

AMENDED Exhibit 377

PLAINTIFFS' OMNIBUS OPPOSITION TO DEFENDANTS' MOTIONS FOR SUMMARY JUDGMENT

Case No.: 4:22-md-03047-YGR

MDL No. 3047

In Re: Social Media Adolescent Addiction/Personal Injury Products Liability Litigation

IG IIC Prevalence Analysis - a/c priv

This document contains the analysis for the IG IIC Prevalence Analysis within Direct Messages.

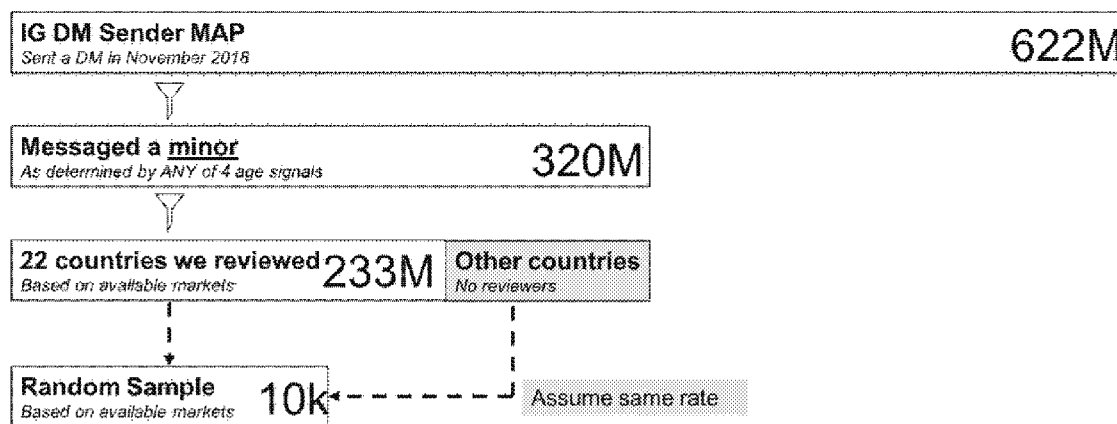
Related Docs

Main research plan doc: [IG IIC Prevalence Estimate](#)

Top-line Results

Top-line

We estimate that there are 3.5M profiles [95% CIs: 2.9M to 4.1M] conducting Inappropriate Interactions with Children (IIC) among all IG DM Sender MAP who sent a message to a minor (a population of 320M). In percent terms, this is 1.1%. (See Figure 1 for the flow chart.)



Assumptions and caveats:

"Potential Groomers" Review. There are a large amount of accounts marked as "potential groomers" (18M [16.6M to 19.5M]). In these cases, the reviewer was either unsure of the profile's user age, or there was ambiguity in the conversation. We will be re-submitting the "potential groomers" to a double review to be extra sure on this number, and it is possible that the IIC numbers could increase with additional review. We have already conducted double review on accounts marked as IIC.

Countries we didn't review. We were only able to review profiles in 22 markets due to labelling resources. This group represents 73% of this population (233M/320M). The above 3.5M number is based on assuming that the countries we did not review have the same rate of IIC activity as the 22 countries we did review. If we do not make this assumption and restrict to the 22 countries we reviewed, our estimate is 2.55M [2.1M to 3.0M] IIC profiles (1.1% of this group).

Disagreement Data. We conducted double review on accounts marked as IIC. We currently look at the last review if there was re-review present or at the expert review and treat this review as the ground truth. In general, we are trusting the expert review and placing a high weight on their decision. This is consistent with Harmful Behavior's methodology on FB (Blue). We are investigating this assumption, and it could revise our estimates.

Missing Labels. There are a small number of missing labels due to technical issues with enqueueing content into Single Review Tool. We assume the missing labels are "missing at random" and not systematically biasing the results. We cannot find any systematic bias after investigating in detail. The number of missing labels is small, and we would have to assume there is a massive amount of systematic bias for this issue to have any effect on the top-line results.

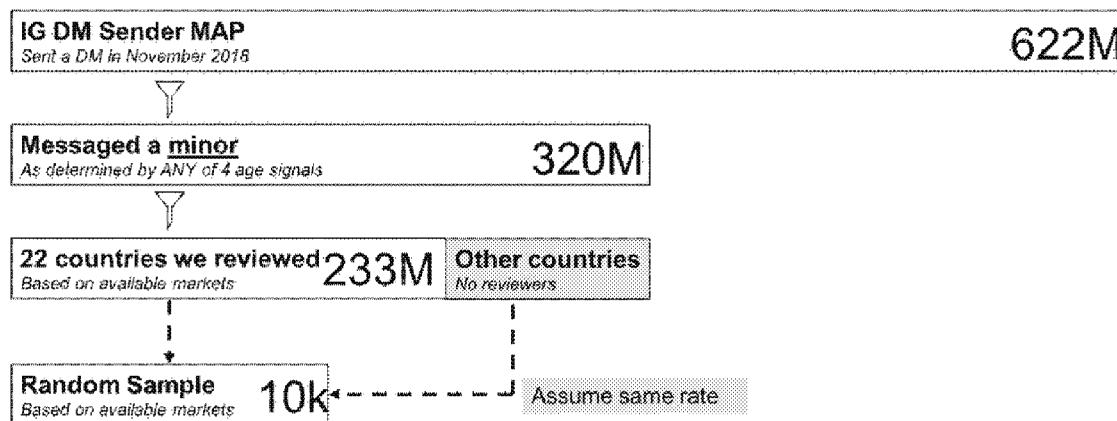
Next Steps:

Additional reviews
 Missing Iraq enqueue (~170)
 re-review of Potential Groomer (~500)
 missed review (13 not reviewed + 71 not enqueued)
 Disagreement labels: incorporate uncertainty when 2 labels disagree
 Breakdown by type of Groomer label and sub-labels

Details

Our population is IG DM Sender MAP who sent at least one DM to a minor. The numbers we report here are unique IG profiles. This is not the same as unique people.

- Total IG DM Sender MAP in November 2018: 622M [query]
- Who sent a DM to a minor: 320M [query]
 - minor = ANY of our age signals identified them as a minor (FB stated age (if IG account linked to FB), FB age prediction (if linked), IG age affinity, IG Adult Classifier).
- in 23 countries where we had reviewers: 236M [query]; 22 countries (excluding Iraq, where we had a bug enqueueing the content): 233M



Results

Of those who sent DMs in Nov 2018 to a minor from the 22 countries we reviewed, we see:

- 2.55M groomers (1.1%) [95% CIs: 2.1M to 3.0M],
- 2.0M missing label (0.85%) [1.6M to 2.4M],
- 13.2M potential groomers (5.6%) [12.0M to 14.2M],
- 75.7M minor age liars (32.6%) [73.5M to 78.2M],
- 139.8M non-groomers (59.9%) [137M to 142M]

<https://fburl.com/daiquery/3z4dnvvg>

We were only able to review 22 countries representing 72.8% of all IG DM Senders who sent to minors (233M / 320M). If we assume that there is the same proportion in each category among this remaining population (e.g., 1.1% of this population is groomers), then we should multiply the above numbers by 1.37 (1 / 0.7281), and our total among the 320M who sent a message to a minor is:

- 3.5M groomers (1.1%) [95% CIs: 2.9M to 4.1M],
- 2.7M missing label (0.85%) [2.2M to 3.3M],
- 18M potential groomers (5.6%) [16.6M to 19.5M],
- 104M minor age liars (32.6%) [101M to 107M],
- 192M non-groomers (59.9%) [189M to 195M]

Since the remaining population did not send a DM to a minor (302M), we can assume that there were no groomers in this population. The above represents our estimate for IG DM MAP Senders in November 2018.

Assumptions and Caveats

Profiles, not users. Our estimate is at the profile-level. We do not have an estimate of the distinct number of people who are conducting IIC on IG.

Countries we didn't review. The rate of each group is uniform between the 22 countries we reviewed (72.8% of IG DM MAP Sender) with the countries we did not review.

We could test this by looking at the 22 countries we already reviewed and checking how uniform grooming activity is across country.

Reporting both numbers (75% vs. 25%)

Missing Labels. We assume the missing labels are missing at random and not systematically biasing the results. In that case, we can effectively add 1.1% groomers of 2.7M missing to the groomer total, which is effectively rounding error, and we can ignore this. If we discover any systematic bias, this may change our estimate.

2 groups of missing data: enqueued but not reviewed, and never enqueued.

Potential Groomers. The potential groomers are cases where the reviewer was either unsure of the profile's user age or where there was ambiguity in the conversation. We will be re-submitting the "potential groomers" to a double review to be extra sure on this number, and it is possible that the groomer numbers could increase with additional review.

500 re-reviews of these profiles

Disagreement Data. We currently look at the last review if there was re-review present or at the expert review and treat this review as the ground truth. In general, we are trusting the expert review and placing a high weight on their decision. We could investigate this assumption more and understand how it would change the top-line numbers. We know that the bias is typically in the direction of more false positives (reviewers have a tendency to see IIC when it is not present).

We could incorporate uncertainty when 2 labels disagree by counting it as 0.5 groomer.

We could also provide upper and lower bounds by only counting groomers when ANY reviewer indicated it was violating (upper bound) or when ALL reviewers agreed it was violating (lower bound).

Gradations/severity of "Groomer" label. Can we differentiate levels of IIC behavior? For example, fraction of the profiles reviewed that were reported to NCMEC. I need the mapping to the below types. There are 2 groomer labels ("Groomer" and "groomer_escalate"). I'm unclear on the difference.

We will break down by additional groomer types.

Sexualized Conversation

Child Exploitation Images

Meeting/Planning to Meet for sex

Position of Trust

Suicidal Ideation

Sibling Abuse

Sextortion

Prostitution of a minor

Bestiality

Notes

- The population is defined in: [D13440310](#)

	grp	map	dap
1	all dm	622,739,079	236,509,643
2	teen_receivers	320,425,593	93,909,415

<https://fburl.com/daiquery/x40r714c>

Next Steps

Disagreement labels: incorporate uncertainty when 2 labels disagree

Breakdown by type of Groomer label and sub-labels

Potential groomers re-review

[Contact Support](#)

Conversation History

Quipbot#15 left the thread Nov 6 at 7:16 pm
