"Schools, parents and social media companies all have a responsibility to promote good character online…"

THE MORAL WEB: YOUTH CHARACTER, ETHICS AND BEHAVIOUR

Peter Harrison-Evans
Alex Krasodomski-Jones

DEM⊙S

Demos is Britain's leading cross-party think-tank. We produce original research, publish innovative thinkers and host thought-provoking events. We have spent over 20 years at the centre of the policy debate, with an overarching mission to bring politics closer to people.

Demos has always been interested in power: how it works, and how to distribute it more equally throughout society. We believe in trusting people with decisions about their own lives and solving problems from the bottom-up.

We pride ourselves on working together with the people who are the focus of our research. Alongside quantitative research, Demos pioneers new forms of deliberative work, from citizens' juries and ethnography to ground-breaking social media analysis.

Find out more at www.demos.co.uk

THE MORAL WEB:
YOUTH CHARACTER,
ETHICS AND BEHAVIOUR
ON SOCIAL MEDIA

Peter Harrison-Evans
Alex Krasodomski-Jones

DEM☉S

Contents

Acknowledgements

Executive summary

Social media has moved from the periphery to the mainstream of public consciousness at a rapid pace, with close to 100 per cent of 16–24-year-olds now having a social media account.[1] From being considered a relatively trivial and narcissistic aspect of online communication, social media is increasingly seen as socially and politically transformative – shaping human interaction from the micro level of individual relationships to the macro level of public discourse and, even, electoral outcomes. As with many disruptive technologies, social media's novelty and pervasiveness has thrown up a host of moral and ethical questions as longer-standing social norms struggle for purchase in new environments that simultaneously accelerate and distance human interaction.

Public anxiety around the impact of social media is arguably strongest in relation to the developmental impacts of young people's experiences on social networking sites. As the earliest and heaviest adopters of the technology, adolescents seem most drawn to social media, at a stage in life when they are most open to developmental influence, and most prone to risk-taking and peer-led behaviour. Understandably parents, school leaders and policy makers feel compelled to intervene to minimise these new sources of risk and harm. This has led policy makers to focus on online safeguarding, aiming to minimise extrinsic risks from radicalisation (eg the Prevent duty in schools) and other forms of abuse, as well as preventing young people from accessing harmful material through filtering and monitoring. Parents are similarly often drawn into restrictive strategies, which attempt to reduce or limit access to online social networks.

While these strategies may have merits, they run up against a number of practical and sociocultural problems. From a practical perspective, the nature of social media

technologies make traditional forms of regulation difficult – for example, constant connectivity through smartphones means that young people are regularly online outside spaces of parental mediation. There's also a danger that overly intrusive interference into young people's digital worlds – seen as spaces of youth and peer independence – becomes counterproductive, limiting the development of digital experience and/or encouraging more covert online behaviour.

Responding to this impasse, this study examines young people's agency in their engagement with social networking sites. Rather than focusing on how to regulate the digital environment that surrounds young people, we aim to understand how young people act on social media, and crucially what motivates them to act in the ways that they do. To do this we take character as our frame of analysis – the personal traits, values and skills that guide individual conduct, and ultimately shape outcomes in school and in wider life. While Demos has built up a significant body of work looking at the impact of character in the offline world, supporting educational, and mental and physical health outcomes, an analysis of what constitutes good character online is a new frontier for research.[2]

Our study takes a particular interest in the moral and civic aspects of character, mapping the prevalence of risky and unethical behaviours on social media, while also assessing engagement prosocial behaviours, and displays of civic virtue. We use a scenario-based approach to drill deeper into young people's online reasoning, to understand what drives their behaviour, and unearth emergent peer-led codes of conduct. Finally, we ask what key stakeholders – schools, parents and social media companies – can do to support positive youth decision making.

Methodology

To explore the behaviour and choices of young people over social media, we used the following primary research methods (see Appendix 2 for a fuller description of the research methodology):

· a survey of 668 16–18-year-olds, administered over Facebook, exploring youth online conduct, as well as responses to three social media scenarios, in which respondents were asked what they would do (and why they would choose that course of action) in three situations of increasing risk of harm:
  – mild: *'you write a post arguing for a cause you believe in. someone comments on it, aggressively disagreeing with your opinion'*
  – moderate: *'one of your classmates writes a post insulting someone else in your class, and tags you in it'*
  – severe: *'a friend shares an explicit image of someone else in your class over social media, and asks you to forward it to another friend'*
· focus groups with 40 16–18-year-olds in Birmingham and London
· an analysis of dynamics and content of trolling attacks on Twitter by Demos' Centre for the Analysis of Social Media
· an expert roundtable bringing together policy experts and practitioners from character education, digital citizenship, online child safety and the tech sector.

Key findings

Social media creates opportunities to display moral and civic virtues, such as honest and empathetic communication, new forms of civic participation and acts of courage that counter online abuse.

Our survey found that the vast majority (88 per cent) of 16–18-year-olds polled say they have given emotional support to a friend on social networking sites, and this figure differs little by gender or frequency of social media use. We also find that just over half (51 per cent) say that they have posted about 'a political or social cause that they care about'.

By creating greater visibility around online communication, including negative online behaviour, social media also presents opportunities for young people to display courage – taking a risk in defence of others, as seen by the prevalence of counter-speech in case study 1.

However, a significant minority of 16–18-year-olds say they have engaged in risky or unethical behaviours on social networking sites, such as cyberbullying or trolling, with boys and high intensity social media users more likely to say they have done this.
Just over a quarter (26 per cent) of those polled say they have 'bullied or insulted someone else' over social media, with a smaller but not insignificant 15 per cent saying that they have 'joined in with other people to "troll" a celebrity or public figure'.
Our survey found that boys are substantially more likely than girls to say that they have bullied or insulted someone (32 per cent compared with 22 per cent) or trolled a public figure (22 per cent compared with 10 per cent). Similarly, individuals who say that they visit social networking sites more than ten times a day are more likely than those visiting social networking sites less frequently to say they have bullied (32 per cent compared with 17 per cent) or trolled (18 per cent compared with 11 per cent) someone.

In responding to our social media scenarios, young people were more likely to choose a positive course of action, although a substantial minority say they would do nothing.
Across two of our three social media scenarios the most commonly selected responses related to some form of 'positive' action, responses that attempt to resolve the situation and minimise harm. However, we also find a substantial proportion of respondents who say they would do nothing in each scenario (mild – 24 per cent; moderate – 17 per cent; severe – 17 per cent).
Our scenario findings show that boys are significantly more likely than girls to say they would do nothing. Some 24 per cent of all responses from boys across the three scenarios

were statements of inaction, compared with just 15 per cent of all responses from girls.

Responses to our scenario questions also suggest that young people are attuned to moral situations over social media, although gender appears to be a significant factor in influencing levels of moral sensitivity.

Our findings show that many young people are sensitive to moral situations on social networking sites – with the proportion of respondents giving reasons that made it clear they did not recognise the situation as one requiring moral judgement (eg, 'it's not a big deal') falling from 40 per cent in our 'mild' scenario to just 8 per cent in our 'severe' scenario.

When we analyse by gender we find that boys are almost twice as likely as girls not to recognise a scenario as requiring any kind of moral judgement. Girls, on the other hand, are more likely to be motivated by concerns for both known and unknown others, and are also more likely to be motivated by anger and emotional distress.

There's evidence of there being an implicit code of conduct among many young people, which fosters action in response to significant harm inflicted on close friends, but engenders the response 'it's not my business' to becoming involved in anything beyond this.

Our survey and qualitative data revealed that the most common reason for stepping in to try to resolve a situation positively was to defend or support someone that respondents knew. However, our focus groups often revealed a narrow conception of 'known other' with young people prepared to act to support close friends or family, but reluctant to help those less well known to them (eg, 'someone at my school'). The focus group participants often said they would act only once a relatively high threshold of harm had been crossed.

Our findings suggest there is a relationship between character and online behaviour and decision making. Young people who admit to engaging in risky or unethical behaviour online also demonstrate lower levels of moral sensitivity to others, and have lower self-reported character capabilities.

We find that individuals with lower 'character scores' have a significantly higher likelihood of engaging in risky or unethical behaviours than those with higher character scores, for example they are more likely to say that they have bullied (38 per cent compared with 16 per cent) or trolled (22 per cent compared with 10 per cent) someone.

When looking at responses to our scenarios, we find that individuals scoring high on our character scale were far more likely to think in relation to known or unknown others, and less likely to focus on themselves in justifying their responses. Some 44 per cent of those with higher character scores said they were motivated by 'moral' or 'ethical' reasons in deciding what to do in our scenarios, compared with just 23 per cent of those with lower character scores.

Certain character traits seem more closely related to specific types of behaviour or reasoning than others. For instance, those who say they have bullied others on social media tend to have lower levels of empathy and self-control, while those who exhibit greater help-seeking behaviours are likely to have higher levels of compassion and civic mindedness.

Individual character traits – of empathy, self-control, compassion, honesty and civic participation – seem generally aligned in their relationship to specific types of behaviour or reasoning. Broadly speaking, those scoring highly on each of our surveyed character traits were less likely to engage in negative behaviours, and more likely to take others into consideration when thinking about what course of action to take.

However, it seems that some traits may be more significant than others for certain types of behaviour or ways of thinking. For example, while 64 per cent of the sample as a whole score highly on empathy, this falls to

49 per cent among those who say they have insulted or bullied someone over social media. Turning to our scenarios, we find that people who score highly for civic mindedness are substantially less likely to justify their responses on the basis of consequences for themselves and, conversely, are more likely to want to seek advice from others when choosing what course of action to take.

Despite a chronic lack of evaluative evidence on what is effective in helping young people to make healthy choices on social media, digital citizenship is a promising approach to support good character online.

School staff have generally been proactive in responding to emerging risks from online social networking, with over 80 per cent of our survey respondents saying that that they had received some form of guidance at school, but it is unclear how effective much of this guidance is. Our survey results suggest there is no relationship between education and improved online conduct.

Despite a lack of evaluative evidence on online safety and skills programmes, digital citizenship education is a promising means of supporting good character online. The benefits of digital citizenship education include focusing on internal competencies and skills, rather than top-down regulation; highlighting the moral issues of online social networking; and using approaches that have been shown to be effective in wider character education and social and emotional learning programmes, such as reflective practice, peer-led learning, and the involvement of parents and guardians.

A number of programme evaluations of digital citizenship and online character schemes and pilots are currently under way, including the Jubilee Centre's pilot Making Wiser Choices Online.[3] Building the evidence base through rigorous programme evaluations will therefore be central to supporting the implementation and scaling up of effective digital citizenship education.

Recommendations

We make the following recommendations for government, schools and social media companies, based on our research:

- *The Department for Education should rejuvenate the character agenda within government by funding a third round of character education grants, focusing on developing good character online.* Over 2015 and 2016 the Government made a strong commitment to the character agenda by providing nearly £10 million to support the delivery and evaluation of character education programmes. Since then, the character agenda in government has somewhat faltered. Our findings demonstrate that there are clear links between positive character traits and reduced online risk taking and heightened prosociality. A clear commitment to a third round of funding, focused on online character, would simultaneously meet policy goals around online safeguarding, while also clarifying the Government's overarching vision for education.

- *The Government should put digital citizenship at the heart of the new digital charter, and use its convening power to secure meaningful cross-sectoral collaboration over digital citizenship education.* Through its commitment to a digital charter set out in the Queen's Speech, the Government has taken a welcome and proactive step to developing a stronger regulatory framework and ensuring that key stakeholders (including social media companies) do more. However, the charter's two core objectives remain too narrowly focused on online child protection and functional digital skills. The Government should therefore broaden the scope of the digital charter, addressing the policy gap around the social and civic aspects of digital life. It presents a golden opportunity to convene keystakeholders, and secure robust commitments and cross-sectoral collaboration around digital citizenship education.

- *The Department for Digital, Culture Media and Sport (DCMS) should work with National Citizen Service providers to expand the digital component of the programme, to promote civic virtues and moral thinking online.* Our findings demonstrate

that there are strong links between propensity for civic participation and volunteering, and prosocial behaviour and attuned moral reasoning on social media. There is scope to do more to develop the links between online and offline civic participation through programmes like National Citizen Service. The Government should consider expanding the scope of the programme to address directly the role of digital skills in supporting civic participation and social action, taking inspiration from the current National Citizen Service pilot with the Raspberry Pi Foundation, which focuses on the employability benefits of digital skills. *Educators drawing up digital citizenship education programmes in schools should strongly emphasise the moral implications of online social networking, focusing on participatory approaches that seek to develop students' moral and ethical sensitivity.* Our survey findings show that individuals who don't see a situation as 'moral' are far more likely to engage in negative or passive behaviours; those who do tend to exhibit greater online prosociality. School teachers should therefore dedicate lesson time to developing students' moral sensitivity as part of broader digital citizenship programmes. Our findings also show that boys have significantly lower levels of moral sensitivity than girls in relation to social networking sites scenarios, so school teachers should find better ways to engage boys in discussing and debating these issues.

· *Schools should look to strengthen school–home links around digital citizenship, encouraging parents to close the digital literacy gap and develop effective parental mediation approaches.* While the home setting is crucial for developing good character, a mismatch in knowledge and attitudes towards social media between parents and children, means many feel unable to effectively mediate their children's online behaviour. Schools need to provide support to parents, both to raise their levels of digital literacy, and adopt mediation approaches that engage positively with social media.

- *Facebook and other social media providers should work with youth charities and digital citizenship campaigns to develop effective ways of disseminating information that supports good character online.* As well as regulating content on their sites, social media companies should enable bottom-up approaches to empowering users to challenge negative behaviours, and co-create online codes of conduct. The Online Civil Courage Initiative is a promising model, leveraging the power of Facebook's network to facilitate counter-speech against extremism. Social media providers should therefore consider ways of applying this approach to other aspects of digital life. *Social media providers should use corporate social responsibility budgets to provide financial and technical support for research into interventions that enable more positive youth engagement with social media.* There is a chronic lack of evidence on how young people can be supported effectively to make positive and prosocial choices on social media. Social media providers can support efforts to build this evidence base by providing technical and financial support. For example, under the Online Civil Courage Initiative, Facebook is providing financial support for academic research into effective responses to online radicalisation. This approach could be applied to other contexts.

Introduction

The depth and breadth of youth engagement with social media has increased significantly in recent years, with the proportion of 16–24-year-olds with a social media profile rising from around 50 per cent a decade ago, to close to 100 per cent today.[4] What is more, young people are now more likely to use a wider range of platforms and engage in a greater variety of online activities than in the past, as the functionality of major sites or apps has increased and the monopolistic dominance of Facebook has waned among younger age groups. Between 2013 and 2016 there was a doubling of the proportion of 16–24-year-olds using Instagram, while use of Snapchat rose threefold over the same period.[5] Overall, the latest statistics show that 11–15-year-olds and 16–24-year-olds now spend an average of 2 hours and 2.5 hours per day respectively on social media platforms.[6]

This growing participation in online social networking has intensified more long-standing concerns over the impact of the online world on child and adolescent development. Academics Sandra Leaton Gray and Andy Phippen have argued that growing media and parental concern around young people and social media constitutes a 'moral panic', describing the UK as an 'extreme case study of collective parental anxiety'.[7]

Arguably, part of this growing public concern reflects a reaction against new, disruptive technology, similar to past anxieties about online chat rooms, computer games and, even, the television. The late sociologist Stanley Cohen, who pioneered research into so-called 'moral panics', argued that public fear is particularly acute in relation to new media forms and content, describing 'media panics' as 'spirals of reaction to any new medium [which are] are utterly repetitive and predictable'.[8]

History will tell whether we view this current concern over social media in a similar vein to past anxieties around new technologies. However, it is difficult for parents, school professionals and other stakeholders concerned with child protection and development to turn a blind eye to statistics that point to rising risk and harm on social network sites, such as the reported doubling of Childline calls related to cyberbullying between 2012 and 2016.[9]

The 'moral panic' label may at the same time both obscure and be driven by deeper questions around the novelty of social interaction on social networking sites, the absence of any well-established moral codes, and the role of social media in proving a new space for offline practices traditionally thought of as right or wrong, ethical or unethical. As they engage in social media more deeply young people are now 'routinely presented with moral and ethical choices'[10] – so there is a growing need to understand the kinds of behaviour exhibited online, and the technological and sociological factors that influence how young people act.

The policy context: youth engagement in the digital world

Government policy making around young people's digital lives has developed into a broadly two-pronged approach, focusing on online safeguarding and digital skills. In the former case, in recent years there has been a raft of new policy documents concerning safeguarding in different contexts, for example the Prevent duty (2015),[11] DfE's 'Keeping children safe in education' (2016),[12] and Ofsted's 'Inspecting safeguarding' (2016).[13] All include statutory guidance on online child protection, primarily focusing on efforts to limit exposure to harmful online content, including through 'filtering and monitoring' software. Plans for a new digital charter announced in the 2017 Queen's Speech suggest that this emphasis on online safeguarding will continue, as the charter aims to make the UK 'the safest place in the world to be online'.[14]

Priorities for a second aspect of government policy – improving digital skills – were set out as part of the wider UK Digital Strategy, published in 2016.[15] The strategy aims to give 'everyone access to the digital skills they need', with a particular focus on upskilling the population to participate in the growing digital economy. Building on the 2014 introduction of computing into the National Curriculum, the strategy set out plans to embed computer science teaching in schools. This includes teaching coding in primary and secondary schools, an extra-curricular cyber schools programme for 14–18-year-olds, and efforts to improve teacher expertise in computer science.

However, despite this relatively intensive period of policy development, a 2017 report by the Children's Commissioner, *Growing Up Digital*, argued that 'at the moment, children are not being equipped with adequate skills to negotiate their lives online'.[16] The report pointed to the gap in policy making and education related to 'soft skills' and 'the "social" elements of life online', and called for a greater focus on supporting young people to be resilient, informed and empowered in the digital world. While the Government has therefore been relatively proactive in developing policy on extrinsic risks to children online, and is beginning to promote greater functional digital skills through education, there has been a relative silence on the social aspects of young people's online lives, although this is often the main concern of parents and the wider public.

The social and developmental context: the pull of social media

It is well known that young people have been the earliest and most prolific adopters of online social networking. The latest statistics from Ofcom show that 11–24-year-olds spend on average more than double the amount of time on social networking sites than any other age group.[17] This relatively heavy use of social media is considered unproblematic, assumed as the natural state of the 'digital native' generation.[18] However, studies have begun to show evidence of clear sociocultural and biological reasons to explain the pull

of social media on this age group. From a sociological perspective, authors such as danah boyd have illuminated the importance of social networking sites as a 'whole new realm for youth', one that is cherished as a largely unregulated space of independence in societies that heavily regulate other aspects of young people's lives.[19] Evidence from neuroscience has also begun to link the heightened sensitivity of social cognition in adolescents to a greater drive to 'connect with others and manage their reputation' over social media.[20]

These tendencies have strong links to key aspects of broader adolescent development. Famously described as a time of 'storm and stress', adolescence has been shown to be a stage in life characterised by heightened vulnerability to risky or harmful behaviour.[21] This tendency is driven by neurological and psychosocial changes that can lead to reduced self-control, increased risk-taking, sharp emotional fluctuations, and crucially (as far as social media is concerned) heightened sensitivity to peer dynamics and influence.[22] It is also, fundamentally, a period of rapidly increasing mental and physical capabilities, and therefore should equally be seen as one of opportunity for healthy development.[23]

American psychologist Lawrence Kohlberg famously described adolescence as a key period in the 'moral development' of individuals.[24] For Kohlberg, adolescence signified a transition from the moral world of childhood driven by self-interest, and questions of punishment and reward, to one in which consideration of known others (close relations and friends) and subsequently unknown others (universal ethical principles) became increasingly important in shaping behaviour. More recent neurobiological studies have shown that adolescence is a particularly sensitive developmental phase in social and cultural processing. While this can increase risk taking, it can also support greater propensity towards prosocial behaviour, such as helping others.[25]

Young people are therefore in the unique position of being most socioculturally and biologically drawn to social media, at a time when they are also most open to developmental influence. It has therefore been argued

that 'internet use and social networking sites dovetail with, facilitate, and perhaps intensify the tasks of adolescent psychosocial development' – increasing space for risk taking, but also providing opportunities to build social understanding and experience.[26]

Our approach – character in the online world

For this study we are interested in investigating young people's *agency* in their interaction with online social networks. While the evidence base on young people's online lives has grown significantly in recent years, many of those involved in obtaining this evidence concentrated on understanding the relative prevalence of different experiences, and exposure to different types of content (eg, young people as the victims of cyberbullying, or at risk from exposure to harmful content). However, when looking at the intersection of adolescent development and risk taking, understanding the choices that young people make is a key component of drawing up effective policy, as explained in a recent evidence review on youth risk:

When adolescent choices and behaviours, driven by development and adaptations, play a part in risk, they cannot then be ignored in attempts to protect and prevent. Harnessing and working with adolescent choices and behaviours is essential to them keeping safe during this life stage.[27]

This focus is not to deny the many extrinsic factors that have serious implications for young people's safety and broader development, or to pathologise adolescent online behaviour. Instead, by limiting the scope of our analysis we want to gain a deeper understanding of those internal capabilities that can help young people to manage risk, embrace opportunities, and generally make healthy choices on social networking sites.

Our frame of analysis draws on Demos' previous work into young people's character – the personal traits, values and skills that guide individual conduct, and ultimately shape outcomes in school and in wider life.[28] While character

has significant overlaps with other theories of educational development, such as social and emotional learning or non-cognitive skills, it tends to place a greater emphasis on moral and civic considerations (see table 1 for summaries of key character traits by two leading organisations). And while these aspects of character are important across youth development broadly, we feel that they are particularly pertinent to online social networking. The risks and opportunities associated with social networking sites have the potential to increase significantly young people's exposure to moral dilemmas at a stage when, following Kohlberg, their moral compass is most susceptible to change.

Table 1 **Summaries of the key components of character set out by the Jubilee Centre and the Department for Education**

Jubilee Centre for Character & Virtues[29]	Department for Education[30]
· *moral virtues* such as courage, honesty, humility, empathy and gratitude	· perseverance, resilience and grit
· *intellectual virtues* such as curiosity and critical thinking	· confidence and optimism
· *performance virtues* such as resilience, application and self-regulation	· motivation, drive and ambition
· *civic virtues* such as acts of service and volunteering	· neighbourliness and community spirit
	· tolerance and respect
	· honesty, integrity and dignity
	· conscientiousness, curiosity and focus

While developing character has historically been seen as a core purpose of education it has received renewed attention as a result of a growing evidence base on the role of character in supporting educational outcomes,[31] moral development,[32] and physical and mental well-being,[33] as well as later life outcomes.[34] Within Government, Nicky Morgan MP during her time as Secretary of State for Education (2014–2016) gave particular support to character. She presided over the introduction of the Character Education Grants Fund, which awarded nearly £10 million to support in- and out-of-school character education programmes.[35] In addition, in 2015 the Conservative Government made it a key policy goal to expand

the National Citizen Service – the out-of-school programme designed to promote youth social action and soft skills development.[36]

The departure of Morgan in 2016 and the more recent snap general election have cast some doubt over the continued prominence of the character agenda within government, however. And while the Department for Education has continued to support research into character education in schools, there have been no announcements about a third round of funding under the Character Grants Programme.[37] Within this context of uncertainty, this study assesses the role of character in influencing young people's online conduct. In doing so it attempts to establish the extent to which an education built around good online character can address the gap in policy making around supporting 'social' digital skills development, and provide a renewed purpose for the wider character agenda.

Methodology

To explore the behaviour and choices of young people over social media, we used the following primary research methods (see Appendix 2 for fuller description of the research methodology):

· A survey of 668 16–18-year-olds, administered over Facebook, explored youth experiences and responses to a series of ethical, social media scenarios. Our approach was designed to situate respondents within the context of their everyday social media interaction – with the aim of eliciting responses that accurately reflected how they think and act on social networking sites. Using this approach we cannot make claims to representivity, although a rough gender balance was sought and achieved (338 boys, 319 girls, 11 other or prefer not to say).
· We held focus groups with 40 16–18-year-olds in Birmingham and London.
· An expert roundtable brought together policy experts and practitioners from character education, digital citizenship, online child safety and the tech sector.

