

誤情報と認知

コグニティブセキュリティ

B. M. Pierce (2021). "Protecting people from disinformation requires a cognitive security proving ground" C4ISRNET

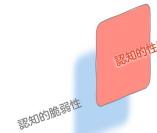
コグニティブセキュリティの目的 個人や集団に対する悪意のある影響を弱めること

サイバーセキュリティはデバイス・コンピュータ、ネットワークなどの保護に重点があるのに対し、**コグニティブセキュリティは人間の保護に重点をおく**。社会科学・行動科学、AI、データサイエンスなどを含む多くの学問分野を統合する社会技術的アプローチが必要となる。

Increase cognitive resilience against malicious influence (悪意のある影響に対する認知的レジリエンスの強化)

例) 誤偽情報による悪影響

Why Johnny can't ...?



How human cognition works?



Step1: 脆弱性をもたらす認知的性質の理解

真実錯覚効果 (illusory truth effect)

確認バイアス (confirmation bias)

誤情報持続効果 (continued influence effect)

Step2: 認知的性質を考慮した介入

フレバシク

デバシク

アメリカ心理学会の声明 2023.11

<https://www.apa.org/pubs/reports/health-misinformation>

目的) 3つの重要な問いについての共通見解を提供し、これらの議論を明確にする

- なぜ人々は誤情報を信じ、それにもとづいて行動しやすいのか、その心理的要因はなにか?
- なぜ、どのように誤情報が広がるのか?
- 誤情報に対抗するためにどのような介入が効果的か?

誤情報による脅威に対処するための8つの推奨事項

Recommendations

Avoid repeating misinformation without including a correction.

The repetition of false claims increases their salience, a phenomenon known as the illusory truth effect. Repeaters of false claims are thus likely to be perceived as credible and their claims as more believable. While this already has implications for knowledge about the topic, when made online and/or in other digital spaces, it can have additional effects. For example, it can perpetuate false beliefs, spreading misinformation, even when only one person is repeating the claim. In these cases, the falsehood should be repeated only with the correction and not more prominently than the falsehood itself.

Collaborate with social media companies to understand and reduce the spread of harmful misinformation.

Most user-generated social media is shared by very few users, and it is often hard to track. This makes it difficult for researchers to understand how misinformation spreads. Social media "echo chambers" limit and isolate information, often leading to the spread of false information and the amplification of harmful content. On social media, identifying and removing harmful content about the "other side" can spread faster than content of one's own side, leading to increased polarization and public health concerns. Research and evidence-based public health responses should work with online platforms to understand and reduce the spread of dangerous misinformation.

Use misinformation correction strategies with links already proven to increase healthy behaviors.

Psychological science research shows that the link between knowledge and behavior is complex. There is growing evidence that linking misinformation correction with the promotion of healthy behaviors can be more effective than simply providing the correction. For example, linking a correction about a health claim about the benefits of exercise to a link about how to get started with exercise can be more effective than simply providing the correction.

Leverage trusted sources to counter misinformation and promote accurate health information.

People have an implicit preference for more credible, trusted sources of information. These trusted sources can be leveraged to counter misinformation and promote accurate health information. For example, a trusted source like a doctor or a trusted organization can be used to provide a correction about a health claim. This can be more effective than simply providing the correction.

Debrief misinformation often and repeatedly using evidence-based methods.

Research shows that debriefing and correction are generally effective across ages and cultures. However, debriefing alone is always more effective than no debriefing. Corrections should be repeated frequently with the reminder that the original information is incorrect. Corrections should be repeated frequently with the reminder that the original information is incorrect. Corrections should be repeated frequently with the reminder that the original information is incorrect.

Provide misinformation to inoculate susceptible audiences by building skills and resilience from an early age.

Research shows that building skills and resilience from an early age can be more effective than simply providing the correction. For example, teaching children how to identify and evaluate information can be more effective than simply providing the correction.

Demand data access and transparency from social media companies for scientific research on misinformation.

Efforts to quantify and understand misinformation on social media are hampered by lack of access to user data from social media companies. Misinformation researchers are able to collect data on user behavior, but not on the full range of user behavior. Demand data access and transparency from social media companies for scientific research on misinformation.

Fund basic and translational research into the psychology of health misinformation, including effective ways to counter it.

Research on the psychology of health misinformation is needed to understand how it spreads and how to counter it. This research should include basic and translational research into the psychology of health misinformation, including effective ways to counter it.



