SENSITIVE

moonshot

Advancing Media Literacy in Indonesia

Building resilience and measuring behaviour change

Final Report
November **2021**

Contents

Project overview

Since 2019, Moonshot has worked in partnership with the University of Notre Dame[1], IREX and GeoPoll to advance media literacy amongst 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 and content produced by IREX.

In late 2020, Moonshot continued these digital campaigns with updated content, including a disinformation inoculation game. This final report summarises the results of those updated campaigns and a behaviour change experiment, both of which ran from April 2021 to October 2021, as well as recommendations for future programming.

Project results

Both the website and game proved effective at taking user attention away from disinformation content towards positive media literacy content. This report provides a blueprint for future programming and priorities to better capture the impact of the media literacy content.

Project insights

Digital campaigns — **Google Search, Google Display, Twitter**

Number of impressions:	Number of times our ads were clicked on:	Average click-through rate:
1,253,722[2]	67,376	1.42%

The website — **Literata.id**

Number of users who visited Literata.id:	Average time users spent on the site:
24,581	26 seconds

1. The University of Notre Dame and UND will be used interchangeably throughout this report.
2. We collected a further 16,658,508 impressions triggered by broad match keyword searches as part of the Google Display campaign.

The game **Gali Fakta**

Number of users who played Gali Fakta:

781

Average time users spent playing the game:[3]

5 minutes 8 seconds

Post-survey campaigns

Number of times our ads were shown:

Treatment: 2,215,281
Control: 284,849
Total: 2,500,130

Number of times our ads were clicked on:

Treatment: 21,302
Control: 1,458
Total: 22,760

Click-through rate

Treatment: 0.96%
Control: 0.51%
Total: 0.91%

Number of control group users who completed our survey:

227

Number of treatment group users who completed our survey:

127

Behaviour change analysis

Main finding:

The impact of the media literacy content on behaviour change was inconclusive. Differences in age and education levels made it impossible to measure whether or not there was any statistically significant behaviour change. When these demographics were controlled for, the sample size was too small to conclude any statistical significance. This programme will be followed up with a further study that will address these outcomes.

