Advancing Media Literacy in Indonesia

Contents

⬰ Project overview

Since 2019, Moonshot has worked in partnership with the University of Notre Dame (UND), IREX and GeoPoll to advance media literacy among new digital arrivals for the USAID Mission in Indonesia.

In the first two years of the project, Moonshot monitored the disinformation environment in Indonesia, developed a database of key disinformation narratives and analysed the audiences consuming disinformation. Moonshot then designed and deployed digital campaigns to reach these audiences with a media literacy website (Literata.id) and content produced by IREX. The website was designed to guide users in how to deal with mis/disinformation on the internet and improve their media literacy skills.

In late 2020, Moonshot continued these digital campaigns with updated content, including a disinformation inoculation game ('Gali Fakta') for an Indonesian audience. The game is designed to inoculate people who are vulnerable to disinformation and equip them with the media literacy skills necessary to identify it in the real world.

The programme was extended for four years to enable further adaptation and testing to advance media literacy in Indonesia. In 2021, Moonshot developed and optimised a new version of the game ('Gali Fakta II'). In this third phase of the project, Moonshot worked to broaden the reach of its digital campaigns to promote both the media literacy website and the game to a wider audience.

Finally, building on insights from the first two phases of the project, Moonshot carried out testing and analysis to determine and evaluate the role that gamification can play in building media literacy in relation to disinformation.

This report presents the findings of the third phase of the project, with a specific focus on the game, and considers recommendations for further research and programming.

Looking back: First phases of the project

From September 2020 to October 2021, Moonshot successfully carried out digital campaigns to redirect those engaging with disinformation online in Indonesia to media literacy content. The media literacy content, housed under the brand Literata, included the Literata.id website produced by IREX and UND, as well as a disinformation inoculation game, Gali Fakta, created by Moonshot.

Moonshot found that both the website and the game were effective at diverting user attention away from disinformation content towards positive media literacy content.

However, the lasting effect of this content on changing users' knowledge and understanding towards disinformation online was inconclusive due to large discrepancies in user age and level of education between the control and treatment groups.

This raised a number of interesting questions for further research and programming - in particular, how to control for age and level of education in an experiment trying to reach people based on their vulnerability to disinformation.

A year on

Since 2021, Moonshot has made several updates to the game to improve the user experience and overall rates of retention.

01

Choice of topic

At the outset of the game, users are now presented with a choice between topics they are most interested in, such as: health and beauty, news and politics, or online investment schemes. At the end of the game, users are given the opportunity to try out the other topics they did not choose the first time round.

02

Length of game

Allowing users to choose their own topic of interest enabled us to decrease the length of the game. With optimising user retention in mind, Moonshot and Brave Factor updated the game to be no more than an average of five minutes of gameplay.

03

Integrated and improved pre-game and post-game surveys

In the 2021 pilot, users who played the game were retargeted with ads for a survey on their experience. This approach, however, led to several limitations which could have impacted the results. For example, we could not control how long it had been since users had played the game. In the 2022 iteration of the game, Moonshot integrated the survey into the game. The survey was also redesigned to include four pre-game questions and four post-game questions to measure users' ability to identify different forms of disinformation.

↘ Methodology

Digital campaigns

Moonshot designed and deployed 409 Redirect Method campaigns across two platforms, Twitter and the Google Display Network, between 1 June and 31 August 2022.

These campaigns were designed to reach individuals expressing interest in or engaging with fake news, debunked myths or conspiracy theories on Twitter or YouTube. The goal was to redirect them to our media literacy content, either on the website or in the game.

THE REDIRECT METHOD

The Redirect Method is an open-source methodology that uses targeted advertising to connect people searching online for harmful content with constructive alternative messages. It was piloted by Jigsaw and Moonshot in 2016 and subsequently deployed internationally by Moonshot in partnership with tech companies, governments, and grassroots organisations.

