Reset.

Evaluation of Instagram's Processes for Risks to Minors

I

Summary

Summary

This report documents an evaluation of systems on Instagram to assess the risks posed to minors, including:

- Instagram’s Content Moderation System;
- Understandability of the platform for younger users;
- Instagram’s safety-by-design settings; and
- Instagram’s ad manager systems.

We discovered multiple issues that potentially do not comply with the Digital Services Act (DSA), including:

- Instagram under-moderates both pro-restrictive eating disorder content, pro-suicide, and/or pro-self-harm materials;
- There is a muted response to these materials when Instagram becomes aware of them via the user-reporting system; Instagram failed to respond to the majority of pro-restrictive eating disorder content and pro-suicide and/or pro-self-harm materials when it became aware of them;
- A 13-year-old would likely not understand the design and functioning of Instagram at the point of signing on because of the length of the policies they would be required to read and the deployment of dark patterns in the sign-up process; and
- Instagram’s safety center is not routinely accessible to young people in their first languages.

Table of contents

II

Introduction

Introduction

The Digital Services Act (DSA) aims to offer children and young people under 18 years old additional protection in the digital sphere.

- Recital 71 states that “the protection of minors is an important policy objective of the Union,” and describes platforms as accessible to minors when:
    - Its terms and conditions permit minors to use the service;
    - Its service is directed at or predominantly used by minors; or
    - Where the provider is otherwise aware that some of the recipients of its service are minors, for example, because it already processes personal data of the recipients of its service revealing their age for other purposes.

- Recital 71 goes on to state, “Providers of online platforms used by minors should take appropriate and proportionate measures to protect minors, for example, by designing their online interfaces or parts thereof with the highest level of privacy, safety and security for minors by default where appropriate or adopting standards for protection of minors, or participating in codes of conduct for protecting minors. They should consider best practices and available guidance, such as that provided by the communication of the Commission on A Digital Decade for children and youth: the new European strategy for a Better Internet for Kids (BIK+). Providers of online platforms should not present advertisements based on profiling using personal data of the recipient of the service when they are aware with reasonable certainty that the recipient of the service is a minor.”

- Recital 81 further indicates that very large online platforms should consider, for example, “how easy it is for minors to understand the design and functioning of the service, as well as how minors can be exposed through their service to content that may impair minors’ health, physical, mental, and moral development.” Such risks may arise, for example, in relation to the design of online interfaces that intentionally or unintentionally exploit the weaknesses and inexperience of minors or which may cause addictive behavior.

- Recital 84 explains that in assessing systemic risk – which includes risks to minors – “providers of very large online platforms and of very large online search engines should focus on the systems or other elements that may contribute to the risks, including all the algorithmic systems that may be relevant, in particular their recommender systems and advertising systems, paying attention to the related data collection and use practices.”

- In addition, Article 34 places additional requirements on Very Large Online Platforms (VLOPS) and Very Large Online Search Engines to assess the risks their services pose to children’s rights. Specifically, Article 34(1)(d) DSA requires VLOPs to undertake risk assessments, including “any actual or foreseeable negative effects in relation to [...] minors.” Article 34(2)(b) DSA explicitly states that algorithmic recommender systems, content moderation systems, enforcement of terms and conditions, and advertising systems be considered.

This report explores Instagram’s compliance with the requirements outlined in these recitals and articles. Specifically, it evaluates four systems on Instagram for compliance:

1. **Content moderation systems:** A method is proposed for testing and evaluating these with regards to creating risks to minors. Specifically, it describes the method used to evaluate if platforms remove content that is harmful to minors when they become aware of it through user-reports. It describes the methods and presents findings from a September 2023 experiment around reporting and monitoring two bodies of content that were assessed by a clinical psychologist and deemed to be harmful to children:
   a. Pro-suicide and/or self-harm content;
   b. Pro-restrictive eating disorder content.

2. **Understandability for young people:** develop a simple method to evaluate *understandability* for young people and assess for dark patterns, meaning platforms' design decisions cumulatively nudge users to accept default choices that may be against their interests. It describes the methods and presents findings from a September 2023 analysis of three platforms, based on an analysis of the user journey when new accounts for minors are created.

3. **Safety-by-design settings:** draws on best practice and the BIK+ strategy. It assesses the user journey on Instagram, and the accessibility of help features on Instagram.

4. **Ad manager system:** a method for testing whether the platform allows advertising to minors based on profiling.

III

Evaluation of Instagram’s Content Moderation Systems in Creating and Perpetuating Risks to Minors

An Evaluation of Instagram's Content Moderation Systems in Creating and Perpetuating Risks to Minors

Research questions:

1. Does Instagram's adequately moderate pro-suicide and/or self-harm material when they become aware of it?

2. Does Instagram's adequately moderate pro-restrictive eating disorder material when they become aware of it?

Methodology

The research involved five steps:

1. **Developing criteria to define harmful material.**

- This research explored two bodies of content posing psychological and physiological risks to minors: pro-suicide and/or self-harm material, and pro-restrictive eating disorder material.
- We used the community guidelines for each platform to develop a coding schema to classify content (see Appendix 1 for more details). This ensures that only content violating Instagram's Terms of Service was included in this research. Each piece of content, according to their guidelines, should warrant a content-moderation action from Instagram.

2. **Identifying pro-suicide and/or self-harm material.**

Using simple searches, we identified content on Instagram that met our criteria and had not been labelled by the platform already. We consulted a clinical psychologist who assessed each piece of content that was identified, confirming that it presented a risk to young people who consume it. Material that was not deemed to be harmful by a psychologist was not included in this research.

In total we identified:

- Pro-suicide and self-harm content: 119 pieces
- Pro-restrictive eating disorder content: 125 pieces

See Appendix 2 for examples of these bodies of content.

3. **Monitoring content pre-reporting.**

We tracked this content for two weeks noting:

- View counts and growth rates;
- Labelling or warning rates, to ascertain whether any of this content was labelled by Instagram during these two weeks. We considered a piece of content labelled if an age-restriction warning, sensitivity filter, or any other sort of flag was placed on it; and
- Take down rates, to ascertain whether any of this content was taken down by Instagram during these two weeks.

4. **Reporting the content.**

We reported each piece of content as suicide and self-harm, or restrictive eating disorder content violating the Terms of Service to the platform.

5 Monitoring content post-reporting.

After reporting, we tracked this content for two further weeks noting:

- View counts and growth rates;
- Labelling or warning rates to observe if any content was labelled by the platforms during these two weeks. Considered labelled if an age-restriction warning, sensitivity filter, or any other flag was placed on it;
- Take down rates, to ascertain whether any of this content was taken down by the platforms during these two weeks.

According to our analysis of the platform's community guidelines (see Appendix 1), Instagram should delete pro-suicide and/or self-harm content, and pro-eating disorder content when they become aware of it. In practice, we often see platforms label and add sensitivity filters or age filters to this body of materials; we therefore also assessed these.

Below, we describe what we found over four weeks of monitoring.

Findings

Instagram's response to pro-suicide and/or self-harm material

Instagram does not appear to adequately label or demote pro-suicide and/or pro-self-harm content.

Removal appears to be the most common response to pro-suicide and/or pro-self-harm material, but the platform's reactions to reporting are inadequate. The majority of content remained available and unlabelled, even after user-reporting.

Over two weeks monitoring	Instagram
Pre reporting removal rate. This is the % of content that was removed during the two weeks before we reported it. It may have been reported by other users, and it is often not clear why content was removed (e.g. users may have deleted the content or their accounts, moved to private, or platforms may have deleted it). However this represents the best estimate of organic removal rates.	0 %
Post reporting removal rate. This is the % of content that was removed within 2 weeks after we reported it.	29.41 %
Effect of reporting on removal rate	**+29.41 %**
Pre reporting labelling or warning rate. This is the % of content that was labelled during the two weeks before we reported it. It may have been reported by other users, but represents the best estimate of organic labelling rates.	0 %
Post reporting labelling or warning rate. This is the % of content that was labelled within 2 weeks after we reported it.	0 %
Effect of reporting on labelling rate	**No change**
Pre reporting growth rate. This is the average growth rate of content over two weeks before we reported it (week-on-week).	0.63% growth week-on week
Pre reporting growth rate. This is the average growth rate of content over two weeks after we reported it (week-on-week).	0.1% growth week-on week
Effect of reporting on growth rate	**-0.5 %**

Instagram’s response to pro-restrictive eating disorder material

Instagram does not appear to adequately label or demote pro-restrictive eating disorder content.

Removal appears to be the most common response to pro-eating disorder material; however, Instagram’s reactions to reporting are inadequate. Most of the content remained available and unlabelled, even after user reporting.

Over two weeks monitoring	**Instagram**
Pre reporting removal rate. This is the % of content that was removed during the two weeks before we reported it. It may have been reported by other users, and it is often not clear why content was removed (e.g. users may have deleted the content or their accounts, moved to private, or platforms may have deleted it). However this represents the best estimate of organic removal rates.	0 %
Post reporting removal rate. This is the % of content that was removed within 2 weeks after we reported it.	10.40 %
Effect of reporting on removal rate	**+10.40 %**
Pre reporting labelling or warning rate. This is the % of content that was labelled during the two weeks before we reported it. It may have been reported by other users, but represents the best estimate of organic labelling rates.	0 %
Post reporting labelling or warning rate. This is the % of content that was labelled within 2 weeks after we reported it.	0 %
Effect of reporting on labelling rate	**No Change**
Pre reporting growth rate. This is the average growth rate of content over two weeks before we reported it (week-on-week).	3.46 % (week-on-week)
Pre reporting growth rate. This is the average growth rate of content over two weeks after we reported it (week-on-week).	3.09 % (week-on-week)
Effect of reporting on growth rate	**-0.38 %**

Limitations

When content is removed, the reasons for its removal can be unclear. The users could have removed it, they may have deleted their account or switched to the private mode, or the platform may have removed their content or accounts.

Therefore, the estimations for removal rates represent the highest-end estimations of removal rates by platforms.

Conclusion

- Instagram under-moderates both pro-restrictive eating disorder content, pro-suicide, and/or pro-self-harm materials.
- There is a muted response to these materials when Instagram becomes aware of them via the user-reporting system; Instagram does not appear to adequately remove, label, or demote pro-restrictive eating disorder content, nor pro-suicide and/or pro-self-harm materials.

IV

Evaluation of Young Users' Understandability of Instagram, Including Dark Patterns

Evaluation of Understandability of Instagram for Young Users, Including Dark Patterns

Research questions:

1. Could younger users understand the design and functioning of Instagram when they choose to use a service at the point of signing on?
2. Do younger users encounter any dark patterns at the point of signing on to Instagram that may cause them to act against their best interests or diminish their understanding of the platform's design or functions?

Methodology

This research involved five steps:

1. **Recording the sign-up process for several accounts with fictional 13-year-old identities, "sock puppet accounts," on Instagram.**

We set up accounts to record the user sign-on journey in:
a. Germany
b. Slovenia
c. The Netherlands

We noted and described the steps involved in this sign-up process, as described in Appendix 1.

2. **Recording and analysing for dark patterns in the sign-up process.**

Using previous research into platforms' sign-on processes,[1] informed by the experience of signing up to these platforms, we developed a six-point typology of dark patterns in sign-on processes, which is described below.

We assessed each step of the sign-on process for identifiable dark patterns.

3. **Recording and analysing policies referenced in the sign-up process for understandability.**

We analysed each policy that was referenced in the sign-on process and determined if it was understandable to younger users. We did this by considering three factors:

- Is the policy available in the first language of the minor?
- What is the length of the policy, and how long would it take to read?
- What is the reading age of the policy and is it possible for 13-year-olds to comprehend?

[1] Reset. Tech Australia 2021 *Did We Really Consent to This?* https://au.reset.tech/news/did-we-really-consent-to-this-terms-and-conditions-young-people-s-data/.

Findings

A typology of dark patterns in the sign-on experience

Dark patterns are design features that are intended to nudge users away from actions that align with their best interests and toward actions that are in the platform's interest.[2] Using previous research into the platform's sign-on processes,[3] and the experience of signing up to these platforms, we developed a six-point typology of dark patterns used in sign-on processes.

1 **Inferring consent by clicking next.** Rather than making it explicit that new users are agreeing to a platform's terms and conditions, they often design the mechanisms by which users consent as the next step in the process. For example, buttons or icons might say "next," "sign up now," or "choose your sign-up method," with small text underneath these buttons that inform new users that "by clicking this you agree to our terms." It may not be immediately obvious to new users that by clicking "next" or choosing their sign-on method they are entering into a contract with the platform.

2 **Obscuring important details.** Rather than attracting attention to and making new users aware that contractual terms and conditions or data processing requirements are involved, these are often obscured. For example, they may be presented in the smallest font, or at the very bottom of the screen.

3 **Presenting options that may not be in a user's best interests as a "better user experience."** Many platforms allow users to choose options that maximise potential data collection, such as syncing the app with contacts or connections to their new social media accounts with old social media accounts. These ensure that more data is collected by the platform, which may not always be in a user's best interest. Likewise, they allow users to choose whether to receive notifications, which may maximise the amount of time a user spends on the platform and habituate use. However, often, these options are presented either visually or using language to provide "a better experience," gently nudging the users to select them. For example, many requests to sync apps with phone contacts claim this makes the platform more entertaining, or requests to allow location data tracking claim this makes the app more effective.

4 **Visual promotion of options that are in a platform's best interests, while demoting options that are in the user's best interests**. Where users are provided with a choice, platforms often use visual techniques to promote one option and demote others. For example, buttons or icons that accept unnecessary data collection are often larger, more colorful, or otherwise more prominent, while options to skip or reject non-essential data collection are presented in smaller and less salient fonts.

5 **Presenting options that are in users' best interests as temporary.** Where users are provided with a choice, platforms often present the choices that might be in users' best interests as only temporary or a choice that the platform may force them to revisit. For example, displaying options to skip or reject non-essential data collection as "not for now" or "maybe later," and/or forcing users to return to these questions repeatedly.

6 **Click twice for no, but only once for yes.** When users are provided with a choice, and they select the choice that might be in their best interests—often declining unnecessary data collection—users are forced to select this twice. For example, if a user chooses to decline syncing apps, they may be presented with an additional step in the sign-on process where they are asked to reconsider or confirm this choice. "Clicking twice" is often not required if users select the choice that is in the platform's best interest.

These dark patterns are not mutually exclusive, and many designs employ multiple dark patterns; nor is this list comprehensive, and different typologies and dark patterns may emerge.

We discovered that dark patterns were prevalent by analysing the sign-on process on each app, as documented in Appendix 3

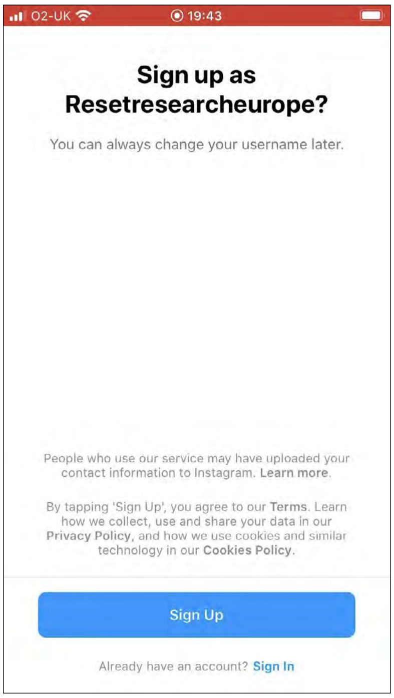
[2] Arunesh Mathur *et al.* 2019 "Dark Patterns at Scale: Findings from a Crawl of 11K Shopping Websites" *Proceedings of the ACM on Human-Computer Interaction* November, p. 81.

[3] Reset. Tech Australia 2021 *Did We Really Consent to This?* https://au.reset.tech/news/did-we-really-consent-to-this-terms-and-conditi ons-young-people-s-data/.

Dark patterns discovered in the sign-on experience

Instagram infers consent. Instagram did not explicitly ask young users to review and agree to the terms and conditions as part of the joining experience; instead, consent was inferred when a user signed up with a username.

(i.e., the app asks the new user if they are ready to sign up with the username they selected in the previous step; this is how they technically consent to the terms and conditions).



Sign up as Resetresearcheurope?

You can always change your username later.

People who use our service may have uploaded your contact information to Instagram. **Learn more**.

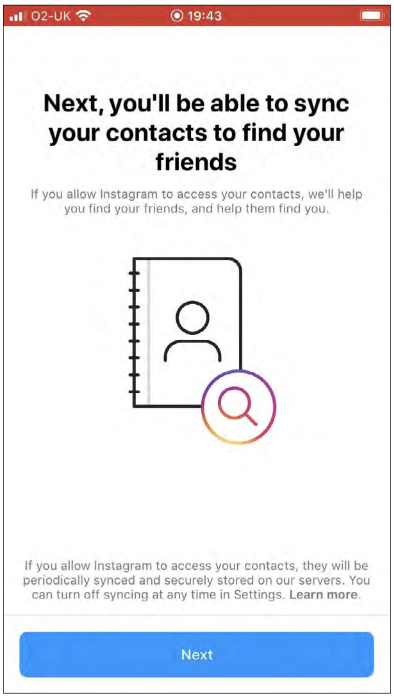
By tapping 'Sign Up', you agree to our **Terms**. Learn how we collect, use and share your data in our **Privacy Policy**, and how we use cookies and similar technology in our **Cookies Policy**.

Already have an account? **Sign In**

Figure 1: Screenshot of the sign-on process on Instagram

Instagram obscures details about terms and conditions.

The terms and conditions the user was asked to agree to were presented once, at the bottom of step 8. The font used to describe the contractual agreement was the smallest and the lightest gray font on the screen, although the names of the policies were in bold. Important details were also obscured when the app requested that users sync their contacts with the app in step 9 (see Figure 2, and also Figure 1).



Next, you'll be able to sync your contacts to find your friends

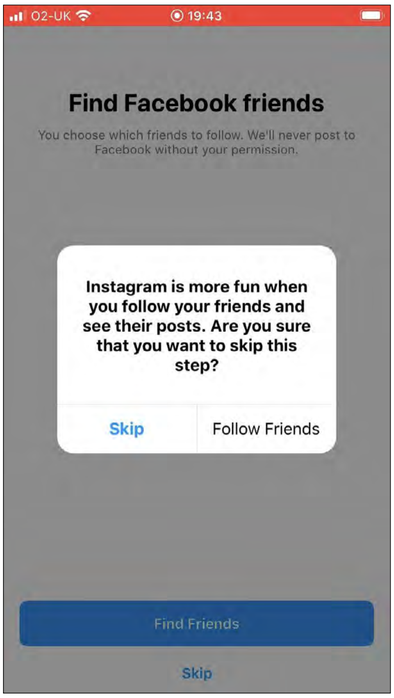
If you allow Instagram to access your contacts, we'll help you find your friends, and help them find you.

If you allow Instagram to access your contacts, they will be periodically synced and securely stored on our servers. You can turn off syncing at any time in Settings. **Learn more**.

Figure 2: Instagram obscuring details about data processing.

Instagram presents options that may not be in a user’s best interests as a “better user experience.”

Instagram presents the option to allow notifications as a better experience. The sign-on process states that allowing the app to sync with the user’s Facebook account makes Instagram “more fun.”


O2-UK 19:43

Find Facebook friends

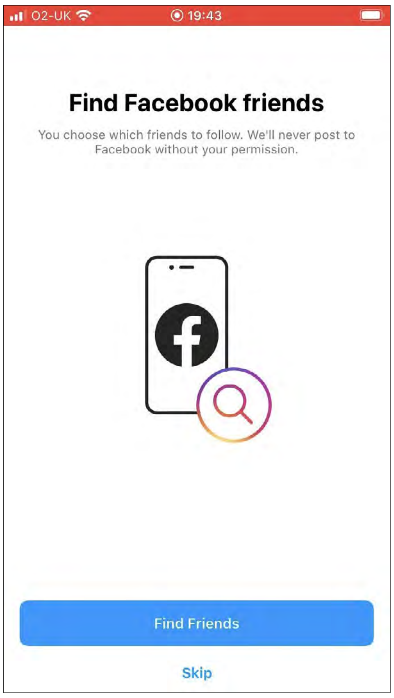
You choose which friends to follow. We'll never post to Facebook without your permission.

Instagram is more fun when you follow your friends and see their posts. Are you sure that you want to skip this step?

Figure 3: Instagram presents options that may not be in a user’s best interest as providing a better user experience.

Instagram visually promotes options that are in the platform’s best interests while demoting options in users’ best interests.

Five steps used visual cues to make choices more prominent for users. This included making the “Find Friends” and sync with Facebook buttons colorful while making the “Skip” option less noticeable and moving the “Next” button and making it less prominent than in other steps to encourage more scrolling.


O2-UK 19:43

Find Facebook friends

You choose which friends to follow. We'll never post to Facebook without your permission.

Figure 4: The more prominent “Find Friends” button.

Instagram did not present options that are in users’ best interests as temporary.

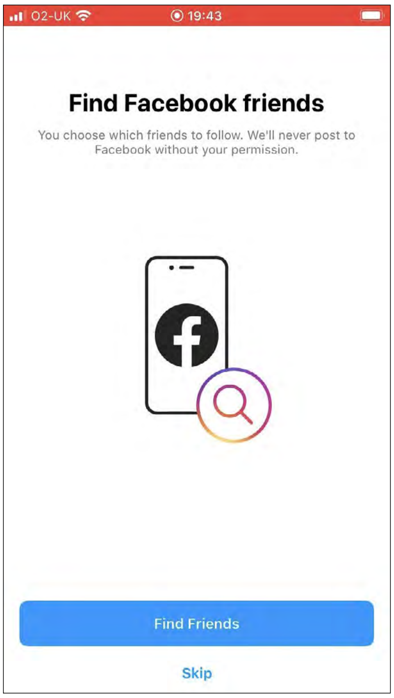
Instagram did not require 13-year-old users to click twice for no, but only once for yes.