- Staff at Demos' Centre for the Analysis of Social Media analysed the dynamics and content of Twitter-based trolling attacks. This is the only part of the research that did not focus exclusively on young people (largely owing to technical constraints identifying user age on Twitter). It aims to provide depth insight into a specific type of unethical behaviour on social media, and one which our findings from chapter 1 show is engaged in by a significant minority of young people.

The report

In the following chapters we present the findings of our research. Chapter 1 explores the moral and behavioural landscape of social media, drawing on primary research and the wider literature to understand the prevalence and implications of risky and healthy behaviours. In chapter 2 we investigate how young people's moral reasoning shapes their actions, presenting evidence from our social media scenarios. In chapter 3 we explore the links between moral reasoning and character development, asking how character influences what young people do and think online, and whether their use of social media has any implication for how their character develops. In each of these chapters we include a case study from our Centre for the Analysis of Social Media analysis, each of which sheds light on online behaviour, moral reasoning and character development.

In chapter 4 we assess the responsibilities of school professionals, parents and managers of social media companies in equipping young people with the knowledge and skills to make healthy choices on social networking sites. In the concluding chapter we present our recommendations.

1 The moral and behavioural landscape of social media

When investigating how young people act and think on social networking sites, there's a danger of being immediately drawn to examples of risk or harm. However, it must be remembered that when young people are asked about their social media and wider online experiences most tend to express broadly positive views. In the 2014 Net Children Go Mobile study (part of the project EU Kids Online), for example, only 2 per cent of British 9–16-year-olds disagreed with the statement 'There are lots of things on the internet that are good for children of my age'.[38] Looking at social media specifically, a study by the NSPCC on experiences of 11–16-year-olds found that 72 per cent of those surveyed reported no upsetting experiences on social networking sites over the previous 12 months.[39]

Evidence from our focus groups and other studies shows that when young people are asked in general terms about what they don't like on social media, they rarely volunteer examples of cyberbullying or harmful content, and instead tend to complain about problems such as advertising and unwanted pop-ups.[40] The day to day experience of social media may therefore be significantly more mundane than press reports can indicate.

This is not to deny the potential harms that social media can create or enhance. For example, a recent evidence review into social media and mental health by the Education Policy Institute highlighted findings that some 37 per cent of British 15-year-olds are now classified as 'extreme internet users' (spending six or more hours on the internet outside school time on a weekday), a figure significantly above the

OECD average (26 per cent). As well as being problematic in itself, this amount of time on screen or social networking sites is likely to increase the chance of exposure to risk or harm online.

As the evidence base has broadened and deepened, most studies have tended to characterise young people's online lives as made up by a constellation of risks and opportunities (see table 2 for an overview of the prevalence of risks and opportunities for young people from social media). Many of these studies have also made it clear that these risks and opportunities are not distributed evenly, with some individuals being more vulnerable than others;[41] and that not all risks and opportunities have the same level of impact (eg, while cyberbullying is less commonly reported than exposure to pornographic material, its effects are far more upsetting for the victims[42]).

Table 2 **Evidence of the prevalence of risks and opportunities for young people from social media and wider online use**

Experience	Prevalence	Perspective	Age	Source
Risks				
Cyberbullying or trolling	69% had experienced some form of cyberbullying	Recipient	12–20	Ditch the Label (2013)[43]
	27% had 'ever said something nasty to someone online'	Actor	12–20	Ditch the Label (2016)[44]
	12% had experienced 'hurtful or nasty' comments or actions online in the past 12 months	Recipient	9–16	Net Children Go Mobile (2014)[45]
	10% had been upset by trolling on social media in the past 12 months[46]	Recipient	11–16	NSPCC (2014)[47]

Experience	Prevalence	Perspective	Age	Source
Risks				
Sexual content	4% had received sexual messages in the past 12 months	Recipient	9–16	Net Children Go Mobile (2014)[48]
	7% had seen sexual images on a social networking site in past 12 months	Recipient	9–16	Net Children Go Mobile (2014)[49]
	14% had taken naked and/or semi-naked images of themselves, and 7% had gone on to share the image(s)	Actor	11–16	Martellozzo et al (2017)[50]
Radicalisation and hate speech	34% had seen hate speech online in the past 12 months	Recipient	12–15	Ofcom (2016)[51]
	23% had seen 'hate messages that attack certain groups or individuals' in the past 12 months	Recipient	9–16	Net Children Go Mobile (2014)[52]
Other harmful content	17% had seen content that discussed 'ways of physically harming or hurting themselves' in the past 12 months	Recipient	9–16	Net Children Go Mobile (2014)[53]
	14% had seen content promoting anorexia	Recipient	9–16	Net Children Go Mobile (2014)[54]
Opportunities				
Civic participation	30% had 'signed petitions, shared news stories on social media, written comments or talked online about the news in past 12 months'	Actor	12–15	Ofcom (2016)[55]
	28% 'encourage others to take action on political or social issues' over Facebook	Actor	11–21	Vromen et al (2016)[56]

Experience	Prevalence	Perspective	Age	Source
Opportunities				
Self-identity and support networks	32% said it is 'easier to be myself on the internet'	Actor	9–16	Net Children Go Mobile (2014)[57]
	68% have had people on social media 'supporting them through tough or challenging times'*	Recipient	13–17	Pew Research Center (2015)[58]
Friendships	33% said 'being able to talk to my friends' is the thing they like most about social networking sites	Actor	11–16	NSPCC (2014)[59]
	70% felt 'better connected to their friends' feelings' through social media*	Recipient	13–17	Pew Research Center (2015)[60]

*US figures

As noted in the introduction, most studies tend to assess risk and opportunity from the largely passive perspective of young people's exposure to positive or negative content. This chapter looks to use the risk and opportunity framework in investigating youth behaviours on social networking sites, exploring the extent to which young people are engaging in both risky or unethical activities, as well as more prosocial online interaction. To do this we present findings from our survey of 16–18-year-old Facebook users, administered through the site, and draw on the wider literature. We conclude by going beyond the risk–opportunity binary in a discussion of the scope for prosocial risk taking on social networking sites.

Risky and unethical behaviour

Multiple studies have explored the ways in which social media, or the online world more generally, may enhance risk taking or unethical behaviour, the most well known of which is American psychologist John Suler's 'online disinhibition

effect', a study of why online 'some people… act out more frequently or intensely than they would in person'.[61] Suler argued that six key features of online interaction enabled greater disinhibition, these include opportunities for anonymity, the lack of facial cues, and the asynchronicity of online communication, enabling greater psychological distance from a situation.

While some of these traits may be less significant in a social media context (eg, reduced opportunity for anonymity, and reduced distance or delay in communication), many still apply. The nature of social networking sites may further enhance inhibition or risk taking, such as the ease of participation (eg, liking and sharing), constant access through smartphones, and the intensification of group dynamics on these sites compared with other forms of online and offline communication.

Ofcom's latest statistics show that 18 per cent of 12–15-year-olds engaged in some form of risky online behaviour in 2016 (up from 15 per cent in 2015).[62] However, Ofcom's definition of 'risky' is relatively narrow and predominantly includes activities that involve interacting with a stranger, or sending a photo or video that you later regret. When looking at the breadth of risky experiences set out more widely in the literature, it is clear that social media provides an opportunity to engage in a far greater range of potentially harmful or unethical practices. We investigate cyberbullying and trolling below, as two sides of online abusive behaviour, one focused at a local level, the other targeting high profile public figures or events.

Cyberbullying

Cyberbullying is a clear example of unethical behaviour. Research has shown that similar to offline bullies, those involved in cyberbullying often 'report a range of social and emotional difficulties', and are likely to exhibit other antisocial behaviours and conduct disorders.[63]

Our survey finds that this kind of behaviour is relatively prevalent among 16–18-year-olds, with just over a quarter of those polled (26 per cent) saying that they had 'bullied

or insulted someone else' over social media.[64] When analysing these responses by gender, we find that boys were substantially more likely to report engaging in bullying behaviour than girls (32 per cent compared with 22 per cent) (figure 1). We also compared the incidence of bullying behaviour by the level of social media use, finding that frequent users of social networking sites[65] were much more likely to exhibit bullying behaviour than lower intensity users (32 per cent compared with 17 per cent).

Figure 1 **The prevalence of bullying and trolling on social networking sites, by gender and social networking site usage**
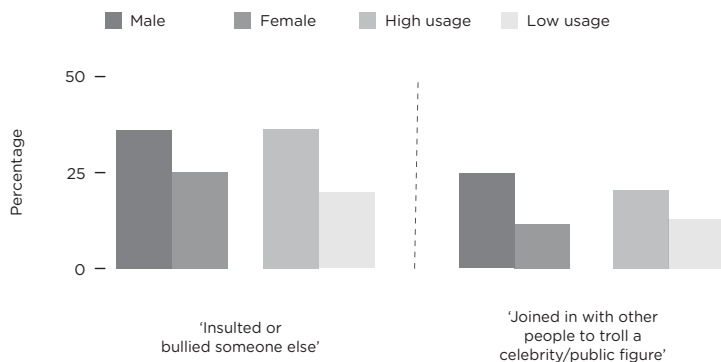
Our survey and focus group findings also suggest why young people may be drawn into these kinds of behaviour. A strikingly high proportion – 93 per cent – of those who said they had insulted or bullied someone else said that they had themselves experienced some form of cyberbullying or abuse. This is significantly higher than respondents who said they had not bullied or insulted someone else (54 per cent of whom said they had experienced online abuse).

Focus group participants discussed a broad range of reasons why people might be drawn to abusive behaviour online. Much of this reflected factors identified in the literature, such as anonymity and asynchronicity (often

expressed as 'hiding behind a screen') or the lack of visual or tonal cues (referred to by one participant as being 'blindfolded'). One of the most commonly cited factors, was the wide visibility of online interaction, intensifying group dynamics, so in certain situations individuals felt compelled to respond in an aggressive way:

If it's in a group chat you feel bad because everyone is watching but you don't want to reply, you don't want to engage with the argument, but you feel like drawn into the argument.
    Boy, Birmingham

You want to make them look like the dickhead of the group and all that because you've got six other people looking at you and laughing... Yeah, you'd be more like 'I want to make this guy look like a twat' instead of 'I want to resolve this and get past this'.
    Boy, Birmingham

Trolling

Another aspect of online abuse that has received increased public attention in recent years is trolling. The term 'internet troll' originally emerged to describe a particular online subculture – 'the self-styled pranksters of the internet', who looked for creative ways of tripping up or provoking (deserving or undeserving) victims.[66] The term has now morphed, into a catch-all word that can be used to define any activity that has caused offence online (and increasingly offline as well). While the evidence base on trolling is limited, research by Canadian academics in 2014 found that regular engagement in trolling[67] correlates with negative personality traits, including sadism and the so-called 'dark triad' of personality – narcissism, psychopathy and Machiavellianism.[68]

In our survey we defined trolling as collective (albeit not necessarily coordinated) public attacks on an individual over social media. We found that some 15 per cent of 16–18-year-olds said that they had 'joined in with other people to "troll" a celebrity or public figure' (see figure 1). As with cyberbullying, boys (22 per cent compared with girls

10 per cent) and higher social networking site users were significantly more likely to engage in this behaviour (18 per cent compared with 11 per cent for lower intensity users).

Perhaps unsurprisingly, people who say they engaged in trolling are also significantly more likely to say they have bullied or insulted someone online (43 per cent, compared with 23 per cent of non-trolls). However, this still leaves 47 per cent of those who have trolled a public figure (plus 10 per cent 'don't knows') who say they have not bullied or insulted someone online. This suggests that many young people find bullying and insulting actions are quite different behaviours from trolling.

Our focus group participants had a variety of personal definitions of trolling – ranging from 'having a little joke' to carrying out 'personal attacks'. Many participants argued that trolling was a rather 'sad' activity, engaged in by people 'people hiding… behind the computer screen'. In three of the four focus groups there was discussion of public figures who had been victims of attacks, but were also considered to be trolls themselves (Katie Hopkins and Donald Trump were mentioned multiple times). Groups were generally split as to whether attacking these figures was a just thing to do and effective in changing their behaviour:

[Talking about Katie Hopkins] So in a way I think it's good that people get on it and comment about it, and voice their opinions, like, she shouldn't be able to get away with writing racist things. If someone else wrote a racist comment then they'd be in trouble.
Girl, London

[In response to someone saying that they had trolled Donald Trump] It's just giving him what he wants – a retaliation – so like why are you doing it in the first place?
Girl, Birmingham

These debates reveal the contested nature of social media's moral landscape, with clear dividing lines between right and wrong often difficult to establish. Due to the relative novelty of social interaction over social media there is little for young

people to draw on when deciding the best course of action to take.

'Drama' and 'banter'

The contested nature of social media's moral landscape is also seen in another subtly different aspect of risky behaviour online, 'drama'. Sociological research from the USA by danah boyd and others into interaction on youth social networking sites has found that young people are far more likely to report experience of, and engagement with, 'drama' than cyberbullying, the latter of which is largely seen as an adult construct.[69] While 'drama' overlaps with bullying, it can span a broad range of activities from a '*moral evaluation* of other people's behaviour to minor disagreements that escalate'.[70] Crucially, 'drama' is seen as bi-directional, rather than the bully–victim dynamic of cyberbullying. This is not to say that 'drama' cannot be harmful, and the labelling of an event as 'drama' can be used by young people as a coping strategy to try to save face, but young people also often describe 'drama' as fun or entertaining.

The authors of US research into 'drama' argue that it is a highly socioculturally contingent concept and may not translate effectively into other contexts. However, evidence from our focus groups suggests that this tit-for-tat, generally low-level, conflict is relatively commonplace in the social media lives of young people in the UK. While the word 'drama' is used (although more commonly as something to avoid), it is often referred to in different terms, such as 'banter' or 'jokes':

Pretty much like getting a rise out of people as well, like if there's something going on, on social media it'd be controversial. Someone who might just post something on there to like start something... It can be just like banter.
> Girl, London

As with the American 'drama', 'banter' can be seen as either a bit of fun or be used to dismiss or cope with otherwise hurtful comments from the points of view of both actor and

recipient. And while 'drama' does not necessarily lead to harm, it arguably helps normalise behaviour that lacks key character traits such as self-control and empathy.

Opportunities for prosocial behaviour

Many of the same characteristics of social media technology that can encourage risk-taking behaviour, such as anonymity and distance, also present opportunities for youth development. For example there is a substantial evidence base on the potential for identity development, with online communication enabling greater selective disclosure and providing opportunities to broaden or deepen identity and association.[71] Looked at from a moral character perspective, social media can open up space to display moral and civic virtues, facilitating honest and empathetic interaction, deepening interpersonal relationships and enabling civic participation. We consider two aspects of this below.

Friendship

By far the most common reason that young people give for using social media is, unsurprisingly, to interact with friends. In the NSPCC's study of the likes and dislikes of 11–16-year-olds, the three most commonly selected 'likes' all involve friendship-related activity – 'being able to talk to my friends', 'keeping up to date with gossip' and 'all my friends [being] on there'.[72]

Our primary research from our survey of Facebook users finds that young people are commonly engaged in empathetic behaviour in support of friends. The vast majority (88 per cent) say they have given emotional support to a friend over a social networking site. In contrast to the findings on abusive behaviour, there is limited difference in the prevalence of this behaviour across genders and by level of social media use, although girls and higher social networking site users are marginally more likely to provide this sort of support (see figure 2).

A key debate, is the extent to which interaction over social media promotes strong, healthy friendships or simply

incentivises a drive to expand contacts in a shallow 'numbers game'.[73] There is certainly evidence to suggest that a significant proportion of young people feel pressure to increase the size of their social network, and curate their profiles to increase their popularity. A study by the US-based Pew Research Center, for example, found that 39 per cent of teenagers felt pressure to 'post content that will be popular and get lots of comments or likes'.[74]

The same study found that 69 per cent of teenagers felt better connected to their friends' feelings, thanks to social media. Shannon Vallor, philosopher of technology at Santa Clara University, has argued that social networking sites have the potential to promote 'friendships of virtue', which support shared healthy development.[75] In particular, Vallor points to the reciprocal nature of much of the interaction on social networking sites (eg, friend requests, liking and tagging), the high frequency of expressions of empathy in social media posts, and the role of distance in supporting mutual self-disclosure.

Figure 2 **The prevalence of empathetic activity and civic engagement on social networking sites, by gender and site usage**
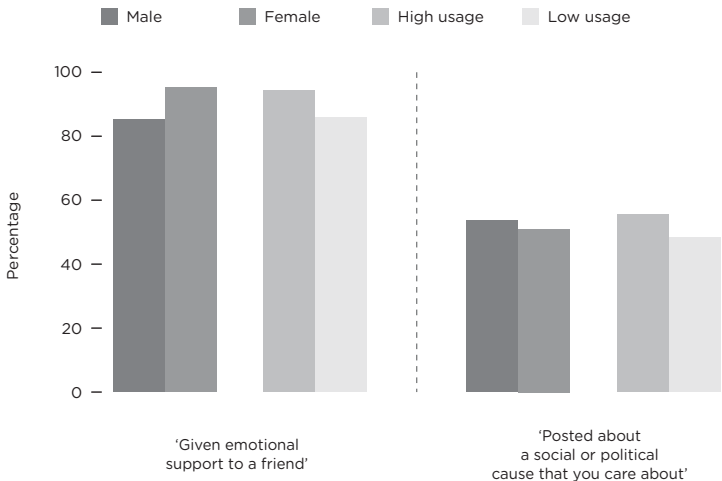
Our focus group participants spoke at length about the nature of friendship on social networking sites, and helped sketch out what a 'good' friend looked like. Most were dismissive of superficial displays of friendship such as unthinkingly liking or commenting on friends' posts – although some revealed the pressure that they felt if they did not receive positive feedback from friends:

If I've seen that she's liked other stuff and not my posts I'm kind of like, oh OK, maybe that's not as nice, she just didn't like my things. And it was only for like two seconds, I'm thinking, oh OK she's liking other stuff, she didn't like mine and we're like best friends, but yeah it's not that deep, but you are aware.
Girl, London

At a more substantive level, many participants thought that being a good friend involved sticking up for your friends if they were being insulted, or helping them to understand how their profile and posts were perceived by others. This latter point came up a number of times, and often involved a friend speaking to them in an offline or more private setting to give advice or support:

I think it depends on what my friend posts, though… If it's an explicit picture then I might message her and say 'you've got to take that down', you know what I mean? Because people could get the wrong idea of her and as a friend I know she's not like that. But it's not in her intention to make them think that.
Girl, London

Thus by increasing the visibility of an individual's behaviour, social media can alert friends to situations that require their support or guidance. In this way social media can support the mutual learning and development that Vallor defines as integral to 'friendships of virtue'.

Civic engagement
Another commonly posited benefit of social media is its capacity to reverse trends of declining civic participation

and political engagement.[76] This has the potential not just to create aggregate societal benefits, but also to provide individual benefits for youth development. In an offline context, the developmental benefits of civic participation have been well documented[77] though, as with friendships, the depth of impact of the online civic participation at a societal and individual level has been questioned. This debate is colloquially framed as 'activism vs slacktivism'. Evgeny Morozov in the *Net Delusion* argued that activities such as signing an online petition or joining an activist page on social media does little to bring about change, and largely serves to enhance personal profiles.[78]

However, this view is increasingly being challenged by a growing recognition of the influence of social media on electoral participation and outcomes. The recent UK general election is a case in point, with the Financial Times reporting that videos from the activist group Momentum reached 13 million unique Facebook users during the election campaign.[79] The impact of social media on political engagement seems particularly significant for young people, with campaigning on social networking sites seen as a major contributory factor for the jump in youth turnout – rising from 43 per cent in 2015 to an estimated 64 per cent in 2017 – reversing years of declining youth engagement.[80]

Our primary research findings show that a significant proportion of young people engage in civic and/or political issues through social networking sites (see figure 2). Just over half (51 per cent) of 16–18-year-olds said that they have 'posted about a political or social cause that they care about'. Again, on this measure there is little difference between genders, and between higher and lower intensity users. This finding is broadly in line with other studies into civic and political participation online, and supports the argument that social networking sites provide 'a low threshold for young people's visible online political engagement, making it easier for them to make statements and interact with others'.[81]

Our focus groups contained a mix of views on the impact of online youth activism in supporting positive change:

When you're speaking about Syria, there's such an importance put on social media… They'll be quick to write a status but that's as far as they'll go to make a change, and I feel like people feel like they've actually done something, whereas before when there wasn't so much social media around, when people did want to make a change they actually like, go and do some fundraising or raise awareness in a more productive way.

Girl, London

But I think political beliefs are a big one. But obviously you get backlash, then you have the chance to actually describe why you believe that and hopefully you convince them to come to your side.

Boy, Birmingham

In comparison with other studies, which have tended to show at best ambivalence towards activism over social networking sites, young people in our focus groups were generally positive about the acceptability of discussing social or political issues on social networks and their capacity to bring about change.

Prosocial risk taking: beyond the risk–opportunity binary

While this analysis of risk and opportunity reflects the weight of evidence on youth social networking sites engagement, it creates a rather Manichean picture of the moral landscape of online social networking, neatly divided between risky and prosocial behaviours. But, looked at from a moral character perspective, are there instances where avoiding risk demonstrates unethical behaviour, or conversely where doing something risky is actually the right thing to do?

Emerging evidence and theoretical work from neuroscience and social psychology on prosocial risk taking suggests a third dimension of online moral behaviour, in which risky and ethical behaviours are intertwined. Do et al define social risk taking as 'the act of engaging in a risky decision with the intention of helping others', and although this field of research has been largely absent from the study of young people's online or social networking experiences,

it adds a vital perspective to the moral dilemmas that may emerge from social networking activity.[82]

Using face-to-face bullying as an example, researchers at the University of North Carolina and University of California have developed a four-way typology, which describes the interaction of prosocial and risk-taking behaviours. When witnessing a bully harass someone, an individual may choose to get involved in the situation (higher risk) or stand back (lower risk), and they may feel empathetic (prosocial) or indifferent towards the victim (antisocial). Combining these elements presents the four main responses to this dilemma – that of the indifferent bystander, the empathetic bystander, the antisocial risk taker (stepping in to worsen the victim's situation) or the prosocial risk taker.[83]

In moral character terms, prosocial risk taking is a clear example of courage, and as with other moral and civic virtues, social media presents a new forum to exhibit this behaviour. For example, while some analyses suggest that the proportion of young people who are victims of persistent cyberbullying is relatively low, the visibility of communication over social networking sites means that many more people are likely to have witnessed the abuse. For example, according to Ofcom, while only 13 per cent of 12–15-year-olds had been bullied in 2016, 35 per cent of this age group knew some who had.[84] With higher visibility of bullying behaviour over social media, there is greater scope for young people to engage with dilemmas that require a courageous or prosocial response.

CASE STUDY 1: Visualising social media's moral landscape – the dynamics of a trolling attack

Our first case study looks at the dynamics of a trolling attack, to provide a window into the moral and behavioural landscape of social media. Clearly by focusing on an overtly negative aspect of online communication – the mass abuse of an individual – this analysis should not be seen as representative of social media communication in general. However, it evidences the contested moral landscape of social

media, with online social networks both facilitating the rapid spread of abuse, while also enabling others to engage in forms of prosocial risk taking, in this case coming to the defence of the victim.

The trolling attack centred on British musician Lily Allen in spring 2017 following a series of tweets from her that aimed to highlight prejudiced tropes by replacing the word 'immigrants' or 'Muslims' with 'pensioners'. Over the next few days, the tweets provoked criticism from other users, much of which was highly abusive. Between 25 February and 3 March 2017, Twitter users sent some 14,000 critical or abusive tweets (classified as trolling) mentioning Lily Allen, either by name or using her Twitter handle @LilyAllen. A substantial, albeit significantly smaller, number of tweets (3,200) were sent supporting Allen in the face of abuse (figure 3).