The Redirect Method uses content made by communities across the globe to challenge narratives that support violent extremism, violent misogyny, disinformation, and other online harms. The Redirect Method places ads in the search results and social media feeds of users who are searching for pre-identified terms that are associated with a particular online harm - in this case, disinformation.

This approach ensures search engine users are given the option of engaging with alternative content based on known behaviours - e.g., they are searching for harmful content.
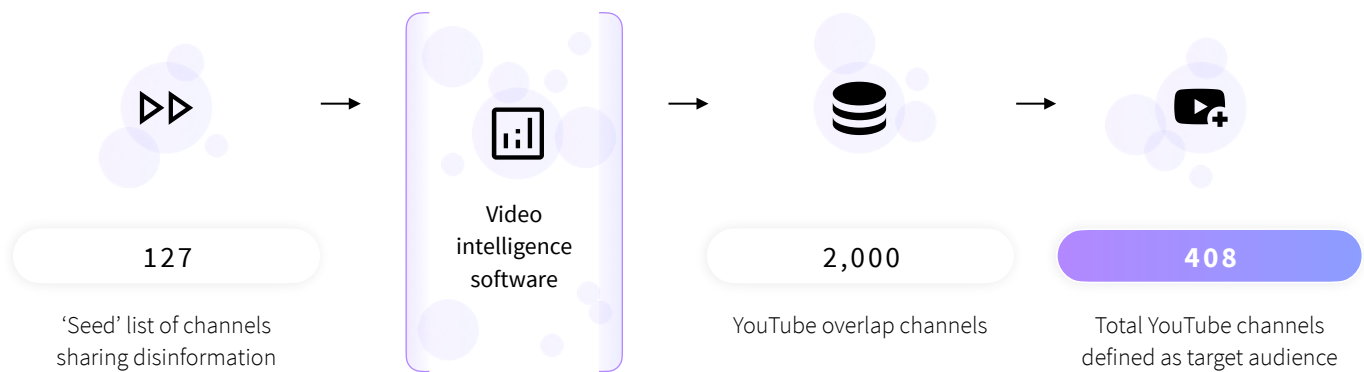
Audience messaging

The text and image in our ads functioned as the entry point to our campaigns and media literacy content. The language, tone and imagery was designed to attract and engage user attention based on the content and emotion in their search term. Images reflecting the colours and branding of Literata were designed and used in the Google Display and Twitter campaigns. Moonshot drew on previous work in the region when defining these messages and images, applying a data-driven approach with quality assurance reviews. In this latest iteration of the project, Moonshot expanded and enhanced the targeting measures on our campaigns.

Implementation and targeting

To reach users at-risk of consuming disinformation on Twitter, Moonshot deployed its database of words and phrases indicative of risk and specific to the Indonesian context (see Appendix I for our Theme and Risk coding matrix).

This year Moonshot also integrated a new targeting methodology on the Google Display Network. Moonshot built a 'seed' list of 173 YouTube channels that shared fake news, myths or conspiracy theories in Bahasa Indonesia. We used our in-house video intelligence software to gather a sample of users who had engaged with one or more of these channels.[1] The resulting audience was defined as having been exposed to disinformation content on the platform.

127	Video intelligence software	2,000	408
'Seed' list of channels sharing disinformation		YouTube overlap channels	Total YouTube channels defined as target audience

Moonshot then ran an overlap analysis, which in this context shows the percentage of the target audience who engaged with the seed list that also engaged with another piece of content. **Overlap analysis reveals other content (disinformation or otherwise) that is popular among the target audience**. This produced a list of a further 2,000 YouTube channels for consideration.

Moonshot then manually coded each of these channels and identified a further 229 channels that had over a 50% overlap with our seed list and shared five or more videos containing fake news, debunked myths or conspiracy theories in Bahasa Indonesia. Our campaigns were designed so that consumers of all of these channels (including our original 'seed' list) were targeted with ads for the game across the Google Display Network. This process enabled access to a larger sample of YouTube users who were engaging with harmful disinformation.

