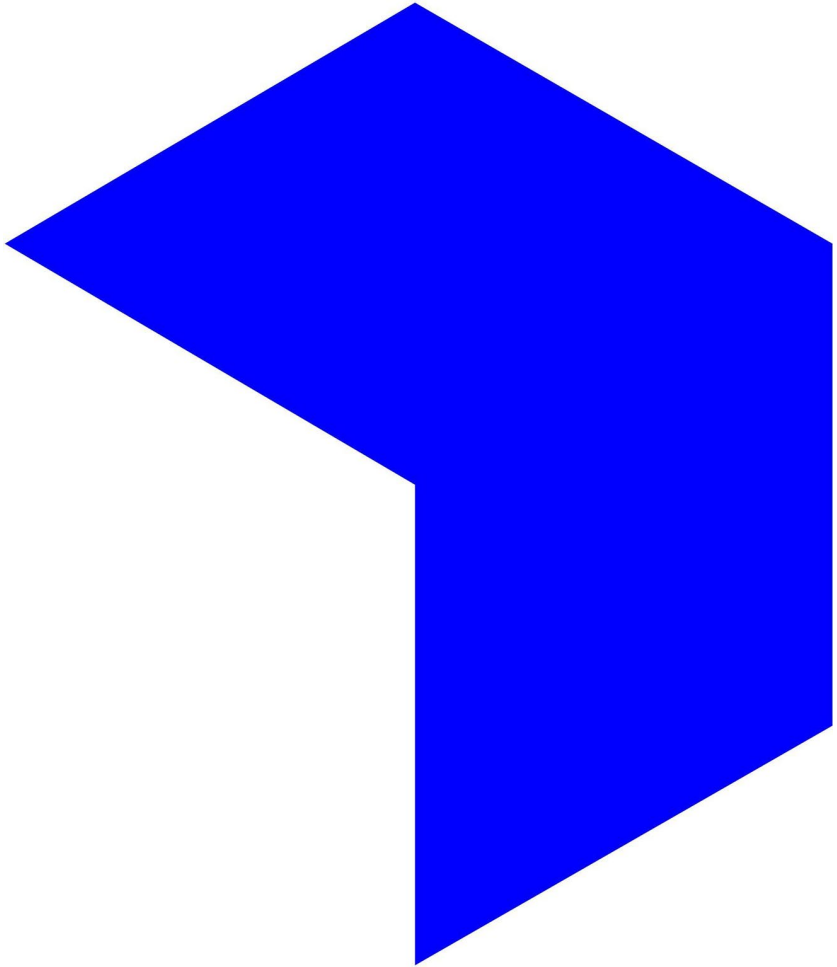
Behavioural Audit of Online Services

Key Findings Report

BIT

Ofcom

Contents



Executive summary

BIT

We examined the potential impact of platform design on user behaviour and online safety via a behavioural audit

BIT was commissioned by Ofcom to conduct a behavioural audit of six major online services[1] to **assess how the ways these platforms were designed may influence user behaviours and online safety**.

We examined **four key Areas of Interest (AoIs)**:

- Signing up to an online service
- Features and functionalities that affect time spent using the service
- Negative sentiment tools
- Reporting mechanisms

We examined these platforms from the **perspective of multiple user personas** including

- Children aged 13-15 years
- Children aged 16-17 years
- Adult users
- Users without an account

What is a behavioural audit?

A behavioural audit involves systematically mapping online design practices and evaluating their potential impact on user behaviours and outcomes.

We analysed platform settings, defaults, and friction points, capturing evidence through structured codebooks, screen recordings, and qualitative assessments to **identify patterns in platform design that may impact user behaviour**.

All information reflects the platform features and settings available **at the time of the audit**, which was conducted between December 2024 and January 2025. Subsequent updates or changes by platforms may mean that some details are no longer current.

[1] The six services chosen for the audit included Facebook, Instagram, Snapchat, TikTok, X and YouTube

We found that services are designed to maximise engagement, but safety measures are not easily accessible

The design of platforms may encourage users to **accept default settings during sign-up**, while offering **limited transparency and few opportunities to customise how their data is used**.

