Digital Resilience
Stronger Citizens Online

Louis Reynolds
Lucie Parker
## Contents

Acknowledgements

1. Executive Summary

2. Building Digital Resilience

3. The *Digital Resilience* Project

4. The *Digital Resilience* Evaluation

5. Conclusions

6. Technical Appendix

References
Acknowledgements

We would like first and foremost to thank the Dutch vocational college teachers and students who took part in our study, and in doing so made this report possible.

We are indebted to the Dutch Ministry of Justice and Security and the Dutch Ministry of Foreign Affairs, whose generous support allowed us to undertake this work, and to the civil servants who facilitated the delivery of this project.

This project was a joint effort between ISD and our project partners Codename Future, whose work on the ground in the Netherlands working with teachers and students formed its core. From Codename Future we would particularly like to thank Aart Hageman and Laura Wiegant, whose efforts were invaluable.

From the Institute of Strategic Dialogue (ISD), we would like to thank Jonathan Birdwell, Rose Beynon, Ruth Boey, Joseph Downy, Ben Holman, Duaa Khalid, Sarah Kennedy, Charlotte Moeyens, Daniela Puska and Henry Tuck. Special mention goes to Cooper Gatewood, Joe McElroy and Josh Philips who played an invaluable role in the analysis of the data.

Any mistakes or omissions are the authors’ own.

Louis Reynolds
Lucie Parker

May 2018
1. Executive Summary

The more important the online aspects of our lives have become, the more vulnerable democratic, pluralistic, societies have found themselves.

In the last three years, extremist groups and state actors have systematically spread disinformation online in order to warp our perceptions, damage social cohesion and influence elections across Europe and North America.¹

Extremists thrive where critical thinking, media literacy and awareness of the dynamics of the online space are lacking. Just this month, an Al Qaeda dossier was released which demonstrated that their recruiters strategically targeted young people who were most naïve and the ignorant and disaffected.² Many former recruiters for extremist organisations will tell you the same thing; that a young person with a grievance who has not been taught critical thinking skills or media literacy is a perfect target.

This phenomenon is deeply worrying for our society. It serves as a reminder that the nature of the extremist threat online is constantly evolving. When our perception of the world is so heavily shaped by social media, social media becomes a more attractive target for extremists.

The bedrock of a democratic state has always been good citizenship, and this requires a considered and nuanced understanding of and response to the world around us. With the rise of the Internet and social media, this aspect of good citizenship is more important than ever.

Yet across the globe, education systems have been slow to react to this challenge. We are failing to prepare our young people for the online world, with potentially disastrous consequences.

For the past three years, ISD has been developing and testing curriculum and pedagogical approaches to teach critical thinking skills and online awareness techniques in an engaging and effective manner. The aim is to develop scalable models to undermine the efforts of extremist groups and hate movements to manipulate our young people, with positive evidence of impact.

In this report, we present the findings from our Digital Resilience pilot, which has been developed and delivered in vocational colleges in three cities in the Netherlands. This pilot forms part of a growing body of evidence collected by ISD, from pilot projects and ongoing programmes across Europe, which suggests that digital citizenship educational approaches can be an effective tool for increasing the resilience of young people to extremist grooming and exploitation online.
The Digital Resilience Pilot

A partnership between ISD and the Dutch citizenship education providers Codename Future, the Digital Resilience project provided a sample of teachers with the training, support and resources to deliver the Digital Resilience curriculum.

The Digital Resilience curriculum aims to provide 16-19 year olds with the knowledge, skills, attitudes and behaviours they need to be positive digital citizens in the 21st century. The curriculum focuses on the online challenges most relevant to grooming and radicalisation online, from effectively dealing with hate speech to spotting manipulation in images and bias in news, from understanding echo chambers to identifying active disinformation.

To carry out this project, ISD and Codename Future re-examined the core nature of digital citizenship, undertook a best practice review of existing resources, developed a theory of change and detailed curriculum content, and produced a teacher guidance manual and a high quality student workbook.

ISD and Codename Future delivered Digital Resilience in six Dutch vocational college classes in Amsterdam, Rotterdam and Den Haag. Six teachers were fully trained in the Digital Resilience curriculum and delivered the curriculum to 135 students. The sessions were subject to an evaluation process, including pre- and post-surveys administered to participating classes and compared against a control group of classes within the same schools. The evaluation also included three focus groups with participating young people, and five interviews with the teachers delivering the sessions.

In addition to the results of the evaluation, this report presents a best practice review of existing resources designed to avoid replication, a description of the curriculum and the resources, and our reflections on the key outcomes of the pilot, as well as a technical appendix.

Key Findings

Our evaluation suggests that the Digital Resilience curriculum had a positive impact, increasing students’ sense of responsibility for their actions online, as well as their self-reported knowledge on a range of topics that are critical to safe and resilient use of the Internet.

Demographics

The classes included in this study were selected through Codename Future’s extensive networks within Dutch vocational colleges. The students who participated in this project came from a range of diverse backgrounds. Just under half of participants (47 per cent) were from a migrant background, with 13 per cent of participants being first-generation migrants, while 34 per cent were second-generation migrants. Participants had been born in
13 different countries, with their parents originating from 26 different countries, and their families speaking 28 different languages or language combinations in the home.

**Impact**

The impact measurement effort for this pilot focused on agreement with confidence-related statements in pre- and post-surveys, delivered before and after the sessions, to both participant and comparison groups. While a relatively small sample size limited our ability to measure statistical significance, eight of the 16 measures returned notable positive impacts in the participant groups, five of which were statistically significant.

Fostering a sense of responsibility among young people is the most critical component of digital citizenship and staying resilient online. The most important finding from the evaluation was a 33 per cent increase in students’ feelings of responsibility for the wellbeing of those they meet on social media, compared to a comparison group.

Moreover, four out of five participants (81 per cent) reported that they gained new skills, and just under half (47 per cent) of participants claimed that they would behave differently online because of the sessions.

Additionally, there were large increases in knowledge and confidence across a range of concepts and measures critical to digital citizenship. These included:

- 85 per cent increase in student confidence that they understand what echo chambers are.
- 56 per cent increase in student confidence that they understand what the ‘filter bubble’ is.
- 47 per cent increase in student confidence that they understand what scapegoating is.
- 10 per cent increase in students’ reporting that they would fact check a story before sharing it if they were unsure it was true.
- 10 per cent increase in students’ confidence that they would recognise when a social media post, article or website is designed to emotionally manipulate people.
- 13 per cent increase in participant confidence that they consider the motivations behind why people post things online.

**Conclusion**

This evaluation produced important positive impact results, demonstrating the efficacy of a digital citizenship education approach that focuses on creating attitudinal change in an effort to influence online behaviour. It also demonstrated that this impact can be achieved with minimal intervention from external organisations, limited to basic teacher training and effective teacher guidance and resources. However, some of the impact measures produced no significant and in some cases no measurable positive change, and while much of this
result can be explained through the small sample size of this pilot study, it also provides clear guidance on areas for future improvement.
2. Building Digital Resilience

While digital citizenship remains an under-developed area of education and resilience-building, the term – attached to a wide array of meanings – has been discussed increasingly since around 2008, as policymakers have sought to respond to the popularisation of social media and its consequences. Digital citizenship education covers a range of knowledge, skills, attitudes and behaviours that vary between curriculums, age ranges, intended outcomes and numerous other factors.

This chapter first outlines the relationship between basic and more advanced digital citizenship skills, and the challenges facing the age group that this project focuses on, and then provides the key insights gathered from our best practice review of 16 digital citizenship resources which were available at the start of this project.