Using Psychological Science to Understand and Fight Health Misinformation

AN APA CONSENSUS STATEMENT

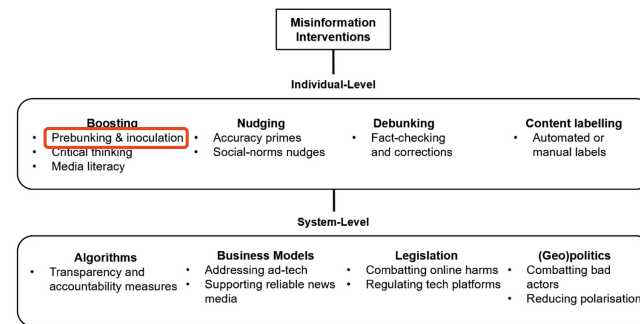
NOVEMBER 2023



43ページの報告書 (p.30-43がReferences)

誤情報に対する介入のレベル

"Prebunking is *not* a one-size-fits-all solution to ending misinformation"



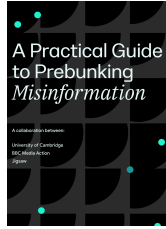
Rozenbeek, J., Gullott, E., & Sultzer, J. (2023). Countering misinformation: Evidence, Knowledge Gaps, and Implications of Current Interventions. *European Psychologist*, 28(3), 189-205. <https://doi.org/10.1027/1016-9040/a000492>
www.oecd-forum.org/posts/prebunking-staying-ahead-of-the-curve-on-misinformation

プレバンキング

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.

事前に誤情報に対して耐性を築くための方法

- 接種理論 (inoculation theory)
 - 1960年代に社会心理学者William McGuireによって提唱された理論
 - 医療用ワクチンが将来の感染に対して生理的な抵抗力を与えるように、心理的な予防接種も、将来の心理操作に対する抵抗力を与えるという考え方もとつく
 - 心理的予防接種の介入によって、誤情報やプロパガンダによる影響を軽減する実証研究が蓄積されている



行動心理学に関する予備知識がなくてもプレバンキングを実践できるよう、学術的研究を実践的なハウツーガイドとしてまとめたもの

ケンブリッジ大学、Jigsaw (Google)、BBCの共同研究

1. 警告

近い将来、誤情報に出くわす可能性がありますがあることを警告する。

受け手を説得しようとする本意を試みに対する受け手の精神的防御を活性化させる

1. Forewarning

A warning activates the viewer's mental defenses against unwanted attempts to persuade them by alerting them that they are likely to encounter misleading messages in the near future.

2. 先制的反論

将来いふかもしれない誤情報に対抗するためのツールを提供

あらかじめ反論の材料を与える。「搬運」の例を与えることで、将来それを認識しやすくなる。

2. Preemptive refutation

An effective rebuttal provides with tools to counter misdeas they may see in the future. It equipping them with counter-advance, it helps to include weakened example of the n that they can more easily refute.

https://interventions.withgoogle.com/static/pdf/A_Practical_Guide_to_Prebunking_Misinformation.pdf

「受動的」プレバンキング

- 受動的な接種
 - テクニックに抵抗するための情報が短い形式 (テキスト、グラフィックス、ビデオ) で提供される
 - 制作・実施が比較的容易 (例: SNSのポップアップでテキストメッセージを提示、Youtubeで広告のような形式で流す)
 - 没入感が少なく、対話がすくないため影響力が小さい可能性がある

EXAMPLES

30-90秒

<https://inoculation.science/>から視聴可能

VIDEO EXAMPLE: FALSE DICHOTOMIES

This video example — produced by Jigsaw and Cambridge University — uses culturally relevant examples to help viewers understand and recognize the use of false dichotomies in the spread of misinformation. [View video >](#)

誤情報の共有意図を軽減

INFOGRAPHIC EXAMPLE: COVID-19 CONSPIRACY THEORIES

This UNESCO infographic explains conspiracy theories by using COVID-19 as an example.²¹

Limitations

Scalability: 実践者は、異なる種類の誤情報・受け手・プラットフォームで行う場合はパイロットスタディが必要

効果は時間とともに薄れる傾向があり、誤情報に対する耐性を維持するためには、定期的な「ブースター」が必要であり、メディアやデジタルリテラシーのトレーニングも必要

効果検証は主に北アメリカや西ヨーロッパ諸国で実施、異文化間での検証が不足。対象者を考慮して設計する必要がある

Youtubeでのフィールド調査 (Google Jigsaw) では、動画ベースの予防接種介入は、情報操作テクニックに対する認識を向上させたものの、他のフィールド調査が不足している。