On Instagram, new users are asked to click twice for:

- "Finding Facebook friends," where clicking "Skip" in the previous step still triggers a phone notification; and
- "Turn on notifications," where clicking "Skip" in the previous step also triggers a phone notification.

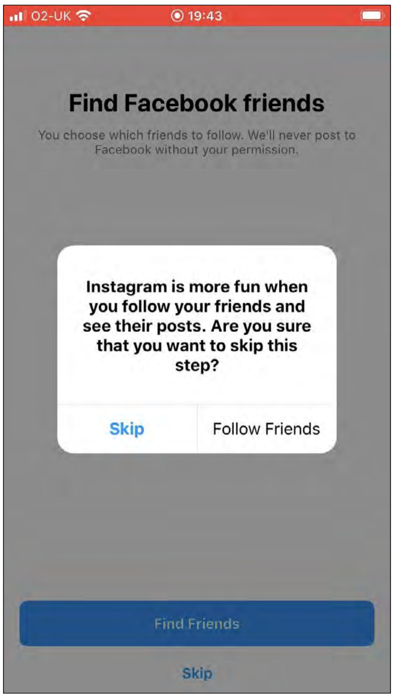
However, we did not include this in our count of dark patterns because it appears to be an example of poor user design rather than a dark pattern. Even if a user clicked on "Find Friends" or "Turn on notifications," the phone notification would be triggered. Other platforms handle this by adding a "Continue" button at the bottom of the previous step.

Figure 5: Examples of steps where new users are required to click twice to decline options but would also be required to click twice to accept them.



Find Facebook friends

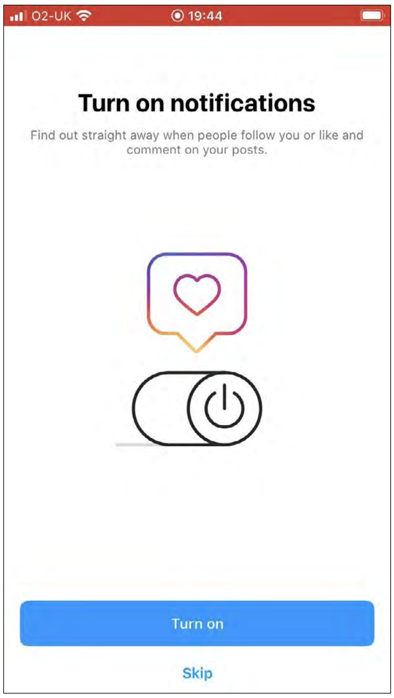
You choose which friends to follow. We'll never post to Facebook without your permission.



Find Facebook friends

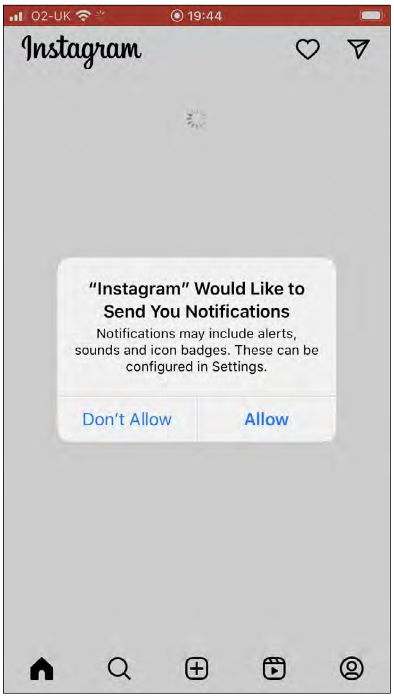
You choose which friends to follow. We'll never post to Facebook without your permission.

Instagram is more fun when you follow your friends and see their posts. Are you sure that you want to skip this step?



Turn on notifications

Find out straight away when people follow you or like and comment on your posts.



"Instagram" Would Like to Send You Notifications

Notifications may include alerts, sounds and icon badges. These can be configured in Settings.

Accessibility and comprehensibility of policies

During the sign-on process, Instagram indicated the policies users agreed to by joining the platform. These include:

1. Terms of service
2. Privacy policy
3. Cookies policy

We explored:

1. If the policies signposted in the process were available in an accessible language;
2. The length of the policies and the time taken to read them assuming an average reading speed of 225 words per minute (possibly an overestimation for a 13-year-old); and
3. The reading age of these documents according to the Flesch–Kincaid Grade Level test for English and Rix Score for non-English. Both the tests provide an interpretation of the school grade where the text would be understandable. Most 13-year-olds are in the seventh or eighth grade, depending on the country, and a grade score of 13+ reflects college or university level. (Note: the Rix Score test is not available for the Greek language.)

Terms of Service:

- Available in all 24 official European languages
- Average length of 3,402 words, it would take an average of 15:06 minutes for a young person to read it
- Average readability: grade 12.8

Privacy policy

- Available in all 24 official European languages
- Based on the sample of full privacy policies in English, Slovenia, and German, the average length of 25,419 words; it would take an average of 1 hr 59 minutes for a young person to read this
- Average readability: above grade 13

Cookies policy

- Available in all 24 official European languages
- Average length of 1,011 words, it would take an average of 4:30 minutes for a young person to read it
- Average readability: grade 12.8

		Instagram		
		Terms of service	**Privacy policy***	**Cookies use**
Bulgarian	Available	Yes	Yes	Yes
	World Count	4,020 17:52 mins	14,149 62:53 mins	1,175 5:13 mins
	Grade	13	11	13
Croatian	Available	Yes	Yes	Yes
	World Count	3,280 14:34 mins	12,234 54:22 mins	995 4:25 mins
	Grade	13	10	13
Czech	Available	Yes	Yes	Yes
	World Count	3,187 14 mins	11,650 51:46 mins	915 4 mins
	Grade	13	11	13
Danish	Available	Yes	Yes	Yes
	World Count	3,515 15:37 mins	12,469 55:25 mins	994 4:25 mins
	Grade	11	9	11
Dutch	Available	Yes	Yes	Yes
	World Count	3,698 16:26 mins	13,315 59:10 mins	1,086 4:49 mins
	Grade	13	10	13
English	Available	Yes	Yes	Yes
	World Count	3,399 15 mins	26,272 words, 1h 57 min	996 4:25 mins
	Grade	12	11	13
Estonian	Available	Yes	Yes	Yes
	World Count	2,839 12:37 mins	12,428 55:14 mins	824 3:39 mins
	Grade	13	9	13
Finnish	Available	Yes	Yes	Yes
	World Count	2,595 11:32 mins	9,578 42:34 mins	725 3:13 mins
	Grade	13	11	13
French	Available	Yes	Yes	Yes
	World Count	3,919 17:25 mins	15,346 68:12 mins	1,255 5:34 mins
	Grade	13	11	13
German	Available	Yes	Yes	Yes
	World Count	3,598 15:59 mins	26,365 1 hr 42 mins	1,060 4:42 mins
	Grade	13	12	13
Greek	Available	Yes	Yes	Yes
	World Count	3,802 17 mins	15,186 67 mins	1,1194 5 mins
	Grade			
Hungarian	Available	Yes	Yes	Yes
	World Count	3,065 13:37 mins	11,757 52:15 mins	989 4:23 mins
	Grade	13	12	13
Irish	Available	Yes	Yes	Yes
	World Count	3,423 15:12 mins	12,747 56:39 mins	1,017 4:31 mins
	Grade	12	9	12
Italian	Available	Yes	Yes	Yes
	World Count	3,625 16 mins	13,415 59:37 mins	1,042 4:37 mins
	Grade	13	10	13

		Instagram		
		Terms of service	**Privacy policy***	**Cookies use**
Latvian	Available	Yes	Yes	Yes
	World Count	3,033 13:28 mins	12,428 55:14 mins	873 3:52 mins
	Grade	13	9	13
Lithuanian	Available	Yes	Yes	Yes
	World Count	2,947 13 mins	12,428 55:14 mins	880 3:54 mins
	Grade	13	9	13
Maltese	Available	Yes	Yes	Yes
	World Count	3,420 15:12 mins	12,428 55:14 mins	1,014 4:30 mins
	Grade	12	9	13
Polish	Available	Yes	Yes	Yes
	World Count	3,328 14:47 mins	12,341 54:50 mins	991 4:24 mins
	Grade	13	11	13
Portuguese	Available	Yes	Yes	Yes
	World Count	3,579 15:54 mins	14,014 62:17 mins	1,058 4:42 mins
	Grade	13	10	13
Romanian	Available	Yes	Yes	Yes
	World Count	3,826 17 mins	14,576 64:46 mins	1,190 5:17 mins
	Grade	13	10	13
Slovak	Available	Yes	Yes	Yes
	World Count	3,199 14:13 mins	11,650 51:46 mins	962 4:16 mins
	Grade	13	11	13
Slovenian	Available	Yes	Yes	Yes
	World Count	3,254 14:27 mins	26,619 2 hr 20 mins	935 4 mins
	Grade	13	Graduate	13
Spanish	Available	Yes	Yes	Yes
	World Count	3,668 16:18 mins	12,747 59:39 mins	1,096 4:52 mins
	Grade	13	9	13
Swedish	Available	Yes	Yes	Yes
	World Count	3,523 15:39 mins	12,085 53:42 mins	994 4:25 mins
	Grade	13	10	12

Conclusion

A 13-year-old would likely not understand the design and functioning of Instagram at the point of signing on. This is because:

- The time it would take younger users to read all the policies they are agreeing to is excessive and may be beyond the legitimate expectations of a 13-year-old.
- Young users could be swayed by dark patterns during the sign-on process. These instances include:
  - Inferring consent rather than explicitly asking users to click to agree to terms;
  - Obscuring details about the terms and conditions;
  - Persuading users to choose options that are not in their best interest, because they are presented as providing a better experience; and
  - Persuading users to choose options that are not in their best interest, because they are presented in visually more prominent ways.

V

Evaluation of Safety-by-Design Settings on Instagram

Evaluation of Safety-by-Design Settings on Instagram

Research questions:

1 Do younger users enjoy the highest levels of privacy?

2 Do younger users enjoy accessible safety tools and features?

Methodology

Our research involved two steps:

1. Recording the sign-up process for the sock puppet account on platform

1 We established several sock puppet accounts for 13-and 16-year-olds on Instagram. We set up accounts in two EU countries to record the user sign-on journey, including:

- Germany
- The Netherlands

We noted the default privacy settings for each account during these sign-up processes.