3. Average among users who started the game. Metric includes multiple sessions played by the same player.

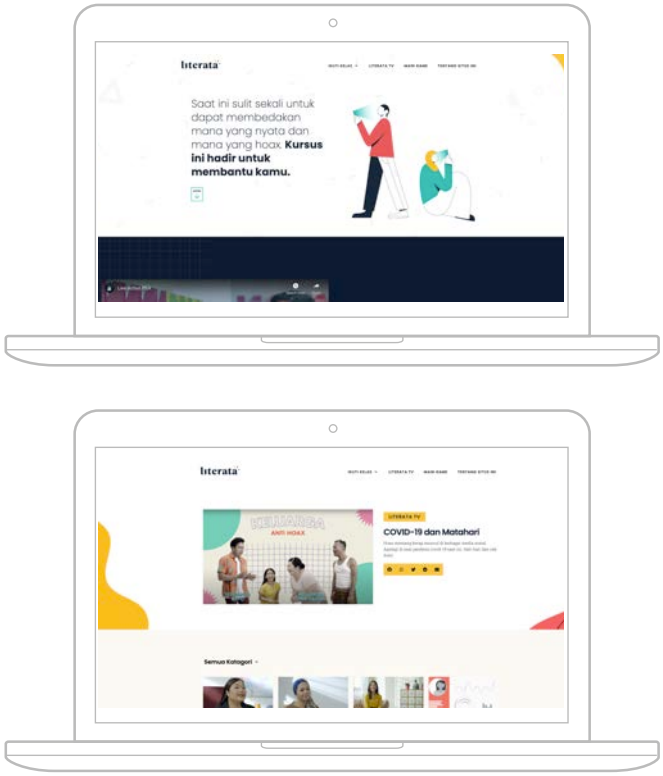
Methodology

Literata.id: The media literacy website

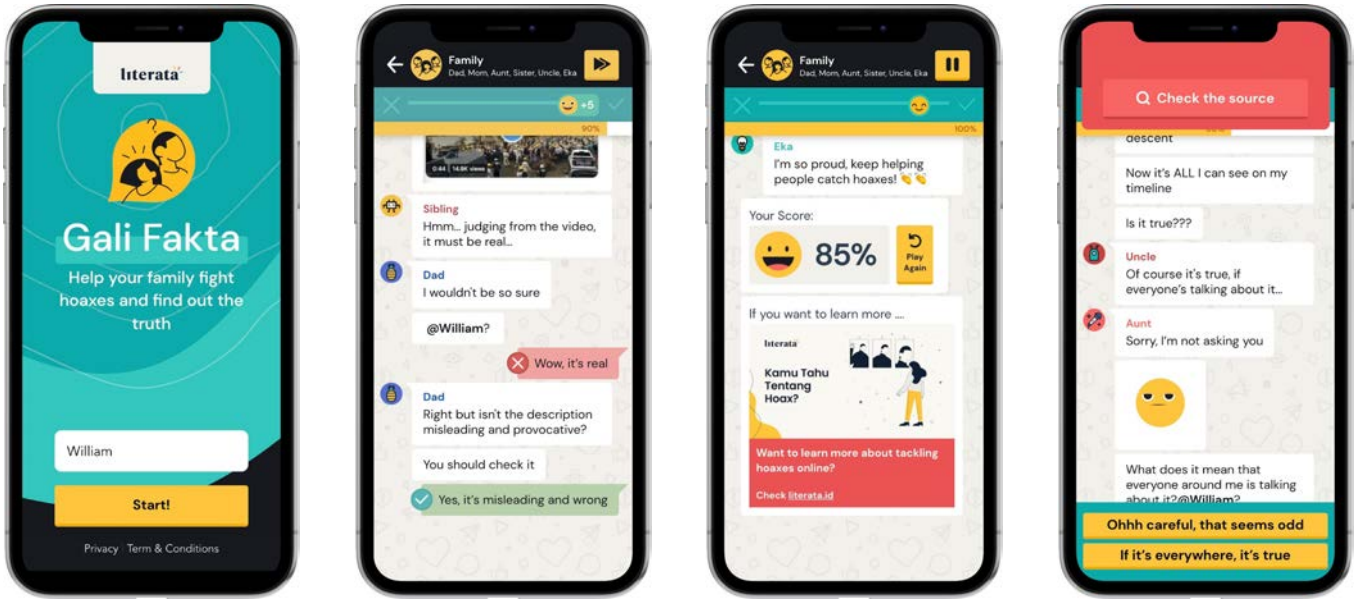
Between September 2020 and March 2021, IREX developed the Literata brand and with it, the Literata.id website, in collaboration with the University of Notre Dame. The website was designed to be an online companion to guide users in how to deal with misinformation on the internet and ultimately improve their media literacy skills.

Literata.id contains eight lessons from the IREX curriculum, including how to spot echo chambers, how articles can use language to mislead and how to identify different types of disinformation. Each lesson page comprises a video detailing the contents of the unit as well as a quiz about its contents to test user engagement and understanding.

The site also contains videos (titled Literata TV) of two digital series produced by IREX. The first of the series is the eight media literacy lesson videos. The second series is 'Keluarga Anti-Hoax' ('Anti-Hoax Family'), a comedy series about a family dealing with COVID-19 disinformation.



Gali Fakta: The disinformation inoculation game



Between September 2020 and April 2021, Moonshot designed, developed and deployed a media literacy game designed for an Indonesian audience. The goal of the game is to inoculate people who are otherwise vulnerable to disinformation and equip them with the media literacy skills necessary to identify it in the real world, both online and off.

Our audience's vulnerability to disinformation was determined by their online behaviour. For example, when they searched for a news story we knew to be false, they were served an advert for the game. This way we can say that everyone who played the game was otherwise actively trying to engage with disinformation.

The content

The game script transposes ten lessons from the IREX curriculum. These include lessons similar to those on the website, such as: 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.

The game centres these media literacy lessons around examples of disinformation. The game only exposed users to select pieces of disinformation in prepared, carefully managed settings in order to build resilience, with a disinformation warning or rapid feedback loop so that users were clear on what was true and what was false.[4]

The intention behind using real disinformation in the game was informed by the inoculation method, developed by researchers at the University of Cambridge. The method is based on the theory 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.

Since the game would only be shown to users who were otherwise searching for or engaging with disinformation (see page 8), this provided Moonshot with a unique opportunity to maximise the chances of inoculation and behaviour change. The game script was reviewed by inoculation theory scholars and Moonshot's internal ethics committee.

The design

The game is designed in the style of the universally recognised family chat group. This is an entertaining and familiar context which also reflects the real-world nature of how disinformation spreads in Indonesia; Whatsapp is the most popular messaging application in Indonesia and family chat groups make up over 70% of Indonesian user activity on the platform.[5] Whatsapp was named by the Indonesian government alongside Facebook as one of the platforms contributing to the spread of disinformation in the wake of the 2019 Gubernatorial elections.[6]

The ten media literacy lessons are infused into the chat through leading 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 player's 'cousin', Eka, is a media literacy expert and functions as a corrective voice should the player get any answers wrong. She is busy teaching, hence the player steps in to help the family, but in a private chat with the player she intermittently relates their answers - right or wrong - back to broader lessons and tips.

