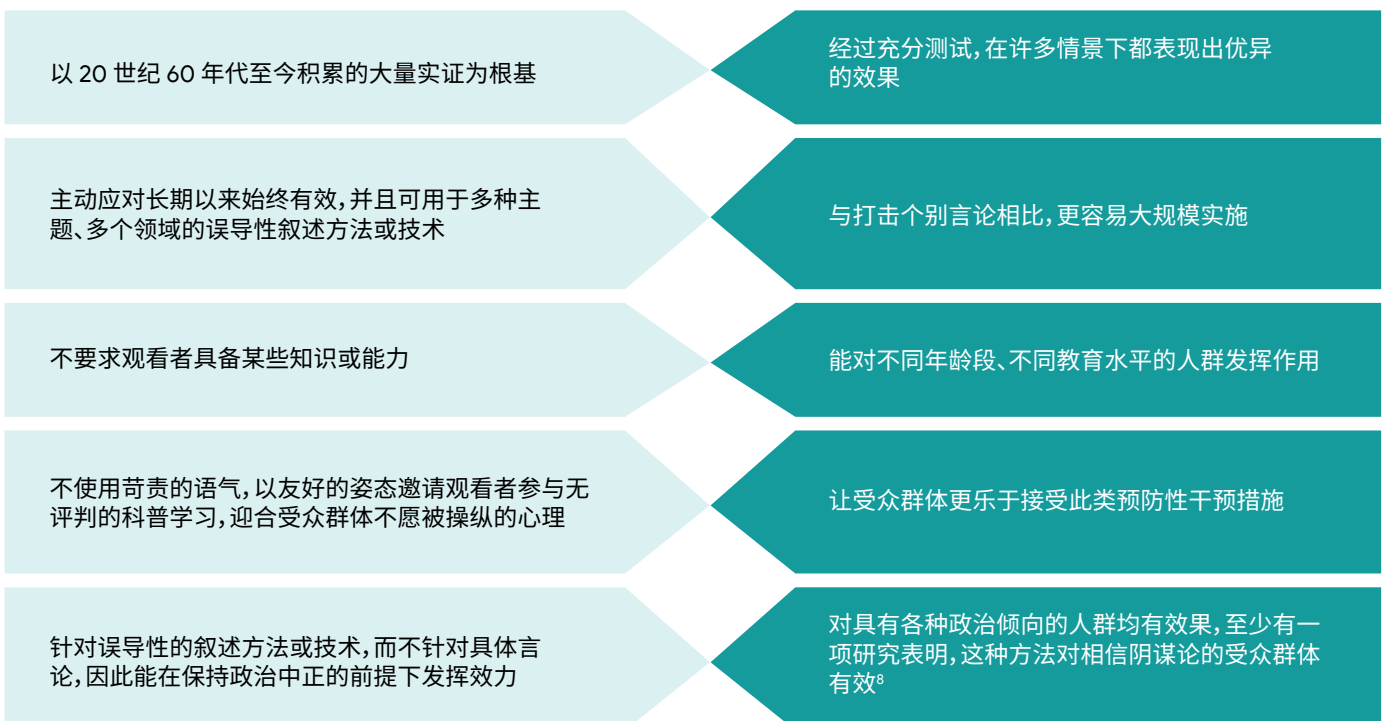
基于虚假信息的信念根深蒂固，极难动摇，有一个日渐兴起的研究领域着力从起点出发，帮助人们避免被虚假信息说服。其中一种方法借鉴了生物医疗科学领域的理论。**免疫**让人们预先接触“剂量”经过弱化的虚假信息，教他们学会识别和驳斥误导性声明，以此来形成对虚假信息的抵抗力。预先揭穿（又称“**心态免疫**”）这种方法提前教会人们识别和抵制操控人心的信息，从而做到防患于未然。可以证明，预先揭穿能有效帮助广大群众形成对误导性信息的抵抗力，包括持有不同政治立场的人们。⁷这种技术关注大众通常被网上的内容操纵、误导的方式，而不会针锋相对地质疑虚假信息传言，也不会指手画脚地告诉人们应该相信些什么。这种方法通常着眼于教育启发，不会妄加评判，也不会苛责他人，因此能在广泛的受众群体中引起共鸣。预先揭穿通常侧重于分享更高层次的技术和叙述方法，致力于赋能个人，让他们能辨别自己是如何被操控的。预先揭穿不会假设受众群体已经具备某个主题的能力或知识，因此广泛适用于不同年龄段、不同背景环境的人群。例如，首开先河的预先揭穿游戏“**Bad News**”旨在供教育工作者使用，帮助年轻的在校学生了解如何辨认心怀叵测者使用的种种伎俩。

示例：预先揭穿操纵技术（伪二分法）



预先揭穿的一个例子是揭穿虚假信息常用的“伪二元法”伎俩，也就是仅给出非黑即白的选择——即便是在现实生活中有众多选择。[观看视频 >](#)

预先揭穿的优点



免疫理论

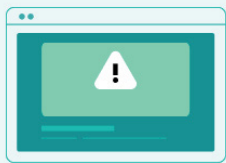
预先揭穿基于 20 世纪 60 年代开发的免疫理论,其开发者为社会心理学家 William McGuire,目的是用作“防洗脑的心理疫苗”。⁹

医学中的疫苗能形成生理抵抗力,从而应对未来的感染;与此相似,免疫方法能形成心理抵抗力(类似于抗体提供的免疫力),应对未来的心理操纵。

过去 60 年间的研究表明,免疫方法在多种文化环境、多种主题(包括环境、公共健康、危机管理、动物权益等)中均切实有效。^{10、11、12、13、14}

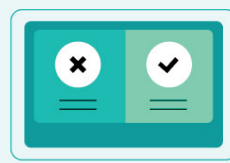
近期,学术界已经证明,免疫宣传信息可以降低网上虚假信息和极端主义宣传的影响。¹⁵

实际上,免疫涉及到两个部分:



1. 预先告诫

提前给出告诫,提醒观看者在近期有可能遇到误导性的信息,从而激活观看者的心理防御机制,对于心怀叵测者说服他们的企图保持一份警惕之心。



2. 先发制人的批驳

有效的批驳可为观看者提供工具,帮助他们抵挡将来可能遇到的误导性信息。除了提前提供反向论证之外,包含“微剂量”或经过弱化的虚假信息示例也很有用,这让人们将来可以更轻松地分辨此类信息。

基于这种基本结构的预先揭穿宣传信息可以加强观看者的心理抵抗力,帮助他们更好地应对未来的游说攻击。¹⁶在 [1.4 预先揭穿的限制性](#) 中进一步探讨了这种抵抗力的局限性。

可以对哪种类型的信息进行预先揭穿？

有两种主要的预先揭穿形式可超越具体虚假信息主张，在更高的层面上应对虚假信息。它们均针对不同类型的虚假信息：

1. 虚假信息言论

在网络世界中遇到的虚假信息通常表现为对特定主题的主张或意见。但个别虚假信息主张往往会被纳入更广泛的叙述之中。基于问题的预先揭穿可应对更广泛且长期存在的虚假信息叙述，而不局限于具体主张。

应对个别虚假信息主张需要耗费大量时间，而且较为被动，如果将更广泛的叙述作为预先揭穿的对象，就能一次性打破许多主张的依据，能有效地帮助观看者建立心理抵抗力，让他们更好地应对从同一种虚假依据中衍生出来的新主张。

示例：人类与气候变化

请看如下叙述：

“31,000 位科学家在请愿书上签名：对于人类排放温室气体是否会对地球气候产生负面影响，许多气候科学家持有不同的看法。”

有一种范畴更广的误导性言论称，科学界对于人类造成的气候变化并未达成共识，气候变化是地球自身自然周期的一部分，而刚刚那条叙述就是这种言论下的诸多谎言之一。预先揭穿可针对这种范畴更广的言论，警告大家在有人企图质疑人类行为助长气候变化这一科学共识时，应该始终保持一分怀疑之心，但不必就有关请愿书的具体主张进行辩驳。

2. 虚假信息技术

基于技术的预先揭穿专注于散布虚假信息的战术。用于操纵和影响网民的信息千变万化，但长期以来，用于产生误导的技术往往是一样的，不论主题如何都是如此。下面简单介绍了几种常用的战术。

示例：常见的虚假信息技术

下一页详细介绍了几种常用的技术。



常见的虚假信息技术

技术	示例
<p>假冒行为</p> <p>为了提高可信度，冒充其他人或其他组织传播信息。</p>	<p>“NASA 承认，气候变化是由地球的太阳轨道变化自然引发的结果，与人类活动因素无关。”</p> <p>解析：这个例子假借 NASA 之名来提高言论的可信度，但 NASA 从未发表过这样的言论。</p>
<p>情感操控</p> <p>巧言令辞，利用强烈的情感语言来激发反应，包括恐惧和愤怒。</p>	<p>“这家航空公司为乘客做的事情好到让人流泪 — 太暖心了。”</p> <p>解析：这个例子展示了如何为激发情感回应而刻意地遣词造句，从而吸引点击量和分享量，同时减少批评。</p>
<p>两极化</p> <p>夸大两个群体之间的差异，刻意树敌，使用“我们”与“他们”这样的对立性措辞。这种技术有时会被政治团体所用，但也能用于其他很多情景。</p>	<p>“人民党：不要相信工人党的谎言。他们过去说过废除学生贷款，可现在负债的人一直在增加。”</p> <p>解析：这个例子将另一个党派描述为说谎者，使用了不友善的“他者化”语言。</p>
<p>阴谋论</p> <p>使用不同的方式解读传统新闻中的事件，着重宣扬有一小群人，通常称某个心怀恶意的秘密精英团体，在控制这些事件。</p>	<p>“打疫苗其实就是植入芯片，那些亿万富翁要通过这个来跟踪我们！我们的身体究竟在被谁控制？”</p> <p>解析：这个例子提出未经证实的主张，宣扬阴谋论，称人们已经失去了对自己的控制权，并指向某个秘密团体，本例中是亿万富翁。</p>
<p>人身攻击</p> <p>人身攻击(拉丁语为 Ad hominem)将矛头对准提出论点的个人，从而让关注点偏离论点的实质，转向个人的细枝末节。此类细节或许有意义(例如，通过细节可以看出那个人并不可信)，但也可能毫不相关，仅仅是用作分散注意力的战术。</p>	<p>“王某脾气暴躁，明显还有人格障碍！我们不能让一个疯子掌权。”</p> <p>解析：这个例子罔顾领导者的政策或领导决策，向领导者的特征发起攻击。</p>

常见虚假信息技术(续)

技术	示例
<p>伪二元法</p> <p>这是一种逻辑诡辩,在辩论或情景中让情况看似仅有两个方面或者两种选择,但实际上还有更多方面。</p>	<p>“要么支持能源抗议,否则你就是不相信正义。”</p> <p>解析: 这个例子将两种意见定为对立面,换句话说,也就是在“支持能源抗议”与“相信正义”之间画上了等号,但人们完全有可能既支持能源抗议又不相信正义、既不支持能源抗议又相信正义,或是持其他多种立场。</p>
<p>假平衡</p> <p>将辩论说成是两种势均力敌的对立观点,但事实上,其中一种论点的支持证据要充分得多。</p>	<p>“专家就地球的形状展开了争论。虽然科学家 Reece Chow 已经发现地球是球形的,但专家 Rene Paul 认为地球是平的。”</p> <p>解析: 在这个例子中,尽管科学家一致认为地球是圆的,但其中提出了一位支持地平论的“专家”,营造出这种论点获得较多支持的表象,但事实上并非如此。</p>

1.3

形式和技术考虑因素

预先揭穿干预措施有主动和被动两种形式，主动是指人们与问题或提示互动，从而了解编造虚假信息的过程，而被动是指人们观察预先揭穿信息。

从可扩展性、有效性、长久性、成本和在线互动性来讲，每一种方法都有自己的优缺点。大体上说，观看者参与一次干预的时间越长、参与度越高，免疫的效果就越大，持久性也越好。



合理匹配内容与平台

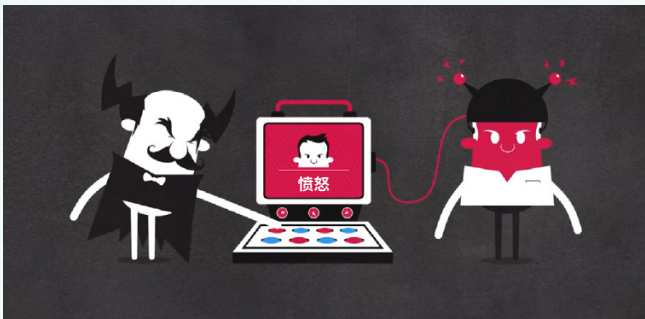
为一个平台(例如: YouTube、TikTok、网站)而设计的内容可能并非总能轻松在其他平台间共享,因此在选择内容形式时,务必仔细考虑预期发布位置。

“被动”预先揭穿

这些干预措施为观看者提供抵制虚假信息所需的全部信息,而且除了消化处理信息之外,不要求他们有其他任何主动参与。例如,解析一种技术如何操控人心的视频就属于被动方法。到目前为止,所研究的被动形式包括**文字、图形和视频**。^{17、18、19、20}

从制作的角度看,被动预先揭穿干预措施可能要简单一些。例如,基于文字的预先揭穿干预措施(比如一系列弹出消息)相对来说可以比较轻松地在社交媒体上实施。但这种形式的沉浸度和交互性都不太高,因此与游戏等互动性更高的形式(也就是主动形式)相比,效果要更小,影响的持久性也略短。²¹

示例

视频示例：
伪二元法

这个示例视频由 Jigsaw 和剑桥大学联合制作,其中使用了与文化相关的例子帮助观看者了解并辨别虚假信息散布者对伪二元法的使用。[观看视频 >](#)



信息图示例新冠阴谋论

由联合国教科文组织 (UNESCO) 制作的这张信息图以新冠为例解释了阴谋论。²²

“主动”预先揭穿

另外一种预先揭穿方法就是主动法,这要求个人采取行动、做出有助于其保留信息的选择,并且更深入地与其看到的内容互动。迄今为止所研究的主动方法以游戏为主。^{23、24}

游戏要更有沉浸感,可以帮助人们免疫虚假信息中用到的多种操纵技术,但也要求观看者投入更多时间和专注力,而这可能减少参与人数。这类内容的制作投资也要更高,但目前已有一些影响力较大的游戏大规模实施,例如《Go Viral》(见下文)。

基于音频的预先揭穿(比如通过无线电或 WhatsApp 等聊天应用广播预先揭穿信息)是一种尚未得到充分利用的媒介,可从进一步的研究中受益(“1.5 未来探索领域”提供了更多相关信息)。



选择形式

从可扩展性、有效性、长久性、成本和互动性来讲,每一种预先揭穿形式都有自己的优缺点。大体上说,受众群体接触预先揭穿信息的时间越长、参与度越高,预先揭穿的效果就越大、持久性也越好。

游戏示例:



BAD NEWS

这是首开先河的预先揭穿游戏。它由 DROG 与剑桥大学联手制作,采用了选择题式的浏览器游戏形式,玩家扮演虚假新闻制作人的角色,通过六个关卡了解和辨别和模拟六种虚假信息技术(例如恶意评论、阴谋推理、冒充他人)。在这款游戏发布后,又有人设计了其他一些具有类似设定的游戏。

[查看游戏 >](#)



HARMONY SQUARE

游戏的背景是一个祥和的社区,池塘里悠闲的天鹅和一年一度的菠萝披萨节让这里远近闻名,玩家要扮演“首席造谣信息官”的角色,分化 Harmony Square 社区里的人们,并在政治选举期间发起恶意评论活动。[查看游戏 >](#)



GO VIRAL!