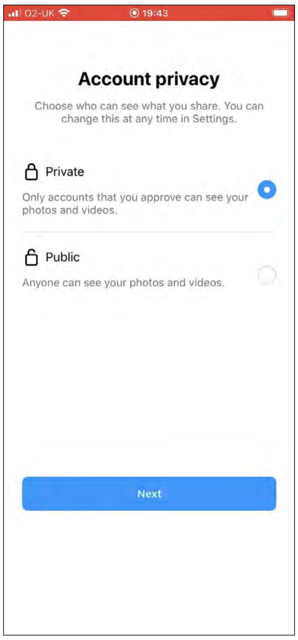
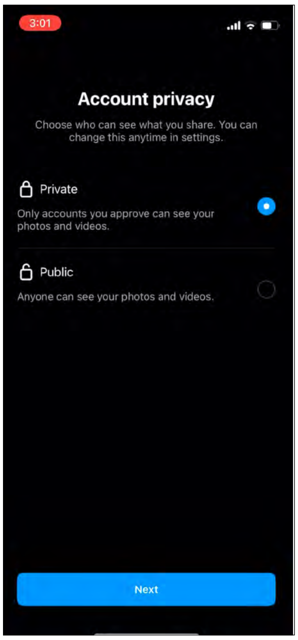
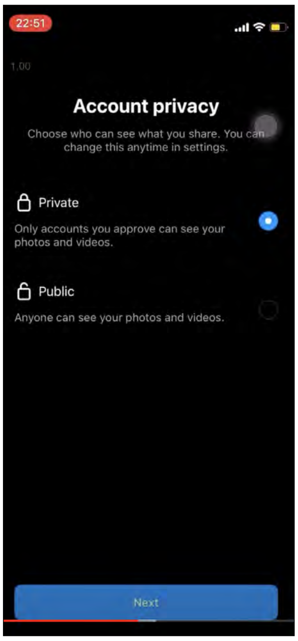
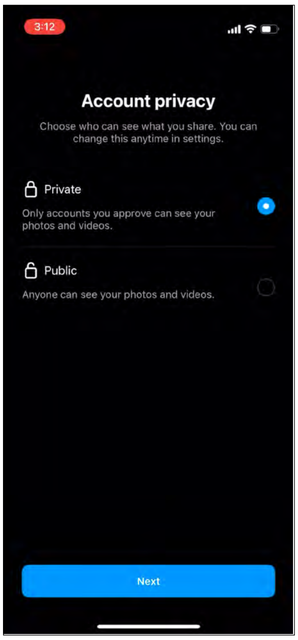
2 **Exploring the availability of safety centers and help tools**

We searched for the available help tools and safety centers on Instagram to confirm if these were available in European languages.

Findings

Default privacy settings

The privacy settings do not default to private, allowing young users to choose their settings. The 13- and 16-year-olds are nudged toward privacy.

	Germany	Netherlands
13 year old	Allows users to choose but nudges them toward the Private setting.  O2-UK 19:43 Account privacy Choose who can see what you share. You can change this at any time in Settings. Private Only accounts that you approve can see your photos and videos. ● Public Anyone can see your photos and videos. ○ Next	Allows users to choose but nudges them toward the Private setting.  3:01 Account privacy Choose who can see what you share. You can change this anytime in settings. Private Only accounts you approve can see your photos and videos. ● Public Anyone can see your photos and videos. ○ Next
16 year old	Allows users to choose but nudges them toward the Private setting.  22:51 1.00 Account privacy Choose who can see what you share. You can change this anytime in settings. Private Only accounts you approve can see your photos and videos. ● Public Anyone can see your photos and videos. ○ Next	Allows users to choose but nudges them toward the Private setting.  3:12 Account privacy Choose who can see what you share. You can change this anytime in settings. Private Only accounts you approve can see your photos and videos. ● Public Anyone can see your photos and videos. ○ Next

Accessibility of safety centers and help tools

Instagram offers a help center,[4] complete with a "Staying safe" section, with quick instructions and guides about:

- Safety tips
- Reporting harassment or bullying
- Reporting DMs or content
- Advice on what to do if people are asking for nudes
- Blocking users
- Quick links to "A parent's guide to Instagram"
- Security tips

Instagram also offers a safety center,[5] which has more specialised guides and advice on blocking and reporting users, guides for parents, and privacy and security.

However, these were not accessible to all young people in their first languages.

[4] Instagram 2023 *Help Centre* https://help.instagram.com/

[5] Instagram 2023 *Keeping Instagram a Safe and Supportive Place* about.instagram.com/safety

	Instagram	
	Instagram’s Help Center	Instagram’s Safety Center
Bulgarian	Yes	–
Croatian	Yes	–
Czech	Yes	–
Danish	Yes	–
Dutch	Yes	–
English	Yes	Yes
Estonian	Yes	–
Finnish	Yes	–
French	Yes	Yes
German	Yes	Yes
Greek	Yes	–
Hungarian	Yes	–
Irish	Yes	–
Italian	Yes	Yes
Latvian	Yes	–
Lithuanian	Yes	–
Maltese	Yes	–
Polish	Yes	–
Portuguese	Yes	Yes
Romanian	Yes	–
Slovak	Yes	–
Slovenian	Yes	–
Spanish	Yes	Yes
Swedish	Yes	–

Conclusion

- Safety-by-design settings nudge young people toward the highest possible defaults.
- Instagram’s safety center is not routinely accessible to young people in their first languages.

VI

Evaluation of Instagram's Ad Manager for Minors

Evaluation of Instagram’s Ad Manager for Minors

Research question:

1 Does Instagram allow ads to reach minors based on profiling?

Methodology

We audited Instagram’s ad manager system, focusing on two aspects:

1 What are Instagram’s ad networks and APIs from the advertisers’ perspective, and do they allow the possibility of targeting minors?

2 From the users’ point of view, what kinds of age propagation occur between a third-party application and the ad network of the platforms, and how is consent gathered or inferred from the underaged users?

Findings

We did not find any evidence of allowing the targeting of minors on Instagram.

- Explicit underaged targeting by selecting the age category 13–17 is not possible on the Meta Ads Manager.
- There is no evident age propagation between Meta accounts and its ad software development kit (SDK).

VII

Appendix I: Instagram's Content Moderation Guidelines

Appendix 1: Instagram’s Content Moderation Guidelines

Instagram’s Community Guidelines on Suicide and/or Self-harm

Instagram’s Community Guidelines[6] outline that the platform aims to:

> *“Maintain [a] supportive environment by not glorifying self-injury. The Instagram community cares for each other, and is often a place where people facing difficult issues such as eating disorders, cutting or other kinds of self-injury come together to create awareness or find support. … Encouraging or urging people to embrace self-injury is counter to this environment of support, and we'll remove it or disable accounts if it's reported to us. We may also remove content identifying victims or survivors of self-injury if the content targets them for attack or humour.”*

Instagram describes self-injury using Meta’s head terms:[7]

> *“While we do not allow people to intentionally or unintentionally celebrate or promote suicide or self-injury, we do allow people to discuss these topics because we want Facebook to be a space where people can share their experiences, raise awareness about these issues, and seek support from one another.*
>
> *We define self-injury as the intentional and direct injuring of the body, including self-mutilation and eating disorders. We remove any content that encourages suicide or self-injury, including fictional content such as memes or illustrations and any self-injury content that is graphic, regardless of context.*
>
> *Content about recovery of suicide or self-harm that is allowed, but may contain imagery that could be upsetting, such as a healed scar, is placed behind a sensitivity screen.*
>
> *Content about recovery of suicide or self-harm that is allowed, but may contain imagery that could be upsetting, such as a healed scar, is placed behind a sensitivity screen. …*
>
> *Do not post:*
>
> - *Content that promotes, encourages, coordinates or provides instructions for*
>   - *Suicide*
>   - *Self-injury*
>   - *Eating disorders*
> - *Content that depicts graphic self-injury imagery*
> - *It is against our policies to post content depicting a person who engaged in a suicide attempt or death by suicide.”*

According to these guidelines, Instagram should remove violative content when they become aware of it.

[6] Instagram 2023 *Community Guidelines* https://help.instagram.com/477434105621119/?helpref=hc_fnav.

[7] Meta 2023 *Suicide and Self Injury* https://transparency.fb.com/en-gb/policies/community-standards/suicide-self-injury/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2Fsuicide_self_injury_violence.

Operationalisation & Definitions for this evaluation

We did not want to include content that is limited to talking about suicide and self-harm, or to documenting the day-to-day life for people who have suicidal ideation or self-harm, nor content about struggling with or trying to stay safe while affected by ideation or self-harming, nor content about recovering after attempts or staying clean. We only included content that is explicitly pro-suicide and/or self-harm.

We define pro-suicide content as content that:

- Shows, promotes or normalises the act of, or preparation for, suicide, e.g.:
    - pictures, videos, memes of people or characters engaging in suicide where there is not also content that suggests this act or preparation is regrettable (for example, images of characters hanging with nooses around their neck, or pictures of people with guns in their mouths);
    - pictures, videos, memes where people express a desire or plan to commit suicide, without expressing regret (for example, a slide show that says “I want to KMS tonight”, or “I want to be dead” with associated suicide terms);
    - pictures, videos, memes about the best ways to die or funny ways to kill yourself, where the best ways to die were described or depicted in realistic terms (for example, by driving your car into a tree). This excluded examples where the best ways to die were potentially tongue in cheek, e.g, by eating too much ice cream.
- Shows, promotes or normalises suicide through humour, e.g.:
    - Pictures, videos or memes with comedic intent but that still depict people engaged in suicide, e.g. videos of children with toilet paper nooses around their necks hanging from a beam and jumping off a chair;
    - videos depicting the suicide of popular characters, such as Kermit the Frog hanging himself in the bathroom.

We do not include content:

- Where people express suicidal ideation but also expressed a desire not to act or wanting to seek help, e.g. posts where people say “I want to KMS, but I couldn’t do it to my family”, or “I think about suicide all the time, but couldn’t go through with it”;
- Where people expressed dark and depressing thoughts, but did not express suicidal ideation, e.g. posts where people described having nothing left to live for, or wanting to go to sleep for a very long time, without explicitly describing suicidal intent;
- Artistic materials where people expressed suicidal thoughts or ideations through art, unless it was a graphic illustration of a suicide method;
- Comedic material that was not graphic, e.g. videos or memes where people describe something cringe-worthy and then talked about wanting to kill themselves.

We define pro-self-harm content as content that:

- Shows self-harm images, e.g. videos of bleeding cuts, the process of cutting or the results of cutting (e.g. bleeding arms, scenes of razors and bathrooms covered in blood, where they are associated with self-harm terms);
- Promotes or normalises self-harm, e.g. pictures, videos or memes about people who self-harm or are self-harming without context that expresses regret (for example, videos of people talking about upgrading their cutters to new, sharper blades, or images of razor blades and blood);
- Shows preparations for self-harm, e.g. images of razors with descriptions or how they were going to cut themselves, or content describing how to use particular self-harm tools;
- Memes or comedy clips that depict people engaging in self-harm, e.g. jokes about cutting yourself on your ankles so your family doesn’t see cuts on your wrists.