4. Another version of the game has also been created with low-risk examples for the purposes of general consumption.
5. Novi, Kurnia & Engelbertus, Wendratama & Rahayu, & Wisnu, Adiputra & Syafrizal, Syafrizal & Zainuddin, Monggilo & Utomo, Wisnu & Aprilia, Monika & Afita Sari, Yuni. (2020). WhatsApp Group and Digital Literacy Among Indonesian Women.. 10.13140/RG.2.2.12351.05289.
6. Cuthbertson, Anthony. "Facebook and WhatsApp Blocked in Indonesia after Deadly Riots." The Independent, Independent Digital News and Media, 23 May 2019, https://www.independent.co.uk/life-style/gadgets-and-tech/news/indonesia-facebook-whatsapp-ban-blocked-election-riots-a8926706.html.

In addition to Eka stepping in to correct them, players 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 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 user testing sessions with a total of 18 participants who served as audience proxies. Participants were selected based on their demographic and geographic characteristics, as well as for their varied experience with messaging apps and social media. Participant feedback was then incorporated into the game navigation, design and script.

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

Digital campaigns

The Redirect Method / Overview

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, including content not created for the explicit purpose of countering harm, to challenge narratives which 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 we have 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 - rather than unreliable and often problematic demographics, such as their age, gender or religion.

Digital campaigns / Design and implementation

Moonshot designed and implemented six separate Redirect Method campaigns on three platforms: Google Search, Google Display and Twitter. These campaigns were designed to engage individuals expressing interest in or engaging with fake news, debunked myths or conspiracy theories. The goal was to redirect them to our media literacy content, either on the website Literata.id or in the game Gali Fakta.

To reach users at risk of consuming disinformation, Moonshot deployed its database of words and phrases indicative of that risk and specific to the Indonesian context. In total, 1,250 risk indicators were coded depending on their theme, with the themes corresponding to one of six narratives commonly found in the Indonesian disinformation space, and listed below.

Theme and risk coding

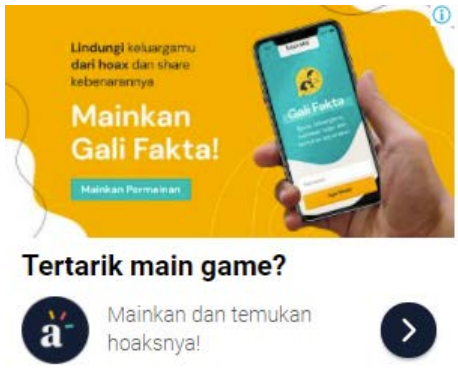
Disinformation Theme	Description	Example
Anti-Chinese	Anti-Chinese sentiment has flared up periodically since independence and has been instrumentalised during contemporary political campaigning. It intersects with disinformation, where malicious actors falsely claim that actors or policies are 'Chinese' 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. Today, nationalist and Islamist parties and activists deploy allegations of sympathy with communism 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 masjid" *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. Although these narratives are similar to the dangerous narratives perpetuated by violent extremist groups, the conservative groups that have been known to spread Islamic chauvinist disinformation do not support or condone violent extremist activities. Searches in this category include those falsely claiming 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 Misinformation	Health misinformation 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 misinformation, 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 (from highest to lowest):

- **At risk:** users engaging with disinformation (e.g. "the pandemic is a hoax")
- **Sceptical:** users inquiring about hoaxes (e.g. "is it true Jokowi is PKI")
- **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.

Audience messages

The text and image in our advertisements functioned as the entry point to our campaigns and their media literacy content. The language, tone and imagery was designed to attract and engage user attention based on the specific need they have expressed (i.e. the search term). Images reflecting the colours and branding of Literata were designed and used in the Display and Twitter campaigns. Moonshot drew on previous work in the region when defining these messages and images, applying a data-driven approach with a series of quality assurance checkpoints.



An example of one of our campaign adverts in Bahasa Indonesia.[7]

Implementation and targeting

Between April and October 2021, we ran our Redirect campaigns on Twitter, Google Search and Google Display, targeting audiences in Indonesia. We ran two campaigns consecutively on each platform: one redirecting to our website's homepage, the other redirecting straight to the game.

Platform	Destination	Dates	Timeframe
Twitter	Gali Fakta	April 6th - May 5th	30 days
	Literata.id	May 7th - June 6th	31 days
Google Search	Gali Fakta	June 7th - July 5th	29 days
	Literata.id	July 6th - August 3rd	29 days
Google Display	Gali Fakta	August 6th - August 31st	26 days
	Literata.id	September 1st - October 18th	48 days*
*This campaign was extended for two weeks to make up for the previous two weeks being lost to technical interruption.			