在这款游戏中,玩家要进入一个网络回音室,那里有许多关于新冠疫情的虚假信息。玩家可以通过三个关卡了解到情绪操纵语言是如何使用的、如何借虚假专家之名给虚假信息增加可信度,以及如何利用阴谋论来播下怀疑的种子。迄今为止,这款游戏的展示次数已超过 2 亿次。²⁵ [查看游戏 >](#)

1.4

预先揭穿的局限性

事实已经证明，预先揭穿在保护人们抵制操纵企图方面尤为有效，我们也已经知道这种方法的一些局限性，不过它还可能存在其他一些未知的局限之处，有待进一步探究才能充分了解。



可扩展性

预先揭穿在广泛的受众群体间都有着经证实的有效性，但在分享涉及到不同类型的虚假信息、受众群体和平台的预先揭穿信息时，从业者应该倍加谨慎，开展小规模测试。如果应用的规模过大，涉及到不具备适当特征或本地背景的受众群体，就会造成参与度较低或过度简化，进而降低效力。另一方面，如果预先揭穿仅针对一条言论或一个问题，消息的相关受众群体就过于狭窄，这会限制可扩展性（例如，向对疫苗心怀疑虑的受众群体有针对性地发送预先揭穿疫苗相关虚假信息的内容）。

内容限制

并非所有预先揭穿都具有同样的可扩展性。有些言论可能包含多项主张，但仍然高度特定于虚假信息的某个主题或某个领域。基于技术的预先揭穿可应用于多种主题，因此与基于问题的预先揭穿相比，可扩展性更高，可覆盖多种类型的虚假信息。

而基于问题的预先揭穿有可能就特定主题和言论发挥更深入的防御效果。在选择方法时，充分了解每种方法的优缺点非常重要。

过度简化的风险

扩大预先揭穿规模的一大挑战源自用户与在线内容交互的方式。若要吸引人们与社交媒体上的内容互动，内容必须尽可能简短，以简明扼要的方式向用户传达信息。随着新媒体平台的兴起，这一要求愈加严格，简短的内容很难加入为保证效果而必需的细节。以简短、互动的方式呈现预先揭穿的三个组成部分尤为艰难。过度简化信息可能导致其失去效果，造成受众群体迷惑不解，甚至有进一步传播虚假信息的风险。

平台限制

不同平台鼓励不同的受众群体交互，在多个渠道中使用相同的内容素材形式可能会限制效力。社交媒体平台采用专门的设计，因此特定的一些内容形式非常适合此类平台，但这些内容在其他平台上可能表现不佳。此外，不同的平台上可能存在不同的虚假信息言论，使用不同类型的信息传播者，比如网红，因此很难同时针对多种平台优化信息。

示例：TRUTH LAB 系列



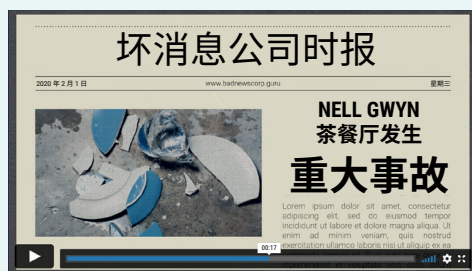
例如，Roozenbeek 等人制作了 5 个 30 秒到 90 秒的简短动画视频，在 YouTube 上向参与者展示相关宣传内容。²⁶他们发现，这些视频让人们能更好地发现操纵企图、区分可信和不可信的内容，并减少用户分享虚假信息的决定。[观看视频 >](#)



效果持续时长

通过教育干预措施掌握的认知会随着时间的推移逐渐消退。研究表明，通过实施类似于“疫苗加强针”的干预措施 — 后续通过简短的提醒来再次预先揭穿虚假信息，就能抵消这种消退效应。这可能涉及到重复最初的预先揭穿，或者提供一个精简版，汇总其中的要点。^{27、28}

示例：加强巩固视频



Jigsaw、剑桥大学和布里斯托尔大学的研究人员共同制作了一系列的加强巩固视频，提醒人们在先前较长版本的预先揭穿视频中看到的内容，就好比一剂数字世界里的“加强针”。实验发现，预先揭穿视频最初可为人们提供大约 10 天的心理抵抗力，第 10 天再次发布 30 秒的加强巩固视频非常有用，可以将抵抗效力延长到至少 30 天。[观看视频 >](#)



非预期效果

在制定预先揭穿干预措施时，从业者应该保持警惕，并考虑到受众可能对宣传信息产生的负面反应。逆火效应（表示干预措施无意中加深了人们对虚假信息的信任）似乎并不是引发担忧的主要理由，²⁹有些人可能会抗拒一切干预措施。例如，在预先揭穿中，特别抗拒他人影响和改变自身态度的人可能就不会喜欢预先揭穿宣传信息。

举一个例子，一次研究发现，关于白人至上主义言论的预先揭穿宣传信息对于极右翼派人群没有效果，由此可以看出，持有强硬观点的人抗拒此类内容。³⁰在设计和分析预先揭穿宣传信息时，务必考虑到受众群体中的离群个体。

1.5

未来探索领域

免疫干预措施早在 20 世纪 60 年代就已经出现,但时至数字时代,预先揭穿干预措施仍处于积极的研究和发展之中。为了充分理解如何才能最好地在全球范围内实施预先揭穿,需要更多的投资、研究和测试。



了解全球状况

虚假消息已成为全球性的问题,但关于预先揭穿的研究大多在北半球开展,例如美国、英国和欧洲。需要开展更多研究,以了解如何在世界各地的其他国家和地区最好地应用预先揭穿,如何让这种方法适应不同国家和地区的环境。

在为预先揭穿这样的方法扩大应用规模时,语言、人口统计学特征、地理位置和文化多样性等因素都会影响到成败,因此需要结合当地实际来解读这些因素。



应对封闭式应用

了解封闭式消息传递平台(如 WhatsApp 和 Telegram)中的虚假信息传播尤为困难。这类技术的设计初衷就是进行私密交流,因此从本质而言,很难了解其中的趋势和习惯。迄今为止,对于如何针对此列信息传播空间运用预先揭穿技术的研究依然较为有限。

比较有价值的做法是开展测试,确定哪些类型的预先揭穿内容可以最好地吸引封闭式聊天应用的用户,他们可能会选择与他人分享哪些形式的内容(分享能让干预措施的影响力成倍增长),以及这会对封闭式消息传递空间中虚假信息的效力和传播产生怎样的影响,例如,运用免疫理论的内容能否降低用户对亲朋好友分享的误导性信息或虚假信息的相信程度?能否降低用户与自己的联系人分享此类内容的可能性?

示例:印度的坏消息



近期的一项研究发现,《Bad News》游戏在印度发挥了很好的预先揭穿效果,在玩过这款游戏后,参与者在评价虚假新闻的可靠性时会给出更低的评级。³¹

BBC Media Action 正尽力改编预先揭穿视频,并通过在北非拥有高覆盖率的现有社交媒体渠道分发这些视频。后续将评估此次宣传活动的有效性,结果预计于 2023 年初期分享。



形式和宣传信息长度

迄今为止的预先揭穿研究主要关注文字、视频和互动式游戏。但大众还会通过其他许多形式消费信息，需要开展更多研究以了解如何针对不同的形式（例如音频或表情包）有效调整预先揭穿方法。

基于音频的干预

在某些背景下，受众群体依然主要依靠音频形式获取和传达信息，例如，在非洲的一些农村社区中，社区无线电广播仍然是主要信息来源；再比如，考虑到数据流量费用，WhatsApp 用户更倾向于在群组中使用音频内容，而非视频内容。

开发基于音频的预先揭穿方法，并通过无线电广播节目或聊天应用测试此类方法的传播效果，这是有待进一步探索并能在此类环境下带来极大好处的一个领域。

碎片化信息

使用在线游戏和动画短视频（每个视频时长约为 2 分钟）形式的预先揭穿干预方法旨在通过数字渠道分发，其影响力已经得到了实践的印证，一部分数字世界中的受众群体更青睐时长更短（例如 30 秒或更短）的数字内容，和/或转为使用支持此类内容的平台（例如 TikTok）。一些初步工作已经证明了 30 秒预先揭穿视频的有效性，但还需要开展进一步的工作，探索预先揭穿能否适应此类“碎片化”数字媒体内容。³²

