Personalized WhatsApp Privacy: Demographic and Cultural Influences on Indian and Saudi Users

Jayati Dev(jdev@iu.edu)
Sanchari Das(sancdas@indiana.edu)
Yasmeen Rashidi(yrashidi@indiana.edu)
L Jean Camp (ljcamp@indiana.edu)

Indiana University Bloomington

Abstract

We offer an analysis of privacy perceptions and behaviors that focus on one of the most popular social networking platforms in cultures where it is heavily used. WhatsApp is the most popular social platform in many parts of the world, specifically on the Arabian peninsula and India. Saudi Arabia has the highest density of WhatsApp adoption, and India has the largest number of WhatsApp adopters. Prior research on security and privacy in social applications have focused primarily on Western users, and thus on Facebook. We report on an analysis of privacy perceptions and behaviors of Arab and Indian WhatsApp users. We examine the results of surveys of a total of 820 Saudis, non-Saudi Arabs, and Indians who use WhatsApp for their daily interactions. Each survey queried participants' use (or avoidance) of features and their concerns about privacy in WhatsApp, as well as asking demographic questions. We report significantly different preferences and influences on behaviors. These differences, especially when it comes to gender, provide a nuanced and complex picture of privacy through the lens of WhatsApp use. We conclude with privacy recommendations for WhatsApp, specifically, and for privacy in social networks generally for Indian and Arabs WhatsApp customers. The recommendations address blocking, location sharing, and group dynamics to empower WhatsApp users in these populations.

1 Introduction

WhatsApp is a multimedia messaging application with a range of capabilities: asynchronous chat, photo sharing, video sharing, synchronous voice and video chat, and location sharing between individuals, ad-hoc groups, or more structured groups. It is an application that is widely adopted in the Middle East and

South-East Asia [30, 37]. While there have been independent qualitative studies on WhatsApp for Saudi Arabian users [30] and Indian users [12] separately, there has been no work that has compared the privacy choices of these two geographically and culturally distinct populations. Having an identical service with the same settings in two very different populations offers an opportunity for a comparison of privacy preferences and behaviors. One of the underlying observations of our results, one reified by the researchers from the region of study, is that WhatsApp functions as the dominant social network for many of our participants.

Prior research on social networking applications have found that privacy preferences of users vary significantly across cultures [40]. However, the design of features and options in such applications are often uninformed by the diverse and complex privacy behavior of non-western populations. Furthermore, the majority of social media privacy studies focus on western populations. Perhaps because there is little use of WhatsApp in the West, research on this specific platform is also sparse. In fact, WhatsApp is a so widely used in India and Saudi Arabia that it is a topic of popular political debate, having been an condemned for being a tool for political propaganda [17] and for instigating mob violence based on forwarded misinformation [16].

In this work we explore how demographics affects privacy perceptions between two cultures. We also examine how demographics correlate with expressed privacy and security concerns as well as with privacy and security behaviors. We implemented highly similar studies that first queried participants about information sharing, feature use, and settings. We then queried the participants about their concerns about security and privacy in WhatsApp, as well as their demographics. Specifically, we extended a previous analysis of Arab WhatsApp use by reproducