



Users' Exposure and Reaction to Problematic Content in India

Do Users in India, Kenya and Ghana React
Differently to Problematic Content?
Details of project available [here](#).



Index

- 03 Project Overview**
Background, Objectives & Envisaged Outcome of the Project
- 06 Summary of Findings**
Key findings from the survey summarised
- 08 Key Recommendations**
Key recommendations based on the findings of the survey
- 09 Broad Findings**
Respondents' perception, reaction and behaviour towards problematic content
- 15 Content Based Findings**
Respondents' exposure to different kinds of problematic content, and harms faced due to problematic content
- 28 Respondent Profile**
Age, Gender, Religion, Geographic Location, Education Level, Occupation and Income Level of Respondents, as well as the amount of time spent by the on social media



Project Overview

Background, Objectives and Envisaged Outcome
of the Project

Project Overview

Social media platforms have the potential to become a vehicle for inclusive, democratic, and free expression. However, their misuse may cause harm to users. Circulating allegedly incorrect, sensational, divisive, hateful, and misleading information (problematic content) is one such misuse.



Objectives of the Survey

The study has the following objectives:

- Understand the perspectives of users towards problematic content; and
- Check if misuse of social media platforms (through problematic content) may cause different kinds of harm to users.

Envisaged Outcome

The study will enable us to understand the rationale behind different kinds of users' reactions, trust and the possibility of harm emanating from exposure to problematic content.

About the Survey

A pan-India survey of 2002 social media users was conducted in the first quarter of 2022. Respondents from different education and income levels and occupation and geographic locations (urban, rural and peri-urban) were randomly selected. However, diversity in the same was ensured. A neutral gender ratio was consciously maintained.

Types of Problematic Content and Different User Perspectives

Types of Problematic Content

Unverified information on historical events, or rumours about local crimes, or unfounded opinions against a person(s)

Hateful content based on political or religious affiliation

Sensational content pertaining to an event or person(s),

Attractive yet suspicious deals/offers, such as for shopping, free gifts, job opportunities, government benefits, availing loans etc.

Different User First Reactions & Actions to Problematic Content

Ignore it

Tempted to trust it

Suspicious for genuineness and accuracy but not wanting to verify it

Eager for verification

Felt like sharing such content with others

Interact with it

Report it to the service provider

Try to verify the accuracy

Follow steps given in content

Factors Affecting Users Trust on Content

Own beliefs

Use of official seals in the content

Form of content, i.e., video, text, image/meme

Availability of similar content in different posts

Who posted the content: like a private blog, news agency, friends or family, big business entity, popular personality etc.

Grammar, spelling and quality of writing, graphics, pictures or illustrations

Select Online & Offline harms due to problematic content

Arguments with people online and/or offline, over differences of opinion

Online hate & bullying

Online harassment & trolling

Emotional/psychological harm

Threatened of offline harm

Hacking of social media or bank or other online accounts

Financially deputed/defrauded

Key Literature Referred

1. Carnegie Mellon University study: Misinformation in Social Media: Definition, Manipulation, and Detection
2. A Survey of Fake News: Fundamental Theories, Detection Methods, and Opportunities
3. UNC-Charlotte College of Computing and Informatics: Human-Misinformation interaction: Understanding the interdisciplinary approach needed to computationally combat false information
4. Real or Fake? User Behaviour and Attitudes Related to Determining the Veracity of Social Media Posts
5. How to Filter Hate Speech, Hoaxes, and Violent Clips Out of Your Social Feeds
6. Proceedings of the National Academy of Sciences of the USA: The spreading of misinformation online
7. Trend Micro study: The Fake News Machine How Propagandists Abuse the Internet and Manipulate the Public
8. Real or Fake? User Behaviour and Attitudes Related to Determining the Veracity of Social Media Posts
9. University of Michigan: Content patterns in COVID-19 related digital misinformation in India
10. Countering Misinformation Fake News In India
11. “I don’t think that’s True, Bro!” An Experiment on Fact-checking Misinformation in India

Find the detailed Literature Review [here](#).

Using the literature a Logic Framework Analysis (LFA) was created. The LFA is available [here](#).



Summary of Findings

Key findings from the survey summarised

Summary of Findings

1

Most respondents consider the source of the content as a factor for trusting it. They tend to trust the government, friends/family and news agencies, more than popular personalities, big business entities and private blogs.

Refer: Slide 10

2

More respondents see problematic content in the form of memes/images and AV files. Most also claim to trust these forms, more than text and weblinks.

Refer: Slide 12

3

Most respondents are heavily exposed to problematic content. Some of them claim to be frequently trying to verify it.

Refer: Slide 11

Summary of Findings

4

When compared to low and moderate-usage users, more percentage of high-usage users of social media were found to verify problematic content.

Refer: Slide 13

6

Trend shows that when respondents engage with problematic content or follow steps as suggested, they tend to face significant harms.

Refer: Slides: 18, 21, 24 & 27

5

Trend shows that when compared to respondents with low exposure, a greater percentage of respondents who have high exposure to problematic content do not ignore but react to such content.

Refer: Slides: 17, 20, 23 & 26

7

Most respondents want social media platforms to be held accountable for the spread of problematic content.

Refer: Slide 15

Key Recommendations

1

There is a need to raise awareness amongst consumers on different types and forms of problematic content.

Consumer groups and civil society organisations may take appropriate initiatives in this regard, with support from the government and social media platforms.

2

Building consumers' capacity to identify and appropriately deal with problematic content is necessary. Consumers must also be empowered to flag problematic content to their peers and social media platforms.

This includes sensitising them to verify online content (and the ways to do so) before relying upon it, to reduce possible harms emanating from it.

3

Social media platforms need to be made more accountable towards problematic content being spread on their platforms, and also encouraged to take more steps to curb it.

The government may explore a self-regulation or co-regulation model in this regard, which keeps consumer interests paramount, while also strengthening transparency in the operations/efforts of social media platforms in this regard.