The Pyramid of Digital Citizenship education

Establishing basic capacities, such as the knowledge of how to create a strong password or how to protect your personal information online, are valuable parts of digital citizenship. However, these capacities might be most appropriately taught to younger audiences as the basic building blocks of effective digital citizenship, rather than being, as they often are, the full extent of digital citizenship education.

More sophisticated learning can build on this basic level of digital citizenship. These capacities, rather than just providing basic protections online, allow young people to safely and effectively use the Internet, and to use it for positive ends as citizens, from an empowered position of enhanced knowledge, skills and responsibility. They include, for example, the ability to critically assess media content online, safeguard peers on social networks, knowledge of how social media influences how we communicate with each other, and an attitude of collective responsibility towards the social media spaces we use, including peer-safeguarding. These capacities are more appropriately taught to pre-teen and teenage students.

The most advanced digital citizenship capacities again build on the previous two levels. If the more sophisticated learning described above allows young people to be effective and resilient digital citizens, this advanced learning allows young people to adopt positions of youth leadership, fuelling active citizenship and activism. This includes the ability to use social media as tools for social campaigning and civil society organisation, to effectively participate in digital democratic processes, or to educate their peers. These capacities are more appropriately taught to teenage and young adult students.
One way to represent the full spectrum of digital citizenship capacities that can be taught through the education system, as described above, is to present digital citizenship education as a pyramid, with more sophisticated capacities built on a foundation of basic aptitudes.

**Fig. 1, The pyramid of digital citizenship education**

Traditionally, the focus of Digital Citizenship education has been safeguarding and basic technical skills, ranging from understanding phishing and online scams to the sharing of personal information and understanding one’s ‘digital footprint.’ In the last few years, the trend in Digital Citizenship educational interventions has been towards the inclusion of some focus on behavioural and attitudinal change and capacities related to responsible behaviours in the online space. Far fewer interventions have had a strong element of ‘citizenship’, relating the rights and responsibilities of citizens online, the responsibility of individuals online to safeguard each other and collectively make the Internet a better place, and other aspects of desirable online behaviour and activity philosophically rooted in the concept of citizenship.
The *Digital Resilience* curriculum is focused on the core knowledge, skills, attitudes and behaviours that allow people to use social media effectively, safely and constructively as effective digital citizens – the middle tier of the pyramid.

**Considerations for this age group**

This project seeks to develop digital resilience among Dutch young people in a vocational college setting. Participants of the age range selected for this intervention, 16-19 year olds, are more likely to be increasingly engaged in the development of their social, political and cultural views, and more actively engaged in those online conversations, through the use of social media as an information source, space for public discourse and avenue for content creation and circulation.

While digital resilience is vital to safeguarding across a large range of topics, and is a fundamental capacity of effective citizenship in the 21st Century, it has particular importance to the problem of violent extremism as well as the broader problem of extremism.

Teenagers coming to the end of their secondary school education can be regarded as entering a particular ‘risk period’ in terms of radicalisation and exposure to extremism. The average age of foreign fighters joining Daesh in 2015 was 25, according to research by the Soufan Group and New America, while the average age of an individual involved in terrorist activities in Europe was 27 in 2016. Within the Netherlands, the average age of those travelling to join Daesh in Syria and Iraq was 23 in 2016. A range of factors, from a desire for adventure to political and social awakening to a search for meaningful identity can help push young people towards extremist ideologies and groups, and ultimately towards violent extremism. By enhancing the digital resilience of this age group, this project seeks to achieve maximum effect in reducing the likelihood of young people uncritically consuming extremist content, joining extremist groups or committing violent extremist acts.

**Best Practice Review**

In addition to establishing clarity of purpose, ensuring that this project does not replicate existing work is vital to its value. To both avoid duplication and learn from the strengths and limitations of existing materials, ISD reviewed 16 existing digital citizenship resources which were public prior to the design of the *Digital Resilience* resources.

These resources were selected on the basis of their availability in English and selected as the highest quality resources from a wider sample. This quality judgement included investigation on the depth and volume of content, relevance to the development of resilience to extremism online, established expertise of the organisation that created them, and, in some cases, the availability of evaluation data. This review highlighted significant
differences between the available resources, with some being of a very high quality and others being more limited.

Some of these resources were basic or focused on fundamental skills, while others sought to inculcate more general attitudinal or behavioural change. Most of these resources do not explicitly mention extremism online or even aim to address it as an issue, while a small number cover it as one of a range of safeguarding aspects. Four of the resources were directly focused on the topic of extremism, even though in some of these extremism is not explicitly mentioned.

These 16 resources represent a diverse range in terms of their quality and the research supporting them, topical focus, delivery format and age range. Many provide effective means by which to develop CVE-relevant digital skills among young people, a reflection which highlights the importance of conducting a needs assessment in the digital citizenship landscape.

A number of resources provide advice, guidance or training for teachers, but do not include lesson plans or a structured intervention for delivery in the classroom. As an alternative, some provide links to third party materials for this purpose, serving principally as a hub for the presentation of resources and guidance on pedagogy or the topic of online safety. While a useful source of general information, this limits their utility for educational delivery, as they do not provide teachers or schools with the complete toolkits required to effectively deliver digital citizenship education in the classroom, with or without a focus on the issue of extremism.

Several of these resources were somewhat dated. In some cases, this merely meant that the websites or the educational materials looked older and less attractive. In a number of cases, however, they either addressed the issues of digital citizenship through less relevant social media formats – reducing the credibility of and engagement with their content in the eyes of students, even if the principles explored in the resources remain relevant – or did not mention social media at all, instead purely discussing static web pages. As a result, the value of these resources has become limited. This limitation is stark where digital citizenship relates to dealing with the threat of online extremism, given the centrality of social media to that threat.

A number of the resources reviewed were up-to-date and provided considered materials for delivery in the classroom, but otherwise lacked a vital aspect required of educational materials in this space. In some cases, these resources lacked sufficient material – perhaps providing a few activities that were linked thematically, but no overarching lesson plan or learning objectives. Other resources provided basic content with structured lesson plans...
and outcomes, but lacked quality classroom materials and guidance for teachers on how to approach these topics.

**Digital Citizenship & Extremism**

The variety of implicit definitions of digital citizenship within these resources significantly influences their activities and learning outcomes. Some focus on ethics or basic skills, while a smaller proportion considers attitudes and behaviours.

In many resources there is a general conflation between ‘information literacy’ – which is generally concerned with the effective and ethical understanding and use of information – and ‘media literacy’, which relates to the critical evaluation of media sources, effective media creation and consideration of issues such as the motivations of content creators and common manipulative techniques. Some of those resources that do fully consider the behavioural and attitudinal aspects of digital citizenship, as well as developing critical thinking and media literacy skills, do not consider extremism. In this context, it should be noted that engaging in CVE education through an implicit approach rather than an explicit approach is a valid method (particularly desirable in certain educational contexts) as long as the learning processes and objectives are developed with extremism in mind.

Even among the highest quality resources, which provide significant depth and volume of multimedia material that encompasses both basic skills and attitudinal and behavioural changes, few consider extremism. Only four of the digital citizenship resources reviewed focus in part on extremism, using either an implicit or an explicit approach to the subject matter. The strengths and limitations of these resources and their evaluations provided the most useful guidance for the development of the *Digital Resilience* resources for this project, the key elements of which are presented below.