Platforms are designed to promote engagement, often limiting users' ability to control how content is personalised, manage notifications, or regulate their time spent online.

Notifications are widely used to encourage re-engagement. Most platforms default notifications to **"on" across multiple categories**, and adjusting these settings often requires multiple steps.

Tools for **indicating preferences about content**—such as "not interested" or "hide post"—are available on all platforms, but their effects are **not always clear to users**. While stronger negative sentiment tools, such as 'block,' are available, **their design may discourage frequent use**.

Significant friction in the reporting process, with complex categories and limited feedback, may make it harder for users to navigate reporting options and understand outcomes.

Time management and well-being features are included to help users monitor their usage, but their **visibility and accessibility vary**. Even where limits exist, they can often be easily **dismissed or bypassed with minimal effort**, and they are **rarely defaulted on for adults**.

There are **minimal differences between child and adult accounts across services**. Many platforms provide **limited signposting to parental oversight tools and tailored support for children**, which may make it difficult for parents to find and activate supervision tools, and for children to access relevant guidance.

Background

BIT

Ofcom commissioned BIT to undertake a behavioural audit of online services

Under the Online Safety Act 2023, Ofcom regulates online services to protect users, particularly children, from online harms. This research is intended to contribute to the evidence base used to **inform the development of Codes of Practice and guidance relating to Illegal Harms, Protection of Children**, and the additional duties on 'categorised' online services under the OSA. In particular, **it helps to build Ofcom's understanding of design practices in relation to specific aspects of online safety measures.**

In this project, Ofcom commissioned BIT to conduct a **behavioural audit of popular social media and video-sharing platforms** (VSPs).

This audit examined how platform design influences user behaviour, with a particular focus on four core areas of interest (AoIs). Establishing a baseline of current practices in these areas offers important contextual evidence for Ofcom's ex-post evaluation of codes of practice.

What is a behavioural audit?

A behavioural audit involves systematically mapping online design practices and evaluating their potential impact on user behaviours and outcomes.

The audit reviewed online choice architecture of six popular online services in the UK

Six online services were selected for the audit by Ofcom based on usage data[1] including the number of users as well as the average time spent on each service.

The audit covered a range of content feeds, including algorithmically-curated home feeds, short-form video feeds, and the search and discovery feed to assess **how online choice architecture (OCA) impacts user behaviour**.

The full details of how the audit was conducted can be found in the Technical Report.

What is online choice architecture?

Online choice architecture (OCA) refers to the design of digital environments that influences how individuals make decisions and interact with online platforms. Based on the Competition and Markets Authority (CMA) definition, this comprises three components[2]:

- **Choice structure:** how options are designed and presented
- **Choice information:** how users receive information about their choices
- **Choice pressure:** indirect influences affecting decisions

OCA impacts user behaviour through design elements such as option order, default settings, and the complexity of accessing controls. It also includes how information is presented (e.g., clear vs. dense terms of service) and features that apply pressure, like time-limited offers.[2]

[1] Ofcom. (2024). Online Nation 2024 Report. Retrieved March 6, 2025 Available here
[2] Competition & Markets Authority. (2022). Online Choice Architecture How digital design can harm competition and consumers. Available here

The audit examined online choice architecture practices, from potentially manipulative to supportive designs

This report examines a wide range of online choice architecture (OCA) practices, including those that may be **dark** or **grey** in nature.

Dark patterns are intentional design choices that manipulate or deceive users into making decisions that are not in their best interests.

Grey patterns have a more ambiguous impact, as their effects depend on the user's preferences and context—sometimes nudging users in directions that may not align with their best interests, while in other cases enhancing their experience.

Bright patterns are designs that foster trust, loyalty, and respect between users and platforms. These practices help users make informed, intentional choices without pressure, promoting a more balanced and user-friendly experience.