Figure 3   **Chart showing tweets mentioning Lily Allen between 25 Feb and 3 Mar 2017, coloured by support (white) and trolling (grey)**

Mostly Supportive

Mostly 'Trolling'

'Dogpile' effect linked to single influential account
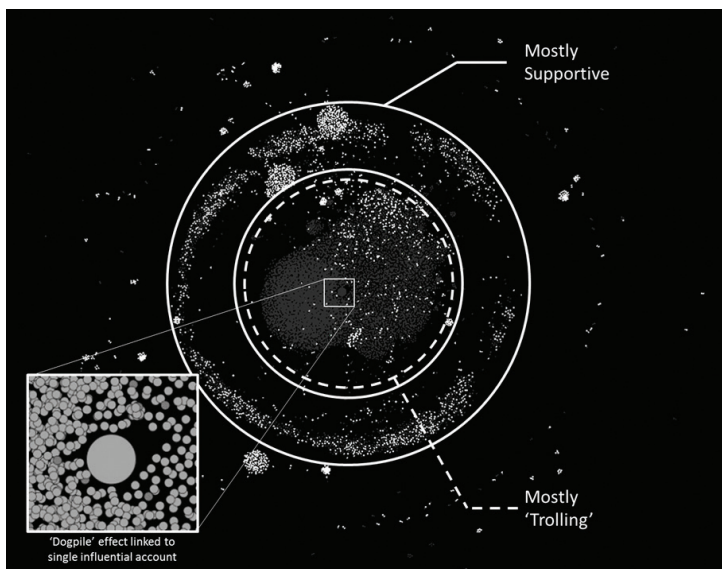
Figure 3 displays the interconnectivity of users who sent messages to Lily Allen during the collection period. Each Twitter account in the collection is shown as a 'node' or point in the chart. Users who mentioned or retweeted each other during the collection period are pulled closer together and appear as tight clusters. It shows the 'dogpile' effect that a single, influential user can have in drawing a large number of other users into attacking someone else. In this case the large red circle at the centre of the chart is Infowars editor Paul Joseph Watson, whose abusive tweet directed at Allen was shared over 3,000 times.

The proximity of users also helps to differentiate between different types of supportive action and the associated risks involved. Large, isolated clusters of blue reflect users who are expressing sympathy and support for Allen, but are in a sense talking among themselves. Supportive and abusive tweets in close proximity illustrate communication between the two sides, such as where supporters actively challenged particular abusive posts or were themselves being trolled as a result of their support.

Overall this analysis demonstrates social media's capacity to draw large numbers of people into displays of unethical or antisocial behaviour, and shows its facilitation of courageous acts through counter-speech – an activity which itself has a significant risk of abuse.

Summary
In this chapter we sketched out the moral and behavioural landscape of social media by assessing the prevalence and implications of risky and prosocial behaviours. We find that as a result of online disinhibition and specific aspects of social media technology – such as ease of participation, constant connectivity, and particularly increased peer group visibility – risky, unethical and potentially harmful behaviour is relatively prevalent among 16–18-year-olds. Just over a quarter (26 per cent) report bullying or insulting others on social networking sites, while a smaller but not insignificant proportion (15 per cent) say that they have trolled a public

figure. And we find that boys and higher intensity social media users are significantly more likely to engage in these negative behaviours than their female or low use counterparts.

However, these same technological characteristics can also open up space for behaviours that display clear moral or civic virtues, such as honest and empathetic communication and new forms of civic participation. For example, some 88 per cent of those surveyed had given emotional support to afriend over social media, while just over half (55 per cent) said that they posted about political or social issues that were important to them. A moral character perspective helps problematise the risk–opportunity binary used in much of the literature on young people's online engagement. By creating greater visibility around online communication, including negative online behaviour, social media presents opportunities for young people to display courage by taking a risk in support of others.

2 Moral and ethical reasoning on social media

As well as investigating *what* young people do over social media, a focus on agency and character also requires an understanding of *why* they might choose one course of action over another. Why, for instance, might one individual decide to join in a WhatsApp group chat mocking someone else in their class, while another stays quiet, and a third leaps to the victim's defence? Similarly, why might someone respond aggressively to a comment disagreeing with their opinion on Twitter, while another may try to engage in a reasoned debate?

To begin to answer questions like these, this chapter aims to explore young people's moral and ethical reasoning online. To do this we draw on primary research using a series of social media scenarios, which ask respondents to think about what they would do in response to an everyday situation that they might encounter on social networking sites. We also ask respondents to give a reason for the action that they decide on. The scenarios aim to uncover the extent to which young people engage in the kind of prosocial risk taking described in chapter 1, and understand their sensitivity to moral or ethical situations in the digital world.

In this chapter we first introduce the three social media scenarios used in the research, before analysing the actions and justifications given by our survey respondents. We conclude by looking at the links between actions and justifications, and drawing on qualitative evidence from our focus groups to better understand the emergent moral and ethical codes that may be influencing young people's behaviour on social networking sites.

Social media scenarios

To gain a deeper insight into the agency, choice and reasoning of young people on social media, we asked respondents to our survey, and our focus group participants, to consider how they would act in the context of three scenarios relating to negative behaviour on social networking sites (figure 4).

Figure 4 **Social media scenarios used in survey and focus groups with 16–18-year-olds**

Senario 1: Mild
You write a post arguing for a cause you believe in. Someone comments on it, aggresively disagreeing with your opinion.

Senario 2: Moderate
One of your classmates writes a post insulting someone else in your class, and tags you in it.

Senario 3: Severe
A friend shares an explicit image of someone else in your class over social media, and asks you to forward it to another friend.

The social media scenarios were designed to reflect three relatively commonplace occurrences on social networking sites, relating to risks identified in the wider literature. Each scenario was classified in accordance with the threat of harm it posed to the victim and respondent, rising from 'mild' for a case of aggressive communication to 'severe' involving the sharing of explicit images of a classmate (these classifications were not referred to in the online questionnaire).

For each scenario, respondents were asked to choose from a list of nine potential *actions*, which were broadly similar across all three cases.[85] Respondents were then asked to choose a *justification* for their decision, this time selecting from 11 potential options (see Appendix 2 for the full list of options, and a detailed description of the methodology).

To conduct a comparative analysis across the three scenarios the action and justification responses were grouped thematically into broader categories. These were the classifications for the action responses:

- *inaction*: responses that signify a lack of action
  (eg, 'I would ignore it')
- *action (negative)*: responses that would increase the potential
  for harm (eg, 'I would insult them back')
- *action (remove self)*: responses where the respondent attempts
  to take themselves out of the situation (eg, 'I would
  'untag' myself')
- *action (positive)*: responses that attempt to resolve the
  situation and minimise harm (eg, 'I would ask them why
  they felt this way')
- *recourse to authority*: responses that try to draw in third parties,
  including parents, teachers, social media providers or police
  (eg, 'I would talk to a teacher or parent about it')

For the justification response classifications, we take
inspiration from Harvard academic Carrie James, whose
research into ethics online focuses on the 'targets' of young
people's thinking – are they primarily concerned with the
implications of their actions on themselves (what James calls
consequence thinking), implications for other people that they
know (what James calls *moral* thinking) or implications for
a wider community of largely unknown others (what James
calls *ethical* thinking).[86] James' typology in some ways reflects
the Kohlbergian analysis of moral development (discussed
in the introduction), with moral thinking gradually expanding
from focusing on the self to considering abstract, universalist
principles. We therefore use James' definitions of 'consequence',
'moral' and 'ethical' thinking when classifying justifications
that refer to the self, or to known or unknown
others (see below).

James' work also draws on American psychologist
James Rest, who emphasised the importance of 'moral
sensitivity' – the capacity to recognise that a situation requires
some kind of moral judgement or reasoning – as a bedrock
of moral and ethical thinking.[87] We classify those responses
that seem not to recognise the situation as requiring moral
judgement as 'not moral' justifications. This is the full list
of justification categories:

- *not moral*: justifications not recognising the situation as requiring moral and ethical judgement (eg, 'It is not a big deal as it's only online')
- *reluctance*: recognition that moral judgement is required, but a reluctance to step in (eg, 'I would be afraid of their reaction towards me')
- *anger*: justifications motivated by anger (eg, 'It would make me angry or upset')
- *consequence*: justifications that focus on consequences for self (eg, 'I wouldn't want to look bad or get into trouble')
- *moral*: justifications that take into consideration people known to the respondent (eg, 'I would want to protect the feelings of the people involved')
- *ethical*: justifications that take into consideration the wider community (eg, 'I would want to prevent them from doing it to other people')
- *advice*: justifications that seek guidance from others (eg, 'I would want to get advice about what to do')

By using coded responses, the picture generated always oversimplifies what people would do, and the often complex and competing motivations that may shape their actions, but our findings suggest that it is possible to make inferences about young people's choices and the reasoning that underpin them at an aggregate level.

How young people respond to social media scenarios

Looking first at what our survey participants say they would do in each of the social media scenarios, we find that there is significant variation across the three scenarios (figure 5). In each case there is a substantial proportion of respondents who say they would do nothing (mild – 24 per cent; moderate – 17 per cent; severe – 17 per cent). Very few – just 11 (or 2 per cent of all respondents) – say that they would do nothing across all three scenarios.

Figure 5   **Types of actions taken by survey respondents in each social media scenario**
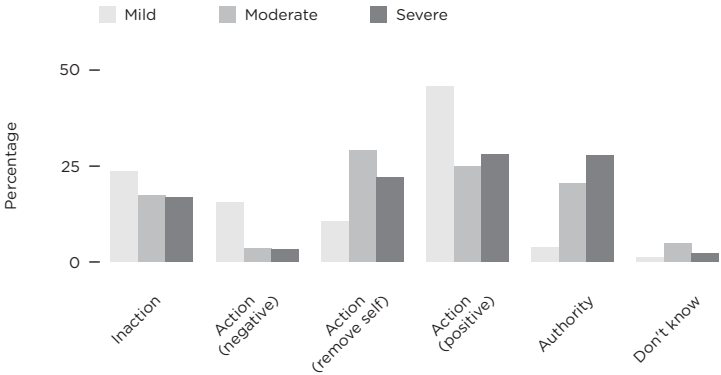
Looking across the rest of the actions taken, it seems that the 'mild' scenario – involving an aggressive response to a post – provokes the most polarised response, with both the highest proportion of positive actions of any scenario (46 per cent) and negative ones (15 per cent). This suggests that in situations on social networking sites that pose less risk of harm young people feel greater freedom to act, although they may not always seek to resolve the situation amicably through the course of action they choose.

Conversely, it seems that respondents are more risk averse in the latter scenarios, being more likely to take themselves out of the situation – by untagging, blocking or unfriending – or to refer the matter to a figure of authority. Responses that relate to authority (telling a parent or teacher, or reporting to the social media provider or police) seem particularly sensitive to the level of severity. While just 4 per cent take this option in the first scenario, 28 per cent chose to do this in the scenario involving explicit image sharing.

Finally, the proportion of respondents saying that they 'don't know' what they would do reaches no higher than 5 per cent across any of the scenarios. This suggests that young people are mostly confident about the decisions they make online.

How actions differ by gender and social networking sites usage
As with the behaviours that we reviewed in chapter 1, our survey data allow us to disaggregate responses by gender and level of social media usage. To make these comparisons we have grouped together all action responses from across all three scenarios. Turning first to gender (figure 6), we find the most commonly chosen responses by young people relate to some form of positive action. Roughly a third of all responses boys and girls chose across the scenarios are those that involve some kind of positive action, but the findings also show that boys are significantly more likely to say that they would do nothing (24 per cent compared with 15 per cent for girls) or to respond in a negative way (10 per cent compared with 5 per cent for girls).

Figure 6   **Types of action taken across all social media scenarios, by gender**
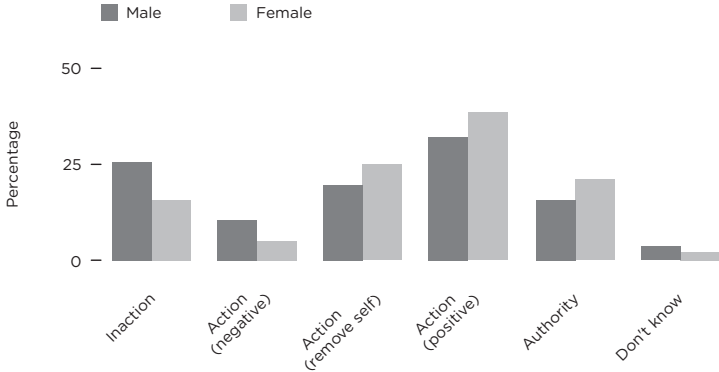
We found that there was little difference in the responses of high and low intensity users of social networking sites for most categories, but higher intensity users are twice as likely to select some kind of negative action (10 per cent, compared with 5 per cent for low users), and lower intensity users are slightly more likely to want to refer the issue to a figure of authority (20 per cent, compared with 16 per cent for high users). While it is difficult to draw causal inferences, this may suggest that a higher level of usage of social networking sites increases young people's confidence in their own decision making, but also serves to normalise negative behaviour for a minority.

Overall our findings on the actions that young people say they would take reveals a greater propensity to take positive actions that map onto the prosocial risk-taking behaviours discussed in chapter 1 than any other domain of response. However, a significant proportion also choose to do nothing, and the relative split between positive and negative actions is partly skewed by gender. Clearly there are caveats about the nature of self-report surveys, with social desirability potentially creating a bias away from negative behaviours, though the findings suggest at least a desire by young people to make positive contributions to their online communities.

Moral sensitivity and reasoning in response to scenarios

While looking at the actions that young people say they would take provides a window into their online decision making, the necessary simplification of responses into inaction, positive actions and negative actions can only ever provide a generalised, and somewhat normative, impression of their actual response in any given situation. Researchers working on the EU Kids Online project note that the reality is that 'online situations are often ambiguous or confusing [and] clear rights and wrongs are difficult to determine'.[88]

To help resolve this problem, and develop a more comprehensive picture of the factors shaping young people's

decision making on social networking sites, we asked survey respondents, to give a reason for their choices. Our focus was less on responses deemed 'right' or 'wrong', or 'risky' or 'healthy', but instead on how the decisions that young people make are shaped by consideration of the implications for themselves and others around them.

As discussed above, our analysis draws on the work of Carrie James, by asking whether our respondents recognise the situation as one requiring moral judgement and, if so, what is the 'target' of their thinking – themselves, people known to them, or the wider community of social media users.

Young people's justifications

First, we find that moral sensitivity is strongly related to the level of severity of the scenario. While nearly 40 per cent of respondents give justifications (eg, classified as 'not moral') that don't recognise the scenario as one requiring moral judgement in the 'mild' scenario, the proportion giving this response drops to just 8 per cent in the 'severe' scenario (figure 7). This gives a strong indication that young people are sensitive to the seriousness of moral situations on social networking sites, and the related levels of risk involved.

Figure 7   **Types of justification given for actions given by survey respondents in each social media scenario**
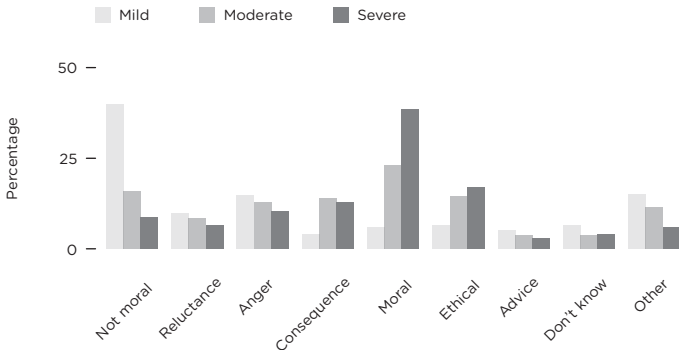
Respondents who recognise that a situation requires some kind of judgement most commonly select justifications classified as 'moral' – focusing on known others (eg, 'I would want to protect the feelings of the people involved') – in both moderate (22 per cent) and severe (36 per cent) scenarios. Again, this demonstrates morally attuned thinking and shows that consideration of known others is generally a much more prevalent driver of behaviour than so-called 'ethical' considerations relating to the broader community (eg, 'I would want to prevent them from doing it to other people').

That said, even when we combine 'moral' and 'ethical' justifications, a majority of respondents only selected these options in the third scenario, and even then the majority was slim (56 per cent). A smaller, but still significant, proportion of respondents across the three scenarios chose options motivated by the implications for themselves, or shallower ones such as being angry or reluctant to engage with the situation.

How justifications differ by gender and use of social networking sites

When we group together all justifications across the scenarios, we again find significant divergence by gender (see figure 8), with smaller differences between higher and lower intensity users. Boys are almost twice as likely as girls not to recognise a scenario as requiring any kind of moral judgement. Just over a quarter (26 per cent) of all justifications selected by boys fall into the 'not moral' category, compared with the equivalent figure for girls of just 14 per cent. Girls on the other hand are more likely to think 'morally' (23 per cent compared with 19 per cent) and 'ethically' (14 per cent compared with 10 per cent). They are also more likely to justify their response as a result of being angry or upset (15 per cent compared with 9 per cent), and this relationship holds across all three scenarios, when analysed individually.

Figure 8   **Types of justifications given for actions in each social media scenario, by gender**
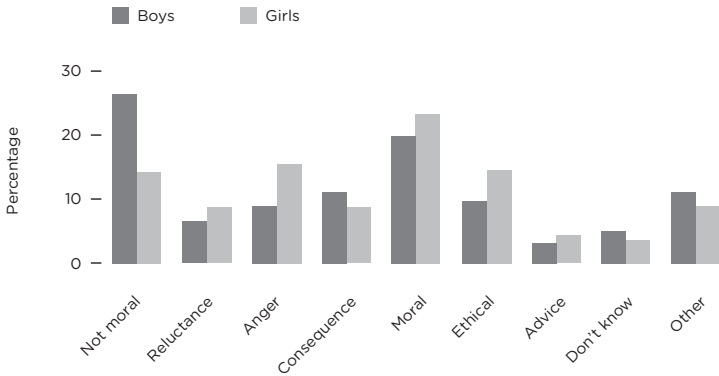
Although divergence in the types of justification given is less pronounced on the basis of intensity of use, higher intensity social media users were slightly more likely to fail to recognise a scenario as requiring any kind of moral judgement than lower intensity users (22 per cent compared with 17 per cent). Conversely, those respondents who use social networking sites comparatively infrequently are more likely to think in 'ethical' terms than more frequent social networking site users (14 per cent compared with 10 per cent). This suggests that there may be a desensitising effect on people from high levels of social media use.

Overall, the findings on young people's justifications – which broadly speaking demonstrate moral sensitivity and a consideration of known others – contrast with other studies, which suggest that self-interest dominates. For example, in a 2012 study Carrie James and Andrea Flores find that 'the most frequent way of thinking about online life was consequence [individualistic] thinking'.[89]

This divergence with the wider literature is partly an issue of classification. For example, respondents most commonly selected an individualistic response that focused on 'me' ('not moral', 'reluctance', 'anger', 'consequence')

as opposed to thinking about others (known or unknown) in both the 'mild' (64 per cent) and the 'moderate' (48 per cent) scenarios, and almost in the 'severe' scenario (35 per cent, compared with 'moral' at 36 per cent).

However, our findings also demonstrate that a significant proportion of young people do exhibit moral judgement that is relatively well attuned to the risk of harm on social networking sites. This supports the broader argument made by researchers behind the iRights campaign on youth digital rights, that 'young people are often very moral in their approach to social interaction in the real world. Therefore, they can be frustrated when the values they try to apply offline don't work online.'[90]

Bringing actions and justifications together: emerging moral codes on social networking sites?

One of the reasons posited for the difficulties of responding to moral dilemmas on social media is an absence of established moral codes or social norms that guide behaviour at a societal level. In the offline world, an individual's actions are at least in part guided by commonly held social norms that 'establish clear expectations of appropriate normative behaviour'.[91] These norms are culturally and historically specific, and generally slow to develop and resistant to change. As social networking sites are relatively new, there has been little time to establish appropriate codes of conduct, and hence less explicit or implicit social guidance on what to do in particular situations.

Research has shown that some social norms have transferred relatively effectively to the digital space. Thanks to the success of internet safety campaigns, most children now say that they are wary of interacting with strangers online.[92] This simplicity of this adult-constructed code of conduct has arguably been key to its success, but fails to deal with the complexities that emerge from peer-based online social interaction.

Recent evidence suggests that peer-constructed codes of conduct on social networking sites may also be emerging (eg, around seeking permission before photos or videos are shared), although this evidence base remains embryonic.[93] By analysing the links between given actions and justifications, and qualitative evidence from our focus groups, we find some degree of peer consensus around certain behaviours, which may constitute emerging codes of conduct that influence how young people respond to moral dilemmas on social networking sites.

Inaction – 'it's not my business'

We find that in the majority of cases (55 per cent) where respondents said they would take no action in a particular situation they took this view because they did not see the scenario as one warranting moral judgement, stating that it 'wasn't a big deal' or 'didn't bother' them. In our focus groups, we also found this to be a common driver of individual inaction, with the emphasis placed on the situation not being their 'problem' or 'business', and therefore not requiring their judgement or involvement:

Because there's no point getting involved with it, it's nothing to do with you.
Boy, London

I'd ignore it, because it has nothing to do with me so… there's no need for me to get involved.
Girl, London

Our survey showed that a smaller proportion of people decided to do nothing for fear of the reaction of the other parties involved ('reluctance' – 11 per cent), or not wanting to look bad ('consequence' – 9 per cent). This again was reflected in the focus groups with a few participants raising the potential personal repercussions of stepping in:

You feel like saying something good or supportive to the person, but you don't want to get involved, or you don't want people to start insulting you. You feel like should I or should I not, and most of the time you just leave it.

Boy, Birmingham

Like the survey, young people in our focus groups were far more likely to take the position of an indifferent bystander ('it's not my business') than an empathetic one (I want to help but am reluctant to) – in justifying non-action. This suggests there is some degree of peer consensus that draws a narrow boundary around the type of events that require young people's engagement or moral judgement.

Negative behaviour – 'you have to act tough'

Where young people said that they would respond in a way that could inflame the situation or heighten risk, this again was largely because they did not see the scenario as one requiring moral judgement (38 per cent) or because the scenario had made them angry or upset (32 per cent). In our focus groups few respondents said that they would react in this way, but many speculated on why someone might – and this too shed some light on the codes of conduct that shaped their behaviour online.

A number of participants spoke about the need for a reciprocal response to the aggressive tone of communication exhibited in the 'mild' scenario. The visibility of communication on social networking sites heightened pressure to respond in this manner, so as not to lose status among friends and followers:

You can't let them say something to you without saying something back, can you?

Boy, Birmingham

You've got an audience so you have to act tough.

Girl, London

A few participants also placed some of the responsibility on the victims of abuse, who they argued could provoke a negative reaction by violating the peer consensus over appropriate and inappropriate behaviour. There was a broad consensus that 'oversharing' was one of the most widely derided behaviours, seen predominantly as 'attention seeking'. This included making too many posts or posting about overly personal subjects:

Yeah some people are rubbish, just spend their whole lives on it, like every hour of the day just chatting shit.
> Boy, Birmingham

A lot of people are saying I'm not trying to attract attention… I'm not doing this for sympathy. But it's kind of like well I have to sympathise with that… So yeah I don't understand why people do it to be honest.
> Girl, London

Prosocial risk taking – 'if it was my mate, then…'

Finally, when looking at the reasons that people gave for stepping in to try to resolve a situation positively, by far the most commonly selected justification was to protect known others. 'Moral' justifications were three times more likely to be used that 'ethical' ones concerning the wider community (32 per cent compared with 11 per cent) (eg, 'I would want to prevent them from doing it to other people'). Protecting people known to you, particularly friends, was also the central reason given in our focus groups for taking action:

I think that if I'd seen a friend in circumstances where someone attacked them verbally, I would like feel bad and back his case. I wouldn't be like oh this boy isn't going to listen to me or it's not going to have any effect. I'm just going to do whatever I feel like doing, because actually if I see my friend being insulted I will feel bad and be like what is this guy doing, and I'm going to back his case.
> Boy, Birmingham

And if it was someone I didn't know… like it's nothing to do with me, but if it's a friend, anything to do with a close friend, then I'm going to let them know.
> Girl, London

Some focus groups revealed a relatively narrow conception of a 'known other', with young people prepared to support close friends or family, but reluctant to act in the interests ofanyone beyond this:

If it was… someone like my sister or a best friend… then I'd obviously, like, tell them about it or report it or something. But if it's somebody that's at my school and it has nothing to do with me then I wouldn't take [any] notice of them.
> Girl, London

A number of participants also suggested they had a relatively high threshold for the kinds of situations that would lead them to act on behalf of a friend:

If they can handle themselves why would you start commenting stuff? It depends on how they're taking it… If someone's there ripping into one of my friends and they're just like, crawling into a ball on the floor, then obviously you're going to be like 'yo!'.
> Boy, Birmingham

These findings depict some degree of normative reasoning behind online behaviour among our survey and focus group participants. Young people generally feel that you *should* act in defence of close friends, but most see anything beyond this as outside their scope of responsibility. And while close friendship may encourage some degree of prosocial risk taking, the dynamics of wider social networks and the visibility of communication may provoke more negative or antisocial behaviour, compelled by a need to be seen to 'act tough' or 'save face'.