较长形式的叙事媒体

电视剧、广播剧、真人秀等较长形式的节目旨在触达广大受众群体。有许多令人信服的证据（包括 BBC Media Action 的工作成果）表明，由本地人经过全面研究精心制作的叙事媒体内容可以大规模吸引受众群体，引发社会变革和行为变化。在资源稀缺的环境中运用故事讲述的力量大规模应对发展问题的做法在不断增加，实证也越来越多，其中包括：艾滋病病毒/艾滋病、性别暴力、性别规范、社会凝聚力、卫生、避孕药具的使用以及儿童生存。^{33、34、35、36、37、38、39}

BBC Media Actio 的经验证明，讲故事的形式非常有用，能够通过非对抗性的方式提出敏感问题，在主要掌权者可能直接造成虚假信息传播的社会中，这尤为有用。但迄今为止尚且无人尝试过将预先揭穿方法整合到此类内容中。一种创新的思路是在考查能否利用影视剧中的情节，向影视剧受众群体传达预先揭穿宣

传消息，让他们透过剧中角色的遭遇接触到预先揭穿警告。这种方法有望覆盖更广泛的受众群体，重点在于，这还有可能覆盖不太有可能接触在线游戏或看到数字免疫理论内容的更多弱势人群，例如不太常使用社交媒体的老年人。



信息传达者的角色

到目前为止,围绕预先揭穿展开的研究大多探索的是预先揭穿宣传消息的内容和形式(本文中提炼了其基本元素),以及这些不同要素会给效果造成怎样的影响。但就目前而言,探讨预先揭穿的传达者或发声者会对其效果产生怎样的影响的研究堪称寥寥。

对于不同来源的信息,人们会产生不同的反应,专业能力、权威程度、信任度和偏见都会影响到我们对于某信息传达者所传达的信息的认知和内化。近期免疫文献的评论已经开始关注信息来源可信度对于抗拒态度的影响。⁴⁰还需要开展更多研究,以了解哪些参与者(例如社交媒体网红、公众人物、权威机构、新闻主播等)在哪些背景下能更有效地面向哪些不同受众群体传达预先揭穿信息。



其他研究领域

为了与不断发展变化的信息环境保持同步,预先揭穿领域也在快速发展。随着研究的推进,虚假信息散布者也在不断调整适应、发展变化,谋求新的操纵人心之道。其他研究领域必定会应运而生,研究人员和从业者等相关认识必须不断推动知识前沿,了解如何对抗虚假信息,更好地保护我们的社会。

02: 如何预先揭穿

2.1 应在何时采用这种技术?应由谁实施?	第 19 页
2.2 方法入门	第 21 页
第 1 步:选择主题	
第 2 步:选择受众群体	
第 3 步:设定目标	
第 4 步:选择一种方法: “基于问题”与“基于技术”	
第 5 步:选择形式	
第 6 步:设计干预措施	
2.3 衡量成效	第 26 页
2.4 内容素材考虑因素	第 29 页
2.5 注意事项	第 31 页
2.6 预先揭穿核对清单	第 32 页

2.1

应在何时采用这种技术?应由谁实施?

若在保证预先揭穿发挥最大效力,最好的时机是受众群体尚未充分了解虚假言论和操纵技术,或者受众对于相关主题的立场还在左右摇摆时。一旦关于某个主题的信念固化或者极端化,预先揭穿就很难发挥实效。

在考虑将预先揭穿作为打击虚假信息的方法时,最好先核对如下条件是否适用:



可以预测到相关言论或技术

虚假信息言论和技术往往在不同主题中反复出现。通过对这些趋势进行周详分析,通常就能预测谣言传播者会将这些言论和技术用于哪些新的虚假信息。例如,选举周期、公共卫生危机和环境灾难等周期性事件都是虚假信息重灾区,在这些时候往往可以观察到一些一再出现的技术或言论。

示例:天花与新冠

长期以来,疫苗一直是虚假信息重灾区。自发明至今,一直有人指责疫苗“有违自然”,关于疫苗的不实主张经常是同一套说辞。

例如,在19世纪初,天花疫苗被谣传为会将人类变成“人和牛的混种生物”,起因就是这种疫苗使用了源自牛痘的配方。时至今日,关于新冠疫苗也有类似的谣言,宣称它会“改变你的基因”。⁴¹这种言论完全可以合理地提前预测到,因此很可能是预先揭穿的有效候选内容。



受众群体尚未完全信服

在设计预先揭穿干预措施时,受众群体的接受度非常关键。理想情况下,干预措施应在受众群体完全认同虚假信息之前接触到他们。虽然有一些证据表明,即便在受众群体接触过虚假信息之后,预先揭穿仍能发挥效果(称为“治疗性免疫”),但在受众群体尚未充分信服相关主张或言论时,预先揭穿的效果要更好。⁴²在设计预先揭穿干预措施时,应该考虑您的受众群体有哪些,他们目前对您想要预先揭穿的虚假信息的认同程度如何,当前媒体和/或政治环境如何,从而确定一种预先揭穿干预措施是否适当。

示例:科学种族主义

Jigsaw 与美国大学开展的研究发现,在美国人中预先揭穿有关白人至上主义的不实言论能有效减少绝大多数受访者对相关信息的散布者及其言论的支持程度。但对于先前已经坚信白人至上主义(根据右翼威权主义量表和社会支配取向量表等调查的衡量结果)的受访者,预先揭穿视频没有任何效果。⁴³

应由谁实施？

人们对于在线信息的信任程度不断下降，因此在预先揭穿时，您必须与受众建立稳固的信任关系，在受众间建立出色的可信度。

确保您的组织具备以下条件：



就相关主题发表权威言论的专业能力

信息世界中充斥着大量建议，有关准确性的争议不计其数。在开始预先揭穿时，请确保您具备必要、充足的专业能力，能够通过可信的方式解析相关虚假信息。如有必要，可与相关专家、学者和权威实体合作，通过这种很好的方式来证明专业能力。



与受众群体的信任和亲善程度

受众群体更有可能信任来自其信赖的分享来源的宣传信息。如果您与尝试触达的受众群体有牢固的关系，或者认为他们对您和/或您的品牌十分亲和，那么您在预先揭穿虚假信息时就更有优势。如果您担心受众群体对您不够信任，不妨考虑与他人合作，选择与相应受众群体拥有更紧密的关系的团体或创作者。



互动能力

预先揭穿不应是单向的“自说自话”。您应做好计划，安排相应资源去监控、衡量工作，并不断迭代完善。此外，在分享打击虚假信息的宣传信息之后，在与受众群体交流时应该保持谦逊的态度。

2.2

方法入门

在制作预先揭穿材料时,应该谨记五个步骤和一些注意事项:

第 1 步:

选择主题: 您要预先揭穿什么虚假信息?

干预主题基于您要打击的虚假信息,其范围可能相当广阔,从气候变化和大流行病等全球危机,到心理健康相关看法等更为个人层面的问题。

如上文“[2.1 应在何时采用这种技术?应由谁实施?](#)”一节所述,在选择主题时,应该考虑以下因素:

- 确保您具备与虚假信息、目标受众群体有关的专业能力,或者与符合这些条件的主题专家合作。
- 研究虚假信息环境,确定困扰您的受众群体的主要和盛行的言论和技术。

第 2 步:

选择受众群体: 您希望自己的预先揭穿信息覆盖哪些人群?

考虑干预措施的受众群体,试着去了解他们目前与您要分享的信息的关系,以及他们可能有兴趣听到您分享哪些信息。

如上文“[2.1 应在何时采用这种技术?应由谁实施?](#)”一节所述,在选择受众群体时,应该考虑以下因素:

- 您能否在某些技术/言论广为传播之前做出相应预测?您能否在信息环境演变之前预测到新的技术/言论?
- 您的受众群体是否已经接触过您尝试对抗的技术和/或言论?他们的信念有多坚定?

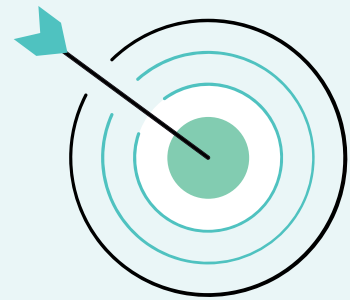
2.2

方法入门(续)

第3步:

设定目标

为干预措施设定目标：
您希望通过分享预先揭穿干预信息
取得怎样的结果？



预先揭穿干预措施可实现一系列的结果，具体可分为三大类：

1. 知识或技能

预先揭穿可以向受众群体传授新知识(例如，准确的统计数据)或技能(例如，区分虚假信息的能力)，让他们能够对抗虚假信息，并形成应对将来操纵企图的抵抗力。

2. 态度

预先揭穿能够改变受众群体对自身“防御”虚假信息的能力的态度，或者改变他们对于虚假信息散布者的看法(例如，虚假信息来源的可信度)。

3. 行为

预先揭穿能改变受众群体的行为，让他们通过不同的方式与虚假信息互动、消费或回应虚假信息(例如，减少对虚假信息的分享)。

结果和目标

此处所列的目标并不全面，您可能还希望实现其他目标。请务必尽早设定清晰的目标，确保您的组织就干预措施的目标保持一致。

您致力实现的结果将影响预先揭穿干预措施的设计方式(请参见[“2.2 方法入门”](#)和[“2.3 衡量成效”](#))。

2.2

方法入门(续)

第 4 步:

选择一种方法:
“基于问题”与“基于技术”

选择您的方法:您要预先揭穿的是某个问题还是某种技术?



基于问题的方法

基于问题(也称为“基于言论”)的预先揭穿针对范围较广、持续存在的虚假信息言论,而非针对特定主张。这让您可以直接应对多种主张的依据,从而更有效地破除虚假信息,而非通过事实核查的方法逐一破除各项主张。如需进一步了解虚假信息言论,请参见[“1.3 虚假信息言论”](#)。

基于问题的预先揭穿适合在何时使用?

如果您应对的虚假信息要求从某个主题的具体事实和解析出发做出反驳,言论预先揭穿就是理想的方法。

基于技术的方法

基于技术的预先揭穿针对的是多种主张和虚假信息言论中普遍采用的技术和策略。这种方法不对操纵人心的内容做出驳斥,而是帮助受众群体了解自己会如何被他人所操纵。如需进一步了解虚假信息传播中使用的操纵策略,请参见[“1.3 虚假信息技术”](#)。

基于技术的预先揭穿适合在何时使用?

如果有多种主张和言论普遍利用了某种技术,那么基于技术的预先揭穿就是有效的方法,能够帮助受众群体建立广泛的抵抗力,即便多次遇到虚假信息也不会盲目相信。基于技术的预先揭穿与特定虚假信息主张或言论无关,因此您的干预措施更容易保持政治中立,在偏政治化的虚假信息主题中,这非常有用。

2.2

方法入门(续)

第5步:

选择形式

哪种媒介最适合传达您的预先揭穿宣传信息？

预先揭穿宣传信息可通过多种形式传达,迄今为止的文献已经探索过的形式有:文字、音频、可视化内容、视频和游戏。

从可扩展性、受众群体互动性、效果大小、效果持续性和成本来看,其中每种形式都有自己的优缺点。“[1.3 形式和技术考虑因素](#)”中简要论述了相关内容。

整体而言,较为“主动”的方法能带来更深层面的操纵抵抗力。但互动性更高的形式(比如视频游戏)往往需要更多的时间和精力,也需要受众群体的高度认同,只有在具备这些条件的情况下,他们才会参与其中。“被动”方法则有着更高的制作和扩展效率,但需要周详设计和部署,确保发挥长久的效果。请注意,这些内容是根据迄今为止的文献资料总结得出的,不同干预措施的具体效果大小可能有所不同。



确定形式时应该考虑的问题:

- 预期受众群体已参与互动的媒体平台和形式有哪些？
- 您在制作中投入了多少时间和精力(或成本)？
- 您是否具备必要的设计能力,能够开发具有视觉吸引力的宣传信息,比如信息图、视频或游戏？
- 您希望实现怎样的可扩展能力和在线参与度? 您选择的格式能否抓住受众群体的注意力？
- 您的宣传消息是长期适用还是需要更多资源定期更新？

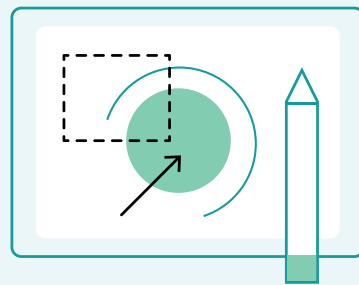
第6步:

设计干预措施

您应始终注意哪些组成部分?

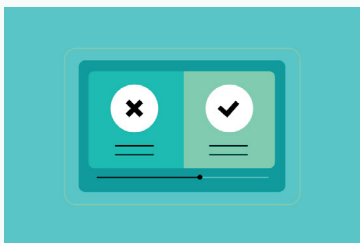
免疫宣传消息能够帮助人们建立心理抵抗力或者“心理抗体”，从而对抗将来遇到的虚假信息，这就像是疫苗在人体内建立能对抗未来病毒感染的抗体一样。但要想成为一项成功的预先揭穿措施，我们的干预需要满足一些条件。

这里展示了成功的预先揭穿宣传信息的三个组成部分：



1. 警告

提醒用户，有人企图操纵他们。



2. 先发制人的批驳

解释言论/技术及其操纵人心的方法



3. 微剂量

提供经过弱化或实际的无害虚假信息示例，这里所说的“无害”是指不会让您的受众群体产生极端化意见或困扰，也不会重申虚假信息



谨怀谦逊之心

有些时候，信息环境瞬息万变，危机期间尤其如此（例如新病毒爆发时）。如有可能，请承认您的解析和辩驳存在局限性，如果信息仍然在发展变化，也要坦诚透明地加以说明。

在设计干预措施时，您有权发挥创意，但确保包含这些关键组成部分对于保持科研操守非常重要。如需有关内容素材的其他指导，请参见“[2.4 内容素材考虑因素](#)”。

2.3

衡量成效

在设计预先揭穿后，您该如何了解它能否成功达成目标？

制定衡量计划非常有用，可以帮您了解干预措施是否达成了预期目标，具体达成程度如何。衡量干预措施的影响可以提供有用的反馈信息，既有助于未来的预先揭穿工作，也能给其他从业者提供实用的借鉴信息。为了衡量成效，您可以遵循三个基本步骤：

第 1 步：

定义关键指标

您选择的指标应该与希望达成的目标直接关联。

如“[2.2 设定目标](#)”所述，常见目标有改变受众群体的**知识/技能、态度和/或行为**。对应于这些结果，一些常见的指标包括：

基于知识或基于技能的结果

- 分辨虚假信息技术的能力
- 区分虚假信息言论的能力
- 区分信息真伪的能力

基于态度的结果

- 对虚假信息辨别力的自信心
- 对信息来源可靠性的信任度
- 在看到虚假信息后产生的情绪 (例如，愤怒、恐惧)
- 相信阴谋论的倾向

基于行为的结果

- 对虚假信息的消费情况 (例如，用于浏览虚假信息来源的时长)
- 与虚假信息的互动情况 (例如评论)
- 虚假信息的分享
- 对虚假信息的支持 (例如点赞)

您可以为干预措施设计自己的指标，无论最终决定使用怎样的指标，都应确保其能够适当、准确地衡量您计划实现的目标。建议结合使用多种指标来衡量目标。



确保指标与宣传信息相符

如果您采用基于问题的预先揭穿方法，指标应该包含与预先揭穿宣传信息中针对的主题 (或问题) 有关的问题。

类似地，如果您采用基于技术的干预措施，在衡量结果的问题中应考查观看者对于预先揭穿中所针对的技术的知识、态度或行为。

第 2 步:**收集数据**

在制定指标之后,您需要哪些数据来衡量这些结果?例如,如果您想帮助某人建立免疫力,学会判断伪二元法,那么哪些信息能帮您判断对方已经知道了伪二元法究竟是什么?