Four key Areas of Interest (AoIs) were identified

AoI	Signing up to an online service	Features and functionalities affecting time spent using the service	Negative sentiment tools	Reporting mechanisms
Research questions	• What data is collected from the user during sign-up, why, and with whom is this shared? • How are terms and conditions (T&Cs) presented to users during sign-up? • Where are community guidelines found? • When are users provided with content control tools? • How do platforms ask users how old they are?	• What features and functionalities that might influence time spent online are present in each platform being audited? What notifications are sent to users within the app? • Are there settings to turn off any of these functionalities and how are they presented? • Are there measures that help users manage their time on the platform?	• What tools are available to publicly and privately express negative sentiment, and how are these presented to users? • What are the mechanics of using negative sentiment tools? • What feedback is provided to the user after using a negative sentiment tool?	• Does a reporting function exist? How is it presented to users? • What are the mechanics of reporting? • What instant feedback is provided to a user about a report?

Across all AOIs		
	Help centre	Is there a Help Centre available? Is there support available within a Help Centre to help users with this feature?
	Discrepancies between child and adult accounts	Are there differences between the experiences of children (13-15 year olds and 16-17 year olds) and adults on the platform with respect to these AoIs?

BIT

The audit used four researcher accounts to examine differences in platform design and user experience

To conduct a comprehensive audit of online services, we created four distinct researcher accounts, each designed to simulate different user experiences and test platform functionalities across various user types. These accounts were systematically used to explore differences in platform design, content delivery, and available features, and included:

- **13–15 year old child account** – simulated the experience of a younger teen user
- **16–17 year old child account** – represented an older teen user
- **Adult account** – examined the experience of an adult registered user
- **A user without an account** – used to analyse what content and functionalities are accessible without an account

These user journeys allowed us to systematically compare platform behaviour across different user types, identifying key differences in browsing experiences, safety features, and engagement strategies.

The audit was conducted on an Android smartphone using each platform's app for the adult user, the 16–17-year-old child, and the 13–15-year-old child. For the user without an account, the audit was conducted using a Chrome web browser.

Signing up to an online service

1/2 Users are encouraged to move through the sign-up process quickly

What we found

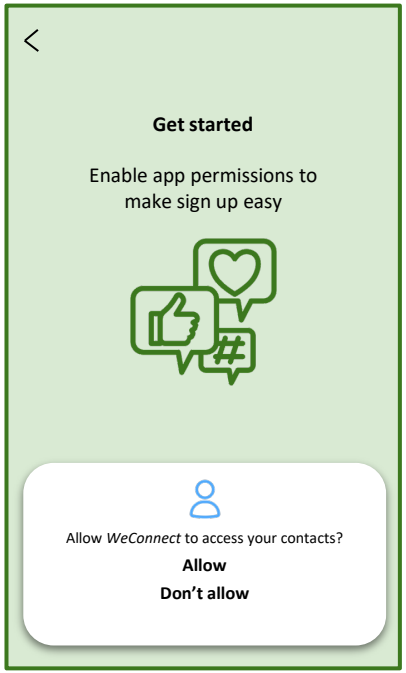
The design of the platforms encouraged users to **complete signing up quickly**, often by reducing friction using passive consent mechanisms, and providing reassurances that settings (such as who can see one's date of birth) could be changed later. Across platforms we observed:

Minimal barriers to proceed - platforms did not require users to view, scroll through or acknowledge Terms of Service (ToS) before continuing.

5/6 **platforms had ToS presented as small hyperlinks with no forced engagement.**

1/6 **platforms were set up such that acceptance of terms was implied by continuing, without a dedicated confirmation step.**

Encouraging users to allow device-level permissions - when it came to device-level permissions such as allowing notifications or syncing contacts, platforms frequently presented this as the required next step in the process, using language such as 'Enable permissions to make sign up easy'. While users were provided a choice (by the device) on whether to allow this or not, it is framed as a requirement by the platform messaging.



Stylised graphic showing a typical request from a platform (called WeConnect for purposes of illustration here) to allow device-level access.

Sign-up

2/2 Users are encouraged to move through sign-up quickly

What we found

Large, clear buttons to continue vs small or less visible opt-outs - The 'Next' or 'Continue' buttons were visually prominent and reinforced progress, whereas options to modify settings (such as ad-personalisation) were not made as salient.

Sequential flow without easy exit – Many platforms structured the sign-up flow in a way that encouraged users to keep moving forward rather than stopping to review settings or explore alternative options.