CASE STUDY 2: Moral and ethical drivers of 'positive' and 'negative' behaviours

While the above findings demonstrate the influence of moral thinking in promoting positive, prosocial behaviour over social media – our analysis of trolling attacks show that defining what is morally 'right' or 'wrong' can itself be highly contested and as a result can influence a range of different types of behaviour.

This case study again uses data from the Lily Allen trolling attack discussed in case study 1, but this time we focus on a qualitative analysis of a random sample of 100 trolling and 100 supportive tweets. We applied a two-stage coding process to the sample of tweets: the first looked at whether they contained any kind of ethical or principled statement, the second took a thematic approach to grouping similar kinds of statements together. While clearly Twitter provides limited scope for making developed moral arguments, this analysis gives a sense of the motivations of individual users.

Our findings show that a majority (63 per cent) of those users tweeting in defence of Lily Allen justified this with some kind of ethical statement (see table 8 in Appendix 2 for a full list of types of supportive and trolling tweets, with examples). This most commonly involved critiques of abuse, bullying or trolling as morally wrong – often with the focus on the sources of the abuse and/or an appeal to authority (mostly Twitter moderators). A smaller proportion of supportive tweets included direct expressions of compassion towards Allen (13 per cent; many referred to her 2010 miscarriage, which had become a focus for some of her abusers) or a defence Allen's right to free speech (7 per cent). The remainder of supportive tweets expressed general support but without clear ethical arguments.

Over half of the trolling tweets (55 per cent) contained purely abusive comments, with no discernible ethical argument, but this still leaves a perhaps surprising 45 per cent of tweets from people who made some kind of principled statement, generally critiquing Allen on the basis of a 'vice'

attributed to her. The majority focused on what they saw as Allen's disrespect towards pensioners, who were the subject of her original tweets (27 per cent of all trolling tweets). A smaller proportion tweeted about what they saw as either Allen's hypocrisy (13 per cent) or dishonesty (5 per cent). Interestingly, many of these tweets argued that Allen was 'virtue signalling' – expressing support for vulnerable groups, not on the basis of genuine concern, but rather to gain personal credibility.

Our analysis therefore finds that ethical or principled reasoning is a key driver of 'counter-speech' online, and many people who direct personal abuse at others justify their position through ethical arguments. More fundamentally, this demonstrates both the complexities and centrality of moral and ethical reasoning and debates to public conversation on social networking sites.

Summary

This chapter explored the links between action and reasoning in shaping youth behaviour on social networking sites. Broadly speaking, we find that young people are attuned to moral situations over social media, with respondents being more likely to recognise a situation as requiring moral judgement as levels of risk increase. Young people were also more likely to choose prosocial courses of action than any other, suggesting a desire to make positive contributions to their online social networks.

However, we still find a significant proportion doing nothing or failing to recognise a situation as moral across all three of our social media scenarios. Gender seems a particularly significant differentiator of behaviour and judgement, with boys far more likely to be passive and to exhibit low levels of moral sensitivity.

By analysing the links between actions and justification we find that inaction and negative behaviours are most commonly driven by low moral sensitivity, while prosocial actions are most likely to be justified through a consideration of known others. While young people who take positive action

demonstrate relatively attuned 'moral' reasoning (thinking of known others), far fewer think in broader 'ethical' terms about the wider online and offline community. This aligns closely to our qualitative findings, which suggest a peer-led code of conduct is emerging, which fosters action in response to significant harm to close friends, but often engenders the response 'it's not my business' to anything beyond this.

3 Character development and social media

So far we have investigated how young people act on social networking sites and the thought processes that guide what they do. In this chapter we ask how this behaviour and reasoning is shaped, if at all, by the attributes and traits that underpin young people's character, and whether there's any countervailing impact of social networking site usage on character development.

A robust and growing evidence base already exists on the positive influence of character and related social and emotional skills in shaping success in the 'offline world', including supporting educational outcomes,[94] moral development,[95] physical and mental well-being,[96] and even later life outcomes, such as financial stability and reduced crime.[97] A key question is to what extent do these traits also influence behaviour (and as a consequence outcomes) in the digital world? To help answer this we draw on our survey findings, analysing differences in behaviours, and responses to our scenarios, by looking at a series of character traits also captured in the questionnaire.

A second related, but arguably trickier, question is whether the digital world presents a setting for character development. Both Aristotelian and Kohlbergian theories of moral character development argue that positive development occurs 'not simply through adults teaching moral principles, but by young people confronting problems in the world, making sense of them, and making choices'.[98] In this respect social media provides not only a space for engaging in the positive developmental opportunities reviewed in chapter 1, but also a new setting for individuals to encounter morally challenging or ambiguous situations, come to decisions about

what course of action to take, and ultimately build moral experience and awareness.

The cross-sectional nature of this study makes causal claims (and particularly the direction of the causal relationship) difficult to establish. However, in concluding this chapter, we draw on our survey and focus group evidence to make inferences about the extent to which social networking sites use may enhance or undermine character development.

Measuring character

Measuring character has been described as the 'profoundest problem' and 'fraught with difficulties' by academics in the fields of character and education research.[99] Perhaps reflecting the range of perspectives on character (from social and emotional learning to virtue ethics), there exists no commonly recognised and comprehensive measurement tool, and all the measures that have been used have specific 'challenges and limitations'.[100] There are also, clearly, more fundamental epistemological problems with self-reported research methodologies, prone to 'problems of social desirability and self-delusion'.[101]

Nonetheless, numerous studies have been able to demonstrate the validity of certain tools in specific research contexts, such as when investigating particular traits – eg, Angela Duckworth's Grit Scale – or when defining character from particular theoretical and philosophical perspectives – eg, virtue ethicists use of the Values in Action survey.[102]

Given this study's focus on character in the context of youth online behaviour and decision making, we have derived our measure from the Strengths and Difficulties Questionnaire (SDQ), which is widely recognised across academic and practitioner fields as 'an international standardised instrument [for] measuring child behaviour'.[103] We used five items from the full 25-item SDQ, selecting those most closely linked to key moral and civic virtues (table 3).

Table 3 **Character trait measures**

Character trait	Statement
Empathy	'I try to be nice. I care about people's feelings'
Self-control	'I get very angry and often lose my temper'*
Compassion	'I am helpful if someone is hurt'
Honesty	'I am often accused of lying or cheating'*
Civic participation	'I often volunteer to help others'

*These items are reverse scored

For each statement respondents selected whether this was 'certainly true', 'somewhat true' or 'not true' of them. These responses were scored from 0 to 2 – the higher number signifying a greater propensity for a particular trait – and these scores were summed across all traits to give a total 'character score' out of 10. Individuals in the upper and lower 20th percentiles of the distribution of scores were classified as having 'higher' and 'lower' character strengths respectively.[104] These two groups are compared in the analysis below, and assessed for how individual traits might relate to particular behaviours and judgements (see Appendix 2 for a more detailed description of the methodology).

Like many of the methods discussed above, there are clearly limitations to our approach. There were restrictions on the length of our survey (to boost completion rates) so we were not able to use the full SDQ scale, or other longer character scales (eg, the one used in the Values in Action survey). As a result our findings can only reflect a very partial measure of character. More fundamentally, self-report surveys are prone to reliability problems, particularly when measuring socially desirable personal traits or attributes like character. Our use of established self-report scales, and a mix of positively and negatively phrased statements, is an attempt to minimise these problems although they can never be fully eradicated.

Character and social networking sites behaviour

When looking at the propensity to engage in the 'positive' and 'negative' behaviour discussed in chapter 1, we find that there does appear to be some divergence in the behaviour of those with lower and higher character scores (see figure 9). In particular, those with lower character scores are significantly more likely to engage in risky or unethical behaviours. These respondents are more than twice as likely to say that they have bullied or trolled someone (38 per cent bullied and 22 per cent trolled) than those with higher character scores (16 per cent and 10 per cent).

Figure 9 **Type of behaviour of respondents on social networking sites, by character score**
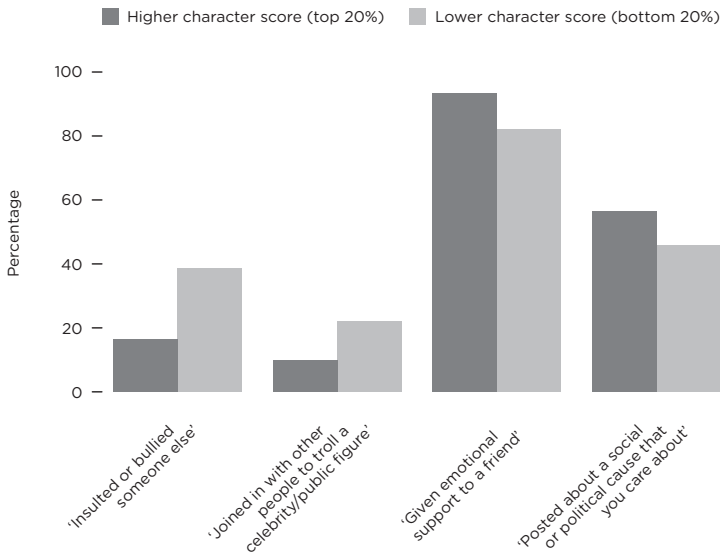
There is less divergence in the behaviour of those with lower and higher character scores when engaging in opportunities to develop friendship and civic participation. However, respondents with higher character scores are slightly more likely to say that they have given support to a friend on social

networking sites or posted about political or social issues
(94 per cent supported a friend and 56 per cent posted about
social issues) than those with lower character scores
(82 per cent and 46 per cent).

We disaggregated the analysis by the individual
character traits measured in our survey (figure 10). To do this
we looked at those individuals who said they have engaged in
a particular type of behaviour (eg, bullying), and calculated
the proportion of this group scoring in the highest category
for each character trait (those scoring 2). We then compared
this to the proportion of individuals with this score for the
sample as a whole, and calculated the percentage point
difference. For example, while 49 per cent of those who say
they have bullied someone score a '2' for empathy, 64 per cent
of the overall sample score a '2' – therefore we find that
individuals who have bullied someone are less likely to score
highly on empathy by 15 percentage points, compared with
the sample as a whole.

Figure 10   **Types of behaviour of respondents on social networking
sites, by character traits**
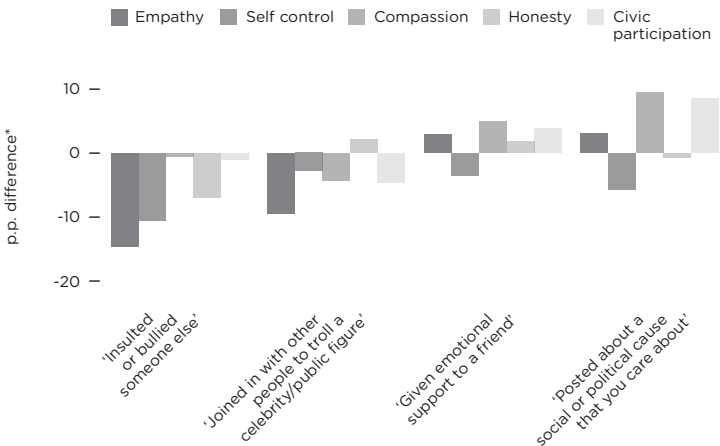
*e.g. 49% of individuals who say they have bullied someone else score highly
on empathy, compared to 64% of the sample as whole (difference of 15
percentage points)

Our analysis finds that individuals who say they have engaged in bullying or trolling also tend to have lower self-reported scores for most of the character traits measured. This trend is particularly pronounced for empathy, with the proportion of 'bullies' and 'trolls' scoring highly on empathy, some 15 and 9 percentages points lower than the sample as a whole. For bullying behaviour, levels of self-control also seem to be important, with the proportion scoring highly on this trait 11 percentage points below the figure for the overall sample.

The pattern is broadly reversed when looking at 'positive' behaviours, with individuals reporting these behaviours having higher than average scores for each of the five character traits. However, divergence in scores is limited when looking at individuals who say they have 'given emotional support to a friend', largely because the vast majority of the entire sample say that they have done this. There is more divergence in the behaviour of those with lower and higher character scores when posting about political or social causes. Individuals who engage in this type of behaviour score particularly highly for compassion and civic participation (9 percentage points above the overall sample). This suggests that respondents who are concerned with helping others and offline volunteering may well show greater political and civic engagement over social media.

Running against this pattern, however, we see that those individuals engaging in prosocial behaviours are slightly less likely to have a high score for self-control. This suggests that lower self-control may lead to a greater propensity to engage in all kinds of online activity, albeit with the impact somewhat skewed towards negative behaviours.

Character and social networking sites reasoning

Our survey findings also enable us to investigate the links between young people's character and their online reasoning and judgement. For this analysis we focus on the justifications that survey respondents gave for their choices when responding to our social media scenarios, discussed in chapter 2. We find that there is substantial divergence in the way those with lower and higher character scores justified their choices (see figure 11).

Individuals scoring high on our character scale were far more likely think in 'moral' terms (in relation to known others) or 'ethical' terms (in relation to a wider community of unknown others) when justifying their response to our scenarios. Some 44 per cent of those with higher character scores said they were motivated by 'moral' or 'ethical' reasons in deciding what to do in our scenarios, compared with just 23 per cent of those with lower character scores.

Figure 11   **Justifications for choices respondents made when responding to social media scenarios, by character score**
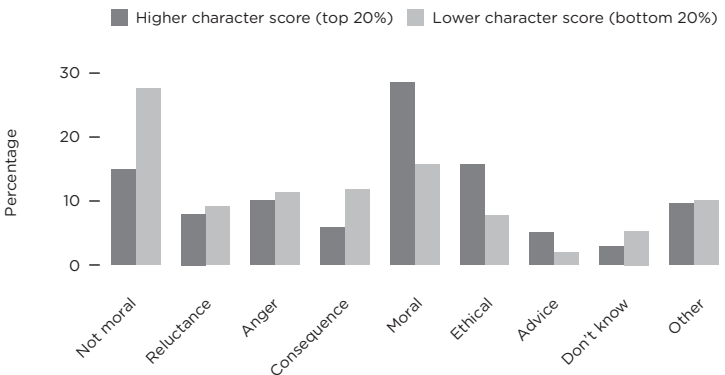
Conversely, the most commonly selected justifications for individuals with low character scores were those classified as 'not moral' (27 per cent) – those responses demonstrating that the respondent did not recognise the scenario as one requiring moral judgement. This group was also twice as likely

to focus on the implications of their actions for themselves, what we have called 'consequence thinking' (eg, 'I wouldn't want to look bad'), than individuals with high character scores. When we disaggregate by individual character traits, we find that 'moral' and 'ethical' thinkers are more likely to score highly across all five measured traits (see figure 12). We see this pattern reversed for those not recognising a situation as moral ('not moral'), with this group less likely to place themselves in the top category for nearly all the traits surveyed.

Differences to the sample as a whole seem particularly pronounced when looking at those individuals who think primarily of themselves, or those who seek advice from others. For example, individuals who justify their actions on the basis of consequences for themselves are 14 percentage points less likely to score in the highest category for 'civic participation'. Conversely, those who want to seek advice about what course of action to take are 16 percentage points more likely to score highly for both compassion and civic participation.

Figure 12 — **Justifications for choices respondents made when responding to social media scenarios, by individual character traits**
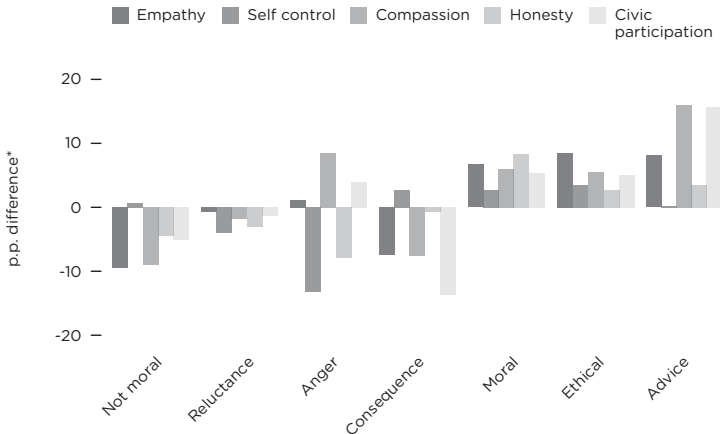
*e.g. 54% of those giving 'not moral' justifications score highly on empathy, compared to 64% of the sample as a whole (difference of -10 percentage points)

While most scenario justifications tend to map onto either higher or lower scores across all five character traits, the findings on 'anger'-driven reasoning present a more complex picture. Young people giving these justifications – such as 'because it made me angry or upset' – unsurprisingly are significantly less likely to score highly for self-control (13 percentage points less than the sample as a whole), but more likely to scoring highly for compassion (9 percentage points). This suggests that individuals who are primarily driven by their emotional reaction to online dilemmas – feeling for the victim of abuse and unable to control the anger that they feel as a result – commonly give an anger-driven response to social media scenarios they encounter.

Character development online?

While the findings discussed above demonstrate the links between character and online behaviour and reasoning, they shed little light on there being any causal relationship between them. Although the tone of much media coverage suggests that young people's use of social media has only a detrimental impact on character development, in a study Vallor postulated that online social interaction also has the potential to enhance character development by providing opportunities to exhibit positive character traits, as well as enabling greater exposure to moral challenges.[105] While the cross-sectional nature of this study makes firm causal claims difficult, some inferences can be drawn from our quantitative and qualitative data.

First, understanding character development as a process of habituation – development through experience of, and learning from, moral challenges – may suggest that greater use of social networking sites could in fact enhance character, as users have greater exposure to online moral dilemmas.[106] Our survey findings do not support this thesis at an aggregate level, however (figure 13). In fact, we find that more frequent users of social media are moderately less likely to have high self-reported character scores (25 per cent) than individuals visiting social networking sites less frequently (31 per cent).

When comparing high and low intensity social networking site users by individual character traits we find little difference in the average scores for most of the surveyed traits. The only trait that differs significantly is self-control, but again it is individuals using social networking sites less frequently who on average have higher levels of reported self-control (10 per cent higher than the whole sample average).

Figure 13   **Total character scores, by social networking site usage**
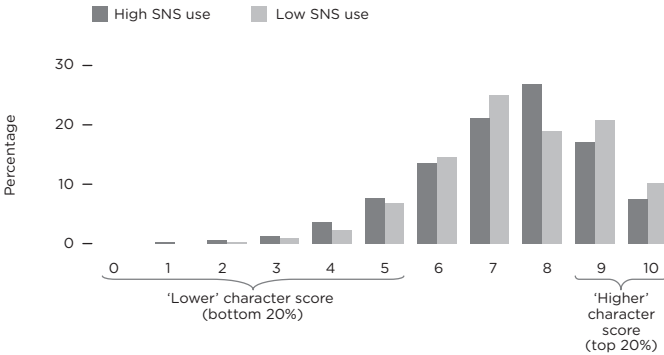
Our survey findings show no evidence for there being any kind of linear relationship between greater use of social networking sites and character development. Instead they suggest there are potential links between high character scores, particularly self-control, and less frequent use, although – again – it is difficult to demonstrate there is a causal relationship between them.

Learning from mistakes

While at a population level there is no clear evidence that social networking sites are a setting for character development, a number of participants in the focus groups suggested they had experienced this. Some reported that positive interactions over social media influenced their attitudes and behaviour. For example, one girl in our London focus group spoke about how the prosocial or empathetic views of others could influence her behaviour:

Definitely, I think social media's shaped a few people. Sometimes if I see one of my favourite YouTubers or, I don't know, like just someone I follow on Instagram, say, I hope everyone has a blessed day or whatever. You kind of feel positive so I'm going to act positive, like I could of woken up in a bad mood, or when someone says that it changes the way I think.

Girl, London

Focus group participants more commonly discussed how negative experiences encountered on social media had changed their outlook and attitudes. For many, this related to their own behaviour, recognising past mistakes in engaging in risky behaviours or responding inappropriately to negative experiences. For some this seemed especially formative, changing perceptions of right and wrong, particularly in relation to online communication:

But you learn from your mistakes then you learn that… if you ignore it, then it's going to keep happening. Like nothing's being solved and ignoring it didn't really work.

Boy, Birmingham

Yeah when I was about 16, 17, I was part of a group chat and we started making fun of people, which I know is wrong, but at the time I didn't really see the problem with it. We all got caught for it, and we all got in trouble… and yeah [it changed the way I act on social media].

Girl, London

While learning from mistakes is clearly not a novel part of growing up, the characteristics of social media technology arguably create more scope for reflection on past behaviour. Both the visibility and relative permanence of interaction on most social networking sites enable individuals easily to revisit previous posts or comments. This was alluded to in some of the focus group discussions:

If every single [one] of us was to look at our posts from two, three years ago we'd shake our heads, the posts are just gonna sound… Cringey… Yeah.

Dialogue between a boy and a girl, Birmingham

The qualitative evidence suggests that some social networking sites can be a forum for behavioural learning and character development, but this learning may often be the outcome of previous engagement in risky or negative behaviour. Our review of the wider literature in chapter 1 found that for a minority this risk may turn to harm, limiting the potential for any developmental benefits to be realised from engaging with social media online. However, social media serves an archiving function for the majority of young people, cataloguing past behaviours and ways of thinking; they can reflect on and use past online experiences when considering alternative courses of action in the future.

CASE STUDY 3: 'Troll communities' – networks of character injury?

Our third trolling case study identified 'communities' of online 'trolls' to understand more about their broader digital lives, and make inferences about the impact of their online worlds on their emotional and character development. In particular, we wanted to establish whether trolling behaviour was in some ways an aberration from the rest of their online activity, or whether these individuals were embedded in networks of unethical practice, which involved an accumulation of risks factors around character injury.

To begin this analysis researchers identified users taking part in a single, very distinctive piece of unethical behaviour: writing fake appeals for help in identifying missing friends orrelatives in the wake of major terrorist attacks. Researchers collected some 327,000 tweets sent by 51 people between 20 and 23 May 2017 related to the Manchester terrorist attack.

Having identified these users, we analysed all the tweets they sent to help understand their interests outside trolling and found five types of conversation where cross-pollination might occur: discussing right-wing ideologies, discussing right-wing politics, sharing comedy and jokes, referring to image boards and referring to gaming. In four of the five cases there was significant overlap. Table 4 shows how many of the 51 people who shared fake images after the Manchester terror attack also tweeted on five subjects identified as being potential areas of cross-pollination.

Table 4 **The number of the 51 people who shared fake images after the Manchester terror attack who tweeted on five other subjects identified as being potential areas of cross-pollination**

Category	No of users	Users (%)
Shared fake images	51	100.0
Discussed right-wing ideological matters	30	58.8
Discussed right-wing political matters	29	56.9
Shared comedy and jokes	29	56.9
Referred to image board content	23	45.1
Referred to gaming content	7	13.7

The ten accounts most frequently retweeted by the group include Donald Trump's account (5,792 retweets), 4chan's account @polNewsForever (5,341 retweets), accounts of right-wing personalities Jack Posobiec (4,118 retweets), Richard Spencer (3,809 retweets) and Paul Joseph Watson (3,501 retweets), and several Twitter accounts sharing right-wing content and comedy (eg @OrwellNGoode, @nontolerantman).

Our findings demonstrate that 'fake image' trolls do show a propensity towards other forms of unethical content, particularly in their engagement with far-right and alt-right political commentators. For example, Richard Spencer, one of the most frequently retweeted users, has been barred from entry into the UK for his white supremacist views. The analysis also shows that these individuals engage in other

seemingly more mundane aspects of online life – internet memes, jokes and cartoons – their interests are often tinged with unethical content. For instance, a widely shared hashtag within the 'image board content' category was #AnimeRight, which hosts content displaying far-right, often Nazi, anime imagery.

Therefore, although these users represent a very small subculture, they demonstrate how social media can enable a voyage through a range of offensive, immoral and possibly harmful content that can come to define the online landscape of some individuals.

Summary

In this chapter we asked to what extent character affects online behaviour, and whether online social networking can support character development. Overall, our analysis of young people's behaviour and reasoning on social networking sites suggests that there is a relationship between character and what they think and do online. Individuals with lower character scores are more likely to engage in risky or unethical behaviours, and they tend to have either reduced sensitivity to online moral dilemmas or a propensity to focus narrowly on implications for themselves. Conversely, those with higher character scores are more likely to make choices that take others into account when confronted with morally challenging situations on social networking sites.