---

1.   Engagement was defined by watching multiple videos to completion from one or more of these channels.

Campaign insights

Moonshot launched two forms of digital campaigns across Twitter, YouTube, TikTok and Facebook.
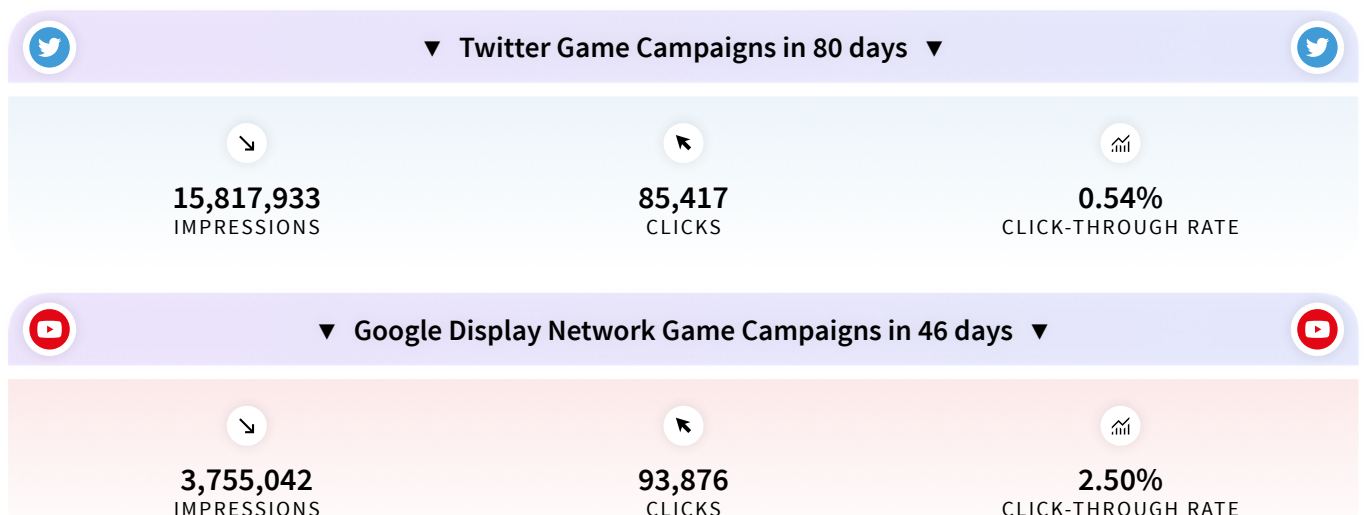
Broad campaigns

The first form of campaigns targeted broad audiences in Indonesia with the Literata website and media literacy content produced by IREX. These were run on TikTok and Facebook.

▼ TikTok Website Campaigns in 38 days ▼

22,409,137	**110,107**	**0.49%**
IMPRESSIONS	CLICKS	CLICK-THROUGH RATE

▼ Facebook Website Campaigns in 94 days ▼

4,522,735	**294,920**	**540,212**	**6.50%**
IMPRESSIONS	CLICKS	POST ENGAGEMENTS	CLICK-THROUGH RATE

Targeted campaigns

The second form of campaigns targeted users searching for and engaging with disinformation in Indonesia with the Gali Fakta 2.0 game. These were run on Twitter and the Google Display Network.

▼ Twitter Game Campaigns in 80 days ▼

15,817,933	**85,417**	**0.54%**
IMPRESSIONS	CLICKS	CLICK-THROUGH RATE

▼ Google Display Network Game Campaigns in 46 days ▼

3,755,042	**93,876**	**2.50%**
IMPRESSIONS	CLICKS	CLICK-THROUGH RATE

▼ Highest Performing Search Term ▼

🔍 **The vaccine was made by the authorities to kill the natives**

Search term that triggered the highest number of impressions (2,427,668) across all platforms.