Reassurances that settings could be adjusted later - Some platforms reassured users that choices made during sign-up (e.g., visibility of date of birth or email address to other users) could be changed later.

Why this matters

Platforms may design sign-up flows to be **quick and seamless** to reduce drop-off rates and encourage more users to complete registration. Using **passive consent mechanisms** and reassuring users that settings can be changed later can make the process feel effortless.

However, this design may discourage users from reviewing or adjusting their privacy and security settings. This could lead to prolonged and excessive data sharing or reduced control over their account. The emphasis on progress and minimal friction may also reduce the likelihood that users explore important settings, limiting their ability to make **informed privacy choices**.

Sign-up

Minimal transparency on data collection and sharing

What we found

Most platforms framed data collection as necessary for improving user experience, but **transparency about how user data was used—and whether it was shared with external entities—was limited**. Users were often given reassurances about privacy, but critical details about data processing, ad targeting, and third-party access were typically only found in the ToS rather than being clearly presented during sign-up.

6/6 **platforms used personal data for targeting advertising[1], but only 2 platforms explicitly note in their messaging during sign-up that they are ad-funded.**

3/6 **platforms share user data with third parties such as advertisers, service providers, or business partners[1].**

Why this matters

When transparency is limited, users may not fully understand how their data is processed, shared, or used for advertising. Users **may assume their data is only used for platform functionality** rather than being shared or used for ad targeting.

The design **reinforces passive consent**—users agree to data processing without being explicitly told how their information is handled.

BIT

[1] based on an AI-assisted review of their ToS

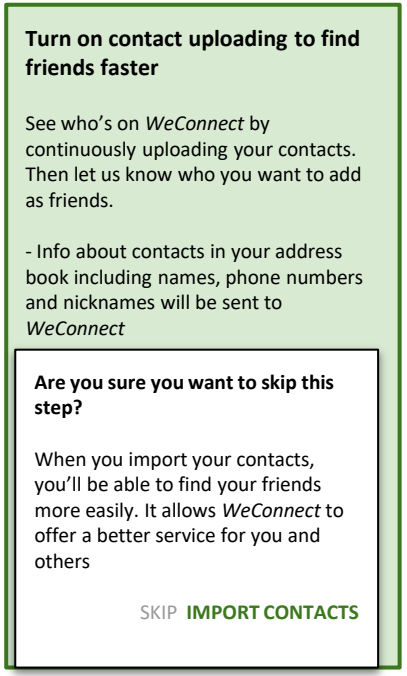
Engagement-focused defaults and friction reduce user control over privacy

What we found

The design of many platforms **steered users towards engagement-driven settings**, such as contact syncing, personalised ads, and notifications, by framing them as beneficial or making them the default.

In contrast, **adjusting settings for greater privacy or disabling certain features often required multiple steps**.

For example, contact syncing was framed as a way to 'find friends easily', and the option to decline this was less visible, or required extra steps such as tapping 'Not now', and then dismissing a further screen inviting them to reconsider their decision.

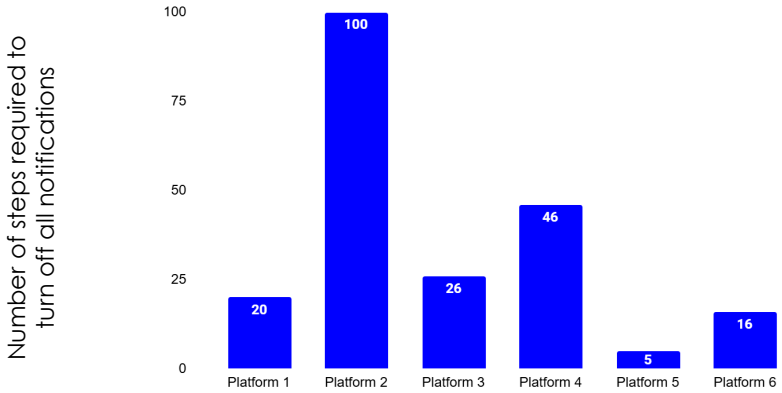


Stylised graphic showing a screen inviting users to reconsider their decision to not allow contact uploading (called WeConnect for purposes of illustration here)

Defaults and friction limit control over notifications

What we found

On all platforms, notifications were defaulted to be turned on at sign-up, and turning these off in the app typically required multiple steps.