We do not include content:

- Where people express self-harm ideation but also expressed a desire not to act or wanting to seek help, e.g. posts where people say “I’ve been clean (from cutting) for 2 days now, but it so hard to keep going”;
- Where people expressed dark and depressing thoughts, but did not express self-harm ideation, e.g. posts where people described being so sad that they can understand why others self-harm, but did not express a desire to self-harm themselves;
- Artistic materials where people depicted self-harm through art, unless it was a graphic illustration of how to cut (e.g. we did not include images or drawings made of people self-harming or the consequences of self-harm).

Instagram’s Community Guidelines on Pro-Eating Disorder Content

Instagram’s Community Guidelines[8] outlines that the platform aims to:

> *“Maintain [a] supportive environment by not glorifying self-injury. The Instagram community cares for each other, and is often a place where people facing difficult issues such as eating disorders, cutting or other kinds of self-injury come together to create awareness or find support. … Encouraging or urging people to embrace self-injury is counter to this environment of support, and we'll remove it or disable accounts if it's reported to us. We may also remove content identifying victims or survivors of self-injury if the content targets them for attack or humour.”*

Instagram describes self-injury using Meta’s head terms:[9]

> *“While we do not allow people to intentionally or unintentionally celebrate or promote suicide or self-injury, we do allow people to discuss these topics because we want Facebook to be a space where people can share their experiences, raise awareness about these issues, and seek support from one another.*
>
> *We define self-injury as the intentional and direct injuring of the body, including self-mutilation and eating disorders. We remove any content that encourages suicide or self-injury, including fictional content such as memes or illustrations and any self-injury content that is graphic, regardless of context.*

[Do not post]

- *Content that focuses on depiction of ribs, collar bones, thigh gaps, hips, concave stomachs or protruding spine or scapula when shared together with terms associated with eating disorders*
- *Content that contains instructions for drastic and unhealthy weight loss when shared together with terms associated with eating disorders.”*

According to these guidelines, Instagram should remove violative content when it becomes aware of it.

[8] Instagram 2023 *Community Guidelines* https://help.instagram.com/477434105621119/?helpref=hc_fnav.

[9] Meta 2023 *Suicide and Self Injury* https://transparency.fb.com/en-gb/policies/community-standards/suicide-self-injury/?source=https%3A%2F%2Fwww.facebook.com%2Fcommunitystandards%2Fsuicide_self_injury_violence.

Operationalisation & Definitions for this evaluation

We did not want to include content that only talks about eating disorders, or documents day-to-day life with them, nor content that about struggling with disorders or roads to recovery.

We only included content that is explicitly pro-restrictive eating disorder.

We describe pro-restrictive eating disorder content as content that:

- Shows, promotes or normalises dangerous weight loss, or dangerous dieting behaviours, such as:
  - Content that focuses on or highlights bone structures, such as ribs, collar bones, thigh gaps, hips, concave stomach or protruding spine or scapulas. This includes content where people are posing or deliberately exposing body parts to highlight extreme thinness;
  - Content that depicts thigh gaps, finger gaps, or other signs of extremely low body mass;
  - Before and after weight loss content, where the after picture shows either bone structures or clinically unhealthy BMI rates;
  - Content that displays or notes BMI rates below 21, or expresses goals for BMI rates below 21;
  - Content regarding weight check ins, or that talks about Goal Weights or Ultimate Goal Weights that would be associated with an unhealthy BMI (such as content where someone describes their weight as 52kg, and their height as 5'8");
  - Showing or promoting unhealthy body measurement and "body checking" trends, such as comparing body part size to household objects;
  - Content that congratulates people for reaching an unhealthy weight, a BMI below 21 or having visible bones, or that encourages them to aspire to this;
  - Content that shows or promotes extremely low-calorie daily food consumption (less than 500 calories per meal, or 1,500 per day) when combined with eating disorder terms, and diets associated with disordered eating terms;
  - Content that congratulates people for restricting their eating to less that 500 calories per meal, or 1,500 per day or that encourages them to achieve this;
  - Content that describes having an eating disorder as a positive outcome or depicts them in a desirable light (e.g. tweets that say 'restricting is easy, will power lets me just eat water').

We do not include content:

- Content that depicts bone structures, thigh gaps or BMIs in association with text or images that describe wanting to recover or gain weight;
- Content that depicts bone structures, thigh gaps or BMIs where associated content (terms etc) indicated that the person in the picture was trying to put on weight or otherwise documenting a successful recovery;
- Content that just features extremely skinny people, who may or may not be affected by restrictive eating disorders, who are just documenting their lives (such as playing guitar, on on a walk), where the content does not explicitly centre around their weight or include associated terms. This does not include images where people are deliberately posing and focusing on their visible bone structures, or thing gaps etc
- Recovery diaries or recovery stories;
- Content that talks about the difficulties of having a restrictive eating disorder, or talked about day-to-day issues (e.g. memes about going to the fridge, losing will power, and eating 1000 calories every night, where it was unclear from the meme if that was all they ate during the day or just a daily 'snack' they regret);

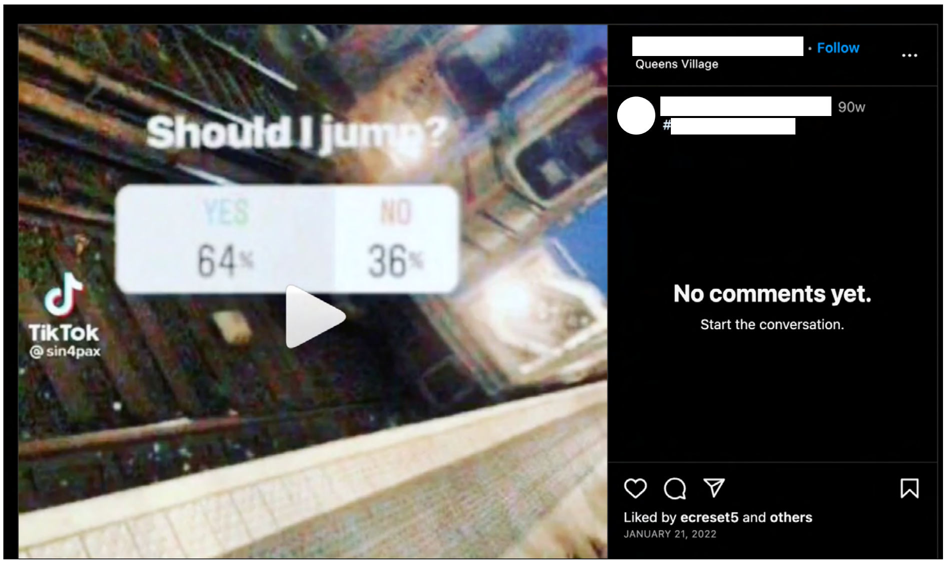
- Content that depicts bone structures, thigh gaps or BMIs in a medical or humanitarian context (e.g. documenting a famine or person ill from non-eating disorder diseases);
- Low calorie diet content that does not include eating disorder terms, such as for content associated with ‘diabetes friendly’ diets, or general weight loss diets
- Images of professional athletes, such as ultra marathon runners or ballerinas;
- Exercise ‘for weight loss’ content.

TRIGGER WARNING
& VIEWER DISCRETION
ADVISED

VII

Appendix II:
Examples of Instagram’s Content Monitored

Pro-suicide and/or self-harm content



[redacted] · Follow ...

Queens Village

[redacted] 90w

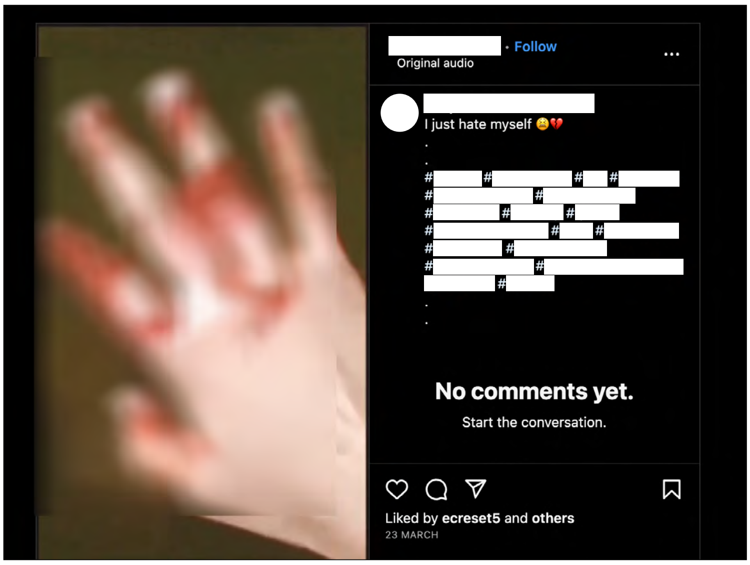
#[redacted]

No comments yet.

Start the conversation.

Liked by ecreset5 and others

JANUARY 21, 2022



[redacted] · Follow ...

Original audio

[redacted]

I just hate myself 😔💔

.

.

[redacted]

.

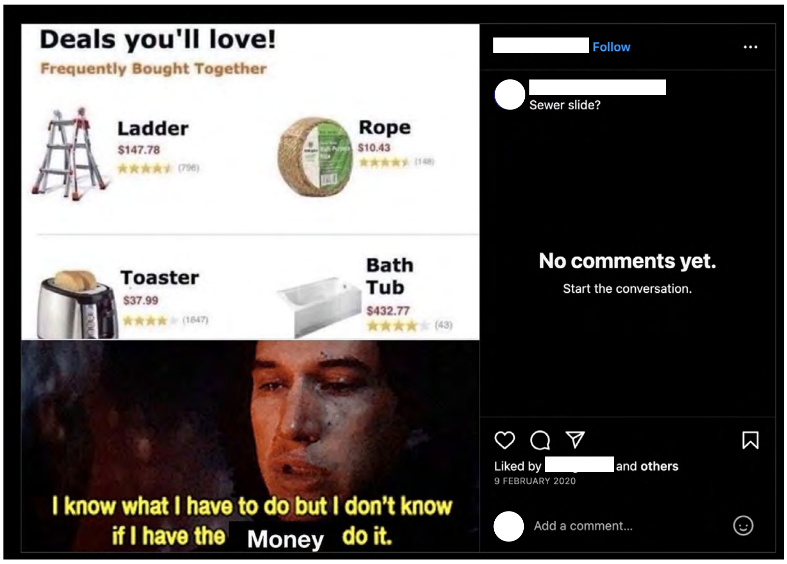
.

No comments yet.

Start the conversation.

Liked by ecreset5 and others

23 MARCH


Deals you'll love!
Frequently Bought Together
Ladder
$147.78
(796)
Rope
$10.43
(148)
Toaster
$37.99
(1647)
Bath Tub
$432.77
(43)
I know what I have to do but I don't know
if I have the Money do it.

[redacted] Follow ...

[redacted]

Sewer slide?

No comments yet.