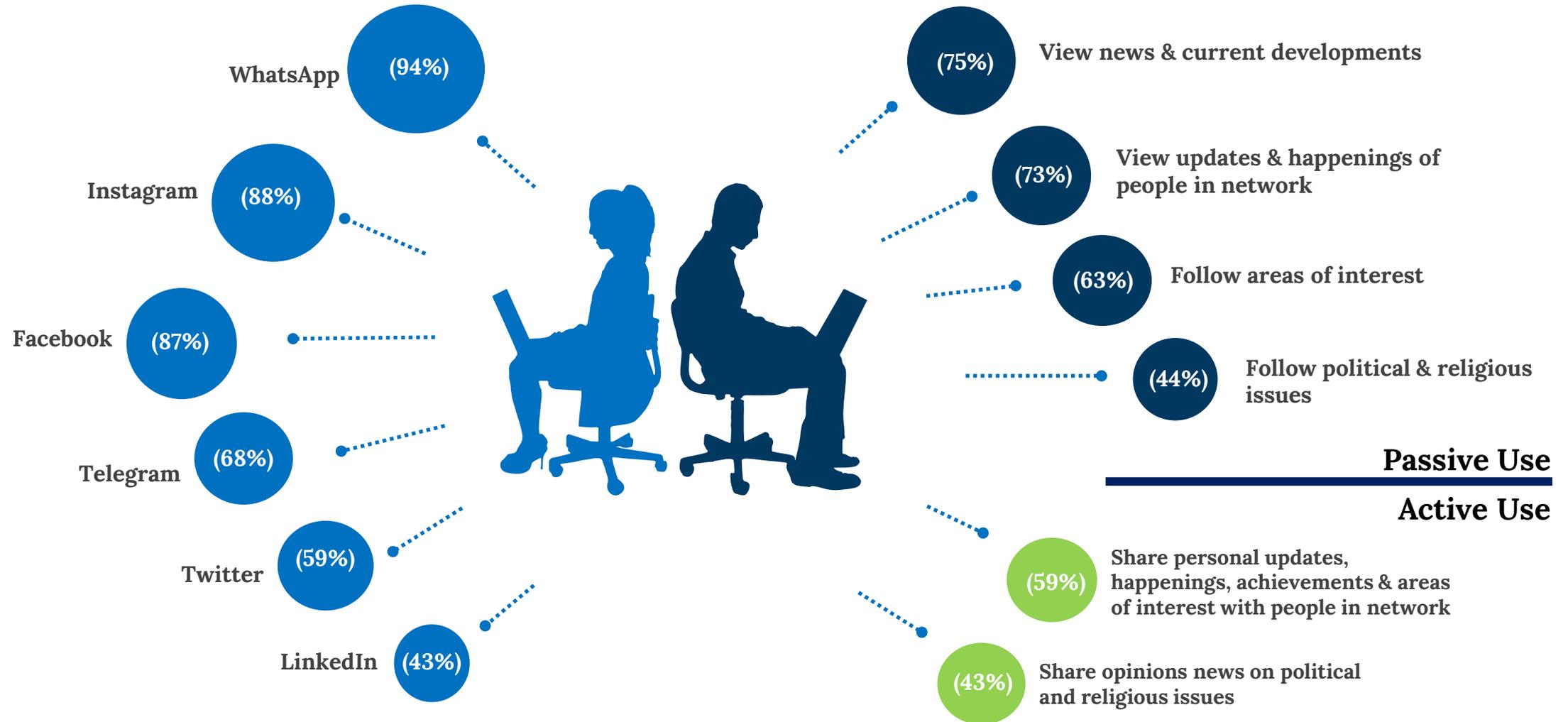


Broad Findings

Respondents' perception, reaction and behaviour towards problematic content

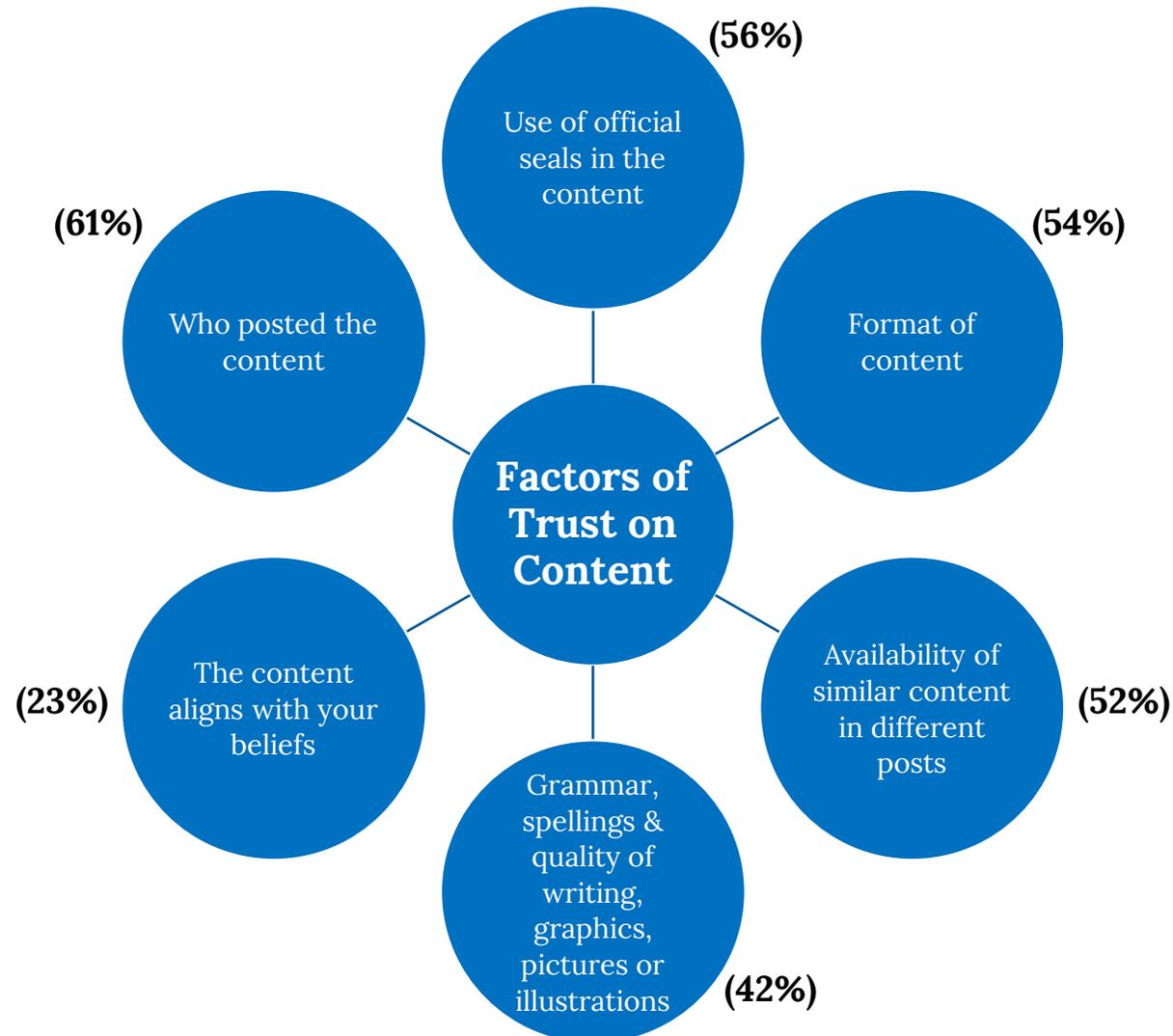
Respondents use multiple social media platforms actively (sharing content) and passively (viewing & following content).

- ✓ Question: Which of the following social media platforms do you use?
- ✓ Question: For which purposes do you use different social media platforms?





Most respondents consider the source of the content as a factor for trusting content. Further, such respondents tend to trust the government, friends/family and news agencies, more than others.



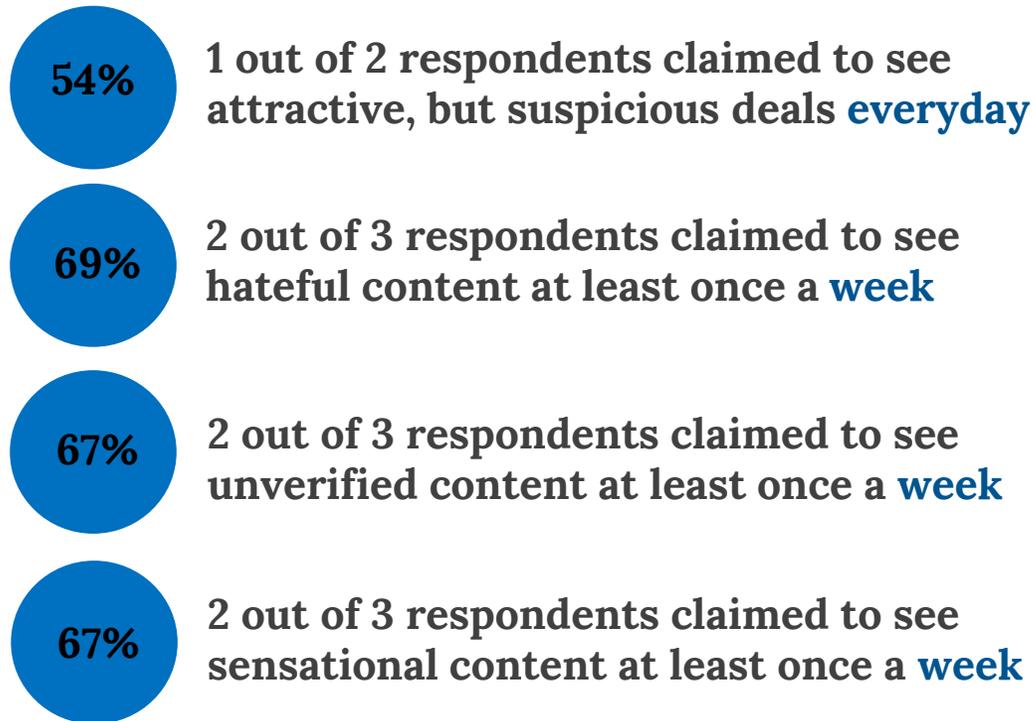
Respondents who consider the poster of content before trusting it, trust the following sources of content.

Government	(88%)
Friends & Family	(82%)
News Agency	(79%)
Popular Personality	(76%)
Big Business Entity	(71%)
Private Blog	(59%)

- ✓ **Question: Which factors do you consider while deciding whether to trust or not the content you see online? (MCQ)**
- ✓ **Question: Which source of content do you usually trust (based on who posted it)? (MCQ)**

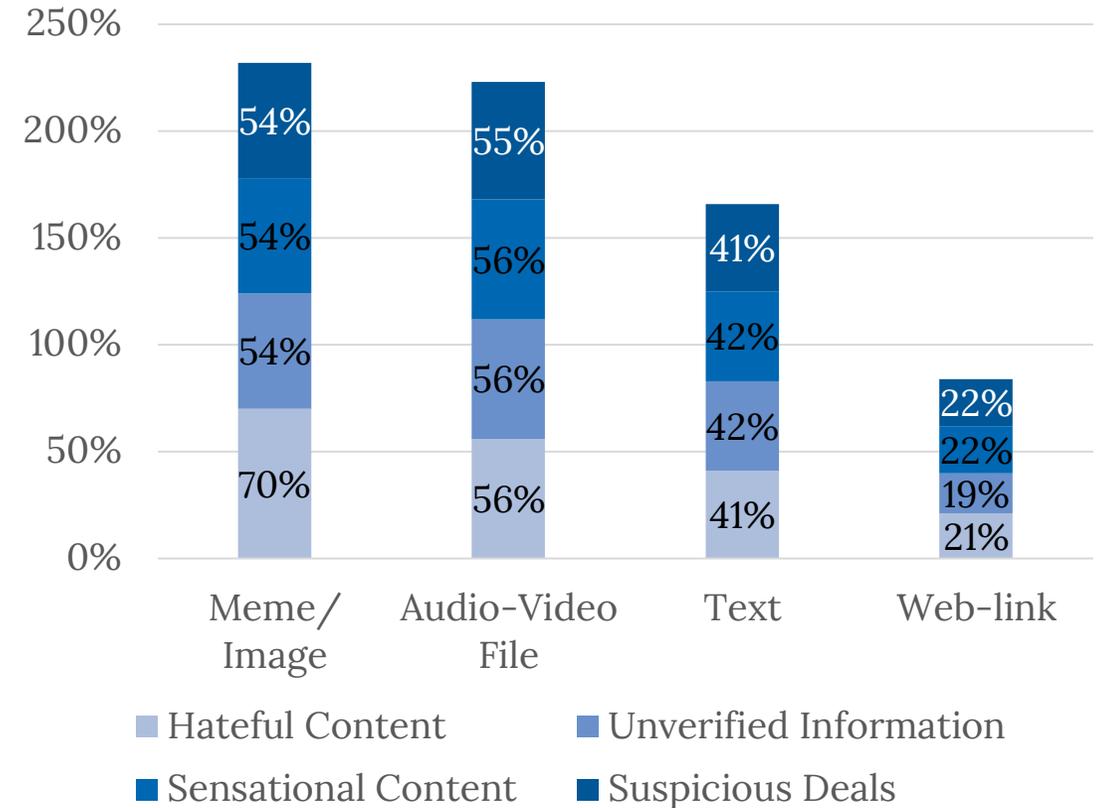
Most respondents are heavily exposed to different kinds of problematic content

✓ Question: How often do you see the following kinds of problematic content on social media platforms?



