Digital Resilience Community - Final Report

Adrien Bouguen, Simon Briole, Quentin Daviot, Charles de Fornel

June 2021

Executive summary

This report presents the results from the experimental evaluation of the Digital Resilience Community training program, an online intervention designed to help the youth increase their resilience to fake news, conspiracy theories and radicalism. 222 youth from deprived areas were recruited to participate to this training and were first invited to complete a baseline survey that included socio-economic measures, psychometric and cognitive tests, and a measure of conspiracy. Youth included in the study typically displayed high levels of adherence to conspiracies and Fake News (60% believed in at least one conspiracy). At the end of the survey, the young people were randomly assigned to a test group and a control group: the analysis of the baseline data confirms the validity of the randomized protocol. After the program, youth from both groups were again invited to participate to a survey in order to measure the impact of the program. Generally speaking, the results point to a positive impact of the program on most of the outcomes measured in the survey (use of traditional media, belief in conspiracy theories, violence disapproval, engagement against misinformation). While the weak statistical power caused by the high attrition and the low active participation to the program make it difficult to detect statistically significant effects, we find a positive and significant effect on the use of fact-checking tools use.

1 Acknowledgements

This document reports the results from the experimental evaluation of a small-scale training program designed to help the youth increase their resilience to fake news, conspiracy theories and radicalism. A pilot version of the program was implemented online by the association ALMA in 2020 in region Ile-de-France. The evaluation was performed in collaboration with the association and the authors of this report are very grateful for all the work performed by its staff members and for the help provided all along the pilot phase, especially Alice Bougnères, Asma Berkane, Aurélia Olmi, Simon Brique and Heithem Houria.

The report has been produced on ALMA's demand. The association asked for the expertise of a research team based at J-PAL to design and implement the evaluation of the program's pilot. The research team is made up of Adrien Bouguen, Assistant Prof. at the University of Santa Clara, US, Simon Briole, Postdoctoral Researcher at J-PAL, Paris School of Economics, and Quentin Daviot, Research Manager at J-PAL, Paris School of Economics. The authors are very grateful to Marin Gillot for his excellent research assistance. The authors of this report declare no conflict of interest and have no link of any kind with the association ALMA.

2 Digital Resilience Community program

This first section of the report provides background information on the rationale of the evaluated program and presents its different steps in details.

2.1 Online radicalism, fake news and social media

The emergence of social media as communication and information systems has been accompanied by an unprecedented increase in social online interactions. However, this recent evolution also caused a sharp increase in individual exposure to extreme political views, fake news and conspiracy theories. Today's society is more polarized than ever, and social and religious tensions have reached a worrying level, in a context of generalized distrust toward institutions and traditional media.[1] Conspiracy theories and political or religious radicalism are particularly popular among the youth ([2]). Yet, educational levels are historically high in most developed economies and most education systems include citizenship education as part of their curriculum. This paradox points the failure of the educational system to instill social cohesion in the society and to educate the youth to the use of social media.

[1] See for example [1]. A complete literature review has been performed by the evaluation team and can be found in the appendix of this report.)

The main objective of the program evaluated in this report, the *Digital Resilience Community* (hereafter DRC), is to increase youth digital skills – in particular their digital citizenship (see the Digital Competence Framework 2.0 developed by the European Commission ([3]) – and to prevent youth radicalism. In particular, it aims to equip the youth with the information processing skills needed to resist to fake news and to raise awareness of the harmful effects of online violence. Given the role of distrust – in others, in public institutions – in the radicalization process,[2] the program also aims to increase participants trust in science and in professional journalists and traditional media as well as tolerance and mutual respect for other groups of individuals.

2.2 The DRC program in practice

DRC is a training program targeted at youth living in deprived areas and designed to increase their resilience to fake news and conspiracy theories and to reduce online violent behaviours. In practice, the young volunteers recruited into the program are first asked to register on the Facebook page of the program to formally enroll. They are then invited to participate in online workshops organized and delivered by ALMA staff members. These workshops were designed to educate participants about the dangers of online hate and conspiracy theories, to inform them about the journalistic profession and the role of traditional and social media in the production of information, as well as to train them to create online media content. In a second phase, participants are invited to take part in 5 contests that consisted of engaging in several actions aimed at fighting fake news, conspiracy theories and online violence (creating online content, sharing this content, photo reporting, etc.). Each time, the winner(s) of the contest receives a prize, worth several hundred euros.