**Key Resource Recommendations**

Digital citizenship education projects which seek to build the resilience of young people to hate and extremism online should include:

- **A clear Theory of Change focused on realistic and achievable objectives**, defined in terms of knowledge, skills, behavioural and attitudinal change, which can be related to an explicit overarching purpose and defined end goals.
- **Flexible, modular and detailed resources for use by and with students in a classroom setting**, which allow teachers to tailor delivery to their class needs, assist in the engagement of the participants and help convey key information and learning points.
- **Detailed lesson plans for teachers**, which ensure that delivery is structured through an appropriate pedagogical method, and that the sessions are delivered in a timely manner in line with key learning objectives.

- **Supplementary materials**, to provide teachers with the specific subject knowledge required to deliver digital citizenship education, including key definitions, to inform them of its purpose and relevance, and to ensure confident delivery.

- **Background material on the profile of the extremist threat in their country and the characteristics of the threat of extremism online**, as well as some basic information of the characteristics of the relevant social media platforms, with which they might not be familiar, and key points of contact for further support or information related to the issue of extremism.

- **Guidance on how to confidently conduct discussions on the sensitive topic of extremism and related issues**, such as racism, hate speech, hate crimes, identity, belonging, grooming and exploitation, and creating a safe space in which to facilitate constructive dialogue on these issues.

- **A focus on attitudinal and behavioural change, not just technical skills.** Ultimately digital citizenship does not consist purely of training in online skills, but in the development of behaviours consistent with the principles of digital citizenship – responsibility over one’s actions online, a desire and ability to assist peers in the online space, a feeling of ownership over one’s social networks. Digital citizenship resources should therefore not merely seek to provide young people with technical knowledge, but seek to develop attitudinal and behavioural change that will empower and motivate them to be better citizens online.
3. The Digital Resilience Project

This chapter outlines the *Digital Resilience* project approach, resources and delivery model.

**Our Theory of Change**

*Digital Resilience* is an educational intervention designed to increase the resilience of young people in the Netherlands to hate and extremism online. The intervention is skills-based and intended to develop participants’ media literacy, critical thinking and digital citizenship skills specific to the challenges of online hate and intolerance, as well as an increased understanding of fake news and propaganda, the arguments and techniques content creators use to manipulate online consumers, and how they react or respond to hate speech online. The project seeks to develop positive attitudinal and behavioural change, using a CVE-implicit and digital citizenship-explicit upstream approach.

**Context**

The project was created in response to the disproportionate targeting of young people by extremist propaganda and grooming efforts online. At the same time, young people online face a number of related challenges that also contribute to the growth of extremism online, including disinformation and hate speech. In order to stay safe and make a positive contribution online, young people need the skills, attitudes, knowledge and behaviours that will make them more resilient to these challenges, including critical thinking, media literacy and digital citizenship skills. However, many young people do not receive sufficient education in these areas within formal or informal education, increasing their vulnerability.

**Participants**

This pilot project was delivered in five Dutch vocational college classes in Amsterdam, Rotterdam and Den Haag. As part of the project, six teachers were trained in the delivery of the resources by Codename Future, and the resources were delivered to 135 participating students aged 16-19 years old, over three one-hour sessions by these teachers. Within the same colleges, 108 students who did not take part in the intervention comprised a comparison group.

**Outcomes**

The *Digital Resilience* project sought to deliver the following outcomes for participants:
- Increased critical thinking skills in an online context, with a particular focus on the Dutch context
- Increased media literacy skills, including the ability to identify fake news more effectively
- Increased digital citizenship skills, including an increased ability to recognise or challenge online emotional manipulation and hate speech
- Improved attitudes towards the online world, including an increased desire to act to tackle hate and extremism online and an increased feeling of responsibility for the wellbeing of their peers online
- Ultimately improved behaviours in online interactions, including flagging hate content for removal and fact-checking news articles online more regularly.

Impact
As a result of participation in this project, Dutch students will be:
- More resilient to extremist grooming and propaganda online
- More able to react effectively to hateful content online
- More active digital citizens.

Fewer Dutch young people will be drawn into extremist groups, movements and ideologies, and they will be less sympathetic to extreme viewpoints. Fewer young people will suffer the negative effects of online hate. The social networks and online spaces in which young people operate and interact will be more positive and healthy environments. This approach ultimately aims to empower more capable and resilient citizens, and in doing so disrupts the ability of extremist groups to influence, exploit and recruit our young people.

Session Plans & Digital Deck
The Digital Resilience materials include three comprehensive and structured session plans that are accompanied by an online digital deck. They are aimed primarily at 16-19 year olds in formal educational settings, with a particular focus on vocational colleges.

The session plans are intended to be flexible in their use, allowing the facilitator to adapt the exercises and session times in accordance with their needs and the needs of their group. They contain a number of suggested ‘scripts’, paragraphs that summarise how to introduce or talk through the exercises, as well as ‘prompts’, discussion questions that can be used to guide the conversation.

The session plans cover a range of social challenges relating to the online world, including fake news and propaganda, echo chambers, emotional manipulation and hate speech. They provide background information and key concepts to help facilitators gain insight into the issues discussed in the session. In addition to outlining how to prepare for success in delivery, they include an overview of the activities and timing of each session, its learning
objectives and the required learning materials, and detailed guidance on how to facilitate the session.

The first session, ‘Fake News’, considers contemporary challenges associated with the consumption of media content and discussions on social media, of significant relevance to the development of resilience to hate and extremism in the online space and good citizenship in the 21st Century.

The purpose of this session is to teach participants basic online fact-checking and media literacy skills, and increase their awareness of misinformation and manipulation in media content in an online context as well as their awareness of echo chambers and their negative consequences. Its aim is to leave participants aware of and interested in these issues, more capable of critically consuming media content, and more likely to consume information from a wide range of sources.

The second session, ‘Impact’, explores the impact of the consumption of information online and the use of social media on young people. It focuses on how the emotions of consumers are manipulated by online content creators, and the negative changes that this can bring about in online behaviour.

This session aims to increase participants’ awareness of the emotionally manipulative tactics employed by content creators online, the divisions these tactics cause, and the impact these tactics can have on an individual’s online behaviour, through a series of activities that address the topics at hand. These activities are participatory to maximise participants’ abilities to recognise and respond to emotional manipulation and divisive rhetoric when they encounter it.

The third session, ‘Your Role’, equips young people to recognise hate speech and respond appropriately when they encounter it online. Increasing participants’ understanding of the causes and manifestations of hate speech, the negative effects hate speech can have, and appropriate responses to hate speech are the focus of this session.

Its objective is to leave participants aware of and engaged with these issues, more capable of recognising hate speech and negative online behaviour, self-aware of their own Internet use, and able to respond appropriately when they encounter hate speech online.

Specific topics include:

- The causes and impact of fake news, and how to reduce its impact
- How to distinguish between fact and opinion
- The use and power of imagery
- How to **fact check and assess information** online
- **Echo chambers**, their impact, and how to avoid being drawn into them
- The use of divisive ‘us and them’ tactics by information manipulators
- How to recognise **emotional manipulation** in online media
- The negative impact of **manipulative tactics** on online behaviour
- How to recognise and respond to **hate speech**
- A review of the **key concepts** which have been introduced throughout the course.

The session plans are accompanied by an online digital deck that provides the facilitator with high quality, interactive multimedia content that chronologically correlates with the specific topics included in the session plans. This content serves to provide key examples of these social challenges specifically in the context of the Netherlands.

**Facilitator Guide**

*Digital Resilience* also provides supplementary delivery and facilitation guidance on the use of the session plans and digital deck with young people. Our handbook is intended to provide facilitators with the content needed to build digital resilience and social inclusion skills in young people against sensitive social issues that can often be challenging to discuss in both formal and informal educational settings. This guidance serves to give the facilitator confidence in delivering the sessions and pedagogical approaches contained within them.