More respondents see problematic content in the form of memes/images and AV files

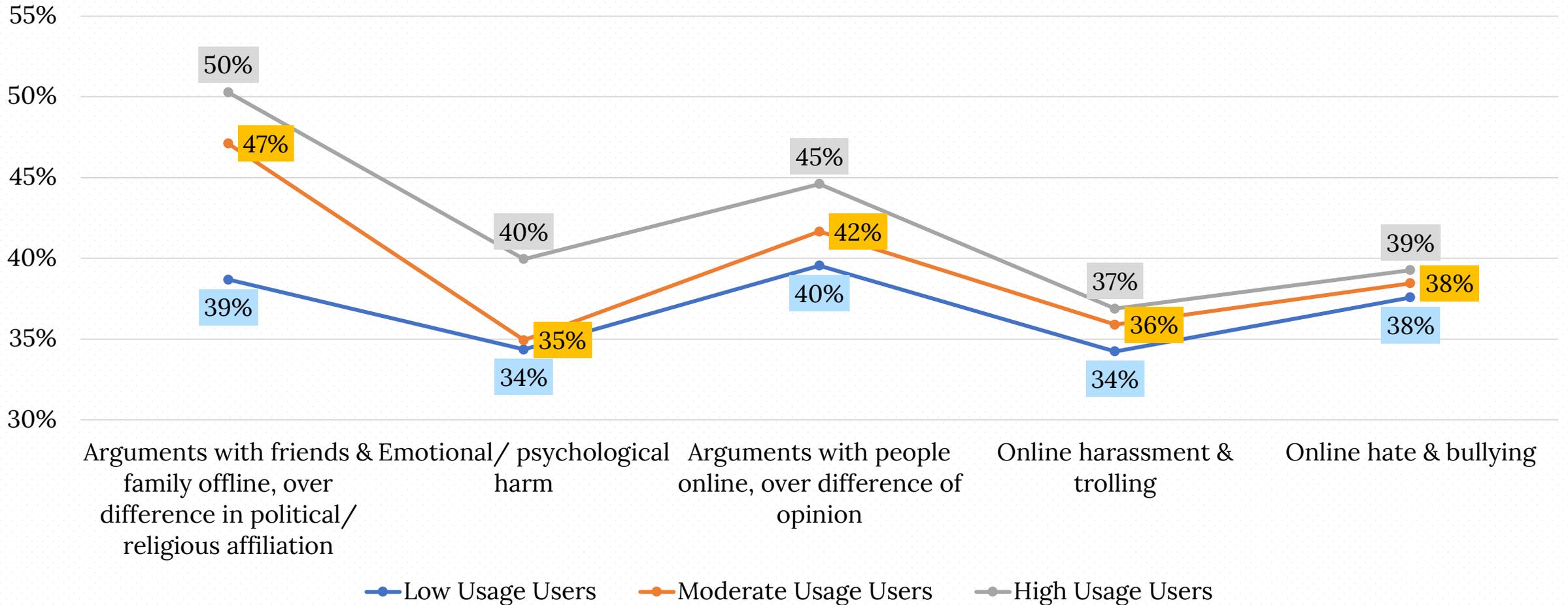
✓ Question: In which format(s) do you usually see different kinds of problematic content?



Note: The percentages given in the graph on the right, represent only those respondents who claim to be exposed to respective types of problematic content. Those who claimed to have not seen such respective content, were not considered for it.



In general, when compared to low and moderate usage users, more percentage of high usage users of social media were found to be facing harms, especially offline harms.



Note:

Low Usage Users: Users who use social media for less than 1 hour 30 mins;
Moderate Usage Users: Users who use social media for 1 hour 30 mins to 2 hours;
High Usage Users: Users who use social media for more than 2 hours.

Example for Interpretation:

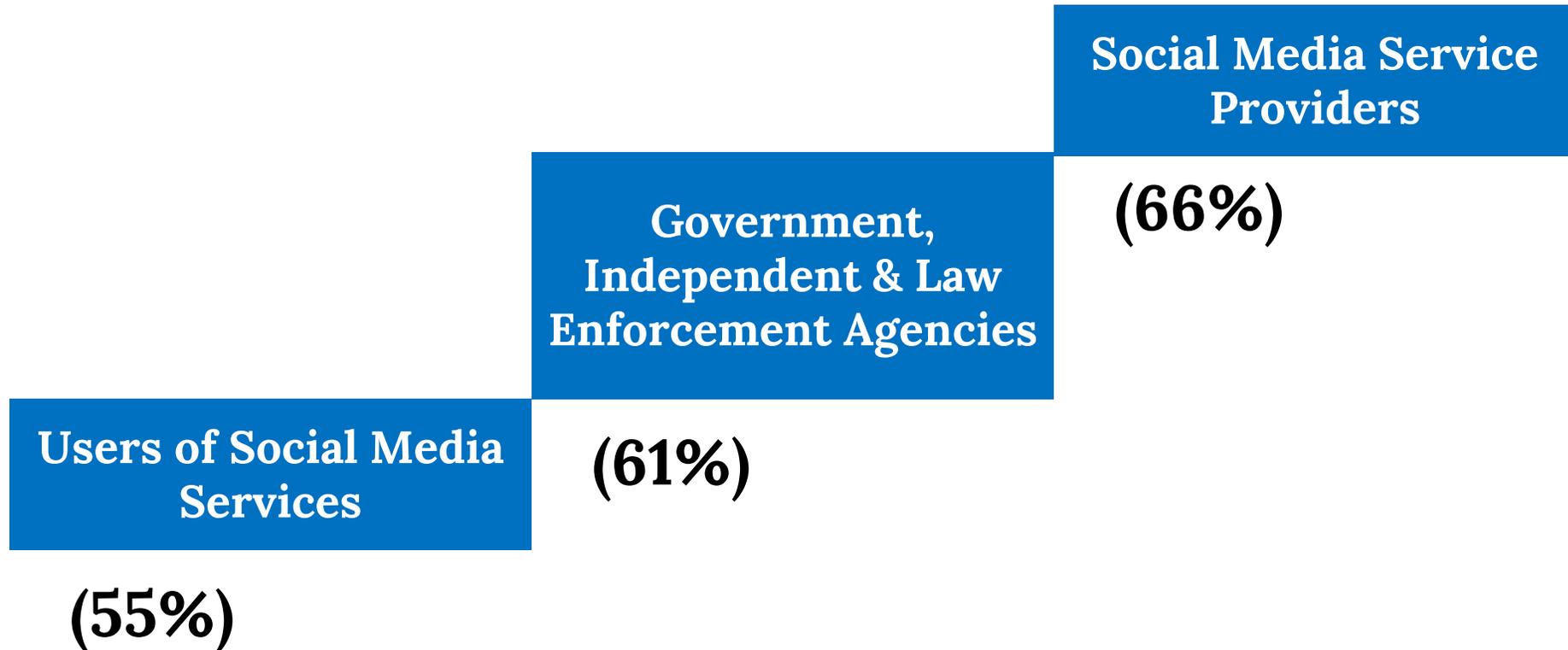
40 percent of all low-usage users experience arguments with people compared to 45 percent of all high-usage users.



Most respondents want social media platforms to be held accountable for the spread of problematic content. Notably, the least chosen options pertained to the government and its agencies.

- ✓ **Question:** Who do you think should be held responsible for curbing the spread of problematic content online?

Note: Only 1601 respondents answered this question.





Content Based Findings

Respondents' exposure to different kinds of problematic content, and harms faced due to problematic content



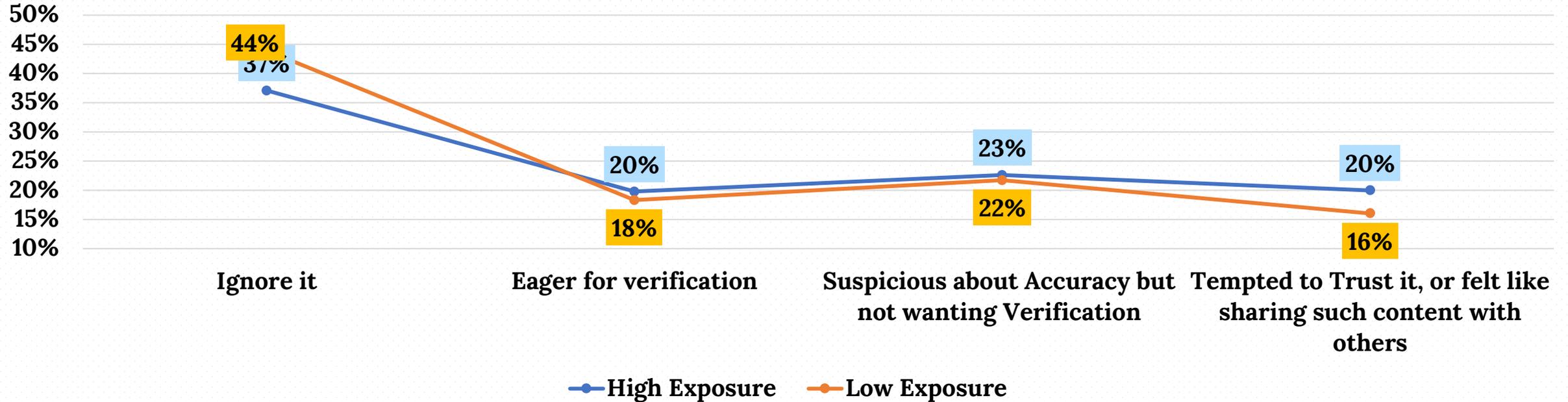
Hateful Content

User Reaction, Behaviour and Harm



Trend shows that out of all respondents with high exposure to hateful content, 63 percent engage with it and out of all respondents with low exposure to hateful content, 56 percent engage with it in one way or another.