3 Evaluation protocol : recruitment, surveys and randomization

3.1 General framework

In order to conduct the evaluation of the training program conducted by ALMA, we opted for the randomized controlled evaluation methodology. This methodology consists of dividing the sample of volunteers into two groups of equal size to create a treatment (or "test") group, which includes youth who will participate in the training program, and a control (or "comparison") group, which includes those who will not participate. In order to create two groups that are exactly comparable on average before the program begins, these two groups were created randomly. Thus, at the end of the program, a simple

[2] See for example [4]. See also the literature review in the appendix of this report.

comparison of the dimensions measured in the two groups makes it possible to determine the impact of the program.

All volunteers are young people aged 15 to 25, recruited by local associations (local missions, youth associations, etc.) and by the ALMA association. These associations identified young people likely to be interested in the program (about 10 to 15 young people per association) and offered them to register online. In total, 222 young participants were invited to register to the online lottery, 111 were assigned to the treatment group and 111 to the control group.

During online registration, volunteers were informed of the purpose of the study, the randomization, and their rights to access and correct their personal information. They also complete an initial survey (baseline) in September 2020. In order to observe the evolution of the behaviors and representations of the young people participating in the draw, a final survey (endline) is also proposed to all the young people in the sample in April-May 2021. The aim of these surveys is to characterize the population of individuals participating to the program and to measure its impact on targeted dimensions. The two following subsections describe the outcomes measured in the surveys as well as the main characteristics of our study sample.

3.2 Outcomes measurement

In order to measure the effects of the DRC program, a battery of cognitive, non-cognitive and behavioral tests was developed and included in the initial and final surveys. We describe here the main dimensions measured:

- **Traditional media use**: measures the level of use of traditional media (newspapers and radio) versus TV and social media;
- **Trust in traditional media**: measures the level of trust in traditional media (newspapers and radio) compared to TV and social media;
- **Fact checking**: measures the frequency of use of fact checking tools;
- **Adherence to conspiracy theories**: measures the degree of adherence to various conspiracy theories, related to recent news (covid, terrorism, etc.) or to historical events (e.g. Kennedy assassination)
- **Fake news belief**: measures the ability to distinguish fake news from real news, based on the presentation of true and false information that have been disseminated on the internet

- **Condemnation of online violence**: measures the degree of tolerance for various online hate behaviors;
- **Tolerance**: measures the degree of tolerance of gender, sexuality, religion and skin color;
- **Engagement against misinformation**: measures the willingness to engage in various actions to fight misinformation

3.3 Sample characteristics

The final sample is composed of 222 individuals, recruited mainly through different local associations and through social networks such as Instagram. This section is based on the baseline questionnaire. 57% of the participants are girls. The average age is around 19 years old (youngest is 13 yo, oldest is 31 yo), and most of the participants (75%) attend school. They mostly live at home with their two parents and more than half have at least three brothers or sisters. They come from low-educated families: the average number of years of education of both their mothers and fathers is around 9 years (in France, the high-school diploma –Baccalauréat– represents 12 years of education). Most of them have some foreign origins, in particular because one of their parents was born outside of France.

Virtually all participants have a phone of their own, and they got it on average at 13 years old. In terms of their use of social networks, only 4% declare not using any of the major social network. Participants mostly use Snapchat (83%) and Instagram (77%). However, the use of Facebook is quite low: less than 40% of the participants have a Facebook account. As ALMA uses Facebook as the platform of the DRC program, it is possible that the low take-up –that we will analyze later on in this report– might be due to the originally low prevalence of the use of Facebook in this sample. We can actually see that the use of Facebook is highly correlated with the age. Young people use less Facebook than older people. For example, in this sample, only 23% of the young people of less than 20 years old are using Facebook, against 65% of the people of more than 20 years old. These figures give great insight on the capacity to involve young people on programs that use Facebook as a vector of an intervention, at least for this type of population.