**Student Workbook**

To support the delivery of these sessions, lend additional structure to them and provide a tangible output for teachers delivering the content, Codename Future created a physical workbook, providing key definitions, guidance on activities and a series of exercises for students.

**Delivery Model**

For this pilot project, the delivery model was based on a ‘train-the-trainer’ model. 6 teachers from a range of vocational colleges in Den Haag, Rotterdam and Amsterdam were selected from within Codename Future’s extensive teacher networks within the Netherlands, and selected for participation on the basis of willingness to take part in the pilot study. These teachers were selected in April-May, and took part in a half day training session in September. These teachers then independently delivered the three sessions to one of their regular classes, as well as administering and collecting the participant pre- and post-surveys.
4. The Digital Resilience Evaluation

This chapter presents the results of our evaluation of the pilot project. This evaluation included both quantitative and qualitative elements. The methods included participant pre- and post- surveys designed to measure changes in skills, knowledge, attitudes and behaviours, through a series of confidence-based Likert measures, as well as questions designed to investigate the experience of the classes. These surveys were complimented by interviews with the trained teachers who delivered the content and focus groups with the young people designed to provide a more detailed insight into their experience. The full description of our evaluation methodology can be found in the technical appendix of this report.

Participant Demographics

The demographic details presented in this section were collected through the pre-surveys delivered to participating students. They give us an insight into who it was that experienced the sessions. These demographic details show that the Digital Resilience project certainly reached its intended target audience in age and gender terms, with a significantly diverse ethnic, religious and gender distribution.

The participants in the sessions were predominantly male, a reflection of the wider gender split in Dutch vocational colleges. 64 per cent of participants were male, while 38 per cent were female, and no participants chose another option.

![Fig. 2 Gender of participants (n=124)](image)
There were a broad range of ages in the participating classes. 72 per cent of the participants were in the 16-19 age range, while 28 per cent were either younger or older, with 12 per cent of participants being either younger than 15 or older than 21. This wide age range might have led to some of the participants, particularly those 12 per cent outside the 15-21 age window, feeling that the sessions were either too basic or too advanced.

![Fig. 3 Age of participants (n=125)](image1)

Participants also came from a wide range of religious backgrounds. The three largest categories were 43 per cent of participants who were atheists, 18 per cent who were Muslim, and 15 per cent who were Catholic.

![Fig. 4 Religion of participants (n=119)](image2)
The 135 participants came from a very wide range of diverse backgrounds. The participants themselves were born in 13 different countries, from Portugal to Afghanistan, Togo to Sudan. Their parents came from an even wider range of backgrounds, originating in 26 different countries, from Turkey to Ethiopia, Kosovo to Morocco. In total, 13 per cent of participants were first generation migrants, and 34 per cent were second generation migrants.

(L) Fig. 5 birthplace of participants (n=125)
(R) Fig. 6 birthplace of participants’ parents (n=118)

Participants came from homes where 28 different languages or language combinations were spoken, and in 38 per cent of participants’ homes, a language other than Dutch was spoken.

This diversity amongst participants increases our confidence that this pilot included a broad cross-section of Dutch society, and participants with a wide range of experiences of cultural difference.

Fig. 7 Languages spoken in participants’ homes (n=119)
One teacher interviewed suggested that for some of the participants, the sessions had a particular value because of their ethnic and cultural backgrounds.

“I have immigrant students and I have some students who come from Surinam and Morocco and some who have grandparents who were immigrants. So, it kind of supported them in their self-esteem. [...] They feel empowered by these lessons because they think, ‘There are other people [who care about this], there are other people giving importance to what I feel at the moment.”

**Impact Summary**

This section presents the key findings of our impact evaluation, accompanied by relevant process question results from the participant surveys. The comparison of the pre- to post-survey change in the comparison and participant groups demonstrates varying levels of impact across measures, from changes that were statistically significant even given the small sample size, to positive changes that were notable but could not be regarded as statistically significant in this context, to measures that showed no positive change.

In total, four impact measures produced statistically significant positive change, a further four produced notable positive change, and eight measures produced no change or no notable change.

The four impact statements that demonstrated a statistically significant positive change between the pre- and post-surveys were:

- I understand what echo chambers are, with an increase in agreement of 85 per cent in the participant group
- I understand what the ‘filter bubble’ is, with an increase in agreement of 56 per cent in the participant group
- I understand what ‘scapegoating’ is, with an increase in agreement of 47 per cent in the participant group
- I feel responsible for the wellbeing of people connected to me through social media, with an increase in agreement of 33 per cent in the participant group.

The four impact statements that demonstrated a notable positive change between the pre- and post-surveys were:

- I consider the motivations behind why people post things online, with an increase in agreement of 13 per cent in the participant group
- If I wasn’t sure a story was true, and I wanted to share it, I’d fact check it first, with an increase in agreement of 10 per cent in the participant group.
• I would recognise when a social media post, video or image is designed to emotionally manipulate people, with an increase in agreement of 9 per cent in the participant group

• I’m comfortable expressing my views online, with an increase in agreement of 47 per cent in the participant group.

**Thematic Analysis**

The impact measures that reported the most success centred on a number of key themes, including attitudinal change related to digital citizenship, the key focus of this project, understanding of key terminology relevant to understanding extremism and hate online, and media literacy.

Those impact measures that were less successful similarly centred on a number of key themes, including the ability to recognise and effectively deal with hate speech online, and the flagging or reporting of social media content.

The thematic grouping of the more and less successful elements of the project means that useful insights can be drawn from these results to improve and refine future efforts. These results are analysed by theme below.

**Media Literacy and Critical Thinking**

Two of the key themes examined in the sessions were media literacy in an online context and critical thinking with regards to persuasive content and individuals. The activities within these themes examined fact checking and responsible sharing of online information, awareness of fake news and emotional manipulation in both propaganda and interactions with individuals. Across three key measures in this subject area, positive impacts could be observed. Participants’ confidence that they would consider the motivations of the creators of online content increased by 13 per cent. This was contrasted with by far the largest change in the comparison group of any measure, a decrease of confidence of 8 per cent, which might be explained by external factors, by a limited survey error, or by the small sample size.

In two other key measures of impact, the survey results indicated notable positive change. Participants registered a 10 per cent increase in confidence that they would fact check a story before sharing it if they were not sure it was true, compared to 1 per cent change in the comparison group. Participants also demonstrated a 9 per cent increase in ability to recognise when a social media post, article or website is designed to emotionally manipulate people, compared to 1 per cent change in the comparison group.
Fig. 8 Impact results related to media literacy and manipulation awareness online, percentage change between pre- and post-surveys (participant n=51-54, comparison n=51)

For two other measures in this subject area, however, no change was recorded. When it came to motivation to seek out views and opinions that differed from their own online, there was no change beyond random variation, with a 3 per cent positive change for the participant group, and a 7 per cent positive change for the comparison group. Similarly, when it came to confidence that they could identify fake news, there was a 6 per cent positive change for the participant group, and a 2 per cent positive change for the comparison group.

This suggests that the more complex or technical aspects of these sessions, related to things like group bias in media consumption and identifying indicators of fake news, were less successfully delivered than those aspects that focused on attitudinal or behavioural change.

**Attitudinal Change**

This conclusion is reinforced by the results of other attitudinal or behavioural measures. For example, there was a 33 per cent increase in the most important attitudinal indicator, participants’ feeling of responsibility for the wellbeing of others on their social media, against an 8 per cent rise in the comparison group. This statistically significant positive increase provides strong evidence of the sessions’ capacity to change participants’ attitudes effectively. A positive change was also evident in the increase in participants’ confidence in...
expressing their views online, with a 9 per cent increase against the comparison group’s 3 per cent increase.