- ✓ Question: What is your usual first reaction when you see hateful problematic content?
- ✓ Question: What do you usually do with hateful problematic content?



Note: High Exposure refers to respondents being exposed to problematic content at least once a day or once a week and Low Exposure refers to respondents being exposed to problematic content at least once a month or few times in a year.

Furthermore, trend shows that users actual behaviour is usually based upon their such first reaction to hateful content

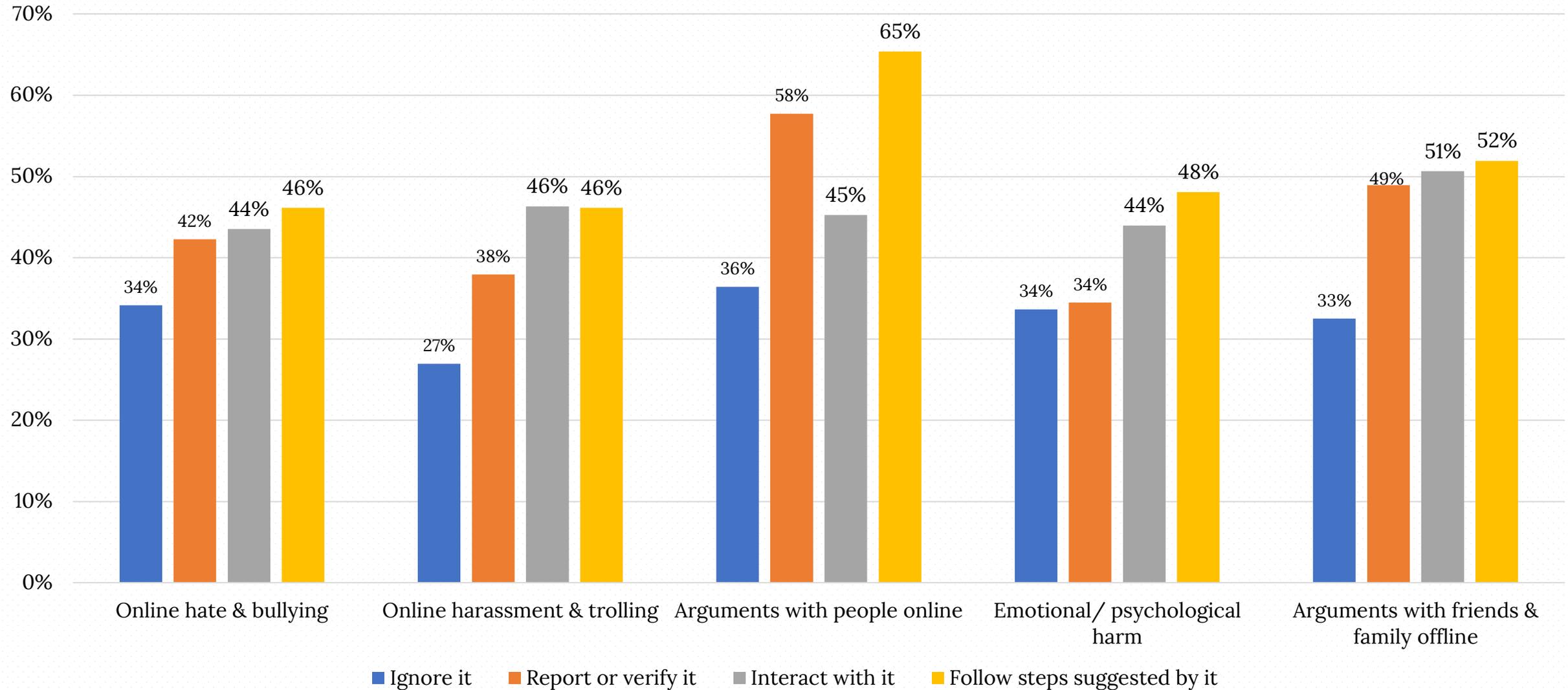
66 percent of those who feel like ignoring hateful content do so.

66 percent of those who are eager for verification, either report hateful content to the service provider or try to verify it

49 percent of those who are tempted to trust hateful content, interact with the post, i.e., comment or like it



In general, the trend shows that respondents who follow steps suggested by hateful content, claim to face more harm than others. Around 56 percent of respondents who report or verify hateful content tend to have arguments with people online.





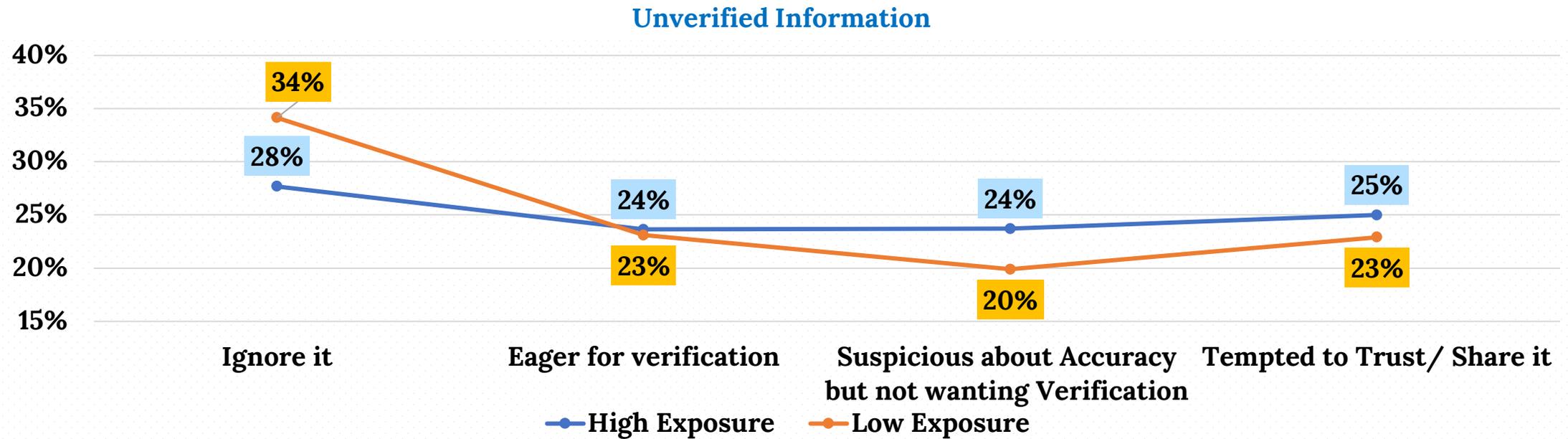
Unverified Information

User Reaction, Behaviour and Harm



In general, trend shows that when compared to respondents with low exposure to unverified information, a greater percentage of respondents who have high exposure to such content do not ignore but react in one way or another. However, irrespective of the exposure, almost an equal percentage of respondents are eager for verification for such content.

✓ Question: What is your usual first reaction when you see unverified information?