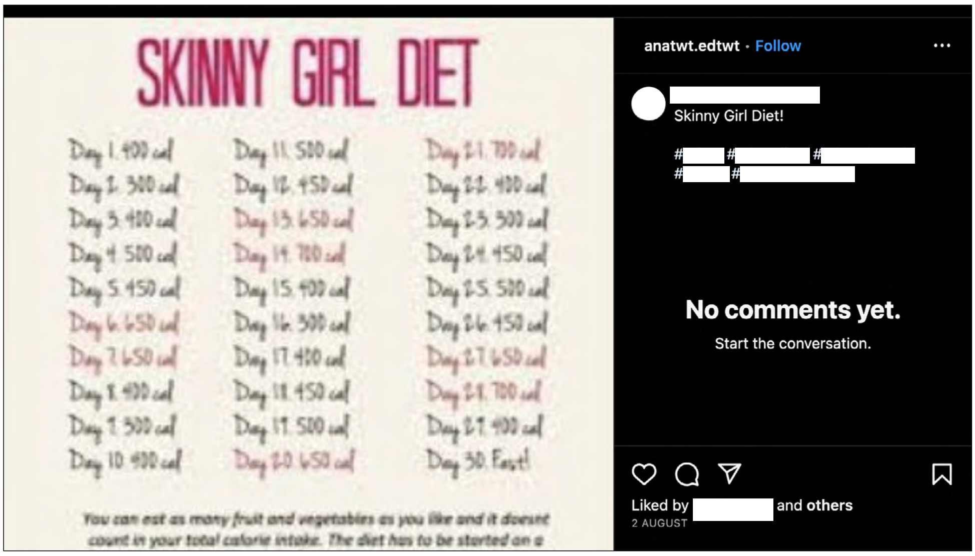
Start the conversation.

Liked by [redacted] and others

9 FEBRUARY 2020

Add a comment...

Pro-restrictive eating disorder content



SKINNY GIRL DIET

Day 1. 400 cal
Day 2. 300 cal
Day 3. 400 cal
Day 4. 500 cal
Day 5. 450 cal
Day 6. 650 cal
Day 7. 650 cal
Day 8. 400 cal
Day 9. 300 cal
Day 10. 400 cal
Day 11. 500 cal
Day 12. 450 cal
Day 13. 650 cal
Day 14. 700 cal
Day 15. 400 cal
Day 16. 300 cal
Day 17. 400 cal
Day 18. 450 cal
Day 19. 500 cal
Day 20. 650 cal
Day 21. 700 cal
Day 22. 400 cal
Day 23. 300 cal
Day 24. 450 cal
Day 25. 500 cal
Day 26. 450 cal
Day 27. 650 cal
Day 28. 700 cal
Day 29. 400 cal
Day 30. Fast!

You can eat as many fruit and vegetables as you like and it doesnt count in your total calorie intake. The diet has to be started on a

anatwt.edtwt • Follow

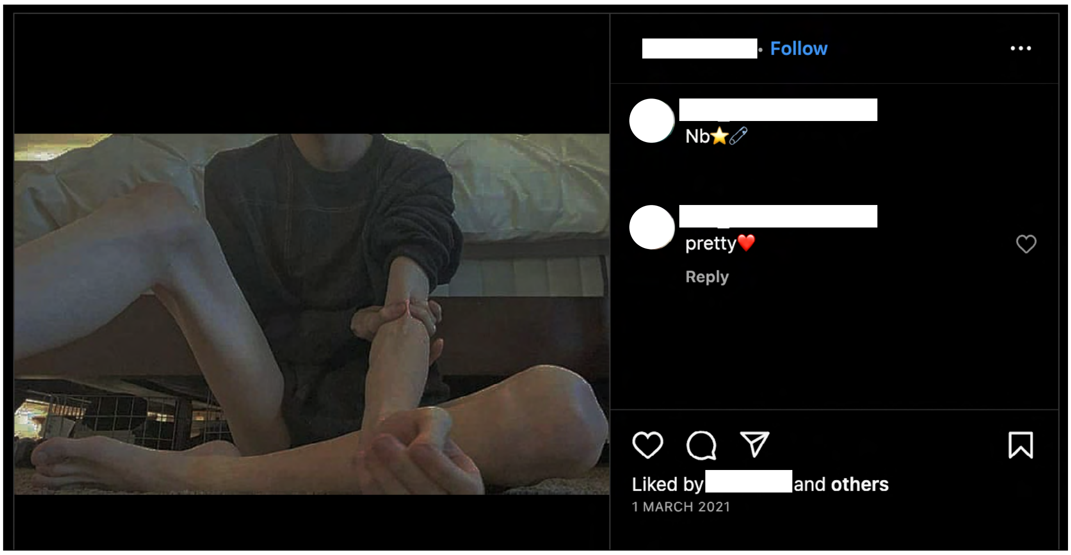
[redacted] Skinny Girl Diet!

#[redacted] #[redacted] #[redacted]
#[redacted] #[redacted]

No comments yet.

Start the conversation.

Liked by [redacted] and others
2 AUGUST

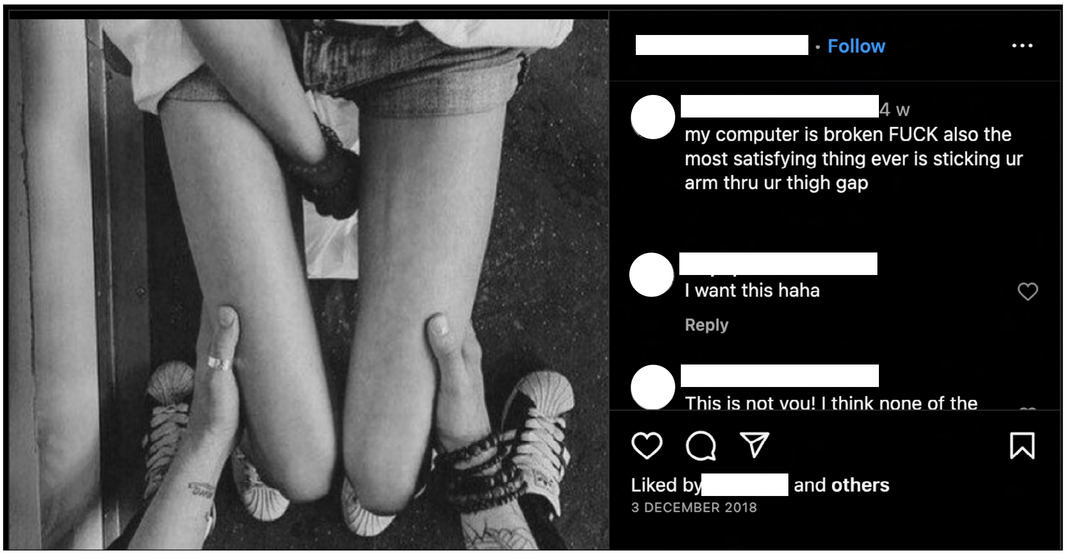


[redacted] • Follow

[redacted]
Nb⭐🩹

[redacted]
pretty❤️
Reply

Liked by [redacted] and others
1 MARCH 2021



[redacted] • Follow

[redacted] 4 w
my computer is broken FUCK also the most satisfying thing ever is sticking ur arm thru ur thigh gap

[redacted]
I want this haha
Reply

[redacted]
This is not you! I think none of the

Liked by [redacted] and others
3 DECEMBER 2018

VII

Appendix III:
Instagram's Sign-on Process

We have broken down Instagram's sign-on process on a mobile app into 17 steps:

1 The app directs the new user to “Create New Account” or “Log in.” The “Create New Account” option is promoted, which could be because there is no account associated with this handset.

2 The app asks the user to enter a phone number or email. It states, “You may receive SMS notifications from us for security and log-in purposes.”

3 Next, the app directs the user to “Enter the confirmation code we sent to [Number entered in step 2].” It provides other options for users to “Change phone number” or “Send SMS message again.” A bold “Next” button underneath, which is functional only when the user has entered the code.

4 Then, the app asks the user to “Add your name.” It explains, “Add your name so that friends can find you,” in a smaller, gray font. It has a bold “Next” button underneath, functional only after the user has entered their full name.

5 The app next asks the new user to “Create a password.” It explains, “We can remember your password, so you don’t have to enter it in your iCloud devices,” in a smaller, gray font.   A bold “Next” button underneath is usable once the user has entered a password.

6 Next, the app guides the new user to “Add your date of birth.” It states, “This won’t be a part of your public profile,” in a smaller, gray font, and has a link to the “Why do I need to provide my date of birth?” section. It displays a bold “Next” button underneath, which is only usable once the user has entered their date of birth.

7 Next, the app directs the new user to “Create a username.” It shows “Choose a username for your account. You can always change this later” in a smaller, gray font. A bold “Next” button underneath functions only after the user has entered a username.

8 The app then asks the new user to “Sign up as [Username selected in step 7]” with “You can always change your name later,” written in a smaller, gray font. At the bottom of the screen, in an even smaller, gray font, the app states, “People who use our service may have uploaded your contact information to Instagram. Learn more” and that “By tapping ‘sign up,’ you agree to our terms. Learn how we collect, use, and share your data in our Privacy Policy and how we collect cookies and similar technology in our Cookies Policy.” There is a large “Sign up” button in blue.

9 The app tells the new user, “Next, you’ll be able to sync your contacts and find your friends.” In a smaller, gray font, it explains, “If you allow Instagram to access your contacts, we’ll help you find your friends, and help them find you.” In an even smaller, gray font at the bottom, it says, “If you allow Instagram to access your contacts, they will be periodically synced and securely stored in our servers. You can turn off syncing at any time in Settings. Learn more.” There is a large blue “Next” button below it.

10 After this, the new user’s phone is notified, “Instagram would like to access your contacts. This helps you and others find people to follow and helps you connect and interact with people you already follow. Your contacts will be synced and securely stored on Instagram’s servers.” It offers “Don’t Allow” and “OK” options. The OK option is the more prominent of the two.

11 The app then asks the new user to “Find Facebook friends.” A smaller, gray font states, “You choose which friends to follow. We’ll never post to Facebook without your permission.” It offers two options, “Find Friends” and “Skip.” The Find Friends option is highlighted.

12 If the new user chooses to “Skip” finding Facebook friends in step 11, they get a notification on their phone saying, “Instagram is more fun when you follow your friends and see their posts. Are you sure that you want to skip this step?” It again offers two options: “Skip,” which is now bold, and “Follow Friends.” This means that it takes two clicks for the new users to decline to find Facebook friends.

13 The app then asks the new user to select “Account Privacy.” In a smaller, gray font, it explains, “Choose who can see what you share. You can change this at any time in Settings.” The user is offered two options: “Private” or “Public”. The Private option is turned on by default, and a large “Next” button at the bottom is automatically enabled.

14 Next, the app prompts the new user to “Add a profile photo.” In a smaller, gray font, it explains, “Add a profile picture so that your friends can know it’s you.” It provides two options, “Add a photo” or “Skip,” with Add a photo made more prominent.