Individuals from our sample mainly get access to news and information through social networks (around 80%). Traditional medias such as the radio or the newspapers are not really used: only 5% of the participants listened to some news at the radio and only 13% read some news from the newspapers in the five days preceding the survey. TV is still quite used to get some information (60%). Around half of

the individuals also get some news through their parents (46%) or their friends (55%). We can contrast the use of media to get access to information and the trust they give to each of these media. Table 1 shows paradoxically that the most used media to access information (i.e social networks) is also the less trusted by the participants. Overall, they do not trust media. In comparison, they significantly trust their parents and friends, more than official medias.

Table 1: Media use and confidence

Media	Use (%)	Not much trust (%)
TV	60	76
Radio	5	79
Newspapers	13	74
Social network	78	89
Parents	46	31
Friends	55	59

The individuals from the sample do not seem to trust much the medias. This is even more stringent when it comes to different institutions. Table 2 shows the global lack of trust that these young individuals have in the French institutions. Overall, they have significantly more confidence in School, scientists and the Army, compared to the government, the Justice, the journalists, the police, or even more the political parties. In contrast, Table 2 also shows that they significantly trust their families and friends.

Table 2: Trust in Institutions and relatives

Institution	Not much trust (%)	Relatives	Trust (%)
Government	86	Family	89
Justice	86	Friends	73
Scientists	57	Neighbors	16
Journalists	89	First-met people	4
Police	88		
Political parties	96		
Army	63		
School	55		

We asked several questions about fake news and conspiracies to the participants. In particular, Table 3 shows how much the individuals from our sample believe or not in some fake news or conspiracies related to the Covid but also some "famous" conspiracies that we can easily find online. We can see that there is a significant share of individuals that believe in some fake news or conspiracies. For example, 17% of the respondents declared that the Covid-19 is a biological weapon that was created to harm the

Chinese economy. 28% of the respondents declared that a powerful group of individuals are organized and influence the different governments in order to rule the world. Interestingly, a large share of respondents just answered that they had no opinion regarding the subject and do not seem to take a side between agreeing or disagreeing.

Furthermore, we asked a few questions about recent events that happened in France, and especially about the Charlie Hebdo attacks that happened in 2015. Only 57% of the respondents totally condemn these attacks. 24% are indifferent, 14% condemn the attack but they share/understand some of the reasons, and finally 5% do not condemn these attacks. Moreover, regarding these attacks, 22% declared that there are some grey areas in this event and that it is not sure that these attacks were planned and led only by Islamist terrorists. 12% actually think that these attacks have been manipulated, especially by the French Secret Service.

Taking together the prevalence of believing in fake news and conspiracies (from the individuals declaring agreeing but also from the individuals that do not take a side, or are indifferent), it seems that programs such as DRC, lead by ALMA, can have great effects on these young individuals.

Table 3: Fake news and conspiracy: prevalence

Fake news & conspiracies	Agree (%)	Disagree (%)	No opinion (%)
Covid is a biological weapon created to harm the chinese economy	17	51	32
5G waves are responsible of the Covid	4	67	29
The Covid vaccin could contain a microchip	10	51	39
Vitamin C could be a Covid cure	9	45	46
Drinking alcohol can kill the Covid	6	67	27
The Health Ministry and the pharma industry hide the truth about the real damages of the Covid	28	25	47
Chemtrails contain chemical products that are spread on purpose for secret reasons	10	43	47
A powerful group of people is secretly organized to influence the governements and rule the world	28	33	39

To complement the profile of the respondents, we determined their self-esteem. Self-esteem is important since the literature showed some significant correlations with academic performance, loneliness, social connectedness, violent behavior or willingness to trust. Considering this, we used the Rosenberg scale: with 10 questions, we can formally assess the level of self-esteem of the respondents. Figure 1 shows the self-esteem of the individuals according to the scoring of the Rosenberg scale. We can see that around half of the respondents have low or really low self-esteem. In our data, we can see for example that there is a

negative and significant correlation existing between self-esteem and believing in fake news: individuals with lower self-esteem tend to believe more in fake news.