Note: High Exposure refers to respondents being exposed to problematic content at least once a day & once a week and Low Exposure refers to respondents being exposed to problematic content at least once a month or few times in a year

Furthermore, trend shows that users actual behaviour is usually based upon their such first reaction to hateful content

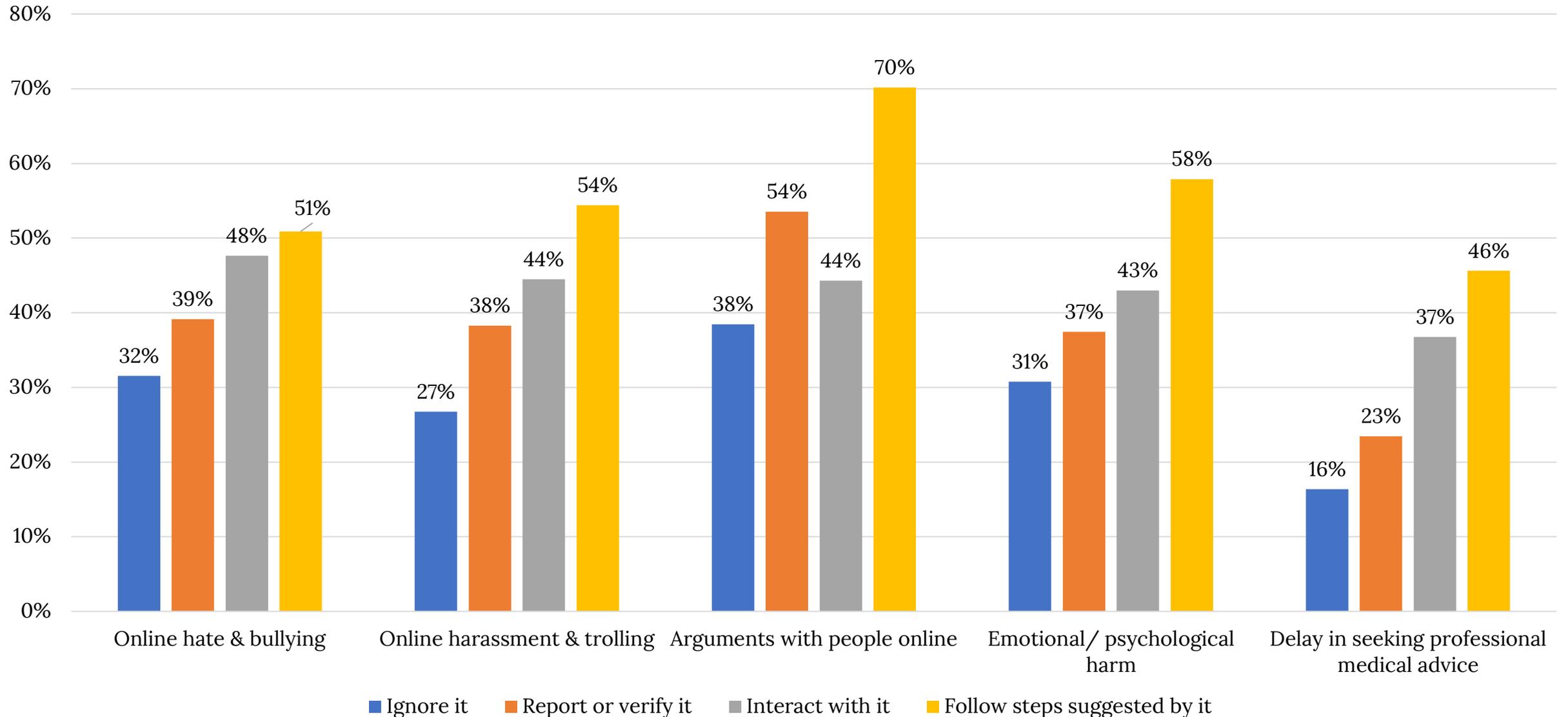
62 percent of those who feel like ignoring unverified information, actually do so.

64 percent of those who are eager for verification, either report unverified information to the service provider, or try to verify it.

46 percent of those who are tempted to trust unverified information, interact with the post, i.e., comment or like it.



In general, the trend shows that respondents who follow steps suggested by unverified information, claim to face more harm than others. Around 54 percent of respondents who report or verify such content tend to have arguments with people online.





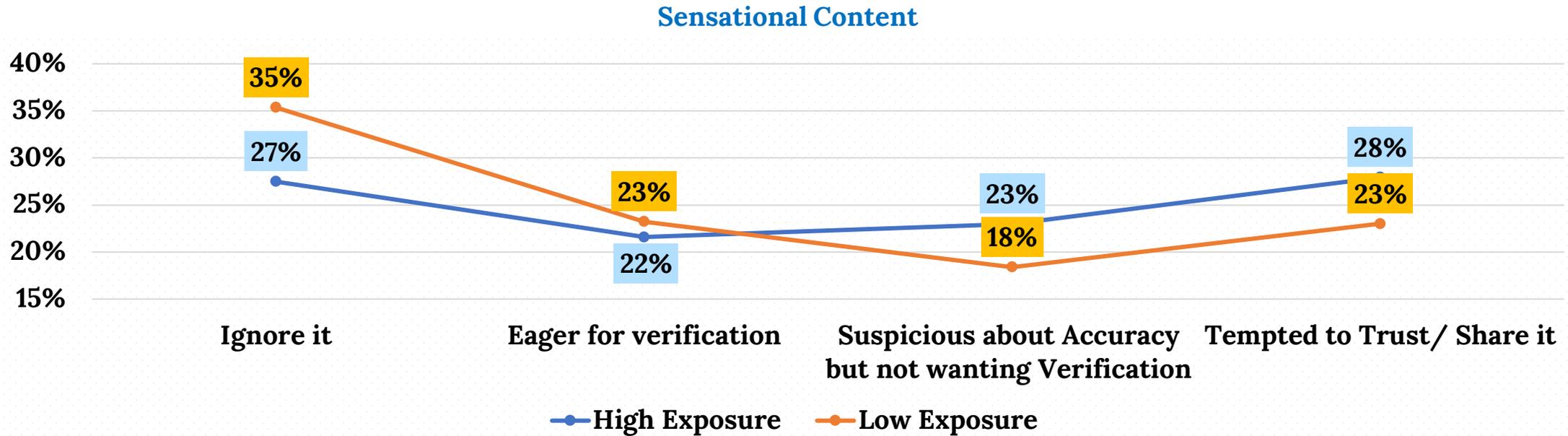
Sensational Content

User Reaction, Behaviour and Harm



In general, the trend shows that compared to respondents with low exposure to sensational information, a lesser percentage of respondents who have high exposure to such content claim to ignore it. 65 percent of respondents with low exposure and 73 percent of respondents with high exposure don't ignore sensational information.

✓ Question: What is your usual first reaction when you see sensational information?



Note: High Exposure refers to respondents being exposed to problematic content at least once a day & once a week and Low Exposure refers to respondents being exposed to problematic content at least once a month or few times in a year

Furthermore, trend shows that users actual behaviour is usually based upon their such first reaction to hateful content

62 percent of those who feel like ignoring sensational information, actually do so.

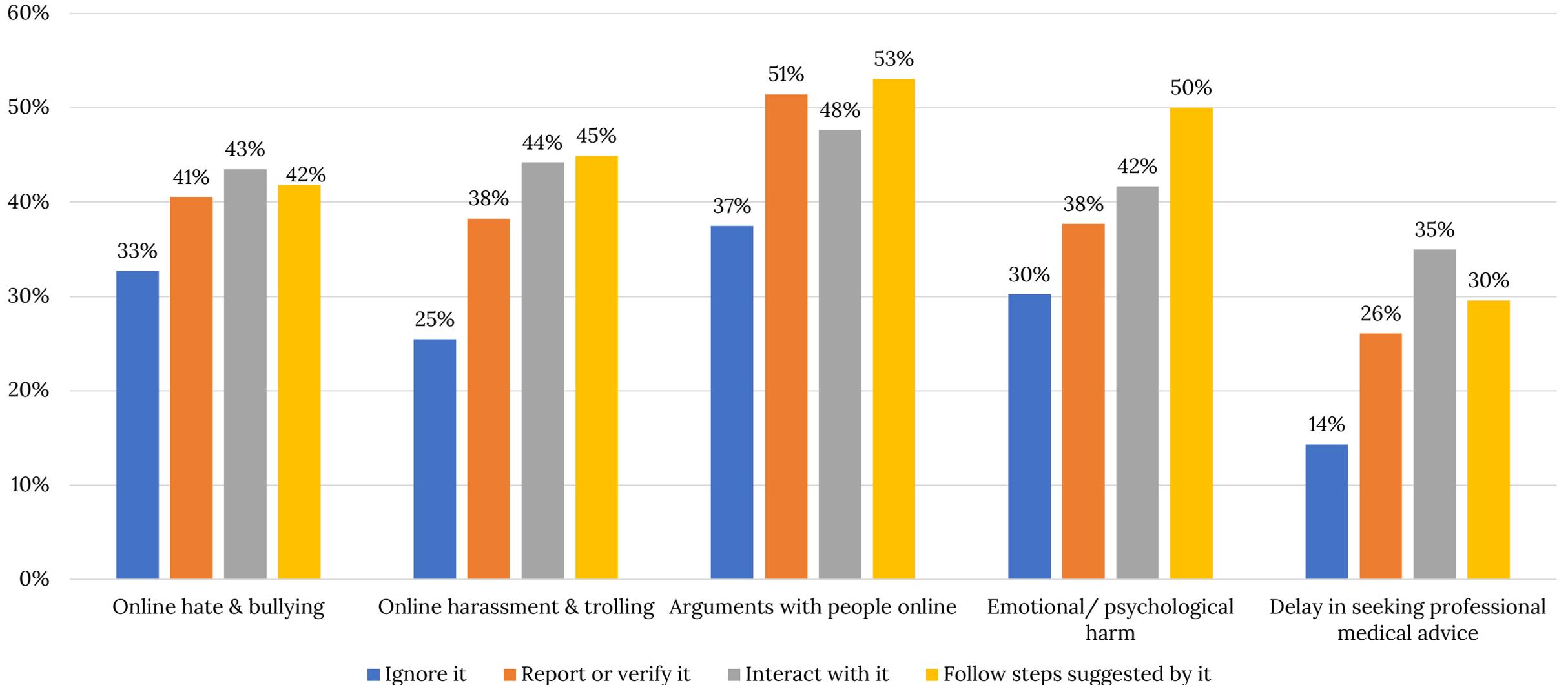
56 percent of those who are eager for verification, either report sensational information to the service provider, or try to verify it.

45 percent of those who are tempted to trust sensational information, interact with the post, i.e., comment or like it.