![Graph showing impact results related to attitudinal change related to digital citizenship](image)

**Fig. 9** Impact results related to attitudinal change related to digital citizenship, percentage change between pre- and post-surveys (participant n=53-54, comparison n=51, *significant at p<0.005*)

In two particular attitudinal measures, no change was recorded probably because of particularly high base lines among the participants. For example, participants reported a 4 per cent increase in comfort talking to people from different backgrounds than their own, while the control group reported an 8 per cent increase; however, participants both started and ended the intervention more comfortable than the comparison group members talking to people from different background than their own, which might account for this variation. Similarly, a high participant baseline measure of willingness to help others online can account for the small rise on that measure. Before the intervention had begun, the average participant Likert measure for this was 6 out of 7, leaving little space to rise following the intervention (after which it rose to 6.2). The comparison group saw a decline of 2 per cent on this measure over the same period, despite having a lower baseline measure.

**Social Media and Extremism**

In addition to attitudinal change, the sessions were most successful in increasing participants’ knowledge confidence with regard to critical concepts about the online world which are relevant to extremism.

Participants’ confidence that they understood what ‘echo chambers’ were increased by 85 per cent, against 28 per cent in the comparison groups, a large statistically significant
change. Similarly, statistically significant positive changes were recorded for participant confidence in their understanding of the ‘filter bubble’ and ‘scapegoating’, two other concepts that are critical to understanding the role the online space can play in facilitating radicalisation. Participant confidence in understanding of the ‘filter bubble’ increased by 56 per cent, against a comparison group change of 27 per cent, while confidence of understanding ‘scapegoating’ increased by 47 per cent against 24 per cent in the comparison group. These large changes in confidence can be explained by a combination of clarity of delivery around these key concepts within the sessions, and on a low initial baseline of understanding, which leaves sufficient room for improvement.

Smaller positive changes in the comparison group across impact measures generally, but particularly in relation to these three measures, are most probably the result of two factors. Firstly, passing familiarity with the concepts as a result of exposure to the pre-survey can have a small effect, and secondly, pre-survey exposure to unfamiliar concepts can precipitate conversations around them, either with peers or with teachers, that increase knowledge confidence in advance of the post-survey.

Fig. 10 Impact results related to extremism-relevant social media knowledge, percentage change between pre- and post-surveys (participant n=51-52, comparison n=51, *significant at p<0.005)

Skills Measures
When it came to some other more detailed skills measures, results were mixed, but no notable positive change was observed. There was an 8 per cent increase in the participant group’s confidence that they could recognise ‘us and them arguments’ against a 1 per cent decrease in comparison group confidence, but variation in the comparison group surveys
means this can’t be confirmed as statistically significant. Across the following confidence measures there was no variation of more than 3 per cent either positively or negatively in either comparison and participant groups: that they knew what to do when confronted with hate speech online, recognising the difference between hate speech and free speech, confidence that they knew how to ‘flag’ content for removal and understanding the difference between free speech and hate speech. Insights from teachers and Dutch policymakers suggest that a significant reason for this less positive impact around hate speech related skills is the wider difficulty of defining and dealing with hate speech effectively in wider society.

**Process evaluation**

In addition to the impact data, the surveys included a number of process questions that focused on participants’ experience of the sessions. The three focus groups with participants and five semi-structured interviews with teachers delivered insights into the strengths and weaknesses of the sessions from the perspectives of both those being taught and those teaching. This section presents those findings of the process questions that related to self-perceived impact based on student observations.

Participants were asked whether they understood the content, whether they felt they gained new skills or knowledge, and whether they thought the sessions would change their behaviour online.

Participants reported a high level of understanding of the content; 67 per cent of participants reported understanding all the content by the end of the sessions, and the rest understood most of it. This suggests that the content was clear and approachable, but it could also mean that it was insufficiently stretching. Feedback from the teachers, however, suggests that the participants were challenged, and that perhaps there was too much information covered, not too little:

“The discussions kind of thing. Discussions, debating, arguing, trying to put into words your own argument. [...] Because it sharpened their opinions on things. They just like to have an opinion on things. That’s what they like [...] if you challenge them and make their opinions sharper. They kind of value it.”

“Maybe next time is to concentrate on some subjects and because now there’s a lot of information and maybe, for some of my students, it was better to make a selection about that.”

“They wanted me to do one hour per session, which was not enough because there was too much.”
“When they have to work individual, some of them had some problems with it. I think the reason for that is in my class it is a different kind of level. I have some students [who are abler and some who are less able].”

In the focus groups, participants described seeing new perspectives on hate and extremism online, and being exposed to new ideas:

“The lesson on how you can influence people towards hate was interesting; I was surprised to see people share that.”

“I have learned more about other students, what they are like and what they think about [these topics]; you’ll think about it and maybe change your own opinion. It’s in your mind. Maybe I should see it from the other side next time.”

In fact, one teacher shared that they had learned new facts from the sessions too:

“The new model of fake news, that was new to me! As in to use it to earn money, not only for politics reasons but also for economic reasons.”

The participants were very positive that they learned new skills, with 81 per cent reporting that they had done so. They were less positive around knowledge gain, with around half suggesting that they had not gained new knowledge, and about half suggesting that they did. This difference supports the idea that the participants gained more from the discursive, attitudinal elements of the sessions than the more detailed, technical aspects. It might also reflect the more practical orientation of the students.

(L) Fig. 11 Participant responses to the survey question “Do you feel like you gained new skills” (n=36)
(R) Fig. 12 Participant responses to the survey question “Do you feel like you gained new knowledge” (n=37)
Both the qualitative and the quantitative analyses suggested that there was a strong preference for the more practical exercises among participants. This observation was echoed by the teacher interviews, which suggested that the participants engaged most effectively with the practical, discursive tasks rather than the reflective sessions:

“The things they have to “do”, these things went very well. The films were very interesting and Session Three really got their attention. They really liked that.”

“The [activity] with the videos, it [provided] a bit of variety. They didn’t get bored. They liked the different activities.”

“They also liked the quiz. It made them active. Because it’s very difficult for my students to focus, to get them to concentrate.”

One teacher suggested that the resources should be focused on skills to a greater extent, to emphasise those elements that the pupils engaged with most:

“Maybe there is a way to concentrate on skills, the skills to recognise fake news, to recognise [online] hate, to recognise the manipulations, and also the attitudes to use those skills.”

The teachers interviewed were positive about the arrangement of the sessions, which started with simpler topics like fake news, and moved on to more complex issues like reactions to hate speech online.

“I thought that the line that they followed was very good. The first lesson was more like what do you recognise, more like the first steps. It was very well-built because, you know, the first lesson was an introduction, so that was good. The second and the third course of lessons went more into the details [...] so I think they built it correctly.”

Perhaps the most important direct questions we asked participants was whether or not they thought that the sessions would influence how they act online. 47 per cent thought that they would, while 47 per cent thought they would not. This question is a purely self-assessed measure, and tracking actual behavioural change is a far harder task. Nevertheless, half of the participants stating that the sessions will change their online behaviour is an encouraging indicator of potential practical impact in their day to day lives.
Fig. 13 Participant responses to the survey question “Do you feel like you will behave differently online as a result of what you learned?” (n=34)

In order to influence behaviour, the sessions had to raise awareness or present new perspectives, and despite the generally positive result, the mix of responses to whether or not that was achieved for individual participants provide an indication of the cause of this divided result:

“Fake news, I knew nothing about it. I had never paid attention to the subject.”