While individual character traits – of empathy, self-control, compassion, honesty and readiness to engage in civic participation – seem broadly aligned in their relationship to specific types of behaviour or thinking, some are potentially more significant than others. For example, individuals exhibiting reduced empathy, and to a lesser extent reduced self-control, seem most prone to negative behaviour, and higher levels of compassion and civic mindedness seem to align closely with prosocial and help-seeking behaviours.

In response to the second question, the study design limits any firm causal claims to be made. However, from the quantitative survey data we find no evidence that greater use of social networking sites correlates with positive character

development. Instead we found that more frequent users of these sites are less likely to have high self-reported character scores. However, our qualitative findings suggest that some young people have formative experiences when using social media, most often related to learning from negative behaviour that they have previously engaged in.

4 Supporting good character online: the role of key stakeholders

Our findings from the previous chapters demonstrate the importance of individual behaviour, moral judgement and character traits in shaping young people's online interactions, an area that policy makers have traditionally either overlooked or struggled to get to grips with. Given high levels of potential harm and public concern around the *extrinsic* risks of young people engaging with social media, policy making has been particularly focused on protection and safeguarding in the digital realm (eg, limiting exposure to harmful content and preventing online child sexual exploitation). In recent years, when *intrinsic* capabilities have been considered, the emphasis has generally been placed on more functional skills and technical know-how, to support the growing digital economy.

Both of these policy imperatives clearly have substantial merit, but the rise of social media – and its open, participatory, often peer-led characteristics – requires an equally developed policy response to help guide young people as they interact with others in the online world. This can't be about invading 'cherished spaces of independence', which – as we found in chapter 1 – are replete with developmental opportunities as well as risks,[107] but we must recognise that a gap in policy risks 'subcontracting out' youth development to peers and social network providers, as one of our expert panel put it.

Part of the reason for this policy gap may relate to the challenges of developing programmes that support individual behaviour change. While online child protection and developing a digitally skilled workforce are huge tasks in themselves, there are at least clear policy levers open

to government in these areas – use of the legal system in the former case, and computer science education in the latter. However, influencing the intrinsic capabilities that shape online behaviour may seem to be an altogether more challenging task.

There is plenty of evidence from the offline world that key character traits and social and emotional skills – those skills that shape behaviour and choice – are malleable and subject to change, however.[108] The Jubilee Centre, for example, argues that good character is 'caught' through role modelling, and peer and adult influence; 'taught' through more direct teaching of knowledge, skills and attributes that support character development; and 'sought' by young people looking to proactively pursue their own character development.[109]

The key challenge is how to apply measures we know are effective in supporting healthy offline behaviour to young people's online social interaction, and then establishing where the responsibility lies in delivering this education and guidance. Our expert roundtable debated these issues and argued for the input of school staff, parents, social media providers and young people themselves. We discuss each of these groups in turn in the chapter below.

Schools

It certainly seems that many schools have been proactive in responding to emerging online risks. For example, over 80 per cent of the young people responding to our survey said that they had received some form of guidance at school (figure 14), and focus groups participants regularly spoke about how their schools had 'drilled in' online safety messages.

However, the extent to which these efforts have been effective in actually influencing behaviour and supporting healthy decision making is difficult to judge. Our survey results did not show any relationship between education and improved outcomes, with the proportion of young people engaging in both positive and negative behaviours almost identical between those who had and had not had school-based training.

Figure 14 **Behaviour on social networking sites, by whether participant had guidance in online safety at school**
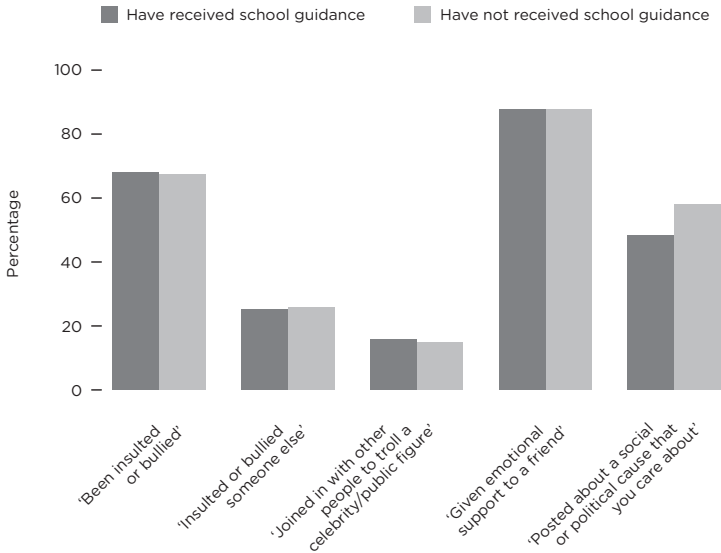
Many of our focus group participants voiced dissatisfaction with or disinterest in the guidance that they had received in school. Some of this criticism was directed at the suitability of teachers in delivering this information:

Yeah because teachers aren't the same age as us, so they don't really understand how to use [social media]. And a lot of kids in school don't listen to their teachers.

Girl, Birmingham

Most people just ignore the teacher. Because they've heard it all before.

Boy, Birmingham

Others criticised the content of the information they received:

It's just constant; block it, ignore it. That's what they teach you, but that's like... Well, it's like teaching a human how to walk, like it's just straightforward.

Boy, Birmingham

Yeah, sometimes they show cheesy videos as well, what happens. Those really tacky videos.
Yeah, it's a bit cringe, yeah.

Dialogue between two girls, London

At our expert roundtable, headteachers and policy professionals argued that while most school leaders and teachers were keen to act, there were significant barriers in place for many to guide young people effectively. Barriers included securing sufficient space within an already crowded and exam-focused curriculum, equipping teachers with the necessary competencies, and overcoming the wider educational and financial inequalities that affect all aspects of learning.

In addition, as a review of US-based review has argued, 'evaluation has not been a priority' for youth internet safety programmes. This dearth of evaluative evidence is apparent on both sides of the Atlantic,[110] and what little evaluative evidence there is has not been overly positive. For example, one of the very few UK-based child internet safety programmes to be independently evaluated, the ThinkUKnow programme, did not support a decrease in risk-taking behaviour.[111] This lack of evidence about what is effective leaves school professionals with little to go on in implementing responses to novel online and social media risks.

Digital citizenship education

One area of promise in helping schools to plug the gap between safeguarding and technical skills, as well as overcoming some of the practical barriers discussed above, is the emergence of externally developed 'digital citizenship' programmes (see table 5 in Appendix 1 for a review of UK-based digital citizenship pilots and programmes).

As well as providing free 'off-the-shelf' materials, limiting the burden on teachers, these programmes attempt to go 'beyond teaching online safety, and [instead] seek to inform and engage pupils in order to give them the skills and dispositions they need to be capable digital citizens'.[112]

In focusing on how 'skills and dispositions' relate to online engagement there are strong overlaps with broader character education. Rather than solely aiming to minimise risk, many of these programmes focus explicitly, or at least implicitly, on supporting good character by promoting 'respectful online behaviour' and 'online civic engagement'.[113]

While evaluative evidence about these programmes is also limited there are reasons to expect that they may be more effective than approaches that focus more directly on safety. Principally this is because many digital citizenship programmes attempt to apply what we know works in wider social and emotional learning and character education to the teaching of online skills and awareness.[114] This includes techniques such as:

- *structured opportunities for personal and group reflection*, eg, Making Wiser Choices Online: use of internet journals to help students record their internet use, and later reflect on challenging situations that they have encountered
- *peer-led learning and mentoring*, eg, Be Strong Online: use of ambassadors in years 9 or above to lead sessions on digital behaviour and awareness to student in years 7 to 9
- *consideration and discussion of moral issues*, eg, Crossing the Line: inclusion of a 'moral compass' component, getting young people to discuss and debate views on statements related to cyberbullying and sexting
- *involvement of parents, guardians and families*, eg, the Digital Citizenship Curriculum: provision of supplementary materials to the main classroom-based programme, to educate parents and families about digital citizenship.

There are therefore reasons to be hopeful about the potential for digital citizenship programmes, and the Children's Commissioner has recommended that digital citizenship education be made compulsory for 4–14-year-olds.[115] The 'lack of evidence around effective practice' remains a major barrier to effective implementation and scale, however.[116] Building the evidence base through rigorous programme evaluations will therefore be central to aiding school staff understand what measures are effective in shaping good character online.

Parents

We know that the home environment is crucial for shaping young people's behaviour and moral judgement. In fact, the weight of neurological, psychological and sociological evidence suggests that 'parents play the principal role in developing... character capabilities in children'.[117] Many parents are particularly concerned about the way young people behave online. Polling by the Jubilee Centre found that a majority of parents (55 per cent) believe that social media is undermining their child's character or moral development.[118]

Parents should therefore have both the means and the motivation to shape how their children think and act on social media, though clearly there are some major complications to this thesis. First, parental intervention in those spaces of youth independence carved out by social media is often fiercely resisted.[119] Second, a 'generation gap in digital literacy' leaves some parents feeling a sense of disempowerment, unable to regulate their child's behaviour on social media in ways that they would do in other areas of their lives.[120] According to the Children's Commissioner, the combination of these two factors can lead to 'a mismatch of knowledge, fears and expectations between parents and their children'[121] – and this was alluded to by a number of participants in our focus groups:

Yeah, maybe if it was of me, then maybe I'd go to my mum, and say you know a picture of me has gone around, but still it would be hard, because my mum will be, like, 'Why did you take that picture in the first place?' I think my mum kind of lacks sympathy for people that do that.

Girl, London

The implications of this mismatch in attitudes may lead parents either to see their child's social media conduct as completely outside their influence, or to take overly intrusive or restrictive parental approaches. Either action may limit the scope for developing good character online. In the former case, this may leave norms of behaviour to be shaped entirely by peer influence. In the latter case, research by danah boyd finds that strong parental regulation often merely encourages young people to find techniques to 'confound' that regulation (eg, through amending privacy settings or creating fake profiles).[122] Furthermore, any regulation that is 'successful' in reducing social media activity and decreasing risk is likely also to limit positive developmental opportunities of engaging with social media.[123]

Research points to the importance of parents mediating actively, but not restrictively, to help their children to engage with social media positively, and to make good choices when using it. This includes activities like talking to children about their social media use and experiences, and engaging in shared activities online.[124] While these approaches are perhaps better suited to younger children, there is evidence that informal but active engagement by parents over social media with teenage children, too, can support prosocial behaviour and healthy development. For example, Sarah Coyne found that online social networking with parents was associated with improved behavioural outcomes, including higher prosociality and lower aggression, for children aged 12 to 17.[125]

This kind of active mediation relies on closing the knowledge and attitude gap between parents and their children, which suggests that efforts by schools to support digital literacy need to be extended to parents. Schools therefore have a role to play in nurturing school–home

links around digital citizenship, enabling mutual and self-reinforcing learning by students and parents.

Social media companies

At our expert roundtable most of the discussion centred on the responsibilities of school teachers and parents in supporting good character online. While this arguably reflects where the bulk of academic, policy and practitioner attention lies, a number of participants argued that this was a skewed focus, which left educators and parents 'picking up the pieces' on behalf of the social media companies. There was also a strong consensus among the young people we spoke to that companies like Facebook, Twitter and Snapchat could be doing more to prevent harm – through either improved user control or provider-driven blocking and filtering:

It really should be down to the people who run Facebook to monitor it... If you're going to make Facebook then you need to make it friendly for everyone. You can't just put it out there.
Boy, Birmingham

I think apps and websites and stuff, there needs to be more user control. So like you need to be able to have the control to be able to stop things from happening on your own account, [but] with apps and stuff you don't really get instructions, you just learn how to use it.
Girl, London

These arguments reflect broader public and political calls for more action on the part of the social media companies, together with demands for greater government regulation of their activities. For example, the digital charter announced in the 2017 Queen's Speech contained pledges to establish an 'effective regulatory framework' and make 'technology companies do more to protect their users and improve safety online'.[126]

These calls for greater action often focus on providers' responsibilities as 'publishers' (a label which many want to avoid), with critics compelling these organisations to do more to take down harmful content either algorithmically or through human moderation. However, the practical and ethical challenges of unilaterally removing content are substantial. First, from a technical perspective, while only algorithms have the capacity to deal with the huge volume of content,[127] only humans (at least for now) are able to make judgement calls about the context and meaning of that content (and even for humans this can be an incredibly difficult and subjective task). Second, the unilateral removal of content is itself ethically problematic, and may not produce the kind of open and pluralistic environment that would enable positive developmental outcomes.

Providers have also attempted to shape conduct from the bottom up. For example, many have introduced or expanded online 'safety centres', providing 'safety education' materials and information on how to filter or report harmful content. Through corporate social responsibility initiatives some providers are also working face to face with local charities to support digital literacy and citizenship. For example, Twitter employees in the USA have developed a workshop series with a San Francisco-based charity to promote 'responsible social networking' among parents and teenagers.[128] While commendable, these initiatives are arguably either too light touch, too small-scale or too narrowly focused on 'safety' to have a significant influence on how the bulk of young people use them.

One project that provides a potential route towards more effective provider intervention is the recently launched Online Civil Courage Initiative (OCCI) – a partnership between Facebook and the anti-extremism non-governmental organisation (NGO) the Institute for Strategic Dialogue. The programme aims to challenge hate speech and radicalisation online, and focuses not on removing extremist content, but on promoting 'the civil courage displayed by organisations and grassroots activists carrying out valuable counter-speech work online'.[129] Rather than focus solely on minimising risk,

the programme aims to enable positive civic action, 'amplifying campaigns which encourage positive dialogue and debate' and challenge extremist attitudes. To do this the OCCI provides support to those engaging in counter-speech, including a dedicated Facebook helpdesk and advertising credits to disseminate campaigns across Facebook's networks. Facebook will also provide financial support for academic research into what is effective in combatting online and offline extremism.

The OCCI therefore provides an alternative model for provider intervention that could be applied to other contexts. Instead of focusing predominantly on 'safety' and the provider-led removal of content, the OCCI approach seeks to empower users, and harnesses the power of online social networks to foster prosocial behaviour and civic action.

Young people

Finally, in supporting young people to make healthy choices online and develop the skills and traits that enable them to do so, it is vital that we develop policy responses that seek youth involvement throughout their design and delivery. Co-productive approaches are needed, as although the label 'digital native' may often mask a lack of functional digital skills, young people are in many ways the 'experts' of their online social worlds. In addition, as our focus groups showed, overly didactic, adult-driven approaches are unlikely to be effective in cutting through and influencing peer-led norms of behaviour.

At our roundtable we heard from teachers and other youth work professionals who argued that young people are often keen to discuss and debate the moral issues arising from online social networking. Those running digital citizenship programmes are harnessing this interest, by using deliberative techniques in the design of course content and delivery methods. For example, Childnet's programme Crossing the Line conducted focus groups with young people to develop the storylines that formed the basis of each module.[130] The programme was piloted across 20 schools to further test and refine the model.

Young people also have a role to play in developing broader policy. In the UK the iRights initiative has pioneered a deliberative approach to policy making in the area of youth digital engagement. The campaign sought the views and input of young people through ten 'youth juries' in which participants 'debated the ethical and practical problems thrown up [by social media] and put forward recommendations'.[131] This culminated in the creation of a framework of digital rights, going beyond protection, and focused on making the digital world an empowering place for young people.

Therefore, while co-productive initiatives such as these can help build internal capabilities and character among those involved, they also have the potential to shape peer culture around norms of behaviour more broadly.

Summary

In this chapter we reviewed what key stakeholders – schools, parents and social media companies – can do to support young people to make healthy choices on social networking sites. While government has provided significant guidance around safeguarding and functional digital skills education, there remains a substantial gap in policy when addressing 'social' digital skills, and those key internal traits that previous chapters have shown shape youth behaviour and decision making.

Digital citizenship is a concept and approach which has the potential to fill this gap by using methods from broader social and emotional learning and character education to engage young people with the civic and ethical implications of social media and wider internet use. However, there is at present a chronic lack of evaluative evidence showing whether current digital citizenship programmes are effective.

A comprehensive, character-focused digital citizenship agenda should bring together all the stakeholders mentioned above. Government can play a central role in supporting the development of an evidence base around what works, and guiding school staff in how to combine broader character goals with digital skills learning. School teachers clearly

should take the lead in delivering much of this work, although they are likely to need support from external providers and practitioners. Parents should be encouraged to engage in their children's digital citizenship learning, helping to boost their own digital literacy, and to develop effective parental mediation techniques. Finally, young people themselves must be at the heart of developing this agenda, which has the potential to shape peer-led codes of conduct to influence online behaviour and civic engagement positively.

5 Conclusion and recommendations

It is understandable to see why a 'discourse of fear' surrounds discussions about young people and social media.[132] From certain perspectives, it can easily seem like the swamp of vice and vitriol that moral panics around social networking sites may imply. Through online disinhibition, easy and near constant access, and the networked visibility of communication, social media in many ways adds fuel to the biologically hardwired tendencies of teenagers to take peer-driven risks. Over a quarter of 16–18-year-olds responding to our survey said that they had bullied or insulted someone over social media, and this arguably underestimates the level of negative interaction, often brushed off as just 'drama' or 'banter'.

However, there is a danger of being technologically determinist in our analysis. While the characteristics of social networking sites can increase risk, they enable behaviours that are just as likely to display virtue as vice. The somewhat paradoxical combination of providing closer connection at greater distance can support a deepening of empathetic relationships, and online networks can provide a gateway into political and civic engagement for young people. While the visibility of group communication may lead some to behave negatively or abusively online, it presents an opportunity for all to display courage – taking a risk to support someone who is being mocked in a group chat, or standing up against abuse on Twitter.

Social media is therefore fundamentally a space to exercise moral judgement and reasoning. Because of the speed, volume and visibility of communication, users are required to confront dilemmas more frequently than in the offline world. Our findings show that young people's moral compass is sensitive to levels of risk, and that many want

to make a positive contribution to their online networks. However, there are countervailing forces, particularly from a growing peer consensus around a very narrow conception of what constitutes 'my business'.

We find that certain character traits seem particularly important in determining the course of action young people take and the reasoning that drives their decision making. Those individuals with lower levels of empathy, and to a lesser extent self-control, seem most prone to negative behaviour. While, those with a propensity to engage in civic activities and volunteering are more likely to exhibit prosocial behaviour and look to parents or teachers for advice.

Character is important not just in thinking about how young people currently think and behave on social media, but also in developing strategies which respond to public and parental concern about their online conduct and safety. Much of the current attention lies in how to develop effective top-down regulation of young people's online interaction and experiences, for example by social media companies removing content, and applying filters and monitoring in school and home. But these adult-led regulatory approaches are in many ways limited in the extent to which they can ensure young people's positive engagement with social networking sites. First, our focus group findings showed that overly intrusive interference in young people's online 'spaces of independence' are likely to be fiercely resisted and may prove counterproductive, encouraging more covert risky behaviour or limiting digital skills development. Second, looked at from a technical perspective it is difficult to see how top-down regulatory approaches can succeed alone, given the sheer volume of content shared daily, the variety and complexity of dilemmas that young people encounter, and the accessibility of social media away from spaces of adult oversight.

Empowering young people by building internal competencies and character traits that help them to manage risk and make positive choices over social media will therefore be key to developing a holistic and effective policy response to the associated risks. At present there is a policy gap around 'social' digital skills. Digital citizenship is a concept and approach

around which to build a character-focused, participatory, online education policy. Here we make recommendations to government, schools and social mediacompanies to help bring this about.

Recommendations

Government
Rejuvenate the character agenda within government through a third round of character education grants, this time focused on developing good character online.

Over 2015 and 2016 the Government made a strong commitment to the character agenda by providing nearly £10 million to support the delivery and evaluation of in- and out-of-school character education programmes. This not only helped deepen the evidence base on effective ways to build character through education, but also gave a clear signal to schools and external providers that the Government was committed to efforts to support broad-based youth development.

Since then, the character agenda in government has arguably faltered. The disruption, caused by ministerial changes and the snap general election, has resulted in lack of clarity on the Government's overarching vision for education. On the one hand, the Department for Education has continued to publish research into character education, including a recent national survey of character education provision in schools.[133] However, the current secretary of state has been less vocal about the importance of character than her predecessor, and there has been no announcement about whether the Character Grants Programme will continue.

The Government has recognised there is growing public concern about the impact of youth engagement in social media on developmental outcomes, but currently schools lack clear guidance in how to respond.[134] Our findings demonstrate that there are clear links between positive character traits and making positive choices online, but there remains a chronic lack of evidence of how good character can be supported successfully online.

The Government should therefore make a clear commitment to a third round of funding through the Character Grants scheme to support programmes that wholly or partially focus on developing good character online. This would give clarity to the Government's overarching vision for education, build the evidence base for effective interventions, and support school staff as they develop an evidence-based response to growing parental concern.

Put digital citizenship at the heart of the new digital charter, and use Government's convening power to secure meaningful cross-sectoral collaboration over digital citizenship education.

Our findings demonstrate the prevalence of both risky and prosocial behaviours among young people over social media, but the current policy agenda does little to shape this moral and behavioural online landscape. Since the 2015 election the Government has made clear it will focus on developing functional digital skills education through its digital strategy, and has issued comprehensive statutory guidance on online child protection and safeguarding. While both of these developments are welcome, they leave a policy gap around the social and civic aspects of digital life.

In the Queen's Speech in June 2017 the Government set out its commitment to a digital charter, a welcome and proactive step to developing a stronger regulatory framework and ensuring that key stakeholders (including the large tech companies) do more to promote a safer and more enabling online environment. However, the charter's two 'core objectives' of making the UK 'the best place to start and run a digital business' and 'the safest place in the world to be online' remain too narrowly tied to functional skills for the digital economy and online safeguarding.

The Government should therefore broaden the scope of the digital charter, addressing the policy gap around the social and civic aspects of digital life. Digital citizenship is a promising and increasingly well-recognised concept around which to build this new policy agenda; it focuses on empowering individuals and developing those internal capabilities and skills that our research finds are crucial to shaping online outcomes.

The digital charter therefore presents a golden opportunity to convene key stakeholders, and secure robust commitments and cross-sectoral collaboration around digital citizenship education. This includes ensuring that managers of social media companies make clear commitments to support this agenda, and focus on empowering younger users to minimise online risk and maximise developmental opportunities.

The Department for Digital, Culture, Media and Sport should work with National Citizen Service providers to expand the digital component of the programme to promote civic virtues and moral thinking online.

One of the strongest and clearest aspects of recent education policy has been the Government's commitment to enabling youth social action – most visibly demonstrated by the expansion of National Citizen Service. Our findings demonstrate there are strong links between a propensity for civic participation and volunteering, and prosocial behaviour and attuned moral reasoning on social media. There is, therefore, scope to do more to develop the links between online and offline civic participation through programmes like National Citizen Service.

Through its digital strategy the Government has already committed to piloting a digital component of National Citizen Service in partnership with the Raspberry Pi Foundation.[135] The pilot aims to equip participants with 'hands-on coding experience' and 'digital entrepreneurship' skills. The Government should consider developing a similar pilot, which focuses on the civic as well as economic significance of digital skills. This could include giving participants greater training on digital campaigning to support their social action projects. More broadly, it should aim to develop civic and ethical awareness in relation to participants' online (as well as offline) lives.

Schools and educators

Deliver digital citizenship education that strongly emphasises the moral implications of online social networking, focusing on participatory approaches that seek to develop students' moral and ethical sensitivity.

Our findings from investigating the choices young people made when responding to our social media scenarios show that the extent to which they recognise online situations as requiring moral judgement is a key determinant in shaping their behaviour. Those who don't see a situation as 'moral' are far more likely to engage in negative or passive behaviours; those who do tend to exhibit greater online prosociality. We also find evidence of there being an implicit code of conduct among many young people, which fosters action when significant harm is inflicted on close friends, but requires very limited ethical consideration beyond this ('it's not my business' being a common refrain).

Schools should therefore dedicate lesson time to developing students' moral sensitivity (the ability to recognise a situation as requiring moral judgement) as part of broader digital citizenship programmes. They should encourage students to think through broader ethical considerations, which extend beyond their close circle of friends, and consider their responsibilities towards the wider online and offline community. Our findings show that boys have significantly lower levels of moral sensitivity than girls when reacting to social networking site scenarios, so educators should think about finding better ways to engage boys in discussing and debating these issues.