In general, the trend shows that respondents who follow steps suggested by sensational information or interact with a content face more harm than others. A large percentage (51 percent) of people who report and verify have arguments with people online.

Figures in Percentage





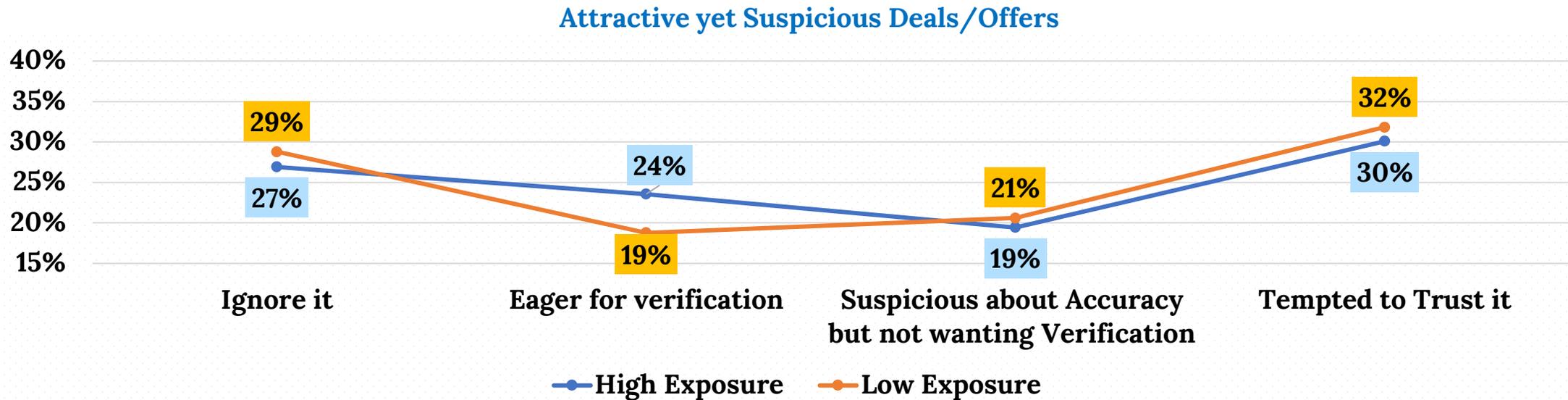
Suspicious Deals / Offers

User Reaction, Behaviour and Harm



In general, the trend shows that out of all respondents with low exposure to attractive yet suspicious deals/offers, 19 percent are eager for verification, as compared to 24 percent in the case of high exposure respondents. Further, around 30 percent of respondents, irrespective of their exposure, are tempted to trust attractive yet suspicious deals/offers.

✓ Question: What is your usual first reaction when you see attractive yet suspicious deals/offers?



Note: High Exposure refers to respondents being exposed to problematic content at least once a day & once a week and Low Exposure refers to respondents being exposed to problematic content at least once a month or few times in a year

Furthermore, trend shows that users actual behaviour is usually based upon their such first reaction to hateful content

62 percent of those who feel like ignoring attractive yet suspicious deals/offers, actually do so.

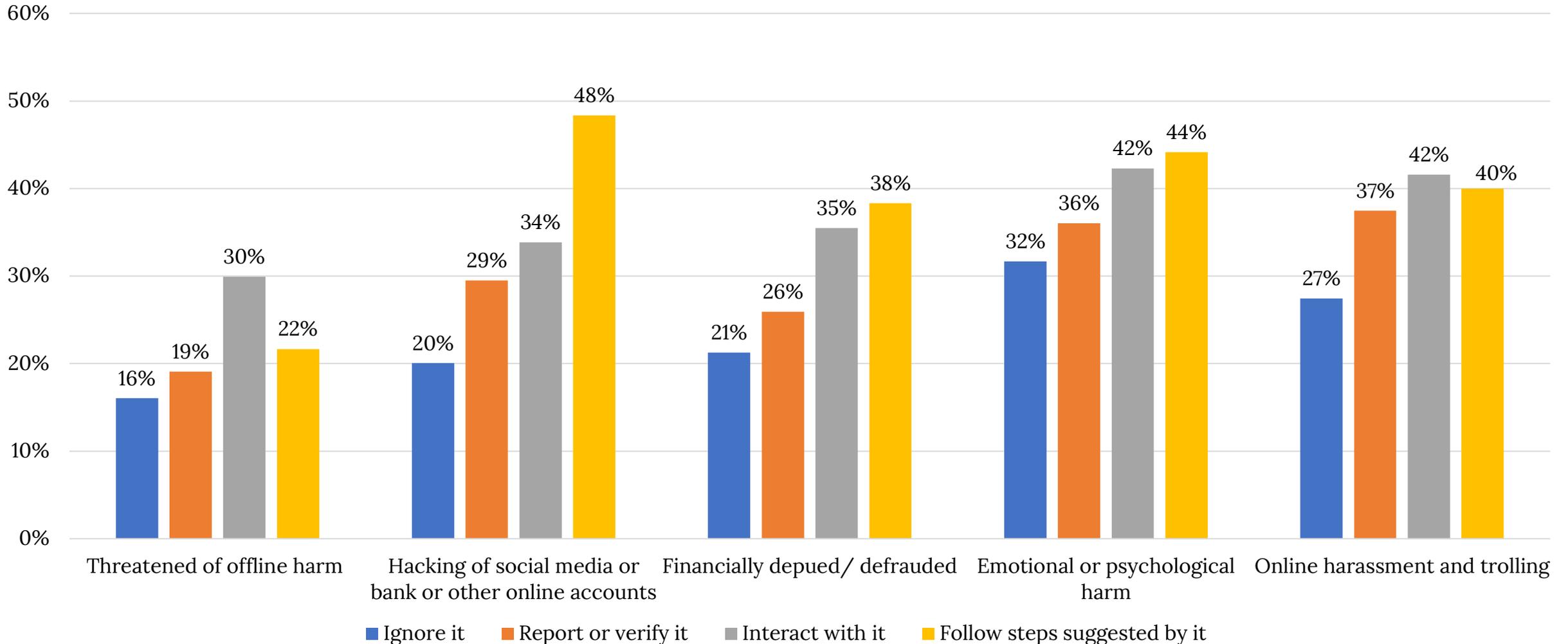
64 percent of those who are eager for verification, either report attractive yet suspicious deals to the service provider or try to verify it

46 percent of those who are tempted to trust attractive yet suspicious deals/offers, interact with the post, i.e., comment or like it



In general, the trend shows that respondents who follow steps suggested with attractive yet suspicious deals/offers or interact with such content claim to face more harm than others. The trend shows that fewer people who ignore such content face harm. Hacking social media or bank accounts happens for 48 percent of people who follow steps suggested by such content.

Figures in Percentage



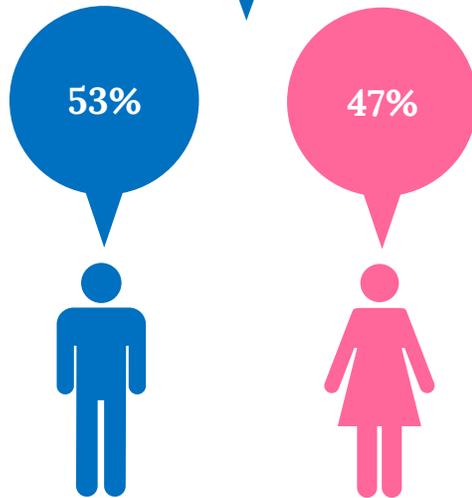


Respondents' Profile

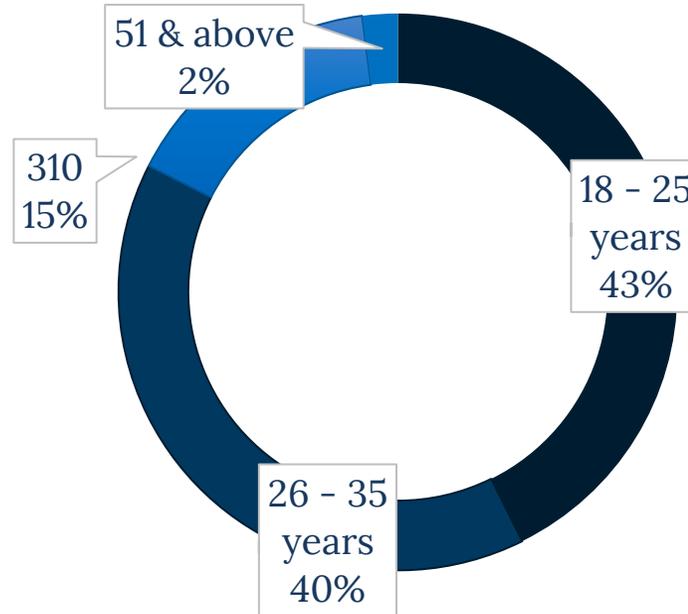
Age, Gender, Religion, Geographic Location, Education Level, Occupation & Income Level of Respondents, as well as the amount of time spent by the on social media

A Total of 2002 Respondents were Surveyed

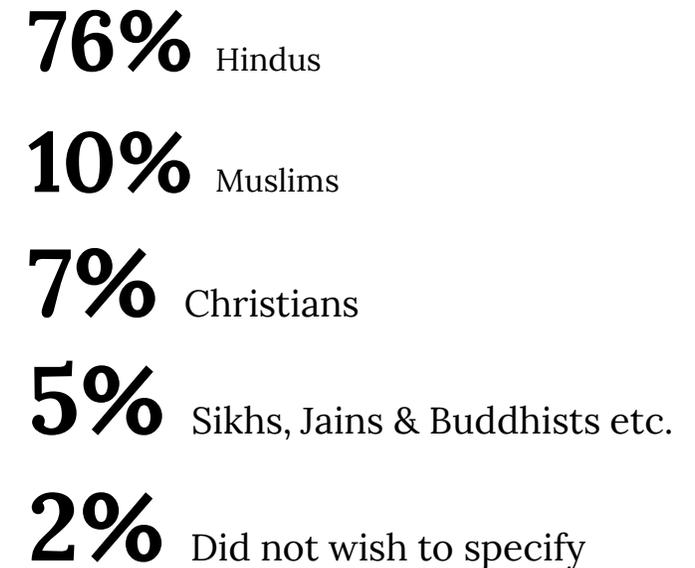
Almost equal number of male-female respondents were surveyed



Respondents from all age groups were surveyed



Respondents from prominent religions were surveyed



- Note:**
- A. A purposive random sampling methodology was adopted for identifying the survey respondents.
 - B. A neutral gender ratio was maintained.
 - C. Conscious efforts were made to ensure a healthy mix of respondents from different age groups, and religions.