在文献中,研究人员通常使用三种方法之一或其中不同方法的组合来收集数据,并衡量预期结果:

1. 任务

任务通常用来考查接触过您的干预措施的个人的知识、技能或特征。这可能简单到一道调查问卷问题,例如,请他们正确辨认一个例子中出现的操纵策略。

2. 自行报告的回复

自行报告的回复使用调查问卷收集,在受调查者与您的干预措施互动之前和/或之后提出一些问题。这可能涉及到他们在接触您的干预措施之后产生的态度或意图。例如,在回复衡量态度变化的自行报告的调查问卷时,受访者可能需要使用从 1 (“完全不可靠”)到 7 (“非常可靠”)的李克特量表评价某示例社交媒体内容的可信度、可靠性、准确性等。^{44、45、46、47、48、49、50、51}

3. 行为观察

行为观察是指在某人与您的干预措施互动之前、期间和/或之后,您收集反映其行为的数据。例如,您可以从某个特定社交媒体平台收集数据,并评估一组用户分享了多少虚假信息。行为数据是衡量真实影响的最直接的指标,但行为数据的获取难度很大,通常要求具备您开展调研的平台上的数据访问权限,或者需要开展大量计算工作,从该平台上爬取数据。

由于对行为数据的访问能力存在局限性,许多学术研究都使用**自我报告式调研替代行为数据**,请受访者提供自我报告的行为意图或判断,以此来表示其行为。例如,要衡量行为意图,您可以请受访者自我报告他们是否会分享一段信息。

**数据访问**

收集这些数据的方式取决于您在什么平台上部署了预先揭穿干预措施。

例如,假设您使用的是社交媒体平台,那么可以通过跟进调查问卷(如果可用)来收集数据。

如果您选择使用自己的平台,那么或许可以访问行为数据(例如,某人是否点击了某个虚假信息的链接)。

第 3 步：

分析数据

在获得数据之后，您要如何知道预先揭穿干预措施是否影响到了关键指标？

数据分析的复杂程度不一而足：

1. 事后衡量(最简单)

在受众群体与您的预先揭穿干预措施互动之后，收集有关预期指标的数据。这可以帮助您了解在衡量之时，受众群体对虚假信息的抵抗力如何，但您无从得知这是不是预先揭穿工作的成果。

2. 事前与事后比较

在受众群体接触您的干预措施之前和之后，收集其在关键指标上的表现数据。这样您就能观察到在他们看到您的干预内容之后，表现发生了怎样的变化，或许有助于获得预先揭穿有效性的一些数据洞见。但还可能有一些其他因素影响到结果的变化。如果没有对照组，您就无从确定您的预先揭穿是不是其知识/技能、态度或行为变化的主要推动因素。

开展随机对照试验(确保科学可靠性)

这是衡量干预成效的最严格、最科学的方法。[此处](#)提供了一份随机对照试验(RCT)的入门指南。请注意：若要真正了解预先揭穿干预措施是否确实有效，唯一的方法就是开展合理的随机对照试验和数据分析。但也要注意，许多组织可能没有能力开展全面的统计分析，我们也在这份指南中介绍了一种更轻量级的替代方案。

2.4

内容素材考虑因素



语气

在撰写预先揭穿宣传信息时，应该确定适合您的信息和受众群体的语气（例如，严肃、幽默、正式、随意、科普等）。合适的语气取决于您的组织与目标受众群体之间的关系，以及您所针对的主题。

在确定语气时，考虑怎样能抓住受众群体的注意力，以及最适合传达您的宣传信息的方法。

有许多预先揭穿干预措施都使用有趣的解释和例子，让幽默自始至终地抓住受众的注意力。^{52、53}有些干预措施利用叙述性的讲故事方法，更好地解析预先揭穿的一个组成部分。但还有许多主题可能不适合用幽默的语气去讲述（比如涉及到人们苦难经历的一些事件）。请合理运用判断力，充分了解受众群体，从而确定怎样的语气最有可能唤起他们的共鸣。



有效性与适当性

在选择语气时，务必仔细考虑对于内容素材的适当性和有效性。如果考虑不够周到，内容很有可能因为错误的原因而爆火。

示例： 轻松活泼的预先揭穿



这段预先揭穿视频由 Jigsaw 和剑桥大学联手设计，旨在解释人身攻击的策略，通过卡通反面角色来帮助观看者辨别网上的操纵技术。 [观看视频 >](#)

示例： 严肃的预先揭穿



这段由 Jigsaw 与 Demagog 联手制作的视频的风格要更为严肃，由演员扮演的几个好友相聚在一起，讨论现实生活中关于乌克兰难民的情景和言论。 [观看视频 >](#)

内容素材考虑因素(续)

在发布宣传信息之前,应该考虑最后一些注意事项:



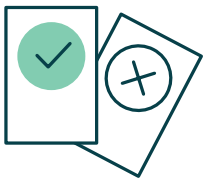
核实信息来源

检查信息来源。务必透明地公示您的信息来自何处,不要让受众群体猜测您的意图。



背景信息

在制作预先揭穿宣传信息时,务必考虑为观看者提供更多有关主题和后续行动的背景信息。您的观看者在看过宣传信息之后会做些什么?您能否指引他们到某个地方去了解更多信息,或者让他们参与宣传,预先阻止虚假信息?不妨考虑使用一个提供更多信息的着陆页,或是制作号召性用语,扩大您的消息的影响力。



测试内容素材

即便有着最好的意图和最完美的创作过程,总会有一些无法预测的变数。我们建议您请目标受众群体的焦点小组代表来测试内容素材(视频、信息图、游戏),了解受众群体会对其做出怎样的回应,然后再与网上更广泛的人群分享这些内容。

2.5

注意事项

可想而知，预先揭穿并非终结虚假信息的万能解决方案。一些局限性和不断变化的趋势都会影响到您的付出。下面仅列出几例供您参考。



它只应作为您工具箱中的工具之一

预先揭穿并不一定优于其他所有干预措施，不过它确实是很好的第一道防线。还应注意，个体层面上的干预措施（包括预先揭穿）会与系统层面上的干预措施（例如对抗两极分化和有组织的造谣信息）协同发挥效力。⁵⁴所有虚假信息干预措施都有自己的优缺点，许多措施都有自己的适用背景环境。务请谨记，预先揭穿只是您工具箱中的工具之一，您甚至可将它[与其他工具配合使用](#)。



小心强硬的观点

如上文“[2.1 应在何时采用这种技术？应由谁实施？](#)”中所述，在受众群体尚未完全相信虚假信息时，预先揭穿确有实效。一旦由于极端化或激进化而导致人们对某个主题的观点固化，预先揭穿就不太可能发挥效果，受众群体的回应也可能不尽人意。这种认知固化的转变节点很难衡量。根据您的受众群体的情况，您可以为不同的受众群体选择不同的宣传消息/渠道。



切勿在受众群体面前摆出高人一等的姿态

在分享信息或向受众群体传授新技能时，有给受众群体留下您好为人师的印象的风险。网上的受众群体都相当精明，能够非常迅速地消化大量信息。请避免在言辞中表现出高人一等的姿态，也不要像对孩子一样对待他们。始终保持理智的谦逊姿态，使用不妄加评判的语气。



避免宣传信息过度简化

在网上分享信息时，内容创作者往往会把自己的信息提炼成更容易吸引人的碎片化内容，以抓住受众群体的注意力，在较新的社交媒体平台上，这已经成为一种日渐兴起的趋势。虽然较为简短的信息更容易广为传播，但通过这种方式讲清细节问题的难度要更大。务必不要过度简化信息，避免其失去应有的效力。如果您使用的平台或媒介不支持复杂的消息，可以考虑通过怎样的方式引导受众群体访问提供更详尽内容的信息来源，让感兴趣的观看者可以深入探索。