「能動的」プレバンキング

- 能動的な接種
 - ゲームまたはクイズの形で提供される
 - 誤情報でよく使われるテクニックに対する抵抗力を高める効果があることが示されている
 - 実証研究のメタレビューの結果、受動的な接種と比べると効果の持続性は高いことが示されている (ブースターが提供されると3ヶ月以上)

所要時間15-20分
教員・生徒向けの解説や教材あり

BAD NEWS

This was the first-ever prebunking game. It is a choice-based browser game created by DfEC and the University of Cambridge in which players take on the role of a fake news producer and learn to identify and resist by misinformation techniques (e.g. trolling, conspiratorial reasoning, impersonation) over six levels. Since then, several other games with similar premises have been designed. [View game >](#)

HARMONY SQUARE

Set in a peaceful community known for its pond swim and annual Pineapple Pizza Festival, this game appoints the player as the 'Chief Disinformation Officer' tasked with poisoning the people of Harmony Square and using trolling campaigns during political elections. [View game >](#)

GO VIRAL!

This is an online game that allows players to identify and resist misinformation techniques (e.g. trolling, conspiratorial reasoning, impersonation) over six levels. Since then, several other games with similar premises have been designed. [View game >](#)

政治的なプロパガンダに対応したゲーム
2019年CISA「The War on Pineapple」を元にしたシナリオ

COVID-19に関する誤情報にフォーカスを当てたゲーム。プレイヤーは感情的に操作される情報の使用や陰謀論的思考などについてまなぶ。2回以上プレイされている。

ケンブリッジ大学とWHOの共同開発 (英語、ドイツ語、フランス語)

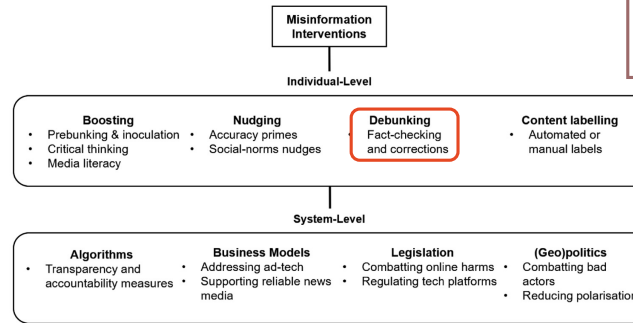
<https://inoculation.science/>から利用可能

誤情報を広めるためによく使われる7テクニック

TECHNIQUE	EXAMPLE
なりすまし Impersonation Spreading information as another person or organization to increase credibility or to cause harm.	"NASA admitted that climate change occurs naturally as a result of changes in Earth's solar orbit and not anthropogenic factors." EXPLANATION: The example uses NASA as a way to increase the credibility of the statement, even though NASA has never made such a claim.
感情操作 Emotional manipulation Using words or images that tap into emotions to influence behavior or to cause harm.	"What this airline did for passengers will make you tear up — 50 tears warning!" EXPLANATION: The example often how information can be presented differently to elicit an emotional reaction, by evoking pity and anger and reduce critical evaluation.
二極化 Polarization Exaggerating existing differences between groups to create a sense of urgency or to cause harm.	"The UK Party: Don't believe the Worker Party lies. They said they would abolish student debt but now they're saying they won't!" EXPLANATION: The example uses hostile "othering" language by describing another party as lies.
陰謀論的な考え方 Conspiratorial ideation Explaining events from traditional uses using alternative explanations that give credence to the idea of a conspiracy.	"The UK Party: Don't believe the Worker Party lies. They said they would abolish student debt but now they're saying they won't!" EXPLANATION: This example encourages conspiratorial thinking to give credence to the idea of a conspiracy, by using the words of the UK Party to suggest that the UK Party is lying.
個人攻撃 Ad hominem attacks Attacking a person's character or reputation to undermine their credibility or to cause harm.	"Bertha has an uncontrollable temper and is generally a personality disorder but she can't have someone crazy in power!" EXPLANATION: The example attacks the character of the leader, instead of focusing on the policy or leadership decisions.
偽の二分法 False dichotomy Presenting two options as the only choices available, when in reality there are many other options available.	"Either you support the energy protests or you don't believe in justice." EXPLANATION: The example presents two options as the only choices available, when in reality there are many other options available.
偽のバランス False balance Presenting a debate as having two equally balanced sides, when in reality one side is much stronger.	"Experts debate the shape of the earth. While scientist Rebecca Chesebrough has found the earth to be spherical, expert Brian Burrows argues that the earth is flat." EXPLANATION: This example presents two equally balanced sides, when in reality one side is much stronger.