For the purpose of analysing the effectiveness of the campaigns, unique UTM links were generated for each campaign. Results can be found in the following section on page 12.

7. In English, the image translates to *Save your family from hoaxes and share the truth. Play Gali Fakta!* The ad text translates to *Love to play games online? Play our free game and become a pro in catching hoaxes!*

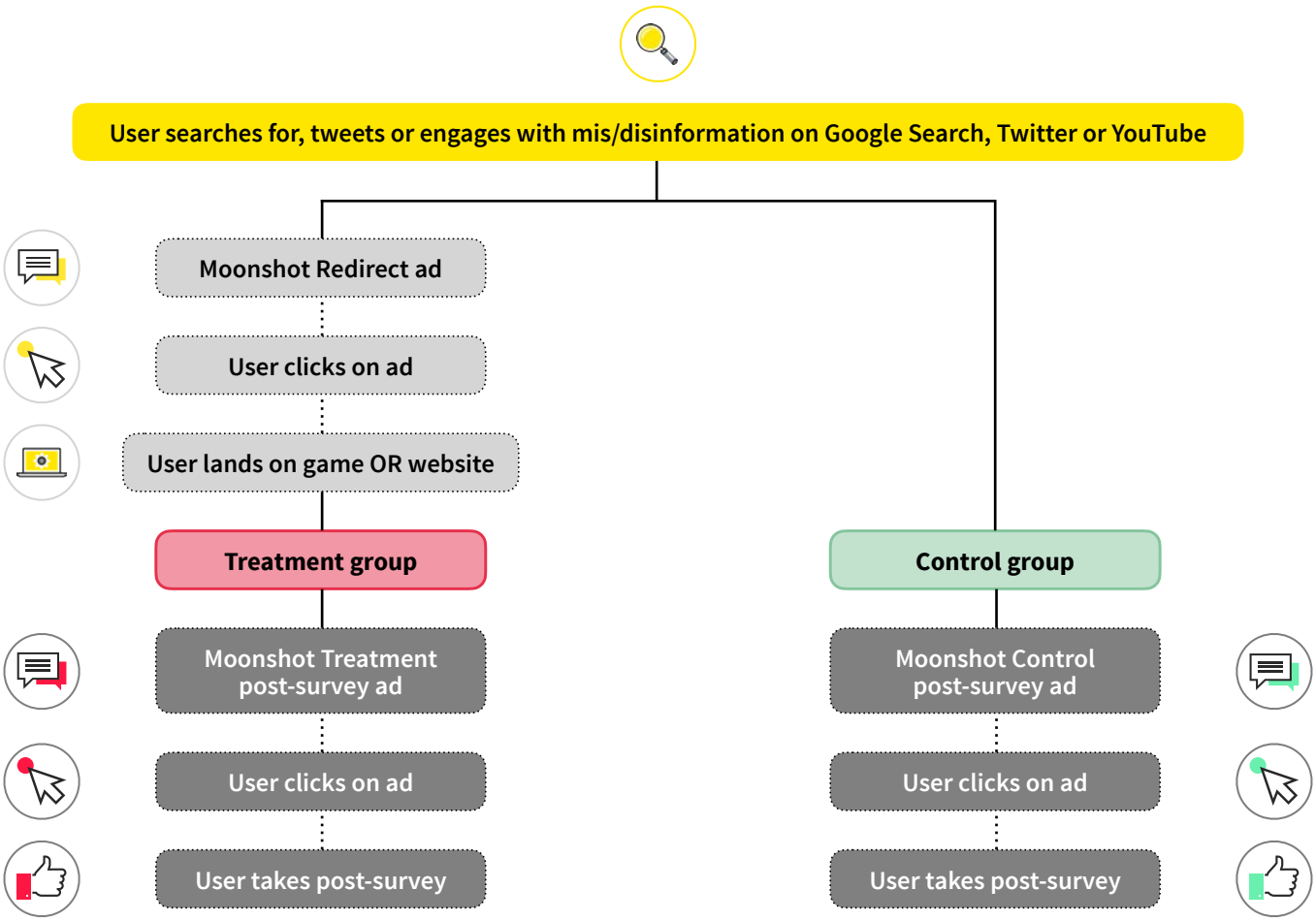
The post-surveys

Once the digital campaigns were complete, Moonshot used post-surveys to measure whether engagement with the media literacy content had impacted participants' self-reported intentions to respond responsibly and proactively to online disinformation.

Moonshot used a between-groups design, where survey participants were assigned to either the treatment group or control group based upon whether they had previously engaged with the media literacy content.[8]

Control group participants were users located in Indonesia who had voluntarily clicked on an ad which had itself been triggered by a particular behaviour - specifically the user had to have searched for, tweeted, or interacted with disinformation stories via Google, Twitter, or YouTube.