误解的可能性

无论您付出怎样的努力去有效、高效、深入地传达信息，受众群体仍然有可能误解您的宣传信息。您应该为误解的情况做好计划，提供链接，让受众群体能获得有关您的工作和目标的更多背景信息。

2.6

预先揭穿核对清单

设计干预措施

选择主题

您要预先揭穿什么虚假信息？

选择受众群体

您的干预措施的目标受众群体是哪些人？

设定目标

您希望取得怎样的结果？

选择一种方法

您的干预措施针对的是虚假信息的内容还是策略？

选择形式

哪种形式最适合您的干预措施？(文字、信息图、视频等)

设计宣传信息

根据文化、策略和受众方面的线索制定干预措施。

部署您的宣传信息

在指定平台上分享您的创作成果

衡量成效

哪些指标与您的预期结果一致？您应如何衡量结果？

参考文献

- 1 Leonardo Bursztyn, Aakaash Rao, Christopher P. Roth and David H. Yanagizawa-Drott, "Misinformation during a Pandemic," National Bureau of Economic Research, June 2020. <https://www.nber.org/papers/w27417>
- 2 Jacob Poushter, Moira Fagan, and Sneha Gubbala, "Climate Change Remains Top Global Threat Across 19-Country Survey," Pew Research Center's Global Attitudes Project (blog), August 31, 2022. <https://www.pewresearch.org/global/2022/08/31/climate-change-remains-top-global-threat-across-19-country-survey/>
- 3 此处我们将虚假信息定义为: 不尽不实、有误导性和/或存在谬误的信息, 可能有意或无意中导致有害的后果; 造谣信息是指刻意编造的虚假信息, 例如在有组织的宣传活动中使用的造谣信息。为便于参考, 在本文中, 我们将所有不尽不实或误导性的信息统称为虚假信息, 包括源自造谣信息或捏造信息的内容。
- 4 Jon Roozenbeek, Jane Suiter, and Eileen Culloty, "Countering Misinformation: Evidence, Knowledge Gaps, and Implications of Current Interventions," *European Psychologist* (September 20, 2022) advance online publication. <https://doi.org/10.31234/osf.io/b52um>
- 5 Stephan Lewandowsky et al., "Misinformation and Its Correction: Continued Influence and Successful Debiasing," *Psychological Science in the Public Interest* 13, no. 3 (December 2012): 106–31. <https://doi.org/10.1177/1529100612451018>
- 6 Fabiana Zollo et al., "Debunking in a World of Tribes," ed. Jose Javier Ramasco, *PLOS ONE* 12, no. 7 (July 24, 2017): e0181821. <https://doi.org/10.1371/journal.pone.0181821>
- 7 Sander van der Linden et al., "Inoculating the Public against Misinformation About Climate Change," *Global Challenges* 1, no. 2 (February 2017): 1600008. <https://doi.org/10.1002/gch2.201600008>
- 8 Jon Roozenbeek et al., "Psychological Inoculation Improves Resilience against Misinformation on Social Media," *Science Advances* 8, no. 34 (August 26, 2022): eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- 9 W. J. McGuire, "Resistance to Persuasion Conferred by Active and Passive Prior Refutation of the Same and Alternative Counterarguments," *The Journal of Abnormal and Social Psychology* 63, no. 2 (September 1961): 326–32. <https://doi.org/10.1037/h0048344>
- 10 Jon Roozenbeek, Sander van der Linden, and Thomas Nygren, "Prebunking Interventions Based on the Psychological Theory of 'Inoculation' Can Reduce Susceptibility to Misinformation across Cultures.," *Harvard Kennedy School Misinformation Review*, (February 3, 2020). <https://doi.org/10.37016/mr-2020-008>
- 11 Sander van der Linden et al., "Inoculating the Public against Misinformation About Climate Change," *Global Challenges* 1, no. 2 (February 2017): 1600008. <https://doi.org/10.1002/gch2.201600008>
- 12 Josh Compton, Ben Jackson, and James A. Dimmock, "Persuading Others to Avoid Persuasion: Inoculation Theory and Resistant Health Attitudes," *Frontiers in Psychology* 7 (February 9, 2016). <https://doi.org/10.3389/fpsyg.2016.00122>
- 13 Bobi Ivanov et al., "Using an Inoculation Message Approach to Promote Public Confidence in Protective Agencies," *Journal of Applied Communication Research* 44, no. 4 (October 2016): 381–98. <https://doi.org/10.1080/00909882.2016.1225165>
- 14 Robin L. Nabi, "'Feeling' Resistance: Exploring the Role of Emotionally Evocative Visuals in Inducing Inoculation," *Media Psychology* 5, no. 2 (May 2003): 199–223. https://doi.org/10.1207/S1532785XMEP0502_4
- 15 John Cook, Stephan Lewandowsky, and Ullrich K. H. Ecker, "Neutralizing Misinformation through Inoculation: Exposing Misleading Argumentation Techniques Reduces Their Influence," ed. Emmanuel Manalo, *PLOS ONE* 12, no. 5 (May 5, 2017): e0175799. <https://doi.org/10.1371/journal.pone.0175799>
- 16 Cecilie S. Traberg, Jon Roozenbeek, and Sander van der Linden, "Psychological Inoculation against Misinformation: Current Evidence and Future Directions," *The ANNALS of the American Academy of Political and Social Science* 700, no. 1 (March 2022): 136–51. <https://doi.org/10.1177/00027162221087936>
- 17 Sander van der Linden et al., "Inoculating the Public Against Misinformation About Climate Change," *Global Challenges* 1, no. 2 (February 2017): 1600008. <https://doi.org/10.1002/gch2.201600008>
- 18 John Cook, Stephan Lewandowsky, and Ullrich K. H. Ecker, "Neutralizing Misinformation Through Inoculation: Exposing Misleading Argumentation Techniques Reduces Their Influence," ed. Emmanuel Manalo, *PLOS ONE* 12, no. 5 (May 5, 2017): e0175799. <https://doi.org/10.1371/journal.pone.0175799>
- 19 Jon Roozenbeek et al., "Psychological Inoculation Improves Resilience against Misinformation on Social Media," *Science Advances* 8, no. 34 (August 26, 2022): eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- 20 Stephan Lewandowsky and Muhsin Yesilada, "Inoculating against the Spread of Islamophobic and Radical-Islamist Disinformation," *Cognitive Research: Principles and Implications* 6, no. 1 (December 2021): 57. <https://doi.org/10.1186/s41235-021-00323-z>
- 21 Jon Roozenbeek and Sander van der Linden, "How to Combat Health Misinformation: A Psychological Approach," *American Journal of Health Promotion* 36, no. 3 (March 2022): 569–75. <https://doi.org/10.1177/08901171211070958>
- 22 Melisa Basol et al., "Towards Psychological Herd Immunity: Cross-Cultural Evidence for Two Prebunking Interventions against COVID-19 Misinformation," *Big Data & Society* 8, no. 1 (January 2021): 205395172110138. <https://doi.org/10.1177/20539517211013868>
- 23 Melisa Basol, Jon Roozenbeek, and Sander Van der Linden, "Good News About Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News," *Journal of Cognition*

参考文献(续)