SENSITIVE

Gali Fakta: The disinformation inoculation game
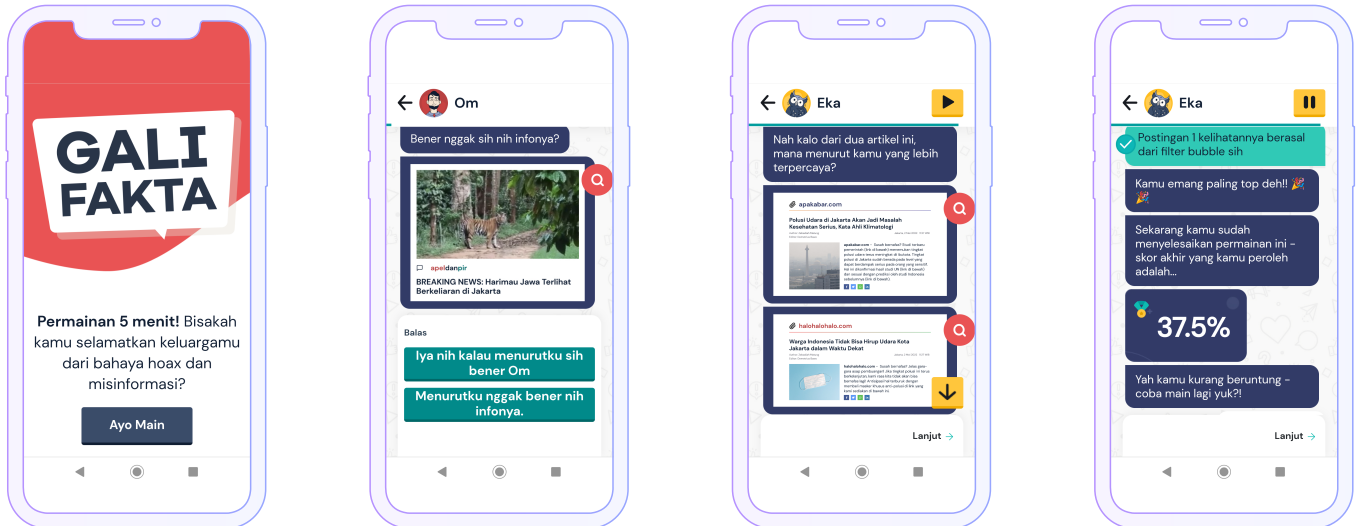
Moonshot designed the 'Gali Fakta' game to equip Indonesian audiences with the media literacy skills to identify disinformation in the real world, both online and offline. Moonshot employed the skills of Indonesian game developers, Brave Factor, to help create the game.

The audiences' vulnerability to disinformation was determined by their online behaviour. For example, users who searched for a news story known to be false were provided with an ad for our digital campaign.

The game included real examples of disinformation, in a carefully designed manner. This approach is based on inoculation theory, which outlines that psychological resistance to disinformation can be developed by exposing individuals to weakened versions of fake or manipulated stories that they will come across in the real world.[2]

WHAT IS INOCULATION THEORY

Inoculation theory was first developed by social psychologist William J. McGuire in 1961. McGuire found that people tend to defend their beliefs by avoiding information that refutes or disproves their views and actively searching for information that supports their beliefs. This behaviour leads to a lack of resilience and critical thinking in the face of disinformation, especially when that disinformation is communicated in a compelling way and aligns with an individual's original beliefs.

Inoculation theory suggests that we can create psychological resistance to disinformation by exposing individuals to forms of disinformation alongside refutations and explanation of manipulation tactics.