The treatment group comprised users who had engaged with our content in any way. Custom audience lists on Facebook and Twitter's advertising platforms allowed Moonshot to retarget users who had visited either the game or the website with ads for the post-test survey.



8. See Appendix for complete list of post-survey questions

Results

Digital campaigns: How our campaigns performed

Twitter Campaigns (April 6 - June 6)

Impressions:	Clicks:	Click-through rate:
1,243,603	3,173	0.26%

Google Search Campaigns (June 7 - August 3)

Impressions:	Clicks:	Click-through rate:
10,119	511	5.05%

Google Display Campaigns (August 6 - October 18)

Impressions:	Clicks:	Click-through rate:
16,658,508	63,692	0.38%

The figures from Twitter and Google Display are significantly higher because those platforms only allow for broad match keyword targeting, whereas Google Ads allows phrase and exact match targeting.

"Covid is a hoax" → Search term which triggered the highest number of impressions on Google Search (8,460)

"The sinovac vaccine doesn't work" → Search term which triggered the highest number of impressions on the Google Display network (16,563,027)

Literata.id and Gali Fakta: How users interacted with our media literacy content

The following data from Google Analytics refers only to users who (a) are in Indonesia and (b) landed directly on either the website or the game having clicked on one of our ads.

Overall, 24,581 individuals visited the website from Twitter, Google Search and Display, of which 72 took a quiz and 289 watched a video.

8,128 individuals viewed the game page, of which 781 started playing it and 98 completed it.

The website had more visitors and a lower bounce rate than the game. On the website, users can access content right away, whereas the game requires the user to type in a username and commit '10 minutes of their time'.

Despite the website's lower bounce rate, the game was more effective at engaging users and maintaining their attention. **Game players spent approximately 12 times (1,185%) as long engaging with the media literacy content than those who visited the website without playing the game**. Results were statistically significant at the 95% confidence level ($p < .05$).

Website

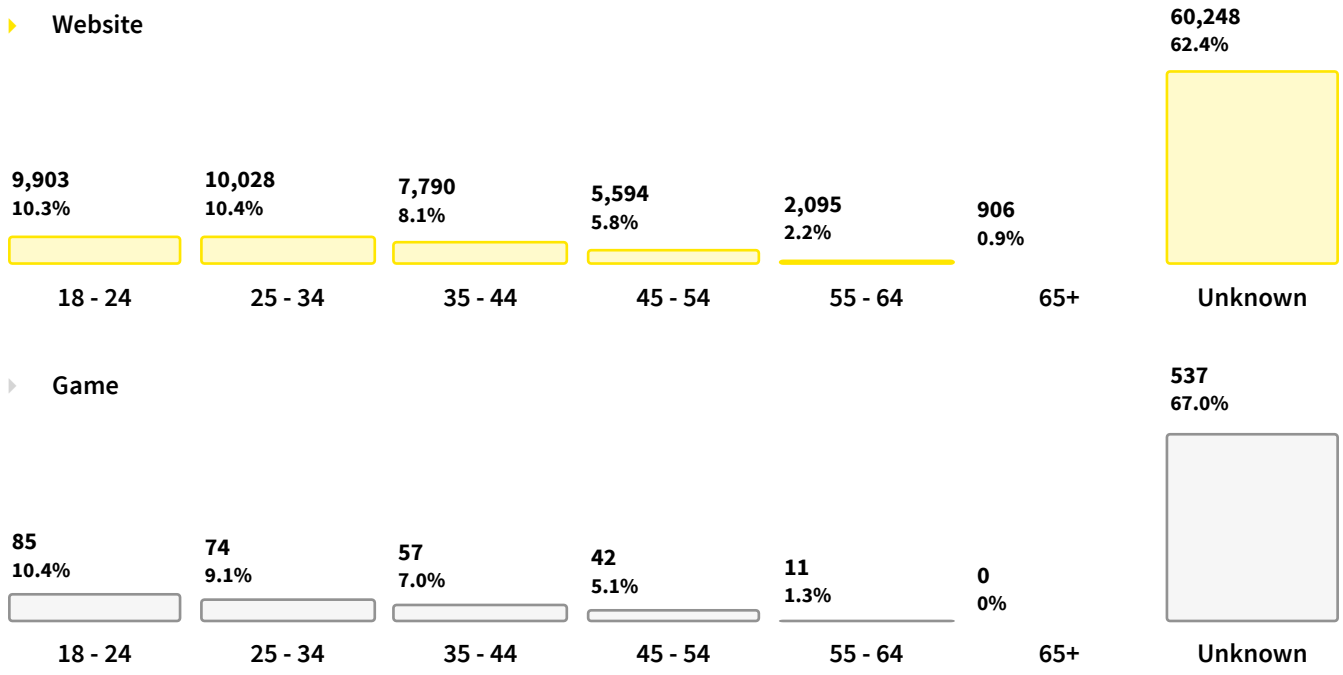
Visitors to site: 24,581	**Average time spent on website:** 26 seconds	**Bounce rate:** 60.9%
Percentage of users who accessed content on website (videos or quizzes): 1.4%	**Total who started quiz:** 72	**Total who watched video content[9]** 289