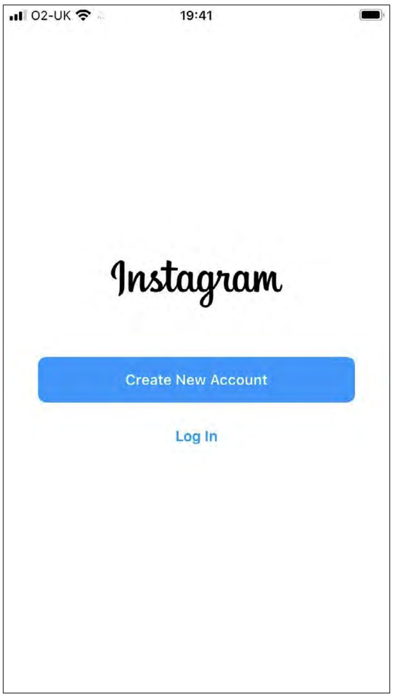
15 The app then offers the new user the option to “Discover People” and presents a list of possible people to follow. Next to each suggested account is a “Follow” button. The list of accounts is long and can be scrolled down. At the top of this step, there is a small “Next” button. The “Next” button has not appeared at the top, nor has it been this less prominent at any other step in the sign-on process. This might confuse the users into scrolling down the long list of suggested accounts to look for the “Next” button at the bottom.

16 Next, the app prompts the new user to “Turn on Notifications.” In a smaller, gray font, it explains, “Find out straight away about when people follow you or like and comment on your posts.” It provides two options: “Turn on” or “Skip,” with Turn on the more prominent option.

17 Finally, a notification appears on the new user’s phone saying, “Instagram would like to send you notifications. Notifications include alerts, sounds, and icon badges. These can be configured in Settings.” It provides two choices, “Allow” or “Don’t Allow’” with the Allow in bold. Choosing “Skip” in step 16 still leads to this notification, making it an example of a two-step to say no process; however, it is also a two-step to say yes process.

Steps in images

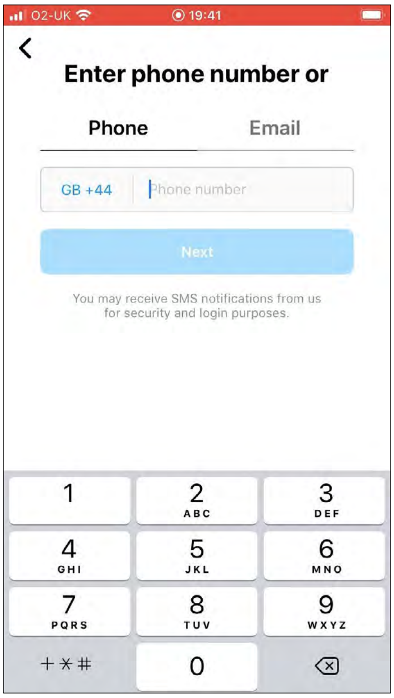
Step 1



Create New Account

Log In

Step 2



Enter phone number or

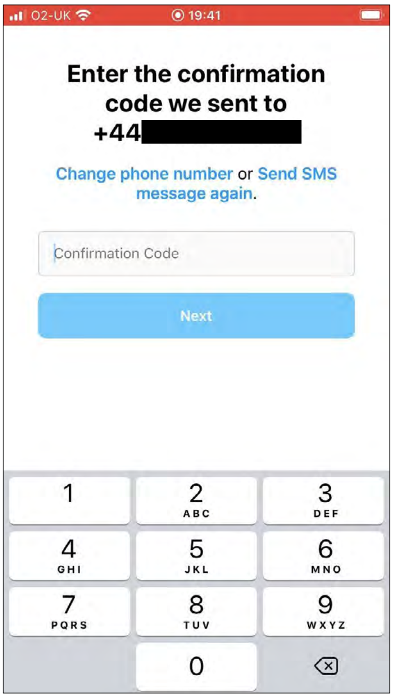
Phone | Email

GB +44 | Phone number

Next

You may receive SMS notifications from us for security and login purposes.

Step 3



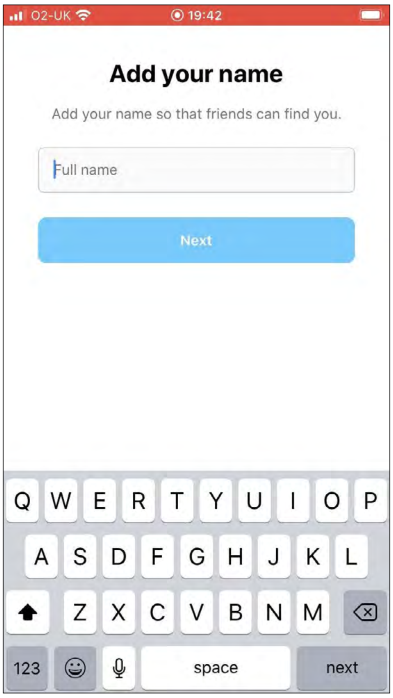
Enter the confirmation code we sent to +44[redacted]

Change phone number or Send SMS message again.

Confirmation Code

Next

Step 4



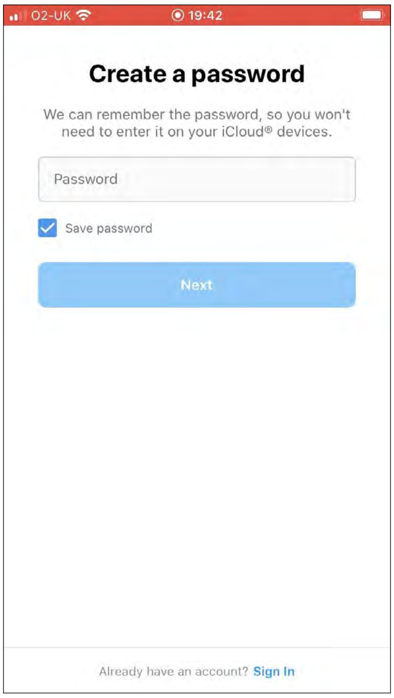
Add your name

Add your name so that friends can find you.

Full name

Next

Step 5



Create a password

We can remember the password, so you won't need to enter it on your iCloud® devices.

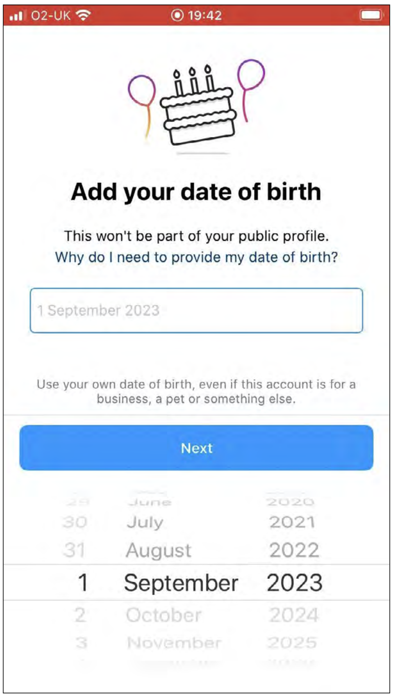
Password

☑ Save password

Next

Already have an account? Sign In

Step 6



Add your date of birth

This won't be part of your public profile.
Why do I need to provide my date of birth?

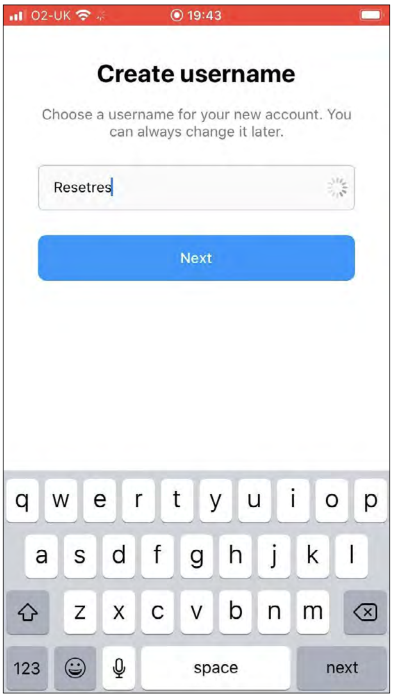
1 September 2023

Use your own date of birth, even if this account is for a business, a pet or something else.

Next

Steps in images

Step 7

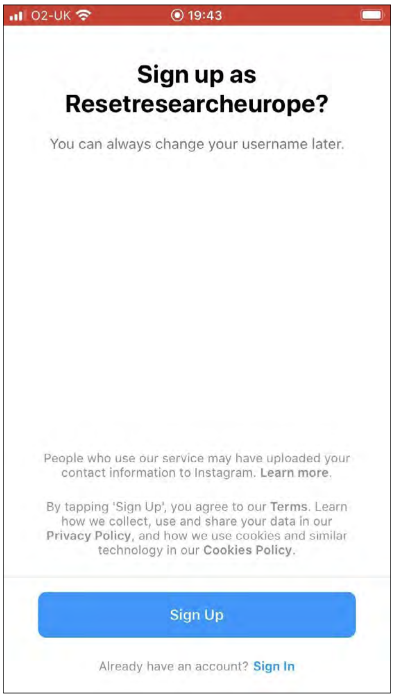


Create username

Choose a username for your new account. You can always change it later.

Resetres

Step 8



Sign up as Resetresearcheurope?

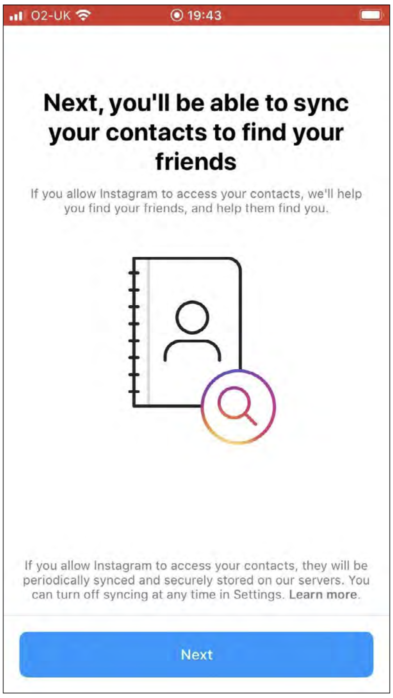
You can always change your username later.

People who use our service may have uploaded your contact information to Instagram. **Learn more.**

By tapping 'Sign Up', you agree to our **Terms**. Learn how we collect, use and share your data in our **Privacy Policy**, and how we use cookies and similar technology in our **Cookies Policy**.

Note: **inferring consent when the user signs up with a username.** By clicking "Sign up as [Username]," the user is actually agreeing to the terms and conditions rather than confirming the username chosen in the previous step. This also demonstrates **obscuring details about terms and conditions**, where the agreements are presented in a tiny, gray font at the bottom of the screen.