2. In 2018, researchers at the University of Cambridge were first to explore how inoculation theory could be developed into online gamified content that could be used to pre-emptively debunk, or 'pre-bunk', fake news by exposing people to the methods used to create and spread disinformation. They built an online game called 'Bad News', where users played the role of 'propaganda producers' to help them identify real world disinformation. The game was found to reduce the susceptibility to fake news headlines by an average of 21%. Similar projects include Harmony Square and Go Viral! - both examples of how inoculation theory has been deployed in engaging formats online. For more information, visit: https://inoculation.science/

Content and design

The game is designed to look like a group chat among family members on a messaging app. This is an entertaining and familiar context, which also reflects the real-world nature of how disinformation spreads in Indonesia.

The game script transposes ten lessons from a curriculum designed by IREX. The lessons include how to identify when you are in an echo chamber; how to spot a fake social media account; and how to tell a reliable source from an unreliable one.

Centred around examples of disinformation, the ten media literacy lessons are incorporated into the chat through conversation, questions and social proof. Correct answers are rewarded by points and the family reaching a consensus. Incorrect answers are docked points and met with general confusion by family members. The user's 'cousin', Eka, is a media literacy expert and functions as a corrective voice should the user get any answers wrong.

In addition to Eka stepping in to correct them, users who answer incorrectly are also immediately given the chance to correct themselves. Should they stick to their original incorrect answer, they lose more points and receive an explanation from Eka.

The game only exposes users to select pieces of disinformation in prepared, carefully managed settings (the 'chat group') to build resilience, with a disinformation warning or rapid feedback loop so that users were clear on what was true and what was false.

The initial design of the game went through multiple reviews. Firstly, the script was audited, translated and transcreated by Indonesian subject matter experts. These experts paid special attention to the casual tone of the script as well as the cultural resonance of examples and jokes. Moonshot then organised testing sessions with a total of 18 users who served as audience proxies. Users were selected based on their varied demographic and geographic characteristics, as well as for their varied experience with messaging apps and social media. User feedback was then incorporated into the game navigation, design, and script.

Moonshot launched the game on 6 April 2021. Throughout the project, Moonshot has continued to make improvements to the game based on user retention and bounce rates.

Game updates and improvements (since 2021)

Several improvements have been made to the game based on user retention and bounce rates.

The first major update gave users a thematic choice at the outset of the game. Users can now pick the topic they are most interested in, from health and beauty, news and politics, and online investment schemes. The topics selected were informed by Indonesian consultants with expertise on the latest disinformation trends occurring on WhatsApp in the country. Regardless of what lesson they choose, users are faced with media literacy challenges from different family members and at the end of the game they are given the opportunity to try out the other topics as well.

This change to the game's structure enabled us to decrease the game's length, which was also the second major improvement introduced. In 2021, gameplay was an average of 5 minutes and 8 seconds of what developers estimated to be a 15-minute game. Moonshot and Brave Factor shortened the game such that it now takes around five minutes to play.

The third update was the integration of a survey to measure self-reported change in user knowledge in identifying disinformation. Last year, users who played the game were later retargeted on social media with ads for a survey. Although this was successful, Moonshot sought to improve the user experience by reducing the number of ads and links they needed to see and click on. We also hoped this would improve user retention and allow us to survey more users.

In the 2021 pilot, Moonshot could not control for the amount of time between a user playing the game and seeing an advert for a survey, which may have impacted our results. In the 2022 phase of the project, Moonshot integrated the survey into the game. We can now also connect data on user behaviour with survey responses, which was one of the priorities highlighted in the 2021 report.

Survey design

The integration of the surveys into the game meant that the survey content needed to be updated so that it 1) would not be too long in duration 2) take too much time away from the lessons in the game and 3) fit with the game's overall tone of voice. Moonshot worked with behaviour change measurement expert Dr. Michael Williams to devise the optimal number and set of survey questions.

Users were presented with four pre-game questions and four post-game questions to measure their knowledge in identifying different forms of disinformation to answer three specific aims:

Did the game improve users' self-reported knowledge in identifying disinformation?

(Pre-game self-reported knowledge vs. post-game self-reported knowledge)

Did the game improve users' self-reported knowledge in any particular topic of the game?