“The assignment about the factual and [misleading] headlines was interesting, I did not know that and I enjoyed learning about it.”

“We know it already, we know that sometimes things are not real. We don’t take it seriously.”

Some teachers said they did not know whether or not the sessions would change their students’ online behaviour, but others suggested they would:

“Yes [it will affect their online behaviour], because I told them a lot of times to be critical and that it’s important to check their sources; I [think] that they will remember that part of the lesson.”
5. Conclusions

This evaluation report provides a number of insights into the efficacy of the Digital Resilience curriculum in Dutch vocational colleges, as well as highlighting areas where improvements in the resources or their delivery might fruitfully be made.

Key Outcomes

There were positive impacts across vital capacities, particularly in key attitudes and knowledge areas.

The greatest positive impacts were observed in two areas. The first was related to attitudes connected to good digital citizenship, such as the 33 per cent increase in feeling responsible for the wellbeing of people connected to the participants on social media, and a 13 per cent increase in confidence that participants consider the motivations behind why people post things online. The second was related to knowledge of key concepts relevant to radicalisation and manipulation online, for example the 85 per cent increase in understanding of echo chambers and a 47 per cent increase in confidence understanding scapegoating.

There were some indicators that the sessions might influence participants’ online behaviours.

Participants’ reflections suggested that they took away practical capacities from the sessions, with 81 per cent suggesting that they had gained new skills. Given that providing learners with new skills is less challenging than actually changing how they behave online – and while self-assessment of future behavioural change is not the same as proof of behavioural change – it is highly encouraging that 47 per cent of participants felt that the intervention would change their behaviour online.

This evaluation shows that with arm’s length support and independent provision, teachers can successfully deliver digital citizenship education for CVE purposes.

This evaluation tested the ability of non-specialist teachers to deliver the Digital Resilience resources in their own vocational college classrooms with basic training, a teacher guide, session plans and no further material provision. Moreover, for each teacher, this pilot was their experience of delivering these resources. This situational context therefore provides a realistic evaluation of the efficacy of the Digital Resilience sessions were they introduced on a wider scale, with positive results.

The diversity of the participants’ ethnic and national backgrounds was significant, which qualitative data suggests might have enhanced their experience.
The 135 participants were born in 13 countries, with parents originating in 26 different countries, speaking 28 different languages or language combinations in the home. 13 per cent of participants were first generation migrants, while 34 per cent were second generation migrants. Some qualitative data suggests that this diversity may have enhanced the value of the sessions, providing an opportunity to discuss identity, hate and tolerance much desired by the participants.

Areas for Improvement

Impact indicators measuring change in confidence on more complex content were less positive.

In some areas, no significant positive or notable positive impacts occurred, with no change in participant confidence in these areas between the pre- and the post-surveys. These areas tended to be related to more complex knowledge and skills outcomes, for example being able to distinguish the difference between free speech (sometimes encompassing highly prejudiced opinions) and hate speech, and being able to react effectively to hate content online. There are a number of potential reasons for this beyond the complexity of these issues, including: lower levels of teacher confidence around these subjects, their more nuanced learning outcomes, a need for simpler, clearer guidance on these issues for teachers, and the difficulty of understanding the lines between free speech and hate speech in wider society in the Netherlands.

The practical activities, discussions and outcomes were most popular with participants and achieved greater efficacy than written, individual work.

Both qualitative and quantitative data from the evaluation suggested that when activities were practically orientated, physically active and discursive, participants were the most engaged and, generally, achieved higher impact. Conversely, more academic, individual activities on more rarefied subject matter achieved lower impact. This might reflect the more practical curricula focus of vocational colleges, but also highlights the importance of ensuring that in future development, a premium is placed on practical, dynamic activities and discussions.

The amount of content delivered within the three sessions needs to be reduced, or the amount of time provided to deliver it needs to be increased.

A consistent theme of the interviews with teachers was that the sessions contained too much content to be delivered in three hours, leading some teachers to condense exercises, and some to extent the amount of time in which to deliver the sessions. In future development, exercise need to be streamlined, or more time provided to deliver the content.
6. Technical Appendix

This appendix provides additional information on how we selected participants and the evaluation methodology for the project.

Methodology

The quantitative aspect of the evaluation centred on the delivery of the pre- and post-surveys to session participants, as well as young people from comparison group classes. These surveys were completed by participant and comparison groups in the week before the series of three classes began, and the week after they concluded, to provide a consistent timeframe from comparative measurement, as well as enough time to reduce the risk of influencing the session or measuring only very short-term change.

These surveys included 16 impact measures, presented as 1-7 Likert scales, which indicated agreement with a series of confidence statements, ranging from strongly agree to strongly disagree. Individuals’ pre- and post-surveys were anonymously matched through the use of an anonymous code generator, so at no point did evaluators have access to participants’ identifying information. The participant surveys also included demographic questions, designed to provide insights into the background of participants, and process questions, designed to provide information about the participant experience of the sessions. Comparison group students were presented only with impact measurement questions.

All 135 participating students were surveyed, as were all 108 comparison group students. However, these surveys were subject to high attrition rates, particularly in terms of the completion and return of the post-surveys. As both pre- and post-surveys at the individual level were required for analysis, only completed pre-post sets were included in the analysis. In total, 51 comparison group students and 54 participant group students returned completed surveys, an attrition rate of 53 per cent and 60 per cent respectively. This attrition rate was particularly high, and while these sample sizes still allow for useful conclusions to be drawn, a larger sample size would have allowed for stronger conclusions.

These surveys were complemented by qualitative research. This included three participant focus groups of six-eight students conducted by Codename Future, each lasting an hour, which took place some time after the delivery of the sessions and focused on the participants’ experience of the sessions, the extent to which they felt they gained new skills or knowledge, whether the sessions changed their attitudes to social media and whether it might change their behaviour online. The participants in these focus groups were selected by their class teachers. Additionally, ISD researchers conducted semi-structured interviews with the five teachers who delivered the sessions to their classes, in order to understand
their perspective on the delivery of the sessions, the content and the response of their classes to them.

In considering these results, it is important to make a number of observations regarding location, sample sizes and the comparability of the data.

This pilot study focused on students in vocational colleges, and so the impact of the session on these students and their experience of them may not be representative of all Dutch young people.

This pilot was based on a hands-off ‘train-the-trainer’ approach, without ongoing support from either Codename Future or ISD to the teachers delivering the content. Additionally, the training preceded the initial delivery in some cases by a period of two months. As such, this represents an ‘acid test’, realistic pilot scenario; teachers given limited training and a copy of the teacher guidance on a larger scale would have the same level of support and assistance that we tested in this pilot.

While the sample size for this project is limited, the use of comparison groups selected from the same school and year as the participating groups represents a significant enhancement of the quality of evidence supporting this type of digital citizenship intervention, characterised by a focus on attitudinal and behavioural development aspects of digital citizenship for the purposes of increasing resilience to extremist grooming and manipulation online. These characteristics typify ISD’s suite of digital citizenship projects and resources.

**Session Locations**
Five classes were selected for participation in the series of three lessons in the winter of 2017. Two of these classes were in Amsterdam, two were in Rotterdam, and one was in Den Haag.