- 3, no. 1 (January 10, 2020): 2.<https://doi.org/10.5334/joc.91>
- 24 John Cook et al., “The Cranky Uncle Game—Combining Humor and Gamification to Build Student Resilience Against Climate Misinformation,” *Environmental Education Research* (June 14, 2022), 1–17. <https://doi.org/10.1080/13504622.2022.2085671>
- 25 “GCS International Joins the Fight against Health Misinformation Worldwide,” Government Communication Service of the United Kingdom (February 18, 2021). <https://gcs.civilservice.gov.uk/news/gcs-international-joins-the-fight-against-health-misinformation-worldwide/>
- 26 Jon Roozenbeek et al., “Psychological Inoculation Improves Resilience against Misinformation on Social Media,” *Science Advances* 8, no. 34 (August 26, 2022): eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- 27 Rakoen Maertens et al., “Long-Term Effectiveness of Inoculation against Misinformation: Three Longitudinal Experiments,” *Journal of Experimental Psychology: Applied* 27, no. 1 (March 2021): 1–16. <https://doi.org/10.1037/xap0000315>
- 28 Bobi Ivanov, Kimberly A. Parker, and Lindsay L. Dillingham, “Testing the Limits of Inoculation-Generated Resistance,” *Western Journal of Communication* 82, no. 5 (October 20, 2018): 648–65. <https://doi.org/10.1080/10570314.2018.1454600>
- 29 Briony Swire-Thompson, Joseph DeGutis, and David Lazer, “Searching for the Backfire Effect: Measurement and Design Considerations,” *Journal of Applied Research in Memory and Cognition* 9, no. 3 (September 2020): 286–299. <https://doi.org/10.1016/j.jarmac.2020.06.006>
- 30 Brian Hughes, Kurt Braddock, Cynthia Miller-Idriss, Beth Goldberg, Meili Criezis, Pasha Dashtgard, and Kesa White, “Inoculating Against Persuasion by Scientific Racism Propaganda: The Moderating Roles of Propaganda Form and Subtlety,” *SocArXiv* (July 31, 2021). <https://doi.org/10.31235/osf.io/ecqn4>
- 31 Ananya Iyengar, Poorvi Gupta, and Nidhi Priya, “Inoculation Against Conspiracy Theories: A Consumer Side Approach to India’s Fake News Problem,” *Applied Cognitive Psychology* (September 14, 2022) acp.3995. <https://doi.org/10.1002/acp.3995>
- 32 Jon Roozenbeek et al., “Psychological Inoculation Improves Resilience against Misinformation on Social Media,” *Science Advances* 8, no. 34 (August 26, 2022): eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- 33 Abhijit Banerjee, Eliana La Ferrara, and Victor Orozco-Olvera, “The Entertaining Way to Behavioral Change: Fighting HIV with MTV,” Cambridge, MA: National Bureau of Economic Research (July 2019). <https://doi.org/10.3386/w26096>
- 34 S. Usdin et al., “Achieving Social Change on Gender-Based Violence: A Report on the Impact Evaluation of Soul City’s Fourth Series,” *Social Science & Medicine* 61, no. 11 (December 2005): 2434–45. <https://doi.org/10.1016/j.socscimed.2005.04.035>
- 35 UNICEF, “Technical Note on Gender Norms” (United Nations, n.d.). <https://www.unicef.org/media/65381/file/GP-2020-Technical-Note-Gender-Norms.pdf>
- 36 Ada Sonnenfeld et al., “Strengthening Intergroup Social Cohesion in Fragile Situations,” *3ie Systematic Review* 46 (2021). <https://www.3ieimpact.org/evidence-hub/publications/systematic-reviews/strengthening-intergroup-social-cohesion-fragile>
- 37 “Creatively Tackling Sanitation in India,” BBC Media Action (September 2020). <https://www.bbc.co.uk/mediaaction/publications-and-resources/research/summaries/executive-summary-navarangi-re-sept-2020/>
- 38 Rachel Glennerster, Joanna Murray, and Victor Pouliquen, “The Media or the Message? Experimental Evidence on Mass Media and Contraception in Burkina Faso,” August 21, 2022. https://www.povertyactionlab.org/sites/default/files/research-paper/working-paper_3835_Mass-Media-and-Contraception_Burkina-Faso_Aug2022.pdf
- 39 Danielle A. Naugle and Robert C. Hornik, “Systematic Review of the Effectiveness of Mass Media Interventions for Child Survival in Low- and Middle-Income Countries,” *Journal of Health Communication* 19, no. sup1 (May 6, 2014): 190–215. <https://doi.org/10.1080/10810730.2014.918217>
- 40 Josh Compton, Sander van der Linden, John Cook, and Melisa Basol, “Inoculation Theory in the Post-Truth Era: Extant Findings and New Frontiers for Contested Science, Misinformation, and Conspiracy Theories,” *Compass* (May 5, 2021). <https://doi.org/10.1111/spc3.12602>
- 41 Renee DiResta, “‘Prebunking’ Health Misinformation Tropes Can Stop Their Spread,” *Wired* (August 28, 2021). <https://www.wired.com/story/prebunking-health-misinformation-tropes-can-stop-their-spread/>
- 42 Josh Compton et al., “Inoculation Theory in the Post-Truth Era: Extant Findings and New Frontiers for Contested Science, Misinformation, and Conspiracy Theories,” *Social and Personality Psychology Compass* 15, no. 6 (June 2021). <https://doi.org/10.1111/spc3.12602>
- 43 Brian Hughes et al., “Inoculating against Persuasion by Scientific Racism Propaganda: The Moderating Roles of Propaganda Form and Subtlety,” preprint: *SocArXiv* (July 31, 2021). <https://doi.org/10.31235/osf.io/ecqn4>
- 44 Jon Roozenbeek, Sander Van der Linden, Rakoen Maertens, Stefan M. Herzog, Michael Geers, Ralf Kurvers, and Mubashir Sultan, “Susceptibility to Misinformation Is Consistent Across Question Framings and Response Modes and Better Explained by Myside Bias and Partisanship than Analytical Thinking,” *Judgment and Decision Making*, Vol. 17, No. 3 (May 2022): pp. 547–573. <https://journal.sjdm.org/22/220228/jdm220228.pdf>
- 45 Rakoen Maertens et al., “Long-Term Effectiveness of Inoculation against Misinformation: Three Longitudinal Experiments,” *Journal of Experimental Psychology: Applied* 27, no. 1 (March 2021): 1–16. <https://doi.org/10.1037/xap0000315>
- 46 Jon Roozenbeek et al., “Psychological Inoculation

参考文献(续)

- Improves Resilience Against Misinformation on Social Media,” *Science Advances* 8, no. 34 (August 26, 2022): eabo6254. <https://doi.org/10.1126/sciadv.abo6254>
- 47 Melisa Basol et al., “Towards Psychological Herd Immunity: Cross-Cultural Evidence for Two Prebunking Interventions against COVID-19 Misinformation,” *Big Data & Society* 8, no. 1 (January 2021): 205395172110138. <https://doi.org/10.1177/20539517211013868>
- 48 Stephan Lewandowsky and Muhsin Yesilada, “Inoculating Against the Spread of Islamophobic and Radical-Islamist Disinformation,” *Cognitive Research: Principles and Implications* 6, no. 1 (December 2021): 57. <https://doi.org/10.1186/s41235-021-00323-z>
- 49 Melisa Basol, Jon Roozenbeek, and Sander Van der Linden, “Good News About Bad News: Gamified Inoculation Boosts Confidence and Cognitive Immunity Against Fake News,” *Journal of Cognition* 3, no. 1 (January 10, 2020): 2. <https://doi.org/10.5334/joc.91>
- 50 Jon Roozenbeek and Sander van der Linden, “How to Combat Health Misinformation: A Psychological Approach,” *American Journal of Health Promotion* 36, no. 3 (March 2022): 569–75. <https://doi.org/10.1177/08901171211070958>
- 51 Jon Roozenbeek and Sander van der Linden, “The Fake News Game: Actively Inoculating against the Risk of Misinformation,” *Journal of Risk Research* 22, no. 5 (May 4, 2019): 570–80. <https://doi.org/10.1080/13669877.2018.1443491>
- 52 John Cook et al., “The Cranky Uncle Game—Combining Humor and Gamification to Build Student Resilience Against Climate Misinformation,” *Environmental Education Research*, June 14, 2022, 1–17. <https://doi.org/10.1080/13504622.2022.2085671>
- 53 Jody C. Baumgartner and Amy Becker, eds., “Political Humor in a Changing Media Landscape: A New Generation of Research,” *Lexington Studies in Political Communication* (Lanham: Lexington Books, 2018).
- 54 Jon Roozenbeek, Jane Suiter, and Eileen Culloty, “Countering Misinformation: Evidence, Knowledge Gaps, and Implications of Current Interventions,” *European Psychologist* (20 September 2022) advance online publication). <https://doi.org/10.31234/osf.io/b52um>

实用指南：预先揭穿虚假信息
(c) 2022

