

Appendix A – Summary of safety measures available on the largest social media networks in the UK

Facebook

- Options to limit who can see your posts and personal information
- Options to choose who can send you friend requests
- Block people – this prevents them from seeing you or you seeing their content.
- Reporting available to have content removed that breaches community standards (including violence, criminal behaviour, bullying and harassment, suicide and self-injury, child abuse, hate speech, false news and spam)
- Digital wellbeing – alerts users who have been on the app for more than two hours

Twitter

- Option to limit who can see your posts
- Options to hide replies to your tweets and filter those who can reply, also to filter which accounts you see in your notifications (e.g. only from those who you follow/follow you)
- Block accounts (so you cannot see each other) or mute accounts (so their posts are hidden from your timeline)
- Mute words so that you do not see tweets with those words in
- Issues a prompt to users who are about to send a message that algorithms believe could be harmful or offensive
- Trialling a range of features including “soft block” (allows people to remove followers without blocking them) and “safety mode” which will temporarily block accounts for seven days if it spots them using harmful language or sending repetitive, uninvited replies and mentions to the user who activated safety mode. Plans are also in place to trial the option to remove yourself from a public conversation.

Instagram

- Option to limit who can see your posts
- Option to hide offensive comments, including blocking certain words, phrases or emojis
- Option to turn off comments or block comments from certain accounts
- Banning @ mentions

- Digital wellbeing – alerts users who have been on the app for more than two hours

YouTube

- Restrict access to age-inappropriate content
- Range of video privacy settings
- Block users – prevent someone from commenting on your videos or sending you messages
- Comment moderation, to control who posts comments on videos

TikTok

- Digital wellbeing – alerts users who have been on the app for more than two hours
- Restricted mode to filter out inappropriate content
- Users under 18 have their accounts set to private by default
- Manage who can comment or direct message
- Minimum age to host a live stream is 16

Google search

- SafeSearch helps to filter explicit content from results
- 'Secure by default' protects from deceptive sites that might steal passwords or infect computers, and keeps data safe