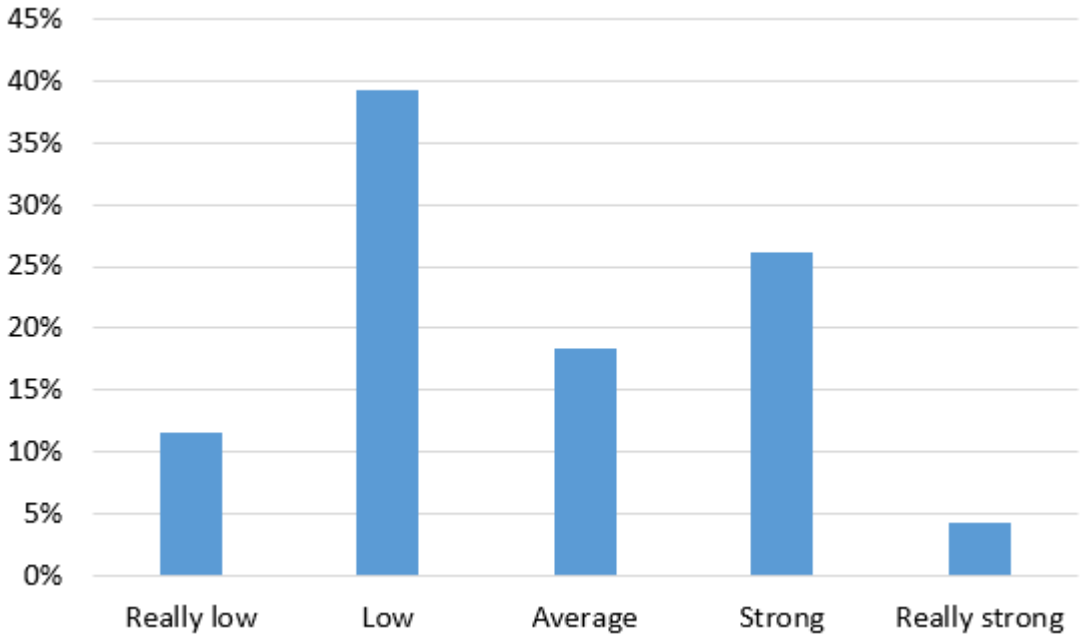

Figure 1: Self-esteem (Rosenberg)

3.4 Program take-up

The evaluation protocol, as mentioned earlier, is based on a sample of 222 individuals, half of which was assigned to participate to the program, and the other half being used as a control group. Both groups are required to fill a baseline and an endline surveys. This section focuses on participants take-up, i.e., the extent to which individuals assigned to participate to the program actually did so.

We build two measures of program take-up, depending on the level of involvement. The first one, defined as *moderate participation*, refers to all individuals who registered on the Facebook webpage of the program and potentially accessed the online material developed by the association. The second one, defined as *active participation*, refers to individuals who attended all of the training sessions and remained in the program until its end. In total, out of the 111 individuals initially assigned to participate to the program, 76 moderately participated (68%) and 24 actively participated to it (22%).

4 The impact of the DRC program

4.1 Preliminary analysis : atttrition and balancing checks

We first describe the attrition between the baseline and endline surveys. These results can be found in Table 4. The attrition rate is quite high, as only 35% of individuals surveyed in endline also answered in endline (column total). Participation rates to the endline survey vary from 39% for individuals in the treatment group to 31% for individuals in the control group. However, the difference between the two groups is not significant at conventional level, as shown in the first column of Table 5.

In order to check that there is no selective attrition between baseline and endline questionnaires that would affect the comparability of the test and the control groups, we implement balancing tests on the subsample of individuals who answered the endline questionnaire, on individual pre-determined characteristics as well as baseline beliefs and attitudes.

Table 6 displays the results of these tests performed on personal characteristics, household composition and social integration. None of the 15 variables available appear to be significantly different between the two groups. This finding confirms that attrition was homogeneous across the two groups and that the two groups remain comparable at endline in terms of personal characteristics.

Tables 7 and 8 reproduce the same analysis on variables related to individuals' activities, online behaviour, beliefs and attitudes at baseline. Overall, little differences appear between the two groups, as only three variables out of the 18 tested are significantly unbalanced. In total, these balancing checks confirm the comparability of the two groups at endline.

Table 4: **Survey answers, by treatment group**

Answers in	Control	Treatment	Total
Baseline	77	67	144
Baseline & Endline	34	44	78
Total	111	111	222

Table 5: **Differential attrition, by treatment group and participation intensity**

	(1) Answer at Endline	(2) Answer at Endline
Assigned to treatment	0.0930	0
	(1.50)	(.)
Active participation		0.359*
		(2.02)
Moderate participation		0.152
		(0.92)
Constante	0.704**	0.169
	(3.02)	(0.45)
Observations	222	111

t statistics in parentheses

* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$

Table 6: Balancing checks on individual characteristics

Variables	Treatment	Control	T-C
Age	19.345	19.439	-0.021
			[0.973]
Gender	0.551	0.588	-0.000
			[0.999]
Household composition			
Single Parent household	0.141	0.118	0.105
			[0.414]
Parents without tertiary education	0.091	0.088	-0.066
			[0.489]
At least one superior SPC parent	0.234	0.333	-0.211
			[0.105]
Number of siblings	2.231	2.206	0.105
			[0.679]
Elder in siblings	0.436	0.500	-0.079
			[0.646]
Rank in siblings	2.128	1.853	0.605
			[0.273]
Respondent Origin			
Respondant or one parent born outside France	0.827	0.765	0.178
			[0.149]
Respondant born in France	0.813	0.882	-0.055
			[0.674]
Social Integration			
Social Integration index	0.046	-0.019	0.308
			[0.369]
Feeling of discrimination	0.397	0.382	0.079
			[0.631]
Self esteem (Rosenberg)	0.006	0.122	-0.118
			[0.629]
Raven score	0.126	0.197	-0.311
			[0.299]
Exposure to racism	0.020	-0.041	0.228
			[0.376]
Observations	44	34	

p-value within brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 7: Balancing tests on social activities, integration and engagement

Variables	Treatment	Control	T-C
Leisure and social media			
In-person leisure share	0.468	0.477	-0.005
			[0.934]
Virtual leisure share	0.201	0.186	0.046
			[0.200]
Alone leisure share	0.332	0.336	-0.040
			[0.547]
Number of social media used	3.436	3.235	1.237**
			[0.029]
Social integration			
Class Representative	0.744	0.882	-0.066
			[0.632]
Social integration	0.046	-0.019	0.308
			[0.369]
Participation in social movements	0.011	-0.152	0.615*
			[0.075]
Relation to social movements			
Knowledge of social movements soc	0.270	0.254	0.045
			[0.799]
Political self-efficacy	0.045	0.267	0.017
			[0.951]
Observations	44	34	

p-value within brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

Table 8: Balancing on baseline outcomes

Variables	Treatment	Control	T-C
Online behavior			
Trolling behavior	0.117	0.062	0.507
			[0.251]
Beliefs in Fake news and Conspiracy and News knowledge			
Belief in covid-related fake news	-0.159	-0.534	0.678***
			[0.006]
Belief in diverse conspiracies	0.050	-0.215	0.365
			[0.220]
News knowledge	0.321	0.473	-0.051
			[0.652]
Belief in conspiracy	-0.047	-0.312	0.453
			[0.183]
Trust			
Trust in institutions	0.154	0.295	-0.237
			[0.430]
Trust in traditional media	0.204	0.287	0.010
			[0.975]
Trust in social media	-0.009	-0.011	-0.007
			[0.978]
Trust in television	0.097	0.191	-0.209
			[0.559]
Observations	44	34	

p-value within brackets

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.2 Main model: Intention-To-Treat

In this section, we present the effects of the DRC program on the main outcomes measured in the surveys. Table 9 displays the impact of the program on these dimensions, as measured by the comparison of the whole treatment and control groups, expressed in terms of standard deviation of the studied variable. As we can see in the table, individuals from the treatment group display better scores on all these dimensions, with the exception of belief in fake news and trust towards traditional media.