Game

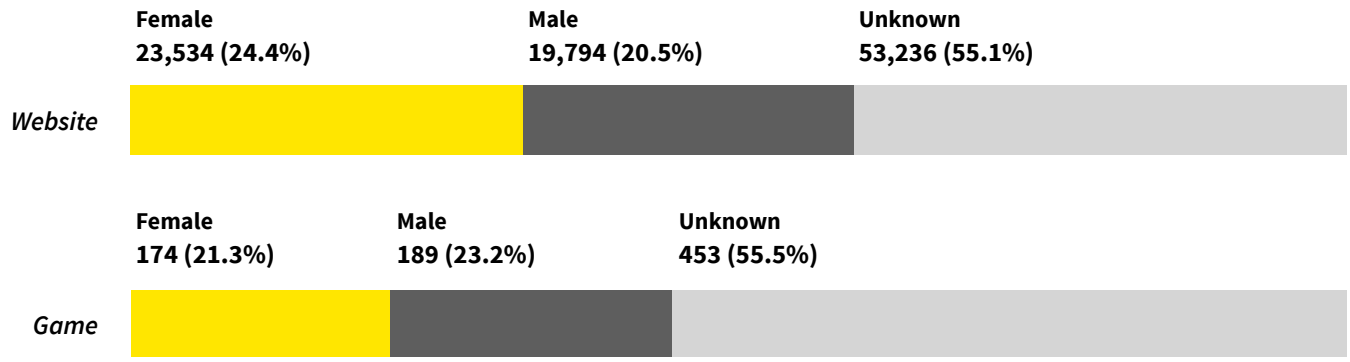
Visitors to game page: 8,128	**Total game players:** 781	**Total who completed game:** 98
Bounce rate[10]: 90.4%	**Average gameplay duration:** 5 minutes 8 seconds	

9. Number of users who started at least one video.
10. Bounce rate here was defined by the percentage of users who landed on the game's welcome page and did not click on "start game".

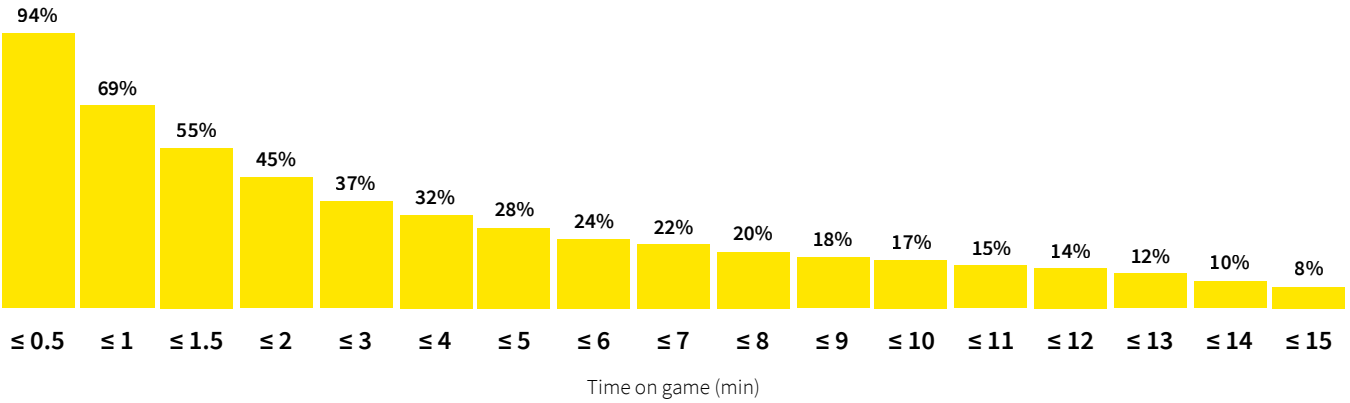
Age of users accessing media literacy content