Our focus groups and expert workshop demonstrated that young people are keen to engage with the moral and ethical dilemmas thrown up by social media, but resistant to being 'told' how or how not to behave in online spaces that they view fundamentally as their own. Schools should therefore ensure that sessions are participatory and deliberative, rather than didactic. This may require them to review their internal capacity to design and deliver these sessions, and consider securing support from external providers.

Strengthen school–home links around digital citizenship, supporting parents to close the digital literacy gap and develop effective parental mediation approaches.

While the home setting is crucial for developing good character, a mismatch of knowledge and attitudes towards social media between parents and children, means some parents feel unable to effectively mediate their children's online behaviour. Schools need to provide support to parents, both to raise their levels of digital literacy, and develop mediation approaches that support positive engagement with social media. This includes educating parents on the risks of social media, but also warning against an overly restrictive approach to monitoring their children, which may limit opportunities or be counterproductive.

Schools can build these links by providing literature for parents that mirrors the classroom-based digital skills and digital citizenship initiatives. This should include guidance on functional skills related to social media, such as account management and privacy settings, but crucially also provide advice on active mediation strategies, such as talking to children about their experiences and sharing activities online. Ultimately, parents should be supported to build relationships of trust, which respect and value their child's independence, but provide an informed outlet for practical and emotional advice and support.

Social media providers

Facebook and other social media providers should work with youth charities and digital citizenship campaigns to develop effective ways of disseminating information that supports good character online.

While there is significant pressure on social media providers to do more to regulate content shared on their sites, this alone appears unlikely to be effective in creating more positive online environments, because of the sheer scale of the task, and the ethical implications of unilateral content removal. Social media companies therefore also need to consider bottom-up approaches to empower users to challenge negative behaviour and co-create online codes of conduct.

The Online Civic Courage Initiative provides a promising model, leveraging the power of Facebook's network to facilitate counter-speech against extremism. It provides NGOs and activists with technical support and ad credits to disseminate counter-speech messages – premised on the understanding that engagement and debate is more effective than censorship in changing behaviour and attitudes.

While the OCCI focuses predominantly on extremism, it is a model that could be replicated to serve broader digital citizenship goals. Facebook and other social media providers should work with those running youth charities and digital citizenship campaigns to help them to leverage the power of online social networks to promote good character online. For example, they could use ad space to engage young people in moral and civic questions and debates linked to social media use. Social media companies should also use existing analytics to help those running youth charities and digital citizenship campaigners understand the effectiveness of their campaigns. This includes measuring the reach of any campaign, and crucially also linking this information to aggregate-level data on user behaviour (eg, intensity of use, or likelihood of posting or reporting hate speech or other abusive content).

Social media providers should use corporate social responsibility budgets to provide financial and technical support for research into interventions that support more positive youth engagement with social media.

There's a chronic lack of evidence on how young people can be supported effectively to make healthy and prosocial choices on social media. While government should do its part in building an evidence base in this area (we suggest through the character education grants), social media companies also shoulder significant responsibility for ensuring that their users (and particularly children) have the knowledge and skills to use their platforms in ways that, at the very least, don't lead to harm.

Social media companies can support efforts to build this evidence base by undertaking internal research and providing

technical and financial support for independent studies. Internally, social media providers should aim to use their data to help build up a better understanding of the behavioural and moral landscape of their platforms (eg, through the dynamics of reported content or textual analyses of posted content). They should also make at least some of this (anonymised) data available to independent researchers to support wider efforts to understand the developmental impacts of social media, and best practice in promoting good online character.

Under the OCCI, Facebook is also providing financial support for academic research into effective responses to online radicalisation. Social media companies should therefore consider providing similar funding for research and evaluation into digital citizenship and healthy decision making on social networking sites. This would encourage a greater focus on evidence-based practice among developers of digital citizenship programmes, and underscore providers' commitment to curate positive discourse on their platforms.

Appendix 1 Digital citizenship programmes

Table 5   **Examples of digital citizenship and online character education programmes**

Programme	Organisation	Age group	Course aims and content	Course delivery	Moral and civic component(s) of course or programme	Evaluation
Be Strong Online[136]	Diana Award[137]	11–13	To help young people explore the digital world safely, become positive online role models and develop skills in areas such as public speaking and leadership. Ten modules focusing on online risks such as cyberbullying, peer pressure and online privacy, and online opportunities, including creativity and coding.	The programme takes a peer-led approach, with sessions designed to be delivered by 'ambassadors' in year 9 or above. Ambassadors receive in-school training before delivering the course to younger or similar age group students. The Diana Award provides guidance on ambassador training (which is designed for teachers to deliver) and course materials for each module.	The cyberbullying module explores the concept of being a bystander and encourages students to discuss when it is right to take action.	Not evaluated (though the wider offline evaluation of the Ambassadors Programme was published in 2016).[138]

Programme	Organisation	Age group	Course aims and content	Course delivery	Moral and civic component(s) of course or programme	Evaluation
Crossing the Line[139]	Childnet	10–14	To facilitate group discussion on online behaviour and its consequences, focusing on talking through the moral dilemmas that young people can encounter online. Four modules focusing predominantly on examples of online risk: cyberbullying, sexting, peer pressure and self-esteem.	Sessions are designed to be teacher-led, delivered during personal, social and health education (PSHE) time or other suitable periods (eg, tutor time). Childnet provides course materials and lesson plans, including introductory videos setting out a particular online situation. The emphasis for most of the session is on facilitating discussion among and between students.	Includes 'moral compass' worksheets and activities where students debate whether a particular scenario is acceptable or not.	Not evaluated

Programme	Organisation	Age group	Course aims and content	Course delivery	Moral and civic component(s) of course or programme	Evaluation
Digital Citizens pilot[140]	Demos	15	To build participants' resilience to the risks of online exploitation and radicalisation through a skills-based teaching approach. Pilot focused on developing knowledge and critical thinking skills around online propaganda, to support participants to critically analyse online material and know what to do if confronted with hate speech.	Course resources were based on a number of anonymised real-life examples of social media conversations that involved various types of hate speech. External facilitators delivered the training, which had been designed to encourage participant-led dialogue.	The pilot was intended to cultivate an understanding of civic responsibility, with students encouraged to take a lead in identifying and arguing against extremism and hate speech online, and in gaining peer-to-peer support.	Evaluation published in 2016.
Digital Literacy & Citizenship curriculum[141]	South West Grid for Learning (SWGfL)	4–18	To build students' knowledge and skills to act responsibly online and use the internet in a better way. Seeks to empower learners through a 21st-century skills approach, focusing on risks around online safety and security, and opportunities around digital citizenship.	SWGfL has produced schemes of learning for all year groups, from foundation stage to key stage 5, designed for teachers to deliver in a classroom setting. The course materials contain supplementary content to educate parents and families about digital citizenship.	Lessons are designed for each age group on what it means to be a good digital citizen, eg, a year 5 session in which the class designs and signs a 'digital citizens pledge'. Many units deal with ethical issues of digital life. One module, 'My Online Code', facilitates a discussion among students on their understanding of ethical behaviour online.	Not evaluated (though US programme is currently undergoing an evaluation[142]).

Programme	Organisation	Age group	Course aims and content	Course delivery	Moral and civic component(s) of course or programme	Evaluation
Making Wiser Choices Online pilot[143]	Jubilee Centre		To develop participants' character and 'cyber-phronesis' – the ability to make wise decisions when using the internet. Helps students understand the importance of character virtues when online, while also increasing their knowledge about moral issues relating to online risks, such as including piracy, plagiarism, bullying.	The four-week, teacher-led programme is designed to be delivered in computer science and/ or PSHE lessons. The programme is taught largely through discussion and reflection, and considering internet-based moral dilemmas. As part of the programme students keep 'reflective journals' to record their daily online experiences.	Develops students' capacities to make moral judgements, such as recognising the moral salience of a situation and making wise judgements about how and when to be honest, loyal and compassionate. To do this the programme is structured around a series of moral dilemmas, which students can reflect on and discuss collectively.	Undergoing programme evaluation.

Appendix 2
Detailed methodology

This appendix provides further information on the methods of data collection and analysis used in the research.

Survey

Respondents

The surveys were disseminated via Facebook advertising, using targeted ads directed at users aged 16 to 18 across the UK. The ads provided links to an online survey hosted by SurveyMonkey. Each respondent was placed into a prize draw to win £100, with a winner selected at random following the closure of the survey. After an initial two-week pilot, the survey ran for seven weeks between 27 March and 12 May 2017.

In total, we received 867 responses. However, a number of responses were removed because they were incomplete or provided clearly unconsidered responses (eg, the respondent had clicked the first option for all questions or given answers to open-response questions that were unrelated to the question). We removed 23 per cent of all responses for these reasons, leaving a total sample size of 668. The gender breakdown of our sample is 338 boys (51 per cent of all respondents), 319 girls (48 per cent) and 11 'other or prefer not to say' (2 per cent).[144]

Scenarios

Our scenario questions formed part of the online survey, and were designed to gain a deeper insight into young people's decision making and moral sensitivity on social media. The three scenarios (table 6) were designed to reflect relatively commonplace situations on social networking sites.

Each scenario was classified in accordance with the threat of harm it posed to the victim and respondent, raising from 'mild' to 'moderate' to 'severe'. This classification was not disclosed to survey respondents.

For each scenario, respondents were asked to choose from a list of nine potential *actions*, which were broadly similar in all three cases. Respondents were then asked to choose a *justification* for their decision, this time selecting from 11 potential options. For the purposes of comparative analysis these responses were then grouped thematically into five action categories and seven justification categories, as set out in table 6.

Table 6 **Action and justification thematic coding for three social media scenarios**

Scenarios	Mild: 'You write a post arguing for a cause you believe in. Someone comments on it, aggressively disagreeing with your opinion.'	Moderate: 'One of your classmates writes a post insulting someone else in your class, and tags you in it.'	Severe: 'A friend shares an explicit image of someone else in your class over social media, and asks you to forward it to another friend.'
Actions			
Inaction	I would ignore it	I would ignore it	I would ignore it
Action (negative action)	I would insult them back	I would like or agree with the post	I would forward it to the other friend or share more widely
Action (remove self)	I would block or unfriend them	I would block or unfriend them	I would block or unfriend them
	I would remove the post	I would 'untag' myself	I would delete the image from my account

Scenarios	Mild	Moderate	Severe
Actions			
Action (positive action)	I would ask them to remove their comment	I would ask them to remove the post	I would ask the person who sent it to delete the image
	I would ask them why they felt this way	I would post something supporting the person who was insulted	I would tell the person who was in the image
Authority	I would talk to a teacher or parent about it	I would talk to a teacher or parent about it	I would talk to a teacher or parent about it
	I would report it to the social media platform I was using	I would report it to the social media platform I was using	I would report it to the social media platform I was using
			I would report it to the police
Don't know	I don't know what I'd do	I don't know what I'd do	I don't know what I'd do
Justifications			
Not moral	It's not a big deal as it's only online	It's not a big deal as it's only online	It's not a big deal as it's only online
	It doesn't really bother me	It doesn't really bother me	It doesn't really bother me
Reluctance	I would be afraid of their reaction towards me	I would be afraid of their reaction towards me	I would be afraid of their reaction towards me
	I wouldn't want to speak to them	I wouldn't want to speak to them	I wouldn't want to speak to them
Anger	It would make me angry or upset	It would make me angry or upset	It would make me angry or upset
Consequence	I wouldn't want to look bad or get into trouble	I wouldn't want to look bad or get into trouble	I wouldn't want to look bad or get into trouble

Scenarios	Mild	Moderate	Severe
Justifications			
Moral	I would want to protect the feelings of the people involved	I would want to protect the feelings of the people involved	I would want to protect the feelings of the people involved
Ethical	I would want to prevent them from doing it to other people	I would want to prevent them from doing it to other people	I would want to prevent them from doing it to other people
Advice	I would want to get advice about what to do	I would want to get advice about what to do	I would want to get advice about what to do
Don't know	I don't know why	I don't know why	I don't know why
Other	Other	Other	Other

Character questions

The online survey also contained five items from the Strength and Difficulties Questionnaire (SDQ), as a proxy measure for self-reported character strengths. For each item respondents were asked whether the statement was 'not true', 'somewhat true' or 'certainly true' of them. Each item was coded according to a specific character trait (table 7). Three of the statements were positively phrased (and consequently positively scored), while two were negatively phrased (referred to an absence of a particular trait) and therefore negatively scored.

Table 7   **Character trait measures and scoring**

Character trait	Statement	Response and scoring		
		Not true	Somewhat true	Certainly true
Empathy	'I try to be nice. I care about people's feelings'	0	1	2
Self-control	'I get very angry and often lose my temper'	2	1	0
Compassion	'I am helpful if someone is hurt'	0	1	2
Honesty	'I am often accused of lying or cheating'	2	1	0
Civic participation	'I often volunteer to help others'	0	1	2

We used the responses to the SDQ items to analyse the data by the level of self-reported character strength in two principle ways. First, we summed respondents' scores across all five items to give a total 'character score' out of 10. We classified individuals in the upper and lower 20th percentiles of the distribution of scores as having 'higher' and 'lower' character strengths respectively, and compared the reported behaviour and moral reasoning of these two groups.

Second, we analysed responses by individual character traits – empathy, self-control, compassion, honesty, civic participation – to assess whether specific traits were more closely linked to particular behaviours or forms of reasoning. To do this we looked at those individuals who said they had engaged in a particular type of behaviour (eg, bullying) or reasoning (eg, 'consequence'), and calculated the proportion of this group scoring in the highest category for each character trait (those scoring 2). We then compared this to the proportion of individuals with this score for the sample as a whole, and calculated the percentage point difference. For example, while 49 per cent of those who said they had bullied someone scored a '2' for empathy, 64 per cent of the overall sample scored a '2' – therefore, we found that individuals who had bullied someone were less likely to score highly on empathy, by 15 percentage points to the sample as a whole.

Limitations

The survey employed a number of innovative and experimental techniques, which were designed to gain insights into the thought processes and character traits that influence how young people act on social media. However, moving away from more 'tried and tested' approaches creates a number of novel methodological challenges, as well as those more commonly encountered limitations from self-report surveys:

· By administering the survey through Facebook we aimed to situate respondents within the context of their everyday social media interaction, but we therefore had to minimise survey length to ensure decent completion rates. As a result, we were not able to collect comprehensive demographic information, so we could not weight responses according to population level distributions. Thus the findings are not statistically representative of the UK population.

· Our scenarios were designed to reflect everyday social media situations, and were partially derived from scenarios previously used in the Demos' Digital Citizens evaluation.[145] The scenarios were also inspired by validated moral dilemma tests (such as the Intermediate Concept Measure for Adolescents, Ad-ICM) used in research into moral character and virtue.[146] However, again partly because of survey capacity constraints, our scenarios are significantly shorter in length than validated dilemma scales (eg, have shorter descriptions of situations and fewer response options), which may limit the extent to which they reflect 'real world' situations or responses.

· A further limitation is one more commonly associated with self-report surveys, which can be prone to reliability problems: such as social desirability bias, which may lead to respondents rating themselves higher for socially desirable traits, or answering according to what they think is the 'right' answer to give. This may well have affected our scenario responses in particular, suppressing negative responses and boosting positive actions.

Trolling analysis

The case studies in the report are based on an analysis of 'trolling' attacks and networks by Demos' Centre for the Analysis of Social Media (CASM). The methodology for this analysis is set out below.

Method52

As data drawn from social media are often too large to fully analyse manually, and also not amenable to the conventional social research methods, the CASM team use a technology platform called Method52. It was developed by CASM Consulting LLP and CASM technologists based at the Text Analytics Group at the University of Sussex and is designed to allow non-technical researchers to analyse very large datasets like Twitter.

Method52 allows researchers to train algorithms to split apart ('to classify') tweets into categories, according to the meaning of the tweet, and on the basis of the text they contain. It does this using a technology called natural language processing, a branch of artificial intelligence research, which combines approaches developed in the fields of computer science, applied mathematics and linguistics. Analysts 'marks up' which category they consider a tweet falls into, and this 'teaches' the algorithm to spot patterns in the language use associated with each category chosen. The algorithm looks for statistical correlations between the language used and the categories assigned to determine the extent to which words and bigrams are indicative of the pre-defined categories.

Trolling attack

To understand the dynamics and content of a high profile trolling attack the CASM team collected 17,100 tweets mentioning Lily Allen between 25 February and 3 March. This data were analysed using Method52, with tweets classified into two broad categories: those that criticised her (categorised as trolling) and those that supported her in the face of abuse. Of the 17,100 tweets, 13,900 were classified as critical (81 per cent), and 3,200 were classified as supportive

(19 per cent). The classifier operated at 74 per cent accuracy on this two way split.

To gain a deeper understanding of the content of these tweets, CASM researchers took a random sample of 100 tweets that expressed negative views towards Lily Allen, and 100 tweets that expressed positive views towards her and/or counter-speech against abusive posts. Researchers applied a two-stage coding process to the sample of tweets: the first looked at whether they contained any kind of ethical or principled statement; the second took a thematic approach to grouping similar kinds of statements together. A summary of this qualitative analysis is set out in table 8.

Table 8   **Analysis of the ethical content of sample of trolling and supportive tweets**

Category	Label	Definition	Proportion of sample	Examples
Supportive tweets				
	Abuse and bullying as wrong	Non-specific hostile behaviour from people	43%	@lilyallen @jack @ twitter i demand you do something about this! because it was herendous !!! no violation? how much more abuse does it take?
				why are people being so horrible to @lilyallen? personally attacking her. you're the scum of the internet.
	Compassion	Not talking about certain topics, eg miscarriage	13%	I often disagree with @lilyallen and am sometimes annoyed but that's it. trolling her over the loss of her child is inhuman.
				@canadapatriots @ lilyallen come on, you can not be so cruel and say that, you don't know what it feels like to lose a son, you're disgusting.

Supportive tweets				
	Free speech	Not allowing people to express their views	7%	what fresh hell is this being aimed @lilyallen ? this is not free speech it is hate speech.
				stay strong @lilyallen ! do not let the haters put you down and silence you! your voice matters!
	Other	Supportive comments but without ethical content	37%	
Trolling tweets				
	Disrespect	Failing to treat other people (eg, old people) with respect	27%	@XXX @XXX @lilyallen old patriots that worked hard to build a country. the elderly should always be first before immigrants.
				@prisonplanet @lilyallen an all time low from lily allen. hope her grandparents if she has any are rightfully disgusted.
	Hypocrisy	Presenting views without acting on them, or acting on views inconsistently	13%	@lilyallen @XXX how many refugees have you taken into your mansion?
				@lilyallen what? i don't believe what i'm seeing, the paragon of virtue lily allen being carried home, pissed and/or high on drugs.
	Dishonesty	Presenting views to shock or perform a certain identity, rather than out of genuine concerns	5%	don't worry, @lilyallen is nothing but a self-important, self-righteous, virtue signalling moron trying to stay relevant.
				@XXX @lilyallen it's funny how suddenly it's hip to be ageist among the left.
	Other	Abuse (no ethical content)	55%	

Troll communities

To understand the networks and wider interests of individuals who engage in trolling activities, CASM researchers identified Twitter users taking part in a very distinctive form of trolling behaviour – users writing fake appeals for help in identifying missing friends or relatives in wake of major terrorists attacks– and attempted to map their wider online networks.

To identify these individuals, researchers collected data after the attack on the Ariana Grande concert in Manchester on 22 May 2017. Using the keyword 'manchester' (which included hashtags #manchester and #manchester attack), analysts collected 327,000 tweets sent between 20 and 23 May. These data were filtered by keywords and key phrases to locate tweets about missing people that analysts judged by to be fake.

An appeal was judged to be fake when the subject of the tweet was either a known subject of trolling in the past or a celebrity who was confirmed not to have attended the concert. Researchers identified 51 users who had engaged in 'fake image' trolling linked to the Manchester attack.

CASM researchers then conducted a qualitative analysis of the retweet activity of these accounts. Using an initial scan of retweet activity, five potentially widely shared categories of content were identified:

· *right-wing political*: tweets supporting right-wing political parties and political figures, particularly in the USA
· *right-wing and far-right ideological*: tweets celebrating right and far-right political ideas, and criticising or insulting groups such as Muslims, ethnic minorities, women and feminists, and the LGTBQ community
· *comedic tweets and meme-sharing*: tweets sharing jokes as images or text, or recycling memes
· *image board content*: image boards are popular, anonymous forums; the most widely visited, 4Chan, has been noted for its role in fostering digital activism, politics and culture; an account was tagged in this category when it employed recognisable tropes, memes or images (eg, Pepe the Frog, greentext) taken from the site
· *gaming content*: tweets about video games and gaming

Researchers then manually coded the last 100 retweets of each 'fake image troll', identifying which accounts shared content falling into each of the five categories set out above. The results of this analysis are set out in case study 3.

Notes

1   Ofcom, *Adults' Media Use and Attitudes*, 2017, www.ofcom.
    org.uk/__data/assets/pdf_file/0020/102755/adults-media-
    use-attitudes-2017.pdf (accessed 17 Aug 2017).

2   J Birdwell, R Scott and L Reynolds, *Character Nation*,
    Demos, 2015, www.demos.co.uk/project/character-nation-2
    (accessed 17 Aug 2017).

3   T Harrison, 'Making wiser choices online',
    2017, forthcoming.

4   Ofcom, *Adults' Media Use and Attitudes*.

5   Ibid.

6   Ofcom, *Digital Day 2016: Results from the children's
    diary study*, 2016, www.ofcom.org.uk/__data/assets/
    pdf_file/0017/94013Childrens-Digital-Day-report-2016.pdf
    (accessed 17 Aug 2017).

7   SL Gray and A Phippen, *Invisibly Blighted: The digital
    erosion of childhood*, UCL IOE Press, 2017.

8   S Cohen, *Folk Devils and Moral Panics*, 3rd edn,
    Routledge, 2002.

9   Childline, *What Children Are Telling Us About Bullying:
    Childline bullying report 2015/16*,
    www.nspcc.org.uk/globalassets/documents/
    research-reportswhat-children-are-telling-us-about-bully
    ing-childline-bullying-report-2015-16.pdf
    (accessed 17 Aug 2017).

10   YoungMinds, *Resilience for the Digital World: Research into children and young people's social and emotional wellbeing online*, 2016, https://youngminds.org.uk/resources/policy/ resilience-for-the-digital-world (accessed 17 Aug 2017).

11   Home Office, 'Prevent duty guidance', 2016, www.gov. uk/government/publications/prevent-duty-guidance (accessed 17 Aug 2017).

12   DfE, 'Keeping children safe in education', Dept for Education, 2016, www.gov.uk/government/publications/ keeping-children-safe-in-education--2 (accessed 17 Aug 2017).

13   Ofsted, 'Inspecting safeguarding in early years, education and skills', 2016, www.gov.uk/government/publications/ inspecting-safeguarding-in-early-years-education-and-skills- from-september-2015 (accessed 17 Aug 2017).

14   Prime Minister's Office, *The Queen's Speech and Associated Background Briefing, on the Occasion of the Opening of Parliament*, Jun 2017, www.gov.uk/government/ uploads/system/uploads/attachment_data/file/620838/ Queens_speech_2017_background_notes.pdf (accessed 17 Aug 2017).

15   DCMS, *UK Digital Strategy 2017*, Dept for Digital, Culture, Media & Sport, 2017, www.gov.uk/government/ publications/uk-digital-strategy/uk-digital-strategy (accessed 17 Aug 2017).

16   Children's Commissioner, *Growing Up Digital: A report of the Growing Up Digital Taskforce*, 2017, www.childrenscommissioner.gov.uk/wp-content/ uploads/2017/06/Growing-Up-Digital-Taskforce-Report- January-2017_0.pdf (accessed 17 Aug 2017).

17   Ofcom, *Digital Day 2016*.

18   M Prensky, 'Digital natives, digital immigrants', *On the Horizon* 9, issue 5.

19   D Boyd, 'Why youth (heart) social network sites: the role of networked publics in teenage social life' in D Buckingham (ed), *Youth, Identity, and Digital Media*, MIT Press, 2007, www.danah.org/papers/WhyYouthHeart.pdf (accessed 17 Aug 2017).

20   D Meshi, DI Tamir and HR Heekeren, 'The emerging neuroscience of social media', *Trends in Cognitive Sciences* 19, issue 12.