**Participant Demographics**
The following tables lay out the key demographic information of the participants.

<table>
<thead>
<tr>
<th>Age</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>15 or younger</td>
<td>6%</td>
</tr>
<tr>
<td>16</td>
<td>18%</td>
</tr>
<tr>
<td>17</td>
<td>32%</td>
</tr>
<tr>
<td>18</td>
<td>20%</td>
</tr>
<tr>
<td>19</td>
<td>12%</td>
</tr>
<tr>
<td>20</td>
<td>4%</td>
</tr>
<tr>
<td>21 or older</td>
<td>8%</td>
</tr>
</tbody>
</table>

*Table 1, Age of participants*
**Gender**

<table>
<thead>
<tr>
<th>Gender</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>male</td>
<td>64%</td>
</tr>
<tr>
<td>female</td>
<td>38%</td>
</tr>
<tr>
<td>other</td>
<td>0%</td>
</tr>
</tbody>
</table>

*Table 2, Gender of participants*

**Religion**

<table>
<thead>
<tr>
<th>Religion</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>No religion</td>
<td>43%</td>
</tr>
<tr>
<td>Catholic</td>
<td>15%</td>
</tr>
<tr>
<td>Protestant</td>
<td>7%</td>
</tr>
<tr>
<td>Buddhist</td>
<td>0%</td>
</tr>
<tr>
<td>Hindu</td>
<td>4%</td>
</tr>
<tr>
<td>Jews</td>
<td>5%</td>
</tr>
<tr>
<td>Muslim</td>
<td>18%</td>
</tr>
<tr>
<td>Ietsist</td>
<td>1%</td>
</tr>
<tr>
<td>Sikh</td>
<td>1%</td>
</tr>
<tr>
<td>Other - Christian</td>
<td>6%</td>
</tr>
<tr>
<td>Other - Myself</td>
<td>1500%</td>
</tr>
</tbody>
</table>

*Table 3 Religion of participants*

**Participants’ Place of Birth**

<table>
<thead>
<tr>
<th>Place of Birth</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Netherlands</td>
<td>87%</td>
</tr>
<tr>
<td>Outside of the Netherlands</td>
<td>13%</td>
</tr>
</tbody>
</table>

*Table 4, Participants’ place of Birth*

**Participants’ Parents’ place of Birth**

<table>
<thead>
<tr>
<th>Place of Birth</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Outside of the Netherlands</td>
<td>47%</td>
</tr>
<tr>
<td>Netherlands</td>
<td>53%</td>
</tr>
</tbody>
</table>

*Table 5, Participants’ parents place of Birth*

**Languages spoken in participants’ home**

<table>
<thead>
<tr>
<th>Language</th>
<th>Percentage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dutch</td>
<td>62%</td>
</tr>
<tr>
<td>Languages other than Dutch</td>
<td>38%</td>
</tr>
</tbody>
</table>

*Table 6, Languages spoken in participants’ home*
Participant Surveys
Figure 19 presents the questions asked in the participant pre- and post-surveys. All demographic questions were asked in the pre-survey, while all process-related questions were asked in the post-survey. All questions were presented in Dutch, and are here translated into English.

1. What is your gender?
   - Male
   - Female
   - Other

2. What is your religion?
   - None/non-religious
   - Catholic
   - Protestant
   - Buddhist
   - Hindu
   - Jewish
   - Muslim
   - Other (please tell us):

4. How old are you?
   - 15 or under
   - 16
   - 17
   - 18
   - 19
   - 20
   - 21 or over

5. Where were you born?
   - The Netherlands
   - Not in the Netherlands (Please tell us):
   - Don’t know

6. Where were your parents born?
   - The Netherlands
   - Not in the Netherlands (Please tell us):
   - Don’t know

7. What language do you speak at home? (Multiple answers possible)
   - Dutch
   - Other (Please tell us):
8. Please tell us how much you agree with the following statements. There are seven options, from 1 which is most 'Strongly Disagree' to 7, which is most 'Strongly Agree.' Tick only once box in each line.

<table>
<thead>
<tr>
<th>Strongly Disagree</th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
<th>Strongly Agree</th>
<th>7</th>
</tr>
</thead>
<tbody>
<tr>
<td>I feel comfortable talking to people from backgrounds different to my own.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel confident expressing my views online.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I feel responsible for the wellbeing of people connected to me through social media.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>If I wasn’t sure a story was true, and I wanted to share it, I’d fact check it first.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I consider the motivations behind why people post things online.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I’d help out a friend if I thought they were in trouble online.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I’m motivated to seek out views and opinions that differ to my own online.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would know what to do if I came across hate speech online.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I know how and why to mark social media content for removal.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I understand what ‘us and them’ arguments are.</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>I would recognise when a social media post, video or image is designed to emotionally manipulate</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
people.

| I understand the differences between hate speech and free speech. |
| I understand what echo chambers are. |
| I understand what the ‘filter bubble’ is. |
| I am able to identify ‘fake news.’ |
| I understand what ‘scapegoating’ is. |

9. Did you enjoy the sessions?
   - [ ] I liked them a lot
   - [ ] I liked them
   - [ ] I neither liked them nor disliked them
   - [ ] I disliked them
   - [ ] I disliked them a lot
   - [ ] I don’t know

10. Which session did you find the most engaging?
    - [ ] The first session
    - [ ] The second session
    - [ ] The third session
    - [ ] I don’t know

11. How relevant do you feel the content of the sessions were to you/your online life?
    - [ ] Highly relevant
    - [ ] Quite relevant
    - [ ] It was neither relevant nor irrelevant
    - [ ] Quite irrelevant
    - [ ] Highly irrelevant
    - [ ] I don’t know

12. Do you feel like you understood the subject matter by the end of the sessions?
    - [ ] Understood everything
    - [ ] Understood some of it
    - [ ] Understood little
    - [ ] Understood nothing.
    - [ ] I don’t know

13. Do you feel like you learned new skills?
    - [ ] Yes, lots
    - [ ] Yes
14. Do you feel like you gained new knowledge?
- Yes, lots
- Yes
- No
- No not at all
- I don’t know

15. Do you think that you’ll behave differently online as a result?
- Yes, significantly
- Yes
- No
- No not at all
- I don’t know

16. Do you feel like the sessions were appropriate for your age?
- Yes, I think they were appropriate for my age group.
- No, I think they were more appropriate for younger people.
- No, I think they were more appropriate for older people.
- I don’t know

17. Would you be interested in being part of any future sessions on the same subject as these ones?
- Yes
- No
- I don’t know

18. If you have any further comments about the sessions, please write them in the box below: [open text].

Fig. 19, Participant pre- and post-survey content

Summary Statistics and Significance Testing
When working with non-parametric data (such as this sample, where answers were limited to responses on a Likert scale), the Mann-Whitney (MW) test has been shown to be the most robust method for significance testing. As such, the change scores between the pre- and post-surveys of the control and participant groups were evaluated using this test. These results are presented in Table 7.

In order to establish that the control group was a valid group for comparison, two tests were conducted prior to the above mentioned Mann-Whitney test:

- Wilcoxon Signed Ranks (WSR) test: The pre- and post-surveys of the control group were analysed with the WSR test to establish that there were no significant changes between the pre- and post-surveys. These results are presented in Table 8.
• Mann-Whitney test: The pre-surveys of the control and participant groups were analysed with a MW test to establish that there were no significant differences between these two groups before the intervention. These results are presented in Table 9.

We use the standard formula for calculating the effect size from MW test, \( \eta^2 \):

\[
\eta^2 = \frac{Z^2}{(N - 1)}
\]

where \( Z \) is the Z test statistic and \( N \) is the sample size.\(^{10} \) These results are presented in Table 9.