Respondents from multiple districts of different states were surveyed

Uttar Pradesh

Agra, Aligarh, Allahabad, Azamgarh, Ayodhya, Baghpat, Bareilly, Etawah, Jhansi, Kanpur, Lucknow, Mathura, Meerut, Muzaffarnagar, Varanasi, Saharanpur etc.

Madhya Pradesh

Ashok Nagar, Balaghat, Betul, Bhind, Bhopal, Satna, Gwalior, Guna, Indore, Jabalpur, Kotma, Ujjain, Morena, Rewa, Ratlam etc.

Assam

Guwahati, Tezpur, Silchar, Barama, Dhemaji, Dhubri, Dibrugarh, Jorhat, Nalbari etc.

Kerala

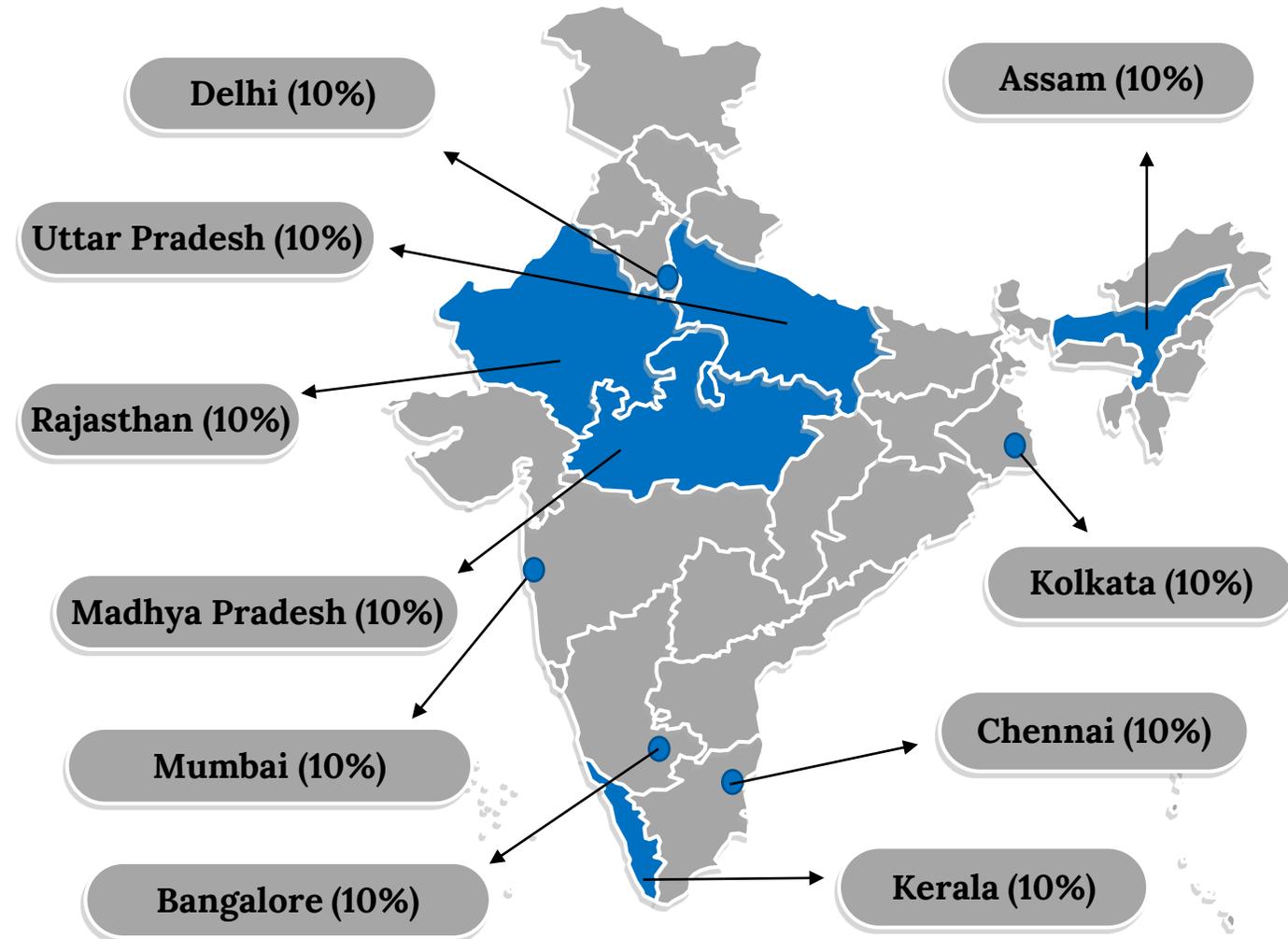
Alappuzha, Kottayam, Malappuram, Trivandrum, Cochin, Kozhikode, Thrissur, Kollam, Palakkad etc.

Rajasthan

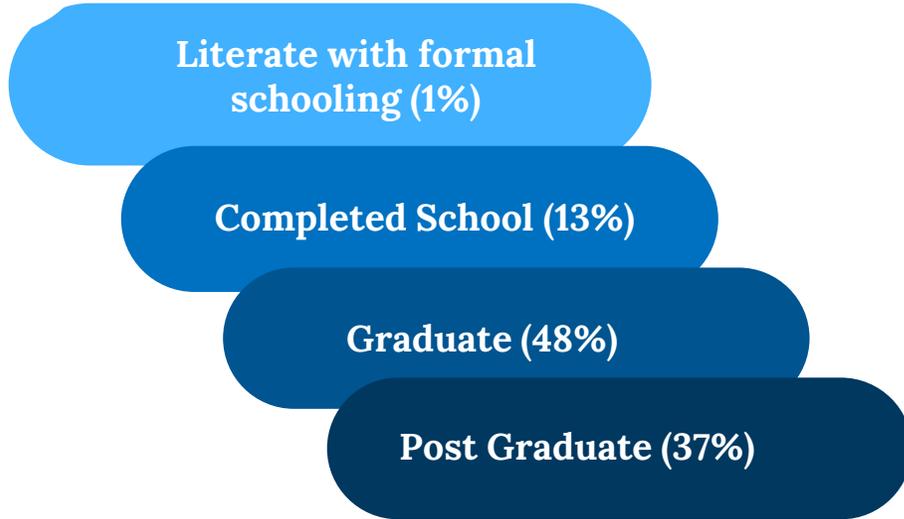
Jaipur, Chittorgarh, Kota, Udaipur, Bhilwara, Bikaner, Jodhpur, Ajmer, Alwar, Sawai Madhopur, Pratapgarh, Barmer, Jhalawar etc.