Gender of users accessing media literacy content



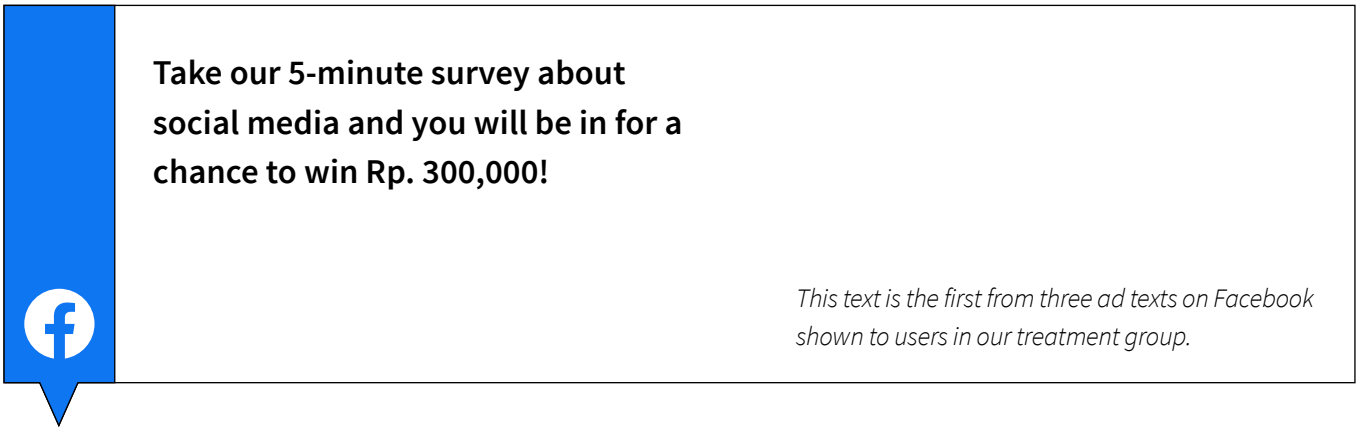
Game retention rate over time



The post-surveys: Did we impact participant behaviour?

Number of control group users who completed our survey:	**Number of treatment group users who completed our survey:**
227	127

In October 2021, Moonshot analysed the results of the control and treatment group surveys to understand the impact of the media literacy content on participant's self-reported intentions to respond responsibly and proactively to online disinformation.



Moonshot found that the impact of the media literacy content, as per the surveys, was inconclusive. The control respondents performed better in the survey and showed, on average, a higher likelihood of responding responsibly to disinformation online. However, the control groups self-reported as being significantly younger and more highly educated than the treatment group. Therefore this apparent counterproductive effect of the media literacy content was found to correlate (in large part, if not completely) by differences in age and education.

When these demographics were controlled for, the sample size was too small to conclude any statistically significant effects.

The large discrepancy in the age and education levels of the two groups may be a factor of how much time each group was prepared to spend engaging with our campaign. Its design demanded more of the users who took the treatment survey than it did of the users who constituted the control. The treatment group was asked to visit a website or play a game and then complete an online survey, whereas the control group only needed to complete the survey.

This outcome raises a number of considerations for future programming. In particular, how to control for age and level of education in an experiment which is trying to reach people regardless of their demographics. Potential solutions and mitigations will be discussed in the following section.

Conclusion and priorities for future programming

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 newly developed brand Literata, included the Literata.id website produced by IREX and UND, as well as a disinformation inoculation game, Gali Fakta, created by Moonshot.

The website was effective at immediately securing the attention of users, whereas the game was more effective at retaining the users who decided to play it.

The lasting effect of this content on user's behaviour towards disinformation online was inconclusive due to large discrepancies in participant age and level of education.

The programme has been extended for four years to enable further adaptation, testing and opportunity to advance media literacy in Indonesia. **Future programme design should focus on the following priorities to better capture the impact of the media literacy content**:

I.

Introduce a placebo in place of treatment for the control group

This priority is based on the assumption that the discrepancy in the age and education levels of the treatment and control groups was caused by the uneven demands our experiment placed on their free time.

Our proposed solution to this problem is to introduce a placebo (or 'sham') treatment to users in the control group. This addition would make the demands of being in the control group equivalent to those of the treatment group.

In this case, the sham treatment could be a similarly engaging or informative online game or website that is not directly related to media literacy.

By making the treatment and control conditions similarly demanding of participants' time, both conditions would be rendered similarly appealing - or unappealing - to potential participants.

II.

More robustly connect user behaviour to their survey responses

The current programme design did not connect data on user behaviour from Google Analytics with the users who responded to surveys. Attributing multiple behaviours to a persistent user in this way is notoriously challenging, but the stronger the link between user behaviour and the results of their post-treatment survey, the more likely it is that we will be able to confidently measure any changes in behaviour.

Moonshot will explore a number of possible solutions, including new survey software which would enable us to pseudonymously connect user behaviour with pre and post surveys in future programming.