<table>
<thead>
<tr>
<th></th>
<th>Control</th>
<th></th>
<th></th>
<th></th>
<th>Participant</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean before</td>
<td>Mean after</td>
<td>Difference</td>
<td>Sample size (N)</td>
<td>Mean before</td>
<td>Mean after</td>
<td>Difference</td>
</tr>
<tr>
<td>Q1</td>
<td>5.2</td>
<td>5.6</td>
<td>0.4</td>
<td>50</td>
<td>5.62963</td>
<td>5.833333</td>
<td>0.203704</td>
</tr>
<tr>
<td>Q2</td>
<td>4.882353</td>
<td>5.019608</td>
<td>0.137255</td>
<td>51</td>
<td>4.759259</td>
<td>5.185185</td>
<td>0.425926</td>
</tr>
<tr>
<td>Q3</td>
<td>3.235294</td>
<td>3.490196</td>
<td>0.254902</td>
<td>51</td>
<td>3.301887</td>
<td>4.396226</td>
<td>1.09434</td>
</tr>
<tr>
<td>Q4</td>
<td>5.372549</td>
<td>5.313725</td>
<td>-0.05882</td>
<td>51</td>
<td>5.259259</td>
<td>5.796296</td>
<td>0.537037</td>
</tr>
<tr>
<td>Q5</td>
<td>4.117647</td>
<td>3.784314</td>
<td>-0.33333</td>
<td>51</td>
<td>3.923077</td>
<td>4.442308</td>
<td>0.519231</td>
</tr>
<tr>
<td>Q6</td>
<td>5.666667</td>
<td>5.568627</td>
<td>-0.09804</td>
<td>51</td>
<td>6.037736</td>
<td>6.226415</td>
<td>0.188679</td>
</tr>
<tr>
<td>Q7</td>
<td>4.88</td>
<td>5.22</td>
<td>0.34</td>
<td>50</td>
<td>5.444444</td>
<td>5.611111</td>
<td>0.16667</td>
</tr>
<tr>
<td>Q8</td>
<td>5.039216</td>
<td>4.901961</td>
<td>-0.13725</td>
<td>51</td>
<td>5.37037</td>
<td>5.425926</td>
<td>0.055556</td>
</tr>
<tr>
<td>Q9</td>
<td>5.607843</td>
<td>5.705882</td>
<td>0.098039</td>
<td>51</td>
<td>5.740741</td>
<td>5.851852</td>
<td>0.111111</td>
</tr>
<tr>
<td>Q10</td>
<td>4.941176</td>
<td>4.901961</td>
<td>-0.03922</td>
<td>51</td>
<td>5.301887</td>
<td>5.735849</td>
<td>0.433962</td>
</tr>
<tr>
<td>Q11</td>
<td>5.607843</td>
<td>5.647059</td>
<td>0.039216</td>
<td>51</td>
<td>5.222222</td>
<td>5.666667</td>
<td>0.444444</td>
</tr>
<tr>
<td>Q12</td>
<td>5.960784</td>
<td>5.901961</td>
<td>-0.05882</td>
<td>51</td>
<td>6.169811</td>
<td>6.075472</td>
<td>-0.09434</td>
</tr>
<tr>
<td>Q13</td>
<td>1.980392</td>
<td>2.529412</td>
<td>0.540902</td>
<td>51</td>
<td>2.173077</td>
<td>4.019231</td>
<td>1.846154</td>
</tr>
<tr>
<td>Q14</td>
<td>2.431373</td>
<td>3.078431</td>
<td>0.647059</td>
<td>51</td>
<td>2.566038</td>
<td>4</td>
<td>1.433962</td>
</tr>
<tr>
<td>Q15</td>
<td>5.72549</td>
<td>5.823529</td>
<td>0.098039</td>
<td>51</td>
<td>6.115385</td>
<td>6.365385</td>
<td>0.25</td>
</tr>
<tr>
<td>Q16</td>
<td>3.078431</td>
<td>3.823529</td>
<td>0.745098</td>
<td>51</td>
<td>3.711538</td>
<td>5.442308</td>
<td>1.730769</td>
</tr>
</tbody>
</table>

Table 7.

Results from Wilcoxon Signed-Rank test on Control group, all matched pairs

<table>
<thead>
<tr>
<th></th>
<th>W statistic</th>
<th>Z statistic</th>
<th>P value (2-tailed)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Q1</td>
<td>-126</td>
<td>-1.79</td>
<td>0.0735</td>
</tr>
<tr>
<td>Q2</td>
<td>-15</td>
<td>-0.25</td>
<td>0.8026</td>
</tr>
<tr>
<td>Q3</td>
<td>-63</td>
<td>-0.61</td>
<td>0.5419</td>
</tr>
<tr>
<td>Q4</td>
<td>-1</td>
<td>-0.01</td>
<td>0.992</td>
</tr>
<tr>
<td>Q5</td>
<td>116</td>
<td>1.13</td>
<td>0.2585</td>
</tr>
<tr>
<td>Q6</td>
<td>30</td>
<td>0.7</td>
<td>0.4839</td>
</tr>
<tr>
<td>Q7</td>
<td>-86</td>
<td>-1.49</td>
<td>0.1362</td>
</tr>
<tr>
<td>Q8</td>
<td>32</td>
<td>0.48</td>
<td>0.6312</td>
</tr>
<tr>
<td>Q9</td>
<td>-17</td>
<td>-0.43</td>
<td>0.6672</td>
</tr>
<tr>
<td></td>
<td>U_A M-W statistic</td>
<td>Z statistic</td>
<td>P value (2-tailed)</td>
</tr>
<tr>
<td>-----</td>
<td>-------------------</td>
<td>-------------</td>
<td>-------------------</td>
</tr>
<tr>
<td>Q1</td>
<td>1127</td>
<td>1.45</td>
<td>0.1471</td>
</tr>
<tr>
<td>Q2</td>
<td>1412.5</td>
<td>-0.22</td>
<td>0.8259</td>
</tr>
<tr>
<td>Q3</td>
<td>1311</td>
<td>0.26</td>
<td>0.7949</td>
</tr>
<tr>
<td>Q4</td>
<td>1368</td>
<td>0.05</td>
<td>0.9601</td>
</tr>
<tr>
<td>Q5</td>
<td>1417</td>
<td>-0.6</td>
<td>0.5485</td>
</tr>
<tr>
<td>Q6</td>
<td>1086.5</td>
<td>1.72</td>
<td>0.0854</td>
</tr>
<tr>
<td>Q7</td>
<td>1091</td>
<td>1.68</td>
<td>0.093</td>
</tr>
<tr>
<td>Q8</td>
<td>1227.5</td>
<td>0.96</td>
<td>0.3371</td>
</tr>
<tr>
<td>Q9</td>
<td>1350.5</td>
<td>0.17</td>
<td>0.865</td>
</tr>
<tr>
<td>Q10</td>
<td>1200.5</td>
<td>0.98</td>
<td>0.3271</td>
</tr>
<tr>
<td>Q11</td>
<td>1524</td>
<td>-0.94</td>
<td>0.3472</td>
</tr>
<tr>
<td>Q12</td>
<td>1184.5</td>
<td>1.08</td>
<td>0.2801</td>
</tr>
<tr>
<td>Q13</td>
<td>1226.5</td>
<td>0.65</td>
<td>0.5157</td>
</tr>
<tr>
<td>Q14</td>
<td>1312.5</td>
<td>0.25</td>
<td>0.8026</td>
</tr>
<tr>
<td>Q15</td>
<td>1062.5</td>
<td>1.73</td>
<td>0.0836</td>
</tr>
<tr>
<td>Q16</td>
<td>1121</td>
<td>1.35</td>
<td>0.177</td>
</tr>
</tbody>
</table>

Table 8.
References