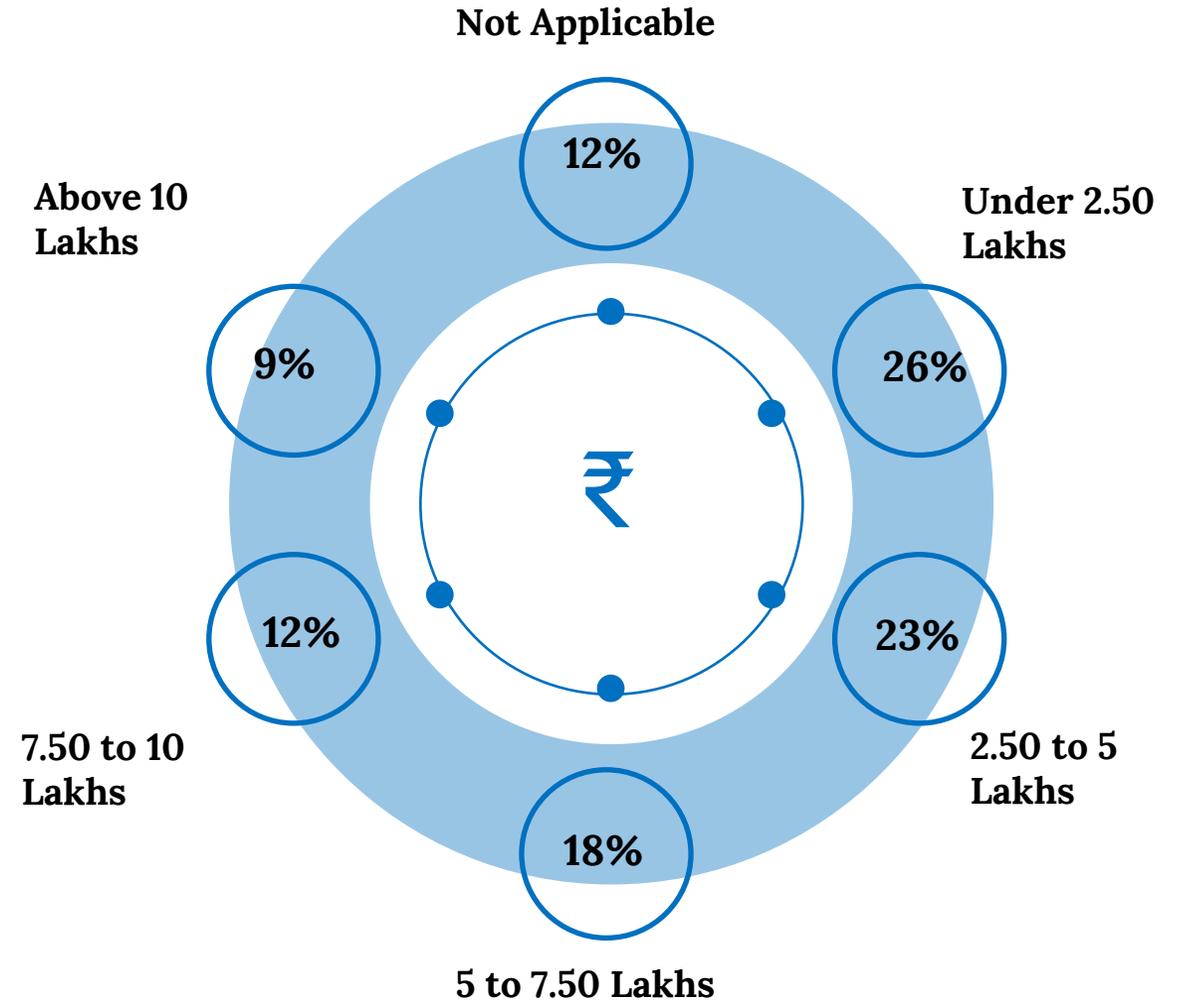
Geographic Diversity of Respondents



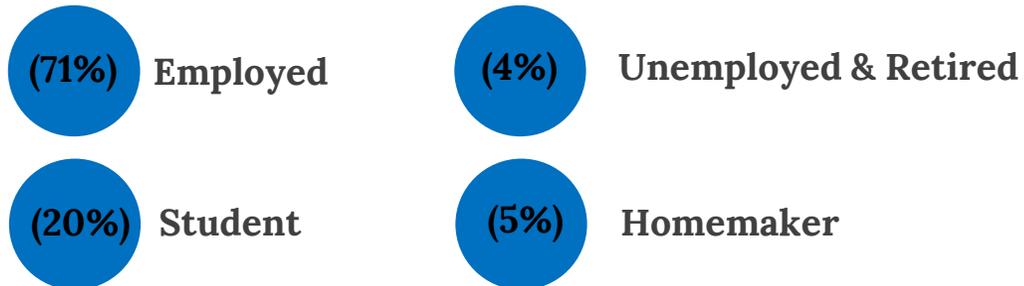
Level of Education of Respondents



Annual Income Level of Respondents

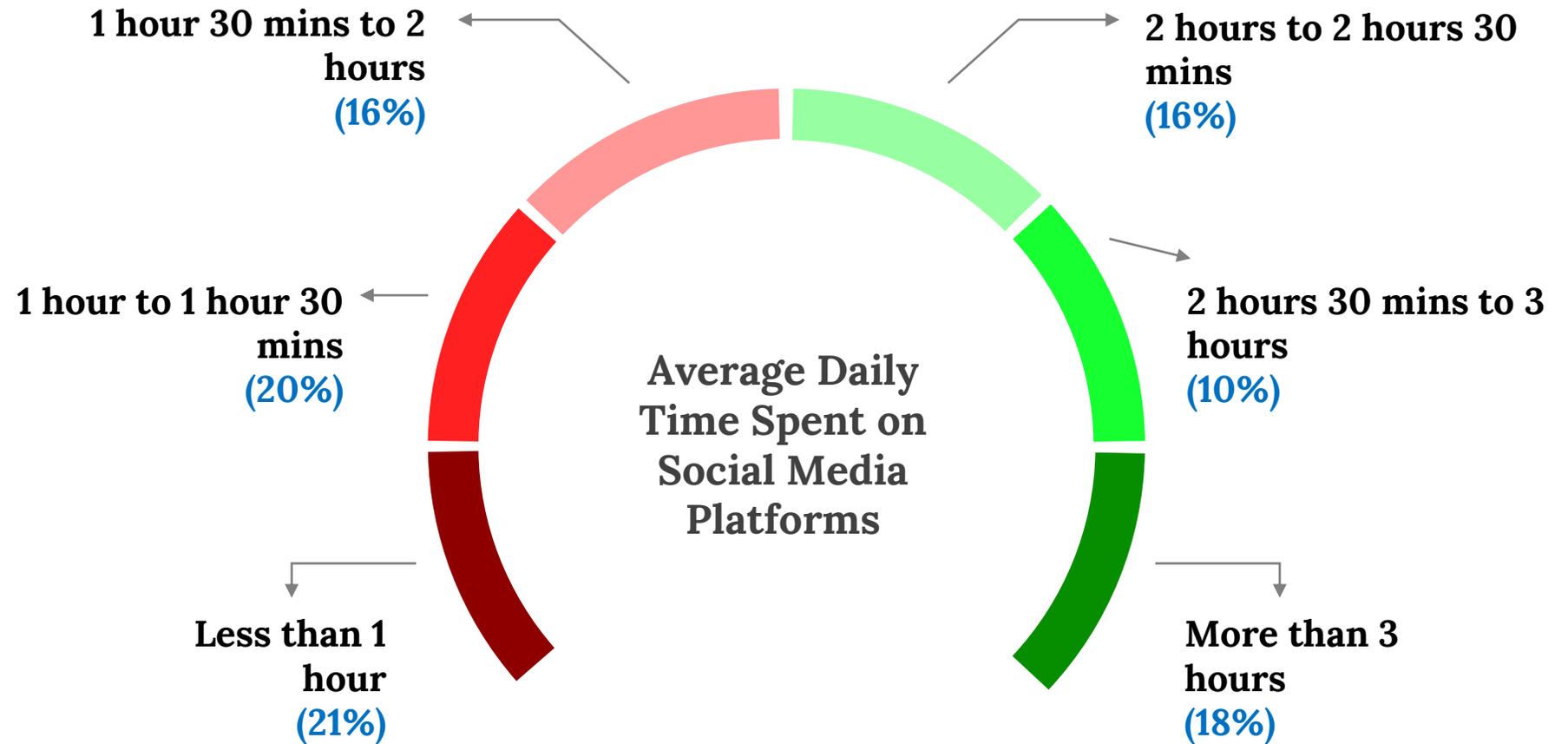
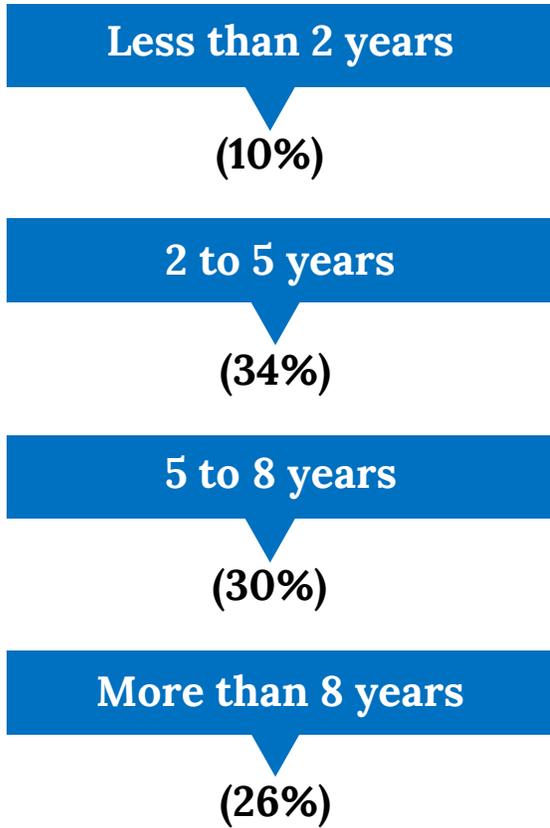


Occupation Status of Respondents



Note: A. Respondents from different education and income levels, as well as occupation were randomly selected. However, diversity in the same was ensured.
B. Income Data was available only for 1931 respondents.

A healthy mix of respondents were reached out to, with respect to average time spent on social media, and experience of using social media



The Team



Sidharth Narayan

Policy Analyst

sid@cuts.org



Prince Gupta

Senior Research Associate

prg@cuts.org



THANK YOU

Comments and suggestions are welcome

Consumer Unity & Trust Society (CUTS)
D-217, Bhaskar Marg, Bani Park, Jaipur
302016, Rajasthan, India

Ph: +91 141 2282821,
Fax: +91 141 2282485,
Email: cuts1@cuts.org