Why this matters

Platforms often design default settings to **encourage engagement and connectivity**, framing features like contact syncing and personalised ads as beneficial. Keeping notifications and other engagement-driven settings **on by default** can help promote platform interaction.

However, these defaults may lead users to **unintentionally share more data** or receive more notifications than they would prefer. High-friction processes—such as requiring multiple steps to opt out—reinforce these defaults, making it harder for users to adjust their settings. This can result in reducing user control over their experience.

Sign-up

Gaps in content controls

What we found

None of the platforms audited gave users the option to limit exposure to sensitive content during sign-up. While some provided content controls in settings, these were not signposted at sign-up. Others lacked clear options to reduce sensitive content exposure altogether, requiring users to navigate multiple menus post-sign-up.

0/6 **platforms offered users the ability to reduce their exposure to sensitive content during sign-up.**

Why this matters

Platforms may prioritise **a broad content experience** during sign-up to keep users engaged and exposed to a wide range of material. Introducing content restrictions later—rather than upfront—can reduce friction and simplify onboarding.

However, without **immediate content controls**, users are **defaulted into broad exposure**, leaving them with **limited ability to tailor their experience from the start**. This may be particularly concerning for **children and vulnerable users**, who could benefit from clearer upfront choices about the type of content they see. Under the provisions of the Online Safety Act (2023) 'categorised' services will need to give users the ability to control the content they see at the earliest possible opportunity[1].

[1] Ofcom. (2024). Implementing the Online Safety Act: Additional duties for 'categorised' online services. Retrieved April 30, 2025 Available here

Sign-up Discrepancies between child and adult accounts

Gaps in communication of age restrictions

What we found

Age restrictions were implemented through **self-reported age verification during sign-up**, with users required to confirm they were above the platform's minimum age to create an account.

The specific minimum age requirements were not clearly communicated to users, and **no further verification measures** were implemented to confirm users' age.

Why this matters

Platforms typically rely on **self-reported age verification** during sign-up to enforce age restrictions while maintaining a smooth onboarding process. However, without **clear communication of minimum age requirements** or additional verification measures, users may **not fully understand these restrictions**.

Relying solely on **self-reported ages** is not a **highly effective** age assurance mechanism, as it can be easily bypassed, potentially exposing younger users to age-inappropriate experiences.

Discrepancies between child and adult accounts

BIT

Discrepancies between child and adult accounts

Variability in protections for under-18 users

What we found

On most platforms, there was **no differentiation between child and adult accounts at sign-up**, meaning younger users did not receive additional privacy or safety prompts.

1/6 **platforms provided additional privacy-focused onboarding content for younger users.**

Parental supervision tools existed across platforms, but were **not signposted during account creation**.

5/6 **platforms offered parental control options; however, none were signposted during sign-up.**

Why this matters

Platforms may choose to keep the **uniformity in sign-up process** for all users to reduce complexity and streamline onboarding. However, without clear **differentiation between child and adult accounts**, younger users may not receive **additional privacy or safety prompts** that could help protect them online.

By **not highlighting parental controls** at sign-up, platforms **miss a key opportunity to support parents** in managing their children's online activity. Instead, **parents must actively seek out these tools later**, which may reduce their uptake.

Additionally, **gaps in protections** leave children more vulnerable to **excessive screen time, age-inappropriate content, and engagement-driven design** that prioritises interaction over safety.

BIT

Discrepancies between child and adult accounts

Variability in privacy for under-18 users

What we found

On some platforms, child accounts were defaulted to private, while adult accounts defaulted to public—though this distinction was not always made clear during sign-up. On the other platforms, both child and adult accounts had the same privacy default settings[1].