誤情報に対する介入のレベル

"Prebunking is not a one-size-fits-all solution to ending misinformation"



デバーク (ファクトチェック・訂正)

誤情報を事後的に修正する介入方法。なぜその情報が正しくないのかを説明したり、正確な情報を提供することも含まれる。

CDCなどが健康情報 (例: COVID-19) などに関してWebサイトやソーシャルメディアでファクトチェックを積極的に行うようになってきている。

CDC Centers for Disease Control and Prevention
COVID-19: Staying Healthy, Protecting People

Bust Myths and Learn the Facts about COVID-19 Vaccines

Getting a COVID-19 vaccine is a safer and more dependable way to build immunity to COVID-19 than getting sick with COVID-19.

FACT: COVID-19 vaccination causes a more predictable immune response than an infection with the virus causes COVID-19.

COVID-19 that cause illness or death. You can also learn more about COVID-19 vaccines at [https://www.cdc.gov/coronavirus/2019-ncov/vaccines/facts.html](#)

Roozenbeek, J., Gulloty, E., & Sultzer, J. (2023). Countering misinformation: Evidence, Knowledge Gaps, and Implications of Current Interventions. *European Psychologist*, 28(3), 189-205. <https://doi.org/10.1027/1016-9040/a000492>
www.oecd-forum.org/posts/prebunking-staying-ahead-of-the-curve-on-misinformation

<https://www.cdc.gov/coronavirus/2019-ncov/vaccines/facts.html>

Quick guide to responding to misinformation



Misinformation can do damage

Misinformation is false information that is spread either by mistake or with intent to mislead. When there is intent to mislead, it is called disinformation. Misinformation has the potential to cause substantial harm to individuals and society. It is therefore important to protect people against being misinformed, either by making them resilient against misinformation before it is encountered or by debunking it after people have been exposed.



Misinformation can be sticky!

Fact-checking can reduce people's belief in false information. However, misinformation often continues—this seems effective—because people's beliefs are resilient to correction. This is because people's beliefs are resilient to correction. This is because people's beliefs are resilient to correction. This is because people's beliefs are resilient to correction.

誤情報は粘着する

ファクトチェックは人々の誤情報への信念を減少させる。ただし、訂正を受け入れた後でも、誤情報はしばしば人々の考えに影響を与え続けることがあり、これは「誤情報持続効果」として知られている。事実による訂正が効果的であるように見えても、人々はしばしば他の文脈で誤情報を利用し続ける。したがって、最大の影響を得るためには、最も効果的な訂正のアプローチを使用することが重要。



Prevent misinformation from sticking if you can

Because misinformation is sticky, it's best preempted. This can be achieved by explaining misleading or manipulative argumentation strategies to people—a technique known as "inoculation" that makes people resilient to subsequent manipulation attempts. A potential drawback of inoculation is that it requires advance knowledge of misinformation techniques and is best administered before people are exposed to the misinformation.



Debunk often and properly

If you cannot preempt, you must debunk. For debunking to be effective, it is important to provide detailed refutations²³. Provide a clear explanation of (1) why it is now clear that the information is false, and (2) what is true instead. When those detailed refutations are provided, misinformation can be "unstuck." Without detailed refutations, the misinformation may continue to stick around despite correction attempts.

Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E. Rand, D. G., Rapp, D. N., Reifer, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, S., Vraga, E. K., Wood, T. J., Zaragoza, M. S. (2020). The Debunking Handbook 2020. Available at <https://skts.io/db2020>. DOI:10.17910/b7.1182

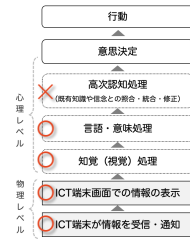
田中優子・大塚美輪・藤本和則 (2022) 誤情報持続効果をもたらす心理プロセスの理解と今後の展望: 誤情報の制御に向けて. 認知科学, 29(3), 509-527. doi.org/10.11225/cs.2022.003

誤情報持続効果

(continued influence effect of misinformation)