21   GS Hall, *Adolescence*, vol 2, Appleton, 1904.

22   BJ Casey, TA Hare and A Galván, 'Risky and impulsive components of adolescent decision making' in MR Delgado, EA Phelps and TW Robbins (eds), *Decision Making, Affect, and Learning: Attention and performance XXIII*, Oxford Scholarship Online, 2011.

23   RE Dahl, 'Adolescent brain development: a period of vulnerabilities and opportunities', *Annals of the New York Academy of Sciences*, Jun 2004.

24   L Kohlberg, 'Moral stages and moralization: the cognitive-developmental approach' in T Lickona, *Moral Development and Behavior: Theory, research and social issues*, Rinehart and Winston, 1976.

25   KT Do, JF Guassi Moreira and EH Telzer, 'But is helping you worth the risk? Defining prosocial risk taking in adolescence', *Developing Cognitive Neuroscience* 25, Jun 2017.

26   LA Spies Shapiro and G Margolin, 'Growing up wired: social networking sites and adolescent psychosocial development', *Clinical Child and Family Psychology Review* 17 no 1, 2014.

27   E Hanson and D Holmes, *That Difficult Age: Developing a more effective response to risks in adolescence*, Research in Practice, 2014, www.rip.org.uk/download/237/evidence-scope-that-difficult-age-main.pdf (accessed 17 Aug 2017).

28   Birdwell et al, *Character Nation*.

29   Jubilee Centre for Character & Virtues, 'A framework for character education in schools', 2017, www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/character-educationFramework%20for%20Character%20Education.pdf (accessed 17 Aug 2017).

30   N Morgan and E Timpson, 'Schools and organisations recognised for instilling character', news story, Dept for Education, 1 Jul 2016, www.gov.uk/government/news/schools-and-organisations-recognised-for-instilling-character (accessed 17 Aug 2017).

31   M Weber and W Ruch, 'The role of a good character in 12-year-old school children: do character strengths matter in the classroom', *Child Indicators Research* 5, issue 2, Jun 2012.

32   J Arthur et al, *Character Education in UK Schools*, Jubilee Centre for Character & Virtues and University of Birmingham, 2015, www.jubileecentre.ac.uk/1571/projects/previous-workcharacter-education-in-uk-schools (accessed 17 Aug 2017); DI Walker et al, 'Adolescent moral judgement: a study of UK secondary school pupils', *British Educational Research Journal* 43, no 3, 2017.

33   SSAT, *Mind the Gap: Collected essays on the development of character, non-cognitive skills, mindfulness and well-being*, SSAT (The Schools, Students and Teachers Network) and the Young Foundation, 2014, https://webcontent.ssatuk.co.uk/wp-content/uploads/2015/10/13124740/Mind-the-Gap.pdf (accessed 17 Aug 2017).

34 A Goodman et al, *Social and Emotional Skills in Childhood and Their Long-Term Effects on Adult Life*, Early Intervention Foundation, UCL Institute of Education, 11 Mar 2015, www.eif.org.uk/wp-content/uploads/2015/03/EIF-Strand-1-Report-FINAL1.pdf (accessed 17 Aug 2017).

35 E Timpson, 'Funding boost for schools helping pupils develop character', press release, Dept for Education, 26 May 2016, www.gov.uk/government/news/funding-boost-for-schools-helping-pupils-develop-character (accessed 17 Aug 2017).

36 A Hillier, 'National Citizen Service to be expanded', Third Sector, 25 Nov 2015, www.thirdsector.co.uk/national-citizen-service-expanded/policy-and-politics/article/1374500 (accessed 17 Aug 2017).

37 DfE, 'Developing character skills in schools', Dept for Education, Aug 2017, www.gov.uk/government/publications/developing-character-skills-in-schools (accessed 17 Aug 2017).

38 S Livingstone et al, *The UK Report: A comparative report with findings from the UK 2010 survey by EU Kids Online*, Net Children Go Mobile project, London School of Economics and Political Science, 2014, www.lse.ac.uk/media@lse/research/EUKidsOnline/EU%20Kids%20III/Reports/NCGMUKReportfinal.pdf (accessed 17 Aug 2017).

39 C Lilley, R Ball and H Vernon, *The Experiences of 11-16-year-olds on Social Networking Sites*, NSPCC, 2014, www.nspcc.org.uk/globalassets/documents/research-reports/experiences-11-16-year-olds-social-networking-sites-report.pdf (accessed 17 Aug 2017).

40 L Haddon and J Vincent, *UK Children's Experience of Smartphones and Tablets: Perspectives from children, parents and teachers*, Net Children Go Mobile project, London School of Economics and Political Science, http://eprints.lse.ac.uk/62125/1/UK%20Childern%20mobile%20experience_author.pdf (accessed 17 Aug 2017).

41 YoungMinds, *Resilience for the Digital World*.

42 S Livingstone, G Mascheroni and E Staksrud, *Developing A Framework for Researching Children's Online Risks and Opportunities in Europe*, EU Kids Online, 2015, http://eprints.lse.ac.uk/64470/1/___lse.ac.uk_storage_LIBRARY_Secondary_libfile_shared_repository_Content_EU%20Kids%20Online_EU%20Kids%20Online_Developing%20framework%20for%20researching_2015.pdf (accessed 23 Aug 2017).

43 Ditch the Label, *The Annual Cyberbullying Survey*, 2013, https://us.ditchthelabel.org/wp-content/uploads/sites/2/2016/07/cyberbullying2013.pdf (accessed 23 Aug 2017).

44 Ditch the Label, *The Annual Cyberbullying Survey*, 2016, www.ditchthelabel.org/research-papers/the-annual-bullying-survey-2016/ (accessed 23 Aug 2017).

45 Livingstone et al, *The UK Report*.

46 NSPCC defines trolling as 'unkind comments or rumours circulated online'. Prevalence figure calculated by author from 28 per cent who reported upsetting experience, of which 37 per cent said this was the result of trolling.

47 Lilley, *The Experiences of 11-16 Year Olds on Social Networking Sites*.

48 Livingstone et al, *The UK Report*.

49 Ibid.

50 E Martellozzo et al, *'I Wasn't Sure It Was Normal To Watch It': A quantitative and qualitative examination of the impact of online pornography on the values, attitudes, beliefs and behaviours of children and young people*, rev edn, NSPCC and the Children's Commissioner for England, May 2017, www.mdx.ac.uk/__data/assets/pdf_file/0021/223266/ MDX-NSPCC-OCC-pornography-report.pdf (accessed 17 Aug 2017).

51 Ofcom, *Children and Parents: Media use and attitudes report*, 2016, www.ofcom.org.uk/__data/assets/pdf_file/0034/93976/Children-Parents-Media-Use-Attitudes-Report-2016.pdf (accessed 17 Aug 2017).

52 Livingstone et al, *The UK Report*.

53 Ibid.

54 Ibid.

55 Ofcom, *Children and Parents*.

56 A Vromen et al, 'Everyday making through Facebook engagement: young citizens' political interactions in Australia, the United Kingdom and the United States', *Political Studies* 64, no 3, 2016.

57 Livingstone et al, *The UK Report*.

58 A Lenhart et al, *Teens, Technology and Friendships: Video games, social media and mobile phones play an integral role in how teens meet and interact with friends*, Pew Research Center, 2015, www.pewinternet.org/2015/08/06/ teens-technology-and-friendships (accessed 17 Aug 2017).

59 Lilley, *The Experiences of 11-16 Year Olds on Social Networking Sites*.

60 Lenhart et al, *Teens, Technology and Friendships*.

61 J Suler, 2004, 'The online disinhibition effect', *CyberPsychology and Behavior* 7, no 3, 2004.

62 Ofcom, *Children and Parents*.

63 H Cowie, 'Cyberbullying and its impact on young people's emotional health and well-being', *The Psychiatrist* 37, 2013.

64 Although the survey was conducted anonymously, self-report questionnaires are susceptible to social desirability bias, which could suppress the number of respondents reporting negative behaviour.

65 Respondents reporting visiting SNS over 10 times per day were classified as 'high' intensity users, while those visiting these sites less than this were classified as 'low' intensity users.

66 R Seymour, 'Schadenfreude with bite', *London Review of Books* 38, no 24, 15 Dec 2016, www.lrb.co.uk/v38/n24/richard-seymour/schadenfreude-with-bite (accessed 17 Aug 2017).

67 The authors of this study define trolling as 'the practice of behaving in a deceptive, destructive, or disruptive manner in a social setting on the internet with no apparent instrumental purpose'.

68 EE Buckels et al, 'Trolls just want to have fun', *Personality and Individual Differences* 67, Sep 2014.

69 A Marwick and d boyd, 'It's just drama': teen perspectives on conflict and aggression in a networked era, *Journal of Youth Studies* 17, issue 9, 2014, www.tandfonline.com/doi/abs/10.108013676261.2014.901493 (accessed 23 Aug 2017).

70 Ibid.

71   See for example, Spies Shapiro and Margolin,
      'Growing up wired'.

72   Lilley et al, *The Experiences of 11-16 Year Olds on Social
      Networking Sites*.

73   See, for exploration of the moral dimensions of friendship
      in children, DI Walker, R Curren and C Jones, 'Good
      friendships among children: a theoretical and empirical
      investigation', *Journal for the Theory of Social Behaviour*
      46, no 3, 2016.

74   Lenhart et al, *Teens, Technology and Friendships*.

75   S Vallor, Flourishing on Facebook: virtue friendship
      and new social media, *Ethics and Information Technology*
      14, issue 3, 2011.

76   See, for example, L Burns, *Bowling Online: How the internet
      is driving the reinvigoration of American social capital*, Working
      Paper01-12, School of Sociology, Politics and International
      Studies, University of Bristol, 2012.

77   Eg, J Arthur et al, 'Habits of service', University
      of Birmingham, 2017 (forthcoming).

78   E Morozov, *The Net Delusion: How not to liberate the world*,
      Public Affairs, 2011.

79   D Bond, 'Labour's slick online campaign outguns Tory
      press: almost 10m Facebook users in the UK watched
      pro-Corbyn videos', *Financial Times*, 9 Jun 2017, www.
      ft.com/content/d1c854f0-4cea-11e7-a3f4-c742b9791d43
      (accessed 23 Aug 2017).

80   Ipsos MORI, 'How Britain voted in the 2017 election',
      20 Jun 2017, www.ipsos.com/ipsos-mori/en-uk/how-britain-
      voted-2017-election (accessed 23 Aug 2017).

81   Vromen et al, 'Everyday making through
     Facebook engagement'.

82   Do et al, 'But is helping you worth the risk?'; see, for an
     example of an emerging application of PSRT literature
     in social media research: JM Fatkin and T Lansdown,
     'Prosocial media in action', *Computers in Human
     Behavior* 48, 2015.

83   Do et al, 'But is helping you worth the risk?'

84   Ofcom, *Children and Parents*.

85   Where different contexts required case-specific responses,
     care was taken to develop broadly equivalent actions across
     all three scenarios.

86   C James, *Disconnected: Youth, new media, and the ethics gap*,
     MIT Press, 2014.

87   JR Rest, *Moral Development: Advances in research
     and theory*, Praeger, 1986.

88   L Haddon and S Livingstone, *The Meaning of Online
     Problematic Situations for Children: The UK report*.

89   A Flores and C James, 'Morality and ethics behind
     the screen: young people's perspectives on digital life',
     New Media & Society 15, issue 6, 2012.

90   iRights, 'Enabling children and young people to access the
     digital world creatively, knowledgably, and fearlessly', 2015,
     www.itu.int/en/council/cwg-cop/Documents/iRights%20
     Full%20Text%20Revised%20Sept%202015%20.pdf
     (accessed 17 Aug 2017).

91   M Hogg and G Vaughan, *Essentials of Social Psychology*,
     Prentice Hall, 2010.

92 S Livingstone, 'Developing social media literacy: how children learn to interpret risky opportunities on social network sites', *Communications: The European Journal of Communication Research* 39, no 3.

93 YoungMinds, *Resilience for the Digital World*.

94 Weber and Ruch, 'The role of a good character in 12-year-old school children'.

95 J Arthur et al, *Character Education in UK Schools*, and J Arthur et al, *Schools of Virtue: Character education in three Birmingham schools: research report*, University of Birmingham, 2017, www.jubileecentre.ac.uk/1589/projects/current-projectsschools-of-virtue (accessed 17 Aug 2017).

96 SAT, *Mind the Gap*.

97 Goodman et al, *Social and Emotional Skills in Childhood and Their Long-Term Effects on Adult Life*.

98 Flores and James, 'Morality and ethics behind the screen'; W Sanderse, 'An Aristotelian model of moral development', *Journal of Philosophy of Education* 49, issue 3, Aug 2015, www.jubileecentre.ac.uk/media/news/art icle/4312An-Aristotelian-Model-of-Moral-Development (accessed 23 Aug 2017); Kohlberg, 'Moral stages and moralization'.

99 K Kristjánsson, *Aristotelian Character Education*, Routledge, 2015; T Harrison, Does the Internet Influence the Character Virtues of 11 to 14 year olds in England?: a mixed method study with particular regard to cyber-bullying, PhD thesis, University of Birmingham, 2014, http://etheses.bham.ac.uk/5488/4/Harrison14PhD.pdf (accessed 17 Aug 2017).

100 E Taylor, 'Thinking about the virtues of character measurement', Centre for Youth Impact, Centre for Youth Impact, 2015, http://youthimpact.uk/sites/www.youthimpact.uk/files/3%20-%20Thinking%20about%20the%20virtues%20of%20character%20measurement.pdf (accessed 17 Aug 2017).

101 D Walker, 'Measuring character strengths and virtue among UK school children aged 14 and 15, 2014', paper given at conference Can Virtue be Measured?, Jubilee Centre for Character & Values, University of Birmingham, 2014, www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/conference-papers/can-virtue-be-measuredwalker-david.pdf (accessed 17 Aug 2017).

102 AL Duckworth and PD Quinn, 'Development and validation of the short grit scale (grit-s)', *Journal of Personality Assessment* 91, no 2, 2009; N Park and C Peterson, 'Moral competence and character strengths among adolescents: the development and validation of the Values in Action Inventory of Strengths for Youth', *Journal of Adolescence* 29, issue 6, 2006.

103 R Goodman, 'The Strengths and Difficulties Questionnaire: a research note', *Journal of Child Psychology and Psychiatry* 38, no 5, 1997; G Russell, LR Rodgers and T Ford, 'The Strengths and Difficulties Questionnaire as a predictor of parent-reported diagnosis of autism spectrum disorder and attention deficit hyperactivity disorder', *PLoS ONE* 8, no 12, 2013.

104 The distribution of character scores was skewed towards the upper end. The bottom 20 per cent of respondents (131 individuals) scored 5 or below, while the top 20per cent (140 individuals) scored 9 and 10.

105 S Vallor, 'Social networking technology and the virtues', *Ethics and Information Technology* 12, 2010.

106 Jubilee Centre for Character & Virtues, 'A framework for character education in schools'.

107 S Coleman et al, *The Internet On Our Own Terms: How children and young people deliberated about their digital rights*, 5Rights, 2017, http://casma.wp.horizon.ac.uk/wp-content/uploads/2016/08/Internet-On-Our-Own-Terms.pdf (accessed 17 Aug 2017).

108 LM Gutman and I Shoon, *The Impact of Non-cognitive Skills on Outcomes for Young People: Literature review*, Institute of Education, 2013, https://educationendowmentfoundation.org.uk/public/files/Publications/EEF_Lit_Review_Non-CognitiveSkills.pdf (accessed 23 Aug 2017).

109 Jubilee Centre for Character & Virtues, 'A framework for character education in schools'.

110 LM Jones, KJ Mitchell and WA Walsh, 'A content analysis of youth internet safety programs: are effective prevention strategies being used?', Crimes Against Children Research Center, 2014, www.unh.edu/ccrc/pdfISE%20Bulletin%202_Contant%20Analysis%20of%20Youth%20ISE%20FINAL-with%20appendix.pdf (accessed 17 Aug 2017).

111 J Davidson, E Martellozzo and M Lorenz, *Evaluation of CEOP ThinkUKnow Internet Safety Programme and exploration of young people's internet safety knowledge*, 2009, Centre for Abuse & Trauma Studies, Kingston University, www.eprints.mdx.ac.uk/5976 (accessed 17 Aug 2017).

112 R Scott and L Reynolds, *Digital Citizens: Countering extremism online*, Demos, 2016, www.demos.co.uk/project/digital-citizens/ (accessed 17 Aug 2017).

113 LM Jones and KJ Mitchell, 'Defining and measuring youth digital citizenship', *New Media & Society*, 25 Mar 2015.

114   See for example Birdwell et al, *Character Nation*.

115   Children's Commissioner, *Growing Up Digital*.

116   PHSE Association, 'Key principles of effective prevention education', 2016, www.thinkuknow.co.uk/Documents/ PSHE%20Association%20report%20to%20CEOP%20 April%202016.pdf (accessed 17 Aug 2017).

117   J Lexmond and R Reeves, *Building Character*, Demos, 2009, www.demos.co.uk/files/Building_Character_Web.pdf (accessed 17 Aug 2017).

118   B Morgan and K Kristjánsson, 'Parents and social media', research paper, University of Birmingham, 2017 (forthcoming).

119   boyd, 'Why youth (heart) social network sites'.

120   G Mascheroni et al, 'Parents: mediation, self-regulation and co-regulation' in B O'Neill, E Staksrud and S McLaughlin (eds), *Towards a Better Internet for Children? Policy pillars, players and paradoxes*, Nordicom, 2013.

121   Children's Commissioner, *Growing Up Digital*.

122   boyd, 'Why youth (heart) social network sites.

123   Livingstone et al, *The UK Report*.

124   Ibid.

125   S Coyne et al, 'A friend request from dear old dad: associations between parent-child social networking and adolescent outcomes', *Cyberpsychology, Behavior and Social Networking* 17, no 1, 2014.

126   Prime Minister's Office, *The Queen's Speech and Associated Background Briefing*.

127   4.75 billion pieces of content are shared daily on
      Facebook, for example. See Facebook, 'Did you know
      that… 4.75 billion pieces of content are shared daily?',
      21 Jun 2013, www.facebook.com/FacebookSingapore/
      posts/563468333703369 (accessed 17 Aug 2017).

128   @UP4LIFE, 'Cultivating digital citizenship among youth
      and parents at Twitter NeighborNest', blog, 8 Nov 2016,
      https://blog.twitter.com/official/en_us/a/2016/cultivating-
      digital-citizenship-among-youth-and-parents-at-twitter-
      neighbornest.html (accessed 17 Aug 2017).

129   Online Civil Courage Initiative, 'Information pack on
      counter-speech engagement', 2012, www.isdglobal.org/
      wp-content/uploads/2017/05/OCCI-Counterspeech-
      Information-Pack-English.pdf (accessed 17 Aug 2017).

130   Childnet International, *Crossing the Line: A practical
      PSHE toolkit for educators to explore online behaviour and
      consequences*, 2016, www.childnet.com/ufiles/Crossing-the-
      Line---full-pdf.pdf (accessed 17 Aug 2017).

131   iRights, 'Enabling children and young people to access
      the digital world creatively, knowledgably, and fearlessly'.

132   Coleman et al, *The Internet On Our Own Terms*.

133   DfE, 'Developing character skills in schools'.

134   DfE, 'Keeping children safe in education'.

135   DCMS, 'Digital skills and inclusion – giving everyone
      access to the digital skills they need', policy paper, Dept
      for Digital, Culture, Media & Sport, 1 Mar 2017, www.gov.
      uk/government/publications/uk-digital-strategy/2-digital-
      skills-and-inclusion-giving-everyone-access-to-the-digital-
      skills-they-need (accessed 17 Aug 2017).

136 Be Strong Online, https://bestrongonline.antibullyingpro.
     com/ (accessed 17 Aug 2017).

137 The Diana Award, 2016, http://diana-award.org.uk/about/
     (accessed 23 Aug 2017).

138 P Kumari, *The Diana Award Anti-Bullying Ambassadors
     Programme: A mixed methods evaluation of the programme*,
     The Diana Award, 2016, https://static1.squarespace.com/
     static/522f7cdbe4b056ca5e8127f6/t/58ede3e82994ca36
     520e334a/1491985428801/DfE+Evaluation+Report.pdf
     (accessed 17 Aug 2017).

139 W Gardner, 'Letter from the CEO', Childnet International,
     [2015], www.childnet.com/ufilesAbout-this-toolkit1.pdf
     (accessed 17 Aug 2017).

140 Scott and Reynolds, *Digital Citizens*.

141 Digital Literacy & Citizenship from the South West Grid
     for Learning, 2015, www.digital-literacy.org.uk/Home.aspx
     (accessed 17 Aug 2017).

142 K Donnelly, 'Common Sense Media employs
     comprehensive evaluation strategy for Digital Citizenship
     Curriculum', Center for Media & Social Impact, 2017,
     http://cmsimpact.org/resource/common-sense-media-
     employs-comprehensive-evaluation-strategy-digital-
     citizenship-curriculum/ (accessed 17 Aug 2017).

143 Personal communication.

144 Over 100% due to rounding.

145 Scott and Reynolds, *Digital Citizens*.

146 Taylor, 'Thinking about the virtues of character measurement'.

References

@UP4LIFE, 'Cultivating digital citizenship among youth and parents at Twitter NeighborNest', blog, 8 Nov 2016, https://blog.twitter.com/official/en_us/a/2016/cultivating-digital-citizenship-among-youth-and-parents-at-twitter-neighbornest.html (accessed 17 Aug 2017).

Arthur J et al, 'Habits of service', University of Birmingham, 2017 (forthcoming).

Arthur J et al, *Character Education in UK Schools*, Jubilee Centre for Character & Virtues and University of Birmingham, 2015, www.jubileecentre.ac.uk/1571/projects/previous-work/character-education-in-uk-schools (accessed 17 Aug 2017).

Arthur J et al, *Schools of Virtue: Character education in three Birmingham schools: research report*, University of Birmingham, 2017, www.jubileecentre.ac.uk/1589/projects/current-projectsschools-of-virtue (accessed 17 Aug 2017).

Be Strong Online, https://bestrongonline.antibullyingpro.com/ (accessed 17 Aug 2017).

Birdwell J, Scott R and Reynolds L, *Character Nation*, Demos, 2015, www.demos.co.uk/project/character-nation-2/ (accessed 17 Aug 2017).

Bond D, 'Labour's slick online campaign outguns Tory press: almost 10m Facebook users in the UK watched pro-Corbyn videos', *Financial Times*, 9 Jun 2017, www.ft.com/content/d1c854f0-4cea-11e7-a3f4-c742b9791d43 (accessed 23 Aug 2017).

boyd d, 'Why youth (heart) social network sites: the role of networked publics in teenage social life' in D Buckingham (ed), *Youth, Identity, and Digital Media*, MIT Press, 2007, www.danah.org/papers/WhyYouthHeart.pdf (accessed 17 Aug 2017).

Buckels EE et al, 'Trolls just want to have fun', *Personality and Individual Differences* 67, Sep 2014.

Burns L, *Bowling Online: How the internet is driving the reinvigoration of American social capital*, School of Sociology, Politics and International Studies, University of Bristol, 2012.

Casey BJ, Hare TA and Galván A, 'Risky and impulsive components of adolescent decision making' in MR Delgado, EA Phelps and TW Robbins (eds), *Decision Making, Affect, and Learning: Attention and performance XXIII*, Oxford Scholarship Online, 2011.

Childline, *What Children Are Telling Us About Bullying: Childline bullying report 2015/16*, www.nspcc.org.uk/globalassets/ documents/research-reports/what-children-are-telling-us-about-bullying-childline-bullying-report-2015-16.pdf (accessed 17 Aug 2017).

Childnet International, *Crossing the Line: A practical PSHE toolkit for educators to explore online behaviour and consequences*, 2016, www.childnet.com/ufiles/Crossing-the-Line---full-pdf.pdf (accessed 17 Aug 2017).

Children's Commissioner, *Growing Up Digital: A report of the Growing Up Digital Taskforce*, 2017, www.childrenscommissioner.gov.uk/wp-content/ uploads/2017/06/Growing-Up-Digital-Taskforce-Report-January-2017_0.pdf (accessed 17 Aug 2017).

Cohen S, *Folk Devils and Moral Panics*, 3rd edn, Routledge, 2002.

Coleman S et al, *The Internet On Our Own Terms: How children and young people deliberated about their digital rights*, 5Rights, 2017, http://casma.wp.horizon.ac.uk/wp-content/uploads/2016/08/Internet-On-Our-Own-Terms.pdf (accessed 17 Aug 2017).