3/6 **platforms defaulted children to private accounts.**

[1] In some cases, both were set to public. In others, there was no binary public/private option, and privacy was instead managed through default audience settings that varied by data type

Why this matters

Platforms do not consistently apply default-on privacy settings, such as limiting the visibility of personal information, restricting interactions with unknown accounts, or setting profiles to private by default. As a result, **potential risks are increased, and the responsibility to create safer online environments often falls to users or caregivers.**

Some platforms implement **stronger privacy defaults for child accounts**, but these protections are not always clearly communicated during sign-up.

BIT

Features and functionalities that affect time spent on online services

BIT

Browsing designed for continuous engagement

What we found

Platforms deployed **algorithmic content delivery, feed refresh mechanisms, and notifications** in the feeds audited to sustain user engagement.

Video and short-form content auto-played as users scrolled, reducing friction to continued engagement.

Feeds frequently refreshed with new content for users who were signed-in (but typically not for users who were not registered for an account).

Why this matters

Continuous engagement design helps to create a seamless and personalised experience, making platforms more dynamic and appealing.

However, these design features can make it harder for users to **regulate their time online**, as content is constantly refreshed and readily available. Users may not always be aware of how these features shape their engagement, potentially leading to longer-than-intended usage.

Limited control over content algorithms

What we found

Users had **few direct tools to shape their content feeds**, so we assume that most platforms rely on indirect signals such as engagement patterns to refine recommendations. This meant that users had to **actively interact** with content for their recommendations to evolve.

Most platforms offered **'Not interested' or keyword filtering** as a way to adjust content, but **many did not provide a clear way to actively select preferred topics**.

2/6 **platforms provided users with positive content selection tools such as 'favouriting' specific topics or themes.**

Several platforms offered **algorithm resets** enabling users to reset their content recommendations. However, there was limited information on how these worked

BIT **3/6** **platforms allowed users to reset their algorithm.**

Why this matters

This approach of using engagement to drive recommendations helps personalise content dynamically, but it also means users must **interact** with content to shape their experience.

Limited direct control over content algorithms makes it harder for users to tailor their feeds beyond reacting to unwanted content. The lack of clarity around **algorithm resets** may discourage users from using them, as the impact of resetting recommendations is often unclear. Additionally, requiring engagement to adjust recommendations can unintentionally **reinforce content bubbles**, reducing exposure to a diverse range of content.

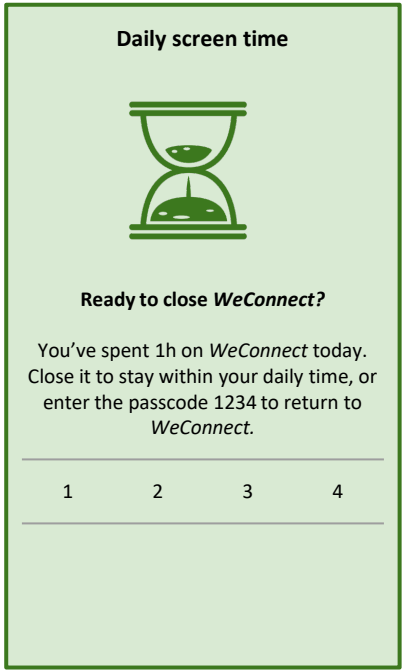
1/2 Time-management tools available but easily bypassed

What we found

Most platforms offered screen-time management features, but these were **typically defaulted 'off' for adult users** and required users to opt-in manually. When it came to children, several platforms automatically enabled screen time management tools.

4/6 **platforms provided screen-time management features, but these had to be manually activated in settings.**

3/6 **platforms had screen-time management features defaulted 'on' for child users, typically set to one hour of screen-time per day.**



Stylised graphic showing a reminder screen that pops up when a user reaches their screen time limit for the day.

2/2 Time-management tools available but easily bypassed

What we found

The screen time limits set on platforms could be bypassed with minimal effort, for example by closing the alert that notified users they had reached their screen time limit, or by entering a simple passcode. As such, these tools functioned as **reminders** rather than **limits** when it came to screen time.

None of the platforms signposted to these tools during sign-up.

Why this matters

Platforms offer **screen-time management tools** to help users regulate their time online, with some enabling these by default for children. These tools are designed to promote **healthier digital habits**, but for adults (and in some cases for children), they require **manual activation**, which may reduce their uptake.