(Pre-game self reported knowledge vs. post-game self-reported knowledge across health and beauty, news and politics, and online investment schemes)

Did the game improve users' self-reported knowledge for any particular training aspect of the game?

(Pre-game self reported knowledge vs. post-game self-reported knowledge for identifying 'fake accounts'; 'confirmation bias'; 'trustworthy sources', and 'filter bubbles').
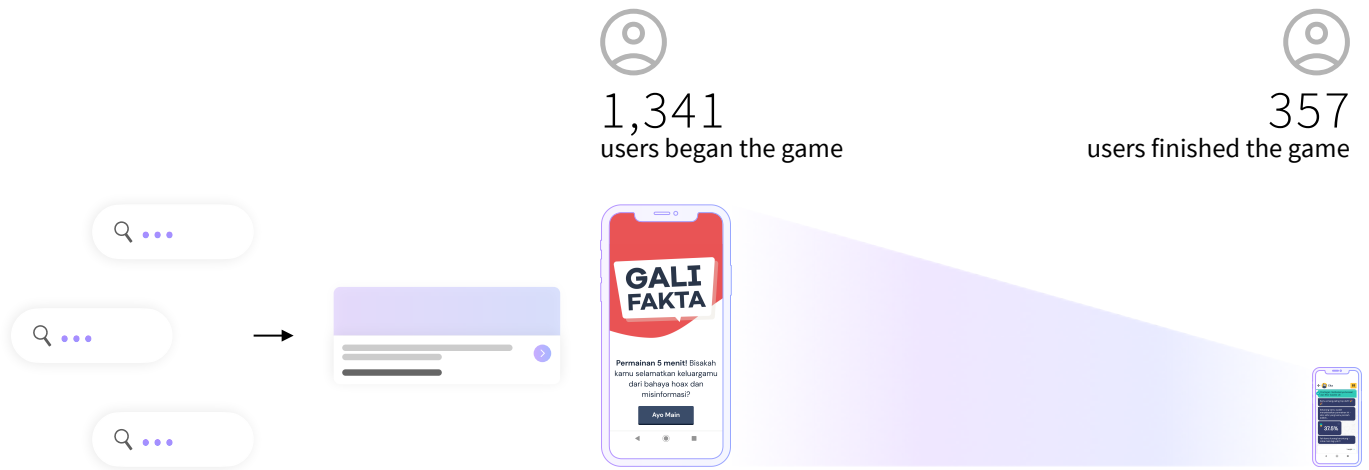
Another significant improvement to the experiment was the shift from a between-groups design to an in-groups design. In 2021, Moonshot used a between-groups design, where survey users were assigned to either the treatment group or control group based upon whether they had previously engaged with the media literacy content. However there were several limitations to this approach, as we could not match our control group and treatment group demographics. In this year's programming, an in-group design was selected to 1) remove this limitation and 2) focus the study on the impact of the game on its users.
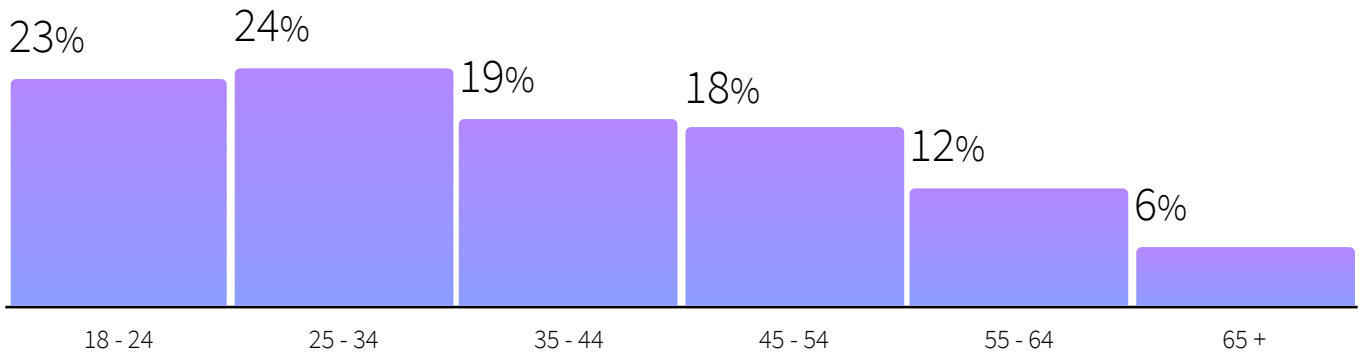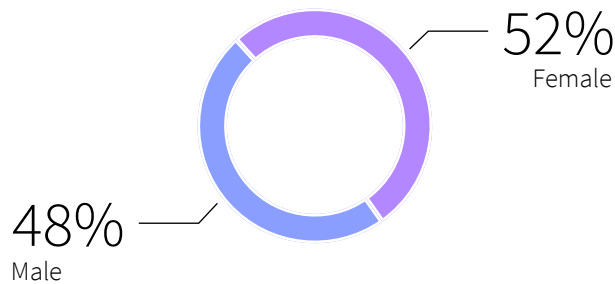
Results