III.

Broaden re-targeting of surveys beyond Facebook and Twitter

Moonshot will explore the possibility of broadening the re-targeting of the post-test beyond Facebook and Twitter. These platforms were chosen based on their ability to retarget users who had accessed the site. However, this meant that the methodology was dependent on users having either a Facebook or Twitter account.

Moonshot will explore whether the same survey software cited above provides the solution, as it is platform agnostic. This method would considerably shorten the time between treatment and survey, which may in turn affect user responses. For greater distance between the treatment and survey, Moonshot could collaborate with our partner GeoPoll to SMS or call users after their treatment (or sham treatment).

Appendices

Appendix I: Post-survey questions (English translation)

GOAL: *Collect endline data for treatment and control group to assess how well audiences identify disinformation and understand how it is spread*

	Questions	Options
Introduction statement	Hi, please take our short 5 minute survey about information online. If you complete it, you will be entered into a prize draw with a chance to win a Telkomsel voucher worth 300,000rp!	—
Question 2	How long have you been using social media?	Never used social media
		Less than 1 month
		About 6 months
		About 1 year
		About 2 years
		3 years or more
Question 3	How long have you been using messaging apps?	Never used messaging apps
		Less than 1 month
		About 6 months
		About 1 year
		About 2 years
		3 years or more
Question 4	How often do you see disinformation shared on social media?	All the time
		Once a week
		Every day
		Once a month
		Very rarely
		Never
		Not sure

	Questions	Options
Question 5	How often does someone send you a fake story on messaging apps?	All the time
		Once a week
		Every day
		Once a month
		Very rarely
		Never
		Not sure
Question 6	If a friend or family member sends you a story on the messaging app you use, do you check it?	No, I don't check anything.
		I trust whatever information my family or friends send me.
		I might check it, but I do not feel comfortable to challenge them if it's a hoax.
		I check it and challenge them if it's a hoax.
Question 7	When you search for something online, how easy is it for you to decide which sources will be the most useful for you?	It's easy, I open the first result that appears.
		It's okay, I open the headlines that make more sense to me.
		Not so easy, I look at the website URL and open those that seem trustworthy.
		Really hard, I fully read different sources to select the most trustworthy option
		I never search for anything online.
Question 8	How do you tell if a story, post or message you see online is fake? -> multiple options possible	When it doesn't contain any sources.
		When it uses stereotypes.
		When the sources it uses are not trustworthy.
		When it doesn't show any proof.
		When the same story doesn't appear in other trustworthy websites
		All of the above.
		None of the above.
		I am not sure.

	Questions	Options
Question 9	How do you tell if an article can be trusted? --> multiple options possible	I check if the article gives any proof.
		I check if the article uses trustworthy sources
		I verify if the same story appears in other trustworthy websites
		If it verifies my existing beliefs
		All of the above
		None of the above
		I don't know
		I don't verify articles
Question 10	How do you tell if a profile on social media is fake? --> multiple options possible	They do not have a realistic number of followers or posts.
		The content of their posts is not realistic.
		They are not officially verified.
		All of the above
		None of the above
		I don't know
Question 11	How accurate is the following statement for you? "When I am interested in a topic, I gather information from a bunch of different sources (like different websites, TV, radio, social media, etc)"	Completely true
		Sometimes true
		It depends
		Rarely
		I never do that
Question 12	If you saw this headline on social media, what would you do? Sudirman said "don't believe everything you see on TV!" --> multiple options possible	I would share it with people who are interested.
		I would check it and then share if it is factual.
		I would do nothing, it doesn't interest me.
		I would check it and if it's a hoax I would do nothing.
		I would check it and if it's a hoax I would tell people.
Question 13	If you saw this headline on social media, what would you do? Prime Minister of Indonesia visits local school --> multiple options possible	I would share it with people who are interested
		I would check it and then share if it is factual
		I would do nothing, it doesn't interest me.
		I would check it and if it's a hoax I would do nothing.
		I would check it and if it's a hoax I would tell people.

	Questions	Options
Break - Note	Tell us a little about yourself in the next question	--
Question 14	How old are you?	18 - 24
		25 - 39
		40 - 60
		60 or over
		Prefer not to say
Question 15	How would you describe your level of education?	No formal education
		Primary school
		Junior High School/Middle School
		High school
		Vocational training
		University
		Prefer not to say
Question 16	Have you played the Gali Fakta game before?	Yes
		No
Question 17	Have you visited the literata.id site before?	Yes
		No
Terms and Conditions	I confirm I am above 19 years old.	
	I accept the Terms and Conditions	
End page	Thank you for your time and attention! Please leave your email below if you want to have a chance to win a Telkomsel voucher worth 300,000rp	—