・ 誤りであると指摘されていることを知った後も、誤情報を信じ続けたり、誤情報の影響を受け続ける心理現象

- ・ 訂正情報に視覚的注意を払っていても、訂正情報の内容を記憶(記銘)していても生じる。
- ・ 高次認知処理レベルの観点から研究が進められている。



真実錯覚効果 (illusory truth effect)

・ 繰り返し同じ情報に接触することで、その情報が正しく感じられるようになること。

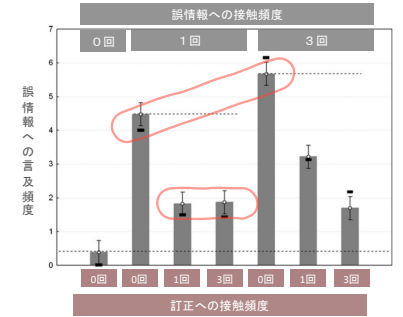
- ・ 情報への「親近性 (familiarity)」や「処理の流暢性 (fluency)」が「正しさ」のシグナルとして利用されるヒューリスティック

「訂正情報」も繰り返し流せばいいのでは？

誤情報の3倍の頻度で訂正情報を出しても、誤情報の影響は消えない (誤情報に触れなかった状態には戻らない)

真実錯覚効果の非対称性

「誤情報の信じられやすさ」と「一度受け入れられた誤情報の影響を事後的に緩和することの難しさ」のギャップ



Ecker, U. K. H., Lewandowsky, S., Swire, B., & Chang, D. (2011). Correcting false information in memory: Manipulating the strength of misinformation encoding and its retraction. *Psychonomic Bulletin and Review*, 18(3), 570-578.

The Debunking Handbook 2020

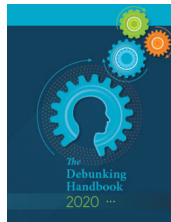
・ 誤情報対策として、学術的知見をもとにした実践的提言

- ・ 対象：市民、政策立案者、ジャーナリスト、その他実務家など
- ・ 方法：
 - ・ 2015年以降に誤情報に関する心理学分野における学術的実績のある研究者22名を選定
 - ・ 心理的特徴と対策案の根拠となるエビデンスを集める
 - ・ それぞれに対し、「エビデンスの強度」と「デバンクにおける重要性」を評定し、評定結果を分析
 - ・ 最終的に、「心理的特徴」17点、「実行可能な対策」10点を選定

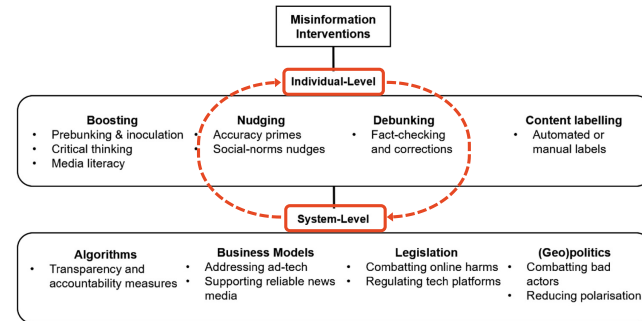
誤情報に関する心理的特徴、それらにもとづく訂正の効果をあげるための留意点が整理されている

19ヶ国語 (ドイツ、イタリア、トルコ、ロシア、チェコ、ポルトガルなど) に翻訳 (日本語訳はなし)

Lewandowsky, S., Cook, J., Ecker, U. K. H., Albarracín, D., Amazeen, M. A., Kendeou, P., Lombardi, D., Newman, E. J., Pennycook, G., Porter, E. Rand, D. G., Rapp, D. N., Reifer, J., Roozenbeek, J., Schmid, P., Seifert, C. M., Sinatra, G. M., Swire-Thompson, B., van der Linden, S., Vraga, E. K., Wood, T. J., Zaragoza, M. S. (2020). The Debunking Handbook 2020. Available at <https://skts.io/db2020>. DOI:10.17910/b7.1182



個人レベルとシステムレベルの相互作用



Roozenbeek, J., Gulloty, E., & Sultzer, J. (2023). Countering misinformation: Evidence, Knowledge Gaps, and Implications of Current Interventions. *European Psychologist*, 28(3), 189-205. <https://doi.org/10.1027/1016-9040/a000492>
www.oecd-forum.org/posts/prebunking-staying-ahead-of-the-curve-on-misinformation