Gali Fakta
How users interacted with our media literacy content

1,341
users began the game

357
users finished the game

▼ Age of game players

23% 24% 19% 18% 12% 6%

| 18 - 24 | 25 - 34 | 35 - 44 | 45 - 54 | 55 - 64 | 65 + |

▼ Gender of game players

52%
Female

48%
Male

Of those who began the game (1341 users), 357 users played it all the way through. 28 users were removed from the sample for completing unrealistically quickly (the data alleged that they completed both the pre-game and post-game questions in less than 2 minutes).

Did we impact user knowledge in identifying disinformation?

Despite the new updates to the game and survey design, the lasting effect of this content on changing users' knowledge and understanding towards disinformation online remains inconclusive. Users' test scores pre-game vs. post-game were significantly different, but not in the direction as intended (e.g., achieving a higher score post-game). However, it is likely that the survey design and/or content were not effective in measuring changes in users' self-reported ability to identify disinformation. The post-game questions were designed to be slightly more difficult than the pre-game questions to more rigorously test improved knowledge. However, there is a chance that the post-game questions were too difficult relative to the pre-game questions. It is also possible that user fatigue may have contributed to poorer post-game performance: questions that appear later in a survey or, in this case, the game, can often be more challenging for users to complete due to waning attention, capacity for reasoning, and motivation. Moonshot will be carrying out further tests to more robustly measure knowledge change.

↘ Conclusions and recommendations

Together with the University of Notre Dame, IREX, and GeoPoll, Moonshot completed the third phase of this project to advance media literacy among new digital arrivals for the USAID Mission in Indonesia. Moonshot developed and optimised a new version of 'Gali Fakta', a disinformation inoculation game. Moonshot also worked to broaden the reach of digital campaigns to promote the media literacy hub and direct more users to the game.

The programme has been extended for a further three years to enable future adaptation, testing and opportunity to advance media literacy in Indonesia.

Future programme design will focus on the following priorities to better capture the impact of the media literacy content:

Recommendations

01

Redesign pre-game and post-game questions to ensure similar difficulty level

Given it is possible that the post-game questions were too difficult compared to their pre-game counterpart questions, which were potentially too easy, we should redesign both sets of survey questions and identify ways to make them the same level of difficulty and ensure users have room to improve.

02

Restructure the game script

In the current version of the game, users are presented with the four pre-game questions prior to playing the game, and all four post-game questions after completing the game. Restructuring the game script so that users are presented with each pre-game question before engaging with the relevant part of the game, and then immediately presenting users with the corresponding post-game question may be effective in addressing potential user fatigue and response rate. This structure may also provide users a more rewarding journey as they are given the opportunity to learn something new and then be immediately tested on what they have learned.