Even when enabled, **time limits are easy to bypass**, often functioning more as reminders than enforced restrictions. Users can often dismiss alerts or override limits with **minimal effort**, reducing their effectiveness. Additionally, as these tools are not signposted during sign-up, many users may remain unaware of them.

Negative sentiment tools

Stronger tools are harder to access

What we found

Only a few platforms provided direct access to all negative sentiment tools from the primary menu (accessible with a single click from the post).

2/6 **platforms allowed users to access all negative sentiment tools from the primary menu.**

Stronger actions, such as **blocking a user, often couldn't be accessed directly from posts** and required multiple steps to access.

Additionally, some platforms introduced confirmation prompts (asking users "Are you sure?") before applying stronger actions, like blocking, adding friction to the process.

Why this matters

Platforms may introduce **frictions in accessing stronger moderation tools**—such as blocking or muting—to prevent accidental actions and encourage users to consider less drastic options first. Confirmation prompts and additional steps can help reduce misuse.

However, making **blocking and muting harder to access** can delay users from managing unwanted interactions effectively. This can lead to a **less controlled and comfortable experience**, particularly when dealing with **harassment or harmful content**. Requiring multiple steps may discourage and stop users from using these tools, leaving them exposed to **unwanted content or interactions for longer**.

Negative sentiment tools

Following is easier than unfollowing

What we found

Most platforms made following or subscribing to content creators a seamless, one-click action.

However, **unfollowing or unsubscribing often required extra steps**, such as visiting the creator's profile instead of removing them directly from the feed.

Some platforms used different icons or placed the buttons in different locations when users wanted to subscribe versus unsubscribe, which could make the process more confusing.

2/6 **platforms used identical processes to follow/unfollow and to subscribe/unsubscribe.**

Why this matters

Platforms may design **following and subscribing** as seamless, one-click actions to encourage engagement and content discovery. By contrast, **unfollowing or unsubscribing** often requires additional steps, which may help prevent accidental removals, but also introduces friction.

This imbalance can make it **easier to accumulate followed accounts** than to remove them over time, making it harder for users to **curate a feed that reflects their evolving interests**. Design inconsistencies—such as different buttons or icon placements for following and unfollowing—can add further confusion. Requiring extra steps to unfollow may **discourage users from refining their feeds**, leading to prolonged exposure to content they no longer wish to see.

BIT

Negative sentiment tools

Lack of detailed feedback on negative sentiment tools

What we found

Most platforms provided instant feedback after using negative sentiment tools like "Not Interested," however this was often limited to confirmation messages rather than clear explanations of how the content would be adjusted or what changes users could expect.

Additionally, feedback was not always consistently available across all feeds.

In some cases, users received no confirmation that their action had taken effect, creating uncertainty about whether their preferences had been registered or would influence their feed.

Why this matters

Platforms may provide **minimal feedback on negative sentiment tools** to keep the user experience simple and avoid overwhelming users with excessive information. However, when platforms offer only basic confirmation messages—or no message at all—users may be **uncertain about whether their content preferences have been applied**.

This lack of transparency can lead to **frustration or distrust** in platform controls, particularly if users feel their actions have no meaningful impact on their feed. Inconsistent feedback across different feeds further complicates user expectations.

BIT

Negative sentiment tools

Follow-up steps provided inconsistently after tool use

What we found

After expressing negative sentiment, some platforms guided **users towards stronger tools**, such as blocking, muting or updating content preferences.

Some platforms offered a "Learn more" link after certain actions, but this was not a standard feature across all tools.

3/6 **platforms did not provide users with guidance on additional steps they could take after using any tool.**

One platform **prompted younger users to consider parental supervision after blocking an account**, demonstrating an opportunity for more platforms to reinforce digital well-being tools and encourage parental involvement in children's online experiences.

Why this matters

Platforms may choose to keep **negative sentiment actions simple** to avoid overwhelming users with too many options at once. However, without clear follow-up suggestions at key moments, users may encounter **unnecessary friction and lack of information on their options when trying to take stronger action**—such as blocking, muting, or adjusting content preferences.