Cowie H, 'Cyberbullying and its impact on young people's emotional health and well-being', *The Psychiatrist* 37, 2013.

Coyne S et al, 'A friend request from dear old dad: associations between parent-child social networking and adolescent outcomes', *Cyberpsychology, Behavior and Social Networking* 17, no 1, 2014.

Dahl RE, 'Adolescent brain development: a period of vulnerabilities and opportunities', *Annals of the New York Academy of Sciences*, Jun 2004.

Davidson J, Martellozzo E and Lorenz M, *Evaluation of CEOP ThinkUKnow Internet Safety Programme and exploration of young people's internet safety knowledge*, 2009, Centre for Abuse & Trauma Studies, Kingston University, http://eprints.mdx.ac.uk/5976/ (accessed 17 Aug 2017).

DCMS, 'Digital skills and inclusion – giving everyone access to the digital skills they need', policy paper, Dept for Digital, Culture, Media & Sport, 1 Mar 2017, www.gov.uk/government/publications/uk-digital-strategy/2-digital-skills-and-inclusion-giving-everyone-access-to-the-digital-skills-they-need (accessed 17 Aug 2017).

DCMS, *UK Digital Strategy 2017*, Dept for Digital, Culture, Media & Sport, 2017, www.gov.uk/government/publications/uk-digital-strategy/uk-digital-strategy (accessed 17 Aug 2017).

DfE, 'Developing character skills in schools', Dept for Education, Aug 2017, www.gov.uk/government/publications/developing-character-skills-in-schools (accessed 17 Aug 2017).

DfE, 'Keeping children safe in education', Dept for Education, 2016, www.gov.uk/government/publications/keeping-children-safe-in-education--2 (accessed 17 Aug 2017).

Digital Literacy & Citizenship from the South West Grid for Learning, 2015, www.digital-literacy.org.ukHome.aspx (accessed 17 Aug 2017).

Ditch the Label, *The Annual Cyberbullying Survey*, 2013, https://us.ditchthelabel.org/wp-content/uploads/sites/2/2016/07/cyberbullying2013.pdf (accessed 23 Aug 2017).

Ditch the Label, *The Annual Cyberbullying Survey*, 2016, www.ditchthelabel.org/research-papers/the-annual-bullying-survey-2016/ (accessed 23 Aug 2017).

Do KT, Guassi Moreira JR and Telzer EH, 'But is helping you worth the risk? Defining prosocial risk taking in adolescence', *Developing Cognitive Neuroscience* 25, Jun 2017.

Donnelly K, 'Common Sense Media employs comprehensive evaluation strategy for Digital Citizenship Curriculum', Center for Media & Social Impact, 2017, http://cmsimpact.org/resource/common-sense-media-employs-comprehensive-evaluation-strategy-digital-citizenship-curriculum/ (accessed 17 Aug 2017).

Duckworth AL and Quinn PD, 'Development and validation of the short grit scale (grit-s)', *Journal of Personality Assessment* 91, no 2, 2009.

Facebook, 'Did you know that… 4.75 billion pieces of content are shared daily?', 21 Jun 2013, www.facebook.com/FacebookSingapore/posts/563468333703369 (accessed 17 Aug 2017).

Fatkin JM and Lansdown T, 'Prosocial media in action', *Computers in Human Behavior* 48, 2015.

Flores A and James C, 'Morality and ethics behind the screen: young people's perspectives on digital life', *New Media & Society* 15, issue 6, 2012.

Gardner W, 'Letter from the CEO', Childnet International, [2015], www.childnet.com/ufiles/About-this-toolkit1.pdf (accessed 17 Aug 2017).

Goodman A et al, *Social and Emotional Skills in Childhood and Their Long-Term Effects on Adult Life*, Early Intervention Foundation, UCL Institute of Education, 11 Mar 2015, www.eif.org.uk/wp-content/uploads/2015/03/EIF-Strand-1-Report-FINAL1.pdf (accessed 17 Aug 2017).

Goodman R, 'The Strengths and Difficulties Questionnaire: a research note', *Journal of Child Psychology and Psychiatry* 38, no 5, 1997.

Gray SL and Phippen A, *Invisibly Blighted: The digital erosion of childhood*, UCL IOE Press, 2017.

Gutman LM and Shoon I, *The Impact of Non-cognitive Skills on Outcomes for Young People: Literature review*, Institute of Education, 2013, https://educationendowmentfoundation.org.uk/public/files/Publications/EEF_Lit_Review_Non-CognitiveSkills.pdf (accessed 23 Aug 2017).

Haddon L and Livingstone S, *The Meaning of Online Problematic Situations for Children: The UK report*.

Haddon L and Vincent J, *UK Children's Experience of Smartphones and Tablets: Perspectives from children, parents and teachers*, Net Children Go Mobile project, London School of Economics and Political Science, http://eprints.lse.ac.uk/62125/1/UK%20Childern%20mobile%20experience_author.pdf (accessed 17 Aug 2017).

Hall GS, *Adolescence*, vol 2, Appleton, 1904.

Hanson E and Holmes D, *That Difficult Age: Developing a more effective response to risks in adolescence*, Research in Practice, 2014, www.rip.org.uk/download/237/evidence-scope-that-difficult-age-main.pdf (accessed 17 Aug 2017).

Harrison T, 'Making wiser choices online', 2017, forthcoming.

Harrison T, Does the Internet Influence the Character Virtues of 11 to 14 year olds in England?: a mixed method study with particular regard to cyber-bullying, PhD thesis, University of Birmingham, 2014, http://etheses.bham.ac.uk/5488/4/Harrison14PhD.pdf (accessed 17 Aug 2017).

Hillier A, 'National Citizen Service to be expanded', Third Sector, 25 Nov 2015, www.thirdsector.co.uk/national-citizen-service-expanded/policy-and-politics/article/1374500 (accessed 17 Aug 2017).

Hogg M and Vaughan G, *Essentials of Social Psychology*, Prentice Hall, 2010.

Home Office, 'Prevent duty guidance', 2016, www.gov.uk/government/publications/prevent-duty-guidance (accessed 17 Aug 2017).

Ipsos MORI, 'How Britain voted in the 2017 election', 20 Jun 2017, www.ipsos.com/ipsos-mori/en-uk/how-britain-voted-2017-election (accessed 23 Aug 2017).

iRights, 'Enabling children and young people to access the digital world creatively, knowledgably, and fearlessly', 2015, www.itu.int/en/council/cwg-cop/Documents/iRights%20Full%20Text%20Revised%20Sept%202015%20.pdf (accessed 17 Aug 2017).

James C, *Disconnected: Youth, new media, and the ethics gap*, MIT Press, 2014.

Jones LM and Mitchell KJ, 'Defining and measuring youth digital citizenship', *New Media & Society*, 25 Mar 2015.

Jones LM, Mitchell KJ and Walsh WA, 'A content analysis of youth internet safety programs: are effective prevention strategies being used?', Crimes Against Children Research Center, 2014, www.unh.edu/ccrc/pdf/ISE%20Bulletin%202_Contant%20Analysis%20of%20Youth%20ISE%20FINAL-with%20appendix.pdf (accessed 17 Aug 2017).

Jubilee Centre for Character & Virtues, 'A framework for character education in schools', 2017, www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/character-education/Framework%20for%20Character%20Education.pdf (accessed 17 Aug 2017).

Kohlberg L, 'Moral stages and moralization: the cognitive-developmental approach' in T Lickona, *Moral Development and Behavior: Theory, research and social issues*, Rinehart and Winston, 1976.

Kristjánsson K, *Aristotelian Character Education*, Routledge, 2015.

Kumari P, *The Diana Award Anti-Bullying Ambassadors Programme: A mixed methods evaluation of the programme*, The Diana Award, 2016, https://static1.squarespace.com/static/522f7cdbe4b056ca5e8127f6/t/58ede3e82994ca36520e334a/1491985428801/DfE+Evaluation+Report.pdf (accessed 17 Aug 2017).

Lenhart A et al, *Teens, Technology and Friendships: Video games, social media and mobile phones play an integral role in how teens meet and interact with friends*, Pew Research Center, 2015, www.pewinternet.org/2015/08/06/teens-technology-and-friendships/ (accessed 17 Aug 2017).

Lexmond J and Reeves R, *Building Character*, Demos, 2009,
www.demos.co.uk/files/Building_Character_Web.pdf
(accessed 17 Aug 2017).

Lilley C, Ball R and Vernon H, *The Experiences of 11-16-year-olds
on Social Networking Sites*, NSPCC, 2014, www.nspcc.org.uk/
globalassets/documents/research-reports/experiences-11-16-
year-olds-social-networking-sites-report.pdf
(accessed 17 Aug 2017).

Livingstone S, 'Developing social media literacy: how children
learn to interpret risky opportunities on social network sites',
*Communications: The European Journal of Communication
Research* 39, no 3.

Livingstone S, Mascheroni G and Staksrud E, *Developing
A Framework for Researching Children's Online Risks and
Opportunities in Europe*, EU Kids Online, 2015, http://eprints.
lse.ac.uk/64470/1/__lse.ac.uk_storage__LIBRARY_
Secondary_libfile_shared_repository_Content_EU%20
Kids%20Online_EU%20Kids%20Online_Developing%20
framework%20for%20researching_2015.pdf
(accessed 23 Aug 2017).

Livingstone S et al, *The UK Report: A comparative report with
findings from the UK 2010 survey by EU Kids Online*, Net Children
Go Mobile project, London School of Economics and Political
Science, 2014, www.lse.ac.uk/media@lse/research/
EUKidsOnline/EU%20Kids%20III/Reports/
NCGMUKReportfinal.pdf (accessed 17 Aug 2017).

Martellozzo E et al, *'I Wasn't Sure It Was Normal To Watch It':
A quantitative and qualitative examination of the impact of online
pornography on the values, attitudes, beliefs and behaviours of
children and young people*, rev edn, NSPCC and the Children's
Commissioner for England, May 2017, www.mdx.ac.uk/__
data/assets/pdf_file/0021/223266/MDX-NSPCC-OCC-
pornography-report.pdf (accessed 17 Aug 2017).

Marwick A and boyd d, 'It's just drama': teen perspectives on conflict and aggression in a networked era, *Journal of Youth Studies* 17, issue 9, 2014, www.tandfonline.com/doi/abs/10.1080 /13676261.2014.901493 (accessed 23 Aug 2017).

Mascheroni G et al, 'Parents: mediation, self-regulation and co-regulation' in B O'Neill, E Staksrud and S McLaughlin (eds), *Towards a Better Internet for Children? Policy pillars, players and paradoxes*, Nordicom, 2013.

Meshi D, Tamir DI and Heekeren HR, 'The emerging neuroscience of social media', *Trends in Cognitive Sciences* 19, issue 12.

Morgan B and Kristjánsson K, 'Parents and social media', research paper, University of Birmingham, 2017 (forthcoming).

Morgan N and Timpson E, 'Schools and organisations recognised for instilling character', news story, Dept for Education, 1 Jul 2016, www.gov.uk/government/news/ schools-and-organisations-recognised-for-instilling-character (accessed 17 Aug 2017).

Morozov E, *The Net Delusion: How not to liberate the world*, Public Affairs, 2011.

Ofcom, *Adults' Media Use and Attitudes*, 2017, www.ofcom.org. uk/__data/assets/pdf_file/0020/102755/adults-media-use-attitudes-2017.pdf (accessed 17 Aug 2017).

Ofcom, *Children and Parents: Media use and attitudes report*, 2016, www.ofcom.org.uk/__data/assets/pdf_file/0034/93976/ Children-Parents-Media-Use-Attitudes-Report-2016.pdf (accessed 17 Aug 2017).

Ofcom, *Digital Day 2016: Results from the children's diary study*, 2016, www.ofcom.org.uk/__data/assets/pdf_file/0017/94013/ Childrens-Digital-Day-report-2016.pdf (accessed 17 Aug 2017).

Ofsted, 'Inspecting safeguarding in early years, education and skills', 2016, www.gov.uk/government/publications/inspecting-safeguarding-in-early-years-education-and-skills-from-september-2015 (accessed 17 Aug 2017).

Online Civil Courage Initiative, 'Information pack on counter-speech engagement', 2012, www.isdglobal.org/wp-content/uploads/2017/05/OCCI-Counterspeech-Information-Pack-English.pdf (accessed 17 Aug 2017).

Park N and Peterson C, 'Moral competence and character strengths among adolescents: the development and validation of the Values in Action Inventory of Strengths for Youth', *Journal of Adolescence* 29, issue 6, 2006.

PHSE Association, 'Key principles of effective prevention education', 2016, www.thinkuknow.co.uk/Documents/PSHE%20Association%20report%20to%20CEOP%20April%202016.pdf (accessed 17 Aug 2017).

Prensky M, 'Digital natives, digital immigrants', *On the Horizon* 9, issue 5.

Prime Minister's Office, *The Queen's Speech and Associated Background Briefing, on the Occasion of the Opening of Parliament*, Jun 2017, www.gov.uk/government/uploads/system/uploads/attachment_data/file/620838/Queens_speech_2017_background_notes.pdf (accessed 17 Aug 2017).

Rest JR, *Moral Development: Advances in research and theory*, Praeger, 1986.

Russell G, Rodgers LR and Ford T, 'The Strengths and Difficulties Questionnaire as a predictor of parent-reported diagnosis of autism spectrum disorder and attention deficit hyperactivity disorder', *PLoS ONE* 8, no 12, 2013.

Sanderse W, 'An Aristotelian model of moral development', *Journal of Philosophy of Education* 49, issue 3, Aug 2015, www.jubileecentre.ac.uk/media/news/article/43 12An-Aristotelian-Model-of-Moral-Development (accessed 23 Aug 2017).

Scott R and Reynolds L, *Digital Citizens: Countering extremism online*, Demos, 2016, www.demos.co.uk/project/digital-citizens/ (accessed 17 Aug 2017).

Seymour R, 'Schadenfreude with bite', *London Review of Books* 38, no 24, 15 Dec 2016, www.lrb.co.uk/v38/n24/richard-seymour/schadenfreude-with-bite (accessed 17 Aug 2017).

Spies Shapiro LA and Margolin G, 'Growing up wired: social networking sites and adolescent psychosocial development', *Clinical Child and Family Psychology Review* 17 no 1, 2014.

SSAT, *Mind the Gap: Collected essays on the development of character, non-cognitive skills, mindfulness and well-being*, SSAT (The Schools, Students and Teachers Network) and the Young Foundation, 2014, https://webcontent.ssatuk.co.uk/wp-content/uploads/2015/10/13124740/Mind-the-Gap.pdf (accessed 17 Aug 2017).

Suler J, 2004, 'The online disinhibition effect', *CyberPsychology and Behavior* 7, no 3, 2004.

Taylor E, 'Thinking about the virtues of character measurement', Centre for Youth Impact, 2015, http://youthimpact.uk/sites/www.youthimpact.uk/files/3%20-%20Thinking%20about%20the%20virtues%20of%20character%20measurement.pdf (accessed 17 Aug 2017).

Timpson E, 'Funding boost for schools helping pupils develop character', press release, Dept for Education, 26 May 2016, www.gov.uk/government/news/funding-boost-for-schools-helping-pupils-develop-character (accessed 17 Aug 2017).

Vallor S, 'Social networking technology and the virtues', *Ethics and Information Technology* 12, 2010.

Vallor S, Flourishing on Facebook: virtue friendship and new social media, *Ethics and Information Technology* 14, issue 3, 2011.

Vromen A et al, 'Everyday making through Facebook engagement: young citizens' political interactions in Australia, the United Kingdom and the United States', *Political Studies* 64, no 3, 2016.

Walker D, 'Measuring character strengths and virtue among UK school children aged 14 and 15, 2014', paper given at conference Can Virtue be Measured?, Jubilee Centre for Character & Values, University of Birmingham, 2014, www.jubileecentre.ac.uk/userfiles/jubileecentre/pdf/conference-papers/can-virtue-be-measured/walker-david.pdf (accessed 17 Aug 2017).

Walker DI, Curren R and Jones C, 'Good friendships among children: a theoretical and empirical investigation', *Journal for the Theory of Social Behaviour* 46, no 3, 2016.

Walker DI et al, 'Adolescent moral judgement: a study of UK secondary school pupils', *British Educational Research Journal* 43, no 3, 2017.

Weber M and Ruch W, 'The role of a good character in 12-year-old school children: do character strengths matter in the classroom', *Child Indicators Research* 5, issue 2, Jun 2012.

YoungMinds, *Resilience for the Digital World: Research into children and young people's social and emotional wellbeing online*, 2016, https://youngminds.org.uk/resources/policy/resilience-for-the-digital-world/ (accessed 17 Aug 2017).

Demos – License to Publish

The work (as defined below) is provided under the terms of this licence ('licence'). The work is protected by copyright and/or other applicable law. Any use of the work other than as authorized under this licence is prohibited. By exercising any rights to the work provided here, you accept and agree to be bound by the terms of this licence. Demos grants you the rights contained here in consideration of your acceptance of such terms and conditions.

1 Definitions
A **'Collective Work'** means a work, such as a periodical issue, anthology or encyclopedia, in which the Work in its entirety in unmodified form, along with a number of other contributions, constituting separate and independent works in themselves, are assembled into a collective whole. A work that constitutes a Collective Work will not be considered a Derivative Work (as defined below) for the purposes of this License.
B **'Derivative Work'** means a work based upon the Work or upon the Work and other pre-existing works, such as a musical arrangement, dramatization, fictionalization, motion picture version, sound recording, art reproduction, abridgment, condensation, or any other form in which the Work may be recast, transformed, or adapted, except that a work that constitutes a Collective Work or a translation from English into another language will not be considered a Derivative Work for the purpose of this License.
C **'Licensor'** means the individual or entity that offers the Work under the terms of this License.
D **'Original Author'** means the individual or entity who created the Work.
E **'Work'** means the copyrightable work of authorship offered under the terms of this License.
F **'You'** means an individual or entity exercising rights under this License who has not previously violated the terms of this License with respect to the Work, or who has received express permission from Demos to exercise rights under this License despite a previous violation.

2 Fair Use Rights
Nothing in this licence is intended to reduce, limit, or restrict any rights arising from fair use, first sale or other limitations on the exclusive rights of the copyright owner under copyright law or other applicable laws.

3 License Grant
Subject to the terms and conditions of this License, Licensor hereby grants You a worldwide, royalty-free, non-exclusive, perpetual (for the duration of the applicable copyright) licence to exercise the rights in the Work as stated below:
A to reproduce the Work, to incorporate the Work into one or more Collective Works, and to reproduce the Work as incorporated in the Collective Works;
B to distribute copies or phono-records of, display publicly, perform publicly, and perform publicly by means of a digital audio transmission the Work including as incorporated in Collective Works; The above rights may be exercised in all media and formats whether now known or hereafter devised. The above rights include the right to make such modifications as are technically necessary to exercise the rights in other media and formats. All rights not expressly granted by Licensor are hereby reserved.

4 Restrictions
The licence granted in Section 3 above is expressly made subject to and limited by the following restrictions:
A You may distribute, publicly display, publicly perform, or publicly digitally perform the Work only under the terms of this License, and You must include a copy of, or the Uniform Resource Identifier for, this License with every copy or phono-record of the Work You distribute, publicly display, publicly perform, or publicly digitally perform. You may not offer or impose any terms on the Work that alter or restrict the terms of this License or the recipients' exercise of the rights granted hereunder. You may not sublicence the Work. You must keep intact all notices that refer to this License and to the disclaimer of warranties. You may not distribute, publicly display, publicly perform, or publicly digitally perform the Work with any technological measures that control access or use of the Work in a manner inconsistent with the terms of this License Agreement. The above applies to the Work as incorporated in a Collective Work, but this does not require the Collective Work apart from the Work itself to be made subject to the terms of this License. If You create a Collective Work, upon notice from any Licensor You must, to the extent practicable, remove from the Collective Work any reference to such Licensor or the Original Author, as requested.

B You may not exercise any of the rights granted to You in Section 3 above in any manner that is primarily intended for or directed toward commercial advantage or private monetary compensation. The exchange of the Work for other copyrighted works by means of digital file sharing or otherwise shall not be considered to be intended for or directed toward commercial advantage or private monetary compensation, provided there is no payment of any monetary compensation in connection with the exchange of copyrighted works.

C If you distribute, publicly display, publicly perform, or publicly digitally perform the Work or any Collective Works, you must keep intact all copyright notices for the Work and give the Original Author credit reasonable to the medium or means You are utilizing by conveying the name (or pseudonym if applicable) of the Original Author if supplied; the title of the Work if supplied. Such credit may be implemented in any reasonable manner; provided, however, that in the case of a Collective Work, at a minimum such credit will appear where any other comparable authorship credit appears and in a manner at least as prominent as such other comparable authorship credit.

5 Representations, Warranties and Disclaimer

A By offering the Work for public release under this License, Licensor represents and warrants that, to the best of Licensor's knowledge after reasonable inquiry:

i   Licensor has secured all rights in the Work necessary to grant the licence rights hereunder and to permit the lawful exercise of the rights granted hereunder without You having any obligation to pay any royalties, compulsory licence fees, residuals or any other payments;

ii The Work does not infringe the copyright, trademark, publicity rights, common law rights or any other right of any third party or constitute defamation, invasion of privacy or other tortious injury to any third party.

B Except as expressly stated in this licence or otherwise agreed in writing or required by applicable law, the work is licenced on an 'as is' basis, without warranties of any kind, either express or implied including, without limitation, any warranties regarding the contents or accuracy of the work.

6 Limitation on Liability

Except to the extent required by applicable law, and except for damages arising from liability to a third party resulting from breach of the warranties in section 5, in no event will licensor be liable to you on any legal theory for any special, incidental, consequential, punitive or exemplary damages arising out of this licence or the use of the work, even if licensor has been advised of the possibility of such damages.

7 Termination

A This License and the rights granted hereunder will terminate automatically upon any breach by You of the terms of this License. Individuals or entities who have received Collective Works from You under this License, however, will not have their licences terminated provided such individuals or entities remain in full compliance with those licences. Sections 1, 2, 5, 6, 7, and 8 will survive any termination of this License.

B Subject to the above terms and conditions, the licence granted here is perpetual (for the duration of the applicable copyright in the Work). Notwithstanding the above, Licensor reserves the right to release the Work under different licence terms or to stop distributing the Work at any time; provided, however that any such election will not serve to withdraw this License (or any other licence that has been, or is required to be, granted under the terms of this License), and this License will continue in full force and effect unless terminated as stated above.

8 Miscellaneous

A Each time You distribute or publicly digitally perform the Work or a Collective Work, Demos offers to the recipient a licence to the Work on the same terms and conditions as the licence granted to You under this License.

B If any provision of this License is invalid or unenforceable under applicable law, it shall not affect the validity or enforceability of the remainder of the terms of this License, and without further action by the parties to this agreement, such provision shall be reformed to the minimum extent necessary to make such provision valid and enforceable.

C No term or provision of this License shall be deemed waived and no breach consented to unless such waiver or consent shall be in writing and signed by the party to be charged with such waiver or consent.

D This License constitutes the entire agreement between the parties with respect to the Work licensed here. There are no understandings, agreements or representations with respect to the Work not specified here. Licensor shall not be bound by any additional provisions that may appear in any communication from You. This License may not be modified without the mutual written agreement of Demos and You.

Social media has moved from the periphery to the mainstream of public consciousness at a rapid pace, with close to 100 per cent of 16–24-year-olds now having a social media account. From being considered a relatively trivial and narcissistic aspect of online communication, social media is increasingly seen as socially and politically transformative – shaping human interaction from the micro level of individual relationships to the macro level of public discourse and, even, electoral outcomes.

This study examines young people's agency in their engagement with social networking sites. Rather than focusing on how to regulate the digital environment that surrounds young people, we aim to understand how young people act on social media, and crucially what motivates them to act in the ways that they do. To do this we take character as our frame of analysis – the personal traits, values and skills that guide individual conduct, and ultimately shape outcomes in school and in wider life.

Our study takes a particular interest in the moral and civic aspects of character, mapping the prevalence of risky and unethical behaviours on social media, while also assessing engagement prosocial behaviours, and displays of civic virtue. We use a scenario-based approach to drill deeper into young people's online reasoning, to understand what drives their behaviour, and unearth emergent peer-led codes of conduct. Finally, we ask what key stakeholders – schools, parents and social media companies – can do to support positive youth decision making.