Generally speaking, the results point to a positive impact of the program on the use of traditional media (+0.199 SD), the use of fact checking tools (+1.202 SD), conspiracy theory beliefs (-0.244 SD), online violence disapproval (+0.317 SD), tolerance (+0.302 SD) and willingness to engage against online misinformation (+ 0.339 SD). It is nevertheless worth noting that, given the weak statistical power caused by the high attrition and the low active participation to the program, the only significant impact found is the one on the use of fact-checking tools use. Moreover, the program doesn't seem to succeed in increasing participant's trust in traditional media, or to better identify fake news, with differences of respectively -0.153 SD and +0.157 SD.

Table 9: Impact of the DRC program

Variables	T-C	S.E.	P-val
Use of tradi media	0.199	(0.215)	0.354
Confidence in tradi media	-0.153	(0.163)	0.346
Fact checking	1.202**	(0.534)	0.024
Belief in Fake news	0.157	(0.257)	0.540
Belief in conspiracy theories	-0.244	(0.328)	0.457
Rejection of violence	0.317	(0.344)	0.357
Tolerance	0.302	(0.300)	0.314
Engagement against fake news	0.339	(0.351)	0.334

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

4.3 Focus on active participants: evolution of beliefs and behaviors

Given the low take-up of the program, it is also interesting to focus on active participants, in order to estimate how actual participation to the program affects participants attitudes and behaviours. To do so, we contrast the evolution of active participants between baseline and endline with those of all other individuals in the sample. The comparison group in this approach is therefore made up of all individuals not actively participating to the program. The underlying assumption here is that, while

active participants may be initially different from other individuals in the sample, their respective evolution would have been the same in the absence of the program.

Table 10: Impact of the DRC program on active participants

Variables	T-	S.E.	P-val
Behavior, beliefs & knowledge			
Use of tradi media	0.557	0.519	0.291
Belief in conspiracy	-0.597	0.550	0.290
Number of media used	0.552	0.404	0.182
Share of trad. media	0.530	0.538	0.333
Self confidence on pol. issues	-0.078	0.252	0.758
News knowledge score	0.498	0.384	0.203
Violence			
Condemns violence	0.484	0.692	0.491
Condemns online violence	0.737	0.656	0.273
Could participate in viol act.	-0.022	0.320	0.945
Troll behavior	-0.117	0.571	0.839
Tolerance			
Global index	0.165	0.326	0.615
Gender issues	0.211	0.150	0.170
Neighbors	-0.559	0.592	0.352
Religion	-0.036	0.360	0.921
Friends	0.759**	0.361	0.043
Trust			
Traditional media	0.558	0.421	0.195
Social Media	-0.285	0.390	0.470
Institutions	0.067	0.326	0.838
Scientists	0.767*	0.432	0.090
Journalists	-0.192	0.458	0.678

* $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

As in the previous estimations, the small size of the sample prevents us from having significant results at conventional levels. Nevertheless, estimated program effects are positive for all of the outcomes tested.

First, we find that the program has a positive impact on the use of traditional media and the number of media used (Table 10). This is encouraging, as diversity of sources, like the use of traditional media - which has a stronger editorial control - should improve news knowledge and reduce belief of fake news. This is indeed what we observe, as these effects translate in an increased news knowledge score and in a decrease in conspiracy belief. On top of this, the program positively affect participants' trust in traditional media and in science and scientists, two key objectives of the program, while trust in social media decreases. Finally, tolerance outcomes provide more mixed results, suggesting that this dimension is not strongly affected by the program.

References

[1] Olivier Galland and Anne Muxel, eds. *La tentation radicale: enquête auprès des lycéens*. Paris: Puf, 2018. 454 pp. ISBN: 978-2-13-079957-3.

[2] Ifop pour la Fondation Jean-Jaurès et Conspiracy Watch. *Enquête sur le complotisme – Vague 2*.

[3] Stephanie Carretero, Riina Vuorikari, and Yves Punie. "The digital competence framework for citizens". In: *Publications Office of the European Union* (2017).

[4] Séraphin Alava, Noha Najjar, and Hasna Hussein. "Étude des processus de radicalisation au sein des réseaux sociaux : place des arguments complotistes et des discours de rupture". In: *Quaderni* 94 (Oct. 2017). Publisher: OpenEdition, pp. 29–40. ISSN: 2105-2956. DOI: 10.4000/quaderni.1106. URL: http://dx.doi.org/10.4000/quaderni.1106.