When pathways from basic negative tools to stronger controls are unclear or hard to access, it can **limit users' ability to manage their online experience**. Poor signposting and limited visibility of available tools **reduce users' control over the content they see and the interactions they have.**

Reporting mechanisms

BIT

Reporting content requires multiple steps

What we found

Reporting options were typically nested within menus, requiring multiple steps to access.

The number of reporting categories and subcategories varied across platforms.

The number of reporting categories and subcategories ranged from **10-42**

Only one platform explicitly reassured users that reports would still be reviewed even if they were misclassified.

2/6 **platforms provided users with an 'Other' category and a free-text box to outline their issue.**

Why this matters

Platforms may structure **reporting options within menus** to streamline the interface and avoid accidental misuse. Providing **detailed reporting categories** can help platforms **process reports more effectively**, but if the system is too complex, it may discourage users from reporting harmful content.

If the **reporting process is difficult to locate or navigate**, users may be **less likely to take action**, weakening moderation efforts. Even when appropriate information is provided, comprehension barriers—such as complex language or unclear instructions—can deter users from reporting. Unclear or overly detailed categories without guidance can also create **uncertainty**, making users hesitant to report or leading them to select incorrect options—potentially **delaying appropriate action**.

Non-registered users have restricted ability to report

What we found

Many platforms required users to be registered in order to report content, preventing external parties from flagging concerning material they encounter.

One platform directed users to ask a friend with an account to report the content on their behalf.

3/6 **platforms required an account to report content, while the others only allowed non-users to report via a Help Centre form.**

Why this matters

Platforms may require users to be **registered to report content** as a way to **reduce spam or bot-driven reports** and ensure accountability. However, this restriction also **limits external oversight** and prevents non-users—such as parents or guardians—from flagging concerning material.

Requiring users to **rely on others to report on their behalf** could **unintentionally expose more people to harmful content**.

Limited feedback on reporting outcomes

What we found

Users did not consistently receive clear instant feedback on the successful submission of their reports.

On some platforms, selecting certain reporting categories redirected users to actions like “Not Interested” rather than submitting a formal report.

2/6 **platforms included reporting categories which did not submit a formal report.**

While all platforms displayed confirmation messages upon report submission, some offered no follow-up in the app, or indication of whether any action had been taken.

Why this matters

Platforms may limit feedback on reporting outcomes to **simplify the user experience** and avoid overwhelming users with updates. However, when users **do not receive clear confirmation** that their report has been submitted or reviewed, they may be **uncertain whether their concerns have been acknowledged or addressed**.

In some cases, selecting certain reporting categories **redirects users to less impactful actions**, such as marking content as "Not Interested," rather than submitting a formal report. This can be confusing, and make users feel that their efforts to flag harmful content are **ineffective**.

Help centre and safety tools

BIT

Limited signposting of support and safety tools

What we found

Although **most platforms offered safety tools, such as parental controls and time management features**, these resources were **not highlighted during sign-up**. The visibility and accessibility of these safety features varied significantly across platforms.

5/6 **platforms offered safety tools.**

0/6 **platforms signposted safety tools during sign-up.**

Additionally, while all platforms had a Help Centre, there was little standardisation in how they were labelled or organised—terms like 'Help Centre','Safety Centre', or 'Family Centre' were used inconsistently, which may confuse users and make key safety information harder to find.

Why this matters

Platforms may choose **not to highlight safety tools during sign-up** to keep the onboarding process simple and minimise friction. Instead, these features are often housed in help centres or settings menus, for users to seek them out when needed.

However, without **clear guidance during sign-up**, users may be **unaware of available safety tools**, leaving them more vulnerable to harmful content. If these features are **hard to find or require active searching**, users may **miss out on essential protections**.

Limited visibility can reduce the effectiveness of these tools, particularly for those who need them most. Inconsistent visibility across platforms may further limit their impact in **promoting user well-being**.

Authors

Dr Sujatha Krishnan-Barman, Isabel Kaldor, Laurence Fenn, Libby Woodhouse, Ruth Persian and Eva Kolker

Get in touch: sujatha.krishnan-barman@bi.team

Ofcom BI Hub: behavioural.insights@ofcom.org.uk