03

Improve the survey design

We should revisit the different designs and other visual aspects of both the pre-game and post-game questions to ensure both parts are equally visually 'appealing and engaging'. The timing of the surveys will also be tested to evaluate if a 'cooling off period' between exposure to inoculation and post-testing would impact survey results.

Appendix

Theme and risk coding matrix

Disinformation Theme	Description	Example
Anti-Chinese	Anti-Chinese sentiment has flared up periodically since independence and has been instrumentalized during contemporary political campaigning. It intersects with disinformation, where malicious actors falsely claim that actors or policies are 'Chinese' or otherwise connected to the Chinese government in order to discourage support for them or instil general fear and distrust.	'Jutaan TKA masuk Indonesia' Millions of Chinese migrant workers enter Indonesia
Anti-Communist	Communism has been a source of tension in Indonesia since the founding of the communist party (PKI) in 1914. Allegations of sympathy with communism are deployed to smear opposition leaders and parties, and other forms of dissent.	'Ibu Jokowi PKI' Jokowi's [President Joko Widodo's] mother is in the communist party
Anti-Papua	Anti-Papua disinformation refers to any stories that are misleading about individuals from Papua, the Free Papua movement, or events occurring in Papua. These are often false reports of anti-Muslim attacks committed by Papuans or that Papua is receiving more state-funding.	'papua bakar masji' Papuans burn down a mosque
Islamic Chauvinism	Islamic chauvinist disinformation refers to the spread of disinformation or smear campaigns, which ultimately advocate that Indonesia should be an Islamic state or that Indonesian Muslims have a greater claim over the land. Such narratives can be non-violent in nature, referring to the belief that the government or politicians are not Islamic enough or that they are directly threatening Muslims in some way.	'Prabowo kafir' Prabowo [Subianto, candidate in the 2014, 2019 Presidential Elections and Minister of Defence of Indonesia] is a kafir
Political Smear	This category includes searches for stories that seriously and falsely besmirch the credibility of politicians or electoral institutions.	'pra-diisi surat suara ditemukan' Pre-filled ballots found
Health Disinformation	Health disinformation refers to information concerning sickness, remedies, preventative measures that are false, inaccurate, or misleading according to best available evidence at the time of reporting.	'vaksin menyebabkan autisme' Vaccines cause autism
COVID-19 Mis/Disinformation	Similar to health disinformation, this category comprises mis and disinformation specific to COVID-19. This includes remedies or prevention techniques, conspiracies about government or global responses, or misleading stories about the nature of the virus.	'bawang putih melawan virus' Garlic fights the virus

To add granularity and nuance to our database, Moonshot coded each keyword according to the level of risk it represented:

At risk: users engaging with disinformation (e.g. 'the pandemic is a hoax')

Sceptic: users inquiring about hoaxes (e.g. 'is it true Jokowi is PK')

Media literate: users searching to fact check their information (e.g. cek fakta virus). These users did not form part of the digital campaigns audience, i.e. they were not surveyed or redirected to the media literacy content. These keywords were used as a baseline as well as for trend monitoring.

Database development and risk coding system

I.   **Reviewing existing databases**
      Moonshot collated and reviewed risk indicators from its proprietary risk-indicator disinformation databases.

III. **Database expansion**
      Moonshot consulted with Indonesian subject matter experts to review and expand its disinformation databases to ensure the relevance and appropriateness of each risk indicator.

V.   **Database coding**
      The Moonshot team and Indonesian subject matter experts categorised all risk indicators according to the broad disinformation narrative they were based upon.

VII. **Keyword maintenance**
      Databases were closely monitored and maintained to identify and remove any inappropriate terms or false positives.

IX.   **Ongoing expansion**
      Throughout the campaigns, Moonshot regularly updated the database of disinformation risk indicators to reflect the findings from research carried out in conjunction with subject matter experts and Indonesian fact-checking organisation, Mafindo.