Step 9



Next, you'll be able to sync your contacts to find your friends

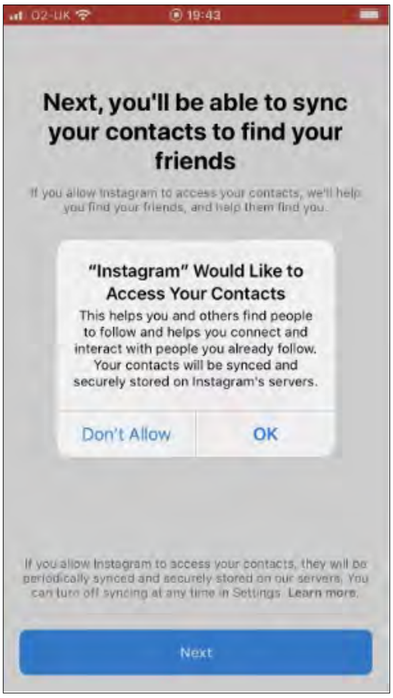
If you allow Instagram to access your contacts, we'll help you find your friends, and help them find you.

If you allow Instagram to access your contacts, they will be periodically synced and securely stored on our servers. You can turn off syncing at any time in Settings. **Learn more.**

This demonstrates **obscuring important details,** as the details about data processing are presented in a small, gray font on the screen.

Steps in images

Step 10



02-UK 19:43

Next, you'll be able to sync your contacts to find your friends

If you allow Instagram to access your contacts, we'll help you find your friends, and help them find you.

"Instagram" Would Like to Access Your Contacts

This helps you and others find people to follow and helps you connect and interact with people you already follow. Your contacts will be synced and securely stored on Instagram's servers.

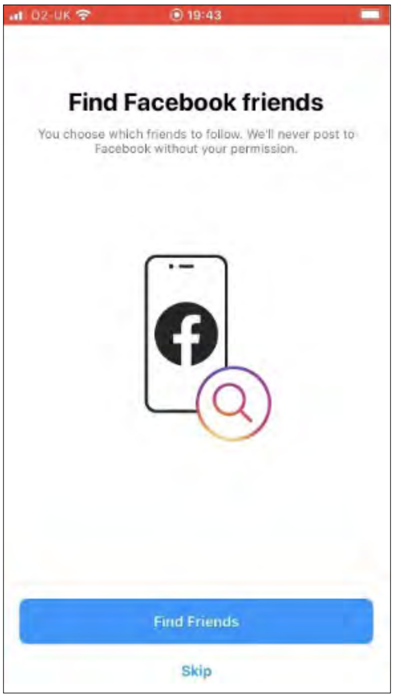
Don't Allow | OK

If you allow Instagram to access your contacts, they will be periodically synced and securely stored on our servers. You can turn off syncing at any time in Settings. Learn more.

Next

Note: **Visual promotion of options that are in the platform’s best interests while demoting options that are in users’ best interests.** “OK” is highlighted in bold.

Step 11



02-UK 19:43

Find Facebook friends

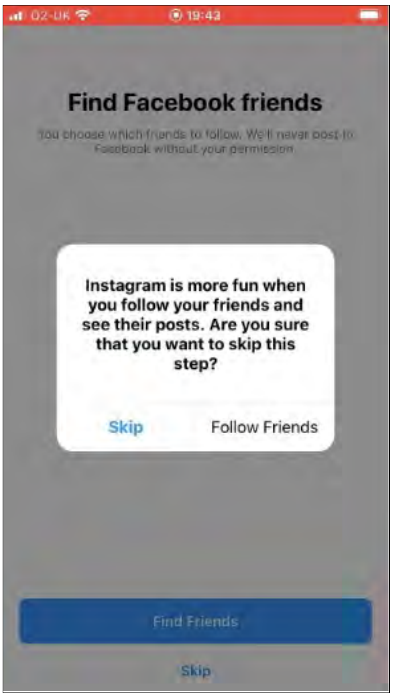
You choose which friends to follow. We'll never post to Facebook without your permission.

Find Friends

Skip

Note: **Visual promotion of options in the platform’s best interests while demoting options in users’ best interests.** “Find Friends” has been made more prominent.

Step 12



02-UK 19:43

Find Facebook friends

You choose which friends to follow. We'll never post to Facebook without your permission.

Instagram is more fun when you follow your friends and see their posts. Are you sure that you want to skip this step?

Skip | Follow Friends

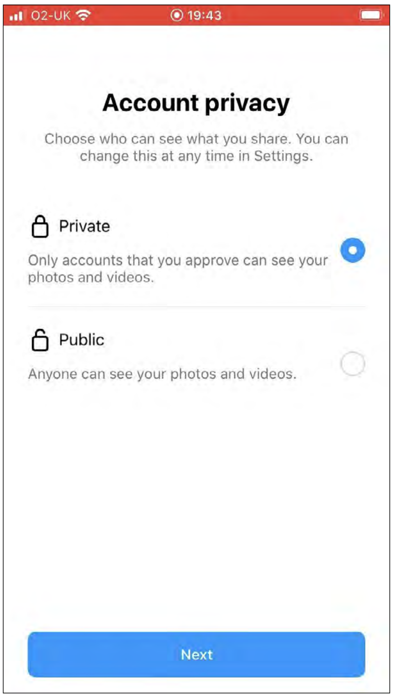
Find Friends

Skip

Note: **Presenting options that may not be in a user’s best interests as a “better user experience.”** Instagram claims to be “more fun” when you follow your friends by syncing accounts. Also, this is the second step you must follow to decline to find friends on Facebook. This could be an example of a **“click twice for no, but only once for yes”** dark pattern; however, it is clear that users would need to click twice for yes as well.

Steps in images

Step 13



O2-UK 19:43

Account privacy

Choose who can see what you share. You can change this at any time in Settings.

Private ●

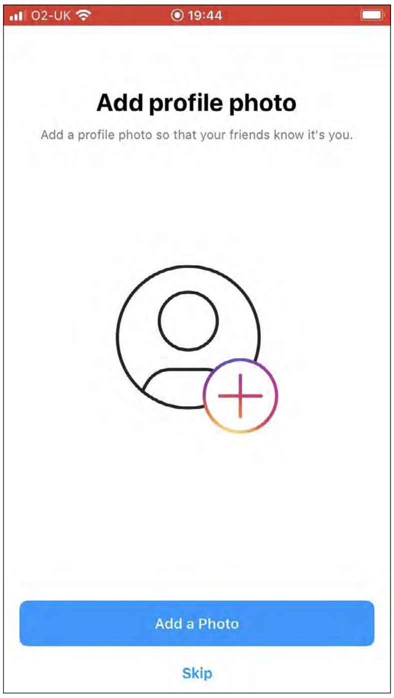
Only accounts that you approve can see your photos and videos.

Public ○

Anyone can see your photos and videos.

Next

Step 14



O2-UK 19:44

Add profile photo

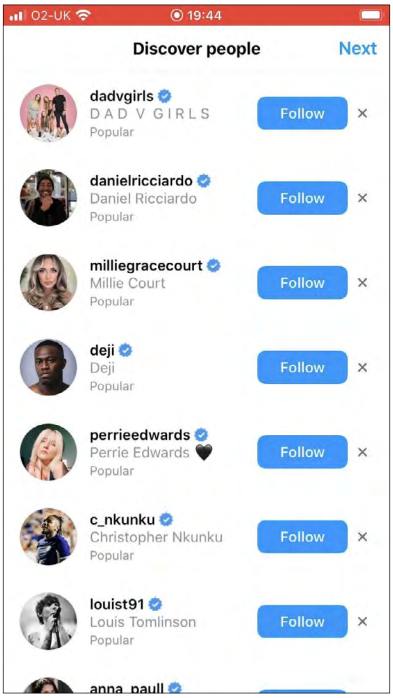
Add a profile photo so that your friends know it's you.

Add a Photo

Skip

Note how "Add a Photo" is more prominent than "Skip." We do not include this in our count of dark patterns as it is unclear whether selecting a profile photo is not in a user's best interests (it may not be an identifying photo of the user, for instance).

Step 15



O2-UK 19:44

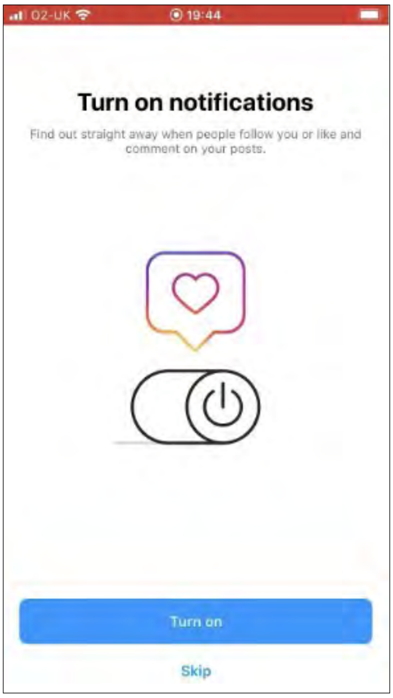
Discover people — Next

dadvgirls DADVGIRLS Popular	Follow	×
danielricciardo Daniel Ricciardo Popular	Follow	×
milliegracecourt Millie Court Popular	Follow	×
deji Deji Popular	Follow	×
perrieedwards Perrie Edwards 🖤 Popular	Follow	×
c_nkunku Christopher Nkunku Popular	Follow	×
louist91 Louis Tomlinson Popular	Follow	×
anna_paull		

Note: **Visual promotion of options in the platform's best interests while demoting options in users' best interests.** The "Next" button is less prominent here than in other steps and is positioned at the top right-hand side. This might confuse users into scrolling down—effectively seeing more possible accounts to follow than they want—while looking for the "Next" button, which, in all the other previous steps, has been located at the bottom of the screen.

Steps in images

Step 16

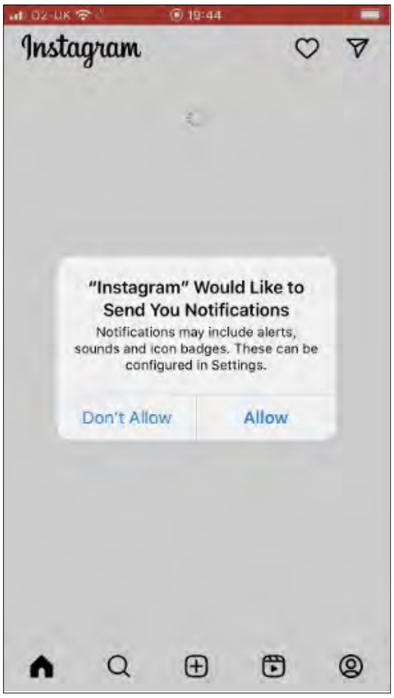


Turn on notifications

Find out straight away when people follow you or like and comment on your posts.

Note: **Visual promotion of options in the platform's best interests while demoting options in users' best interests.** "Turn on" is more prominent than "Skip."

Step 17



"Instagram" Would Like to Send You Notifications

Notifications may include alerts, sounds and icon badges. These can be configured in Settings.

Note: **Visual promotion of options in the platform's best interests while demoting options in users' best interests.** "Allow" is in the bold font. Also, this is the second step you must follow to decline notifications. This could be an example of a **"click twice for no, but only once for yes"** dark pattern; however, it is clear that users would need to click twice for yes as well.

Policies referenced:

1. Instagram terms
2. Privacy policy
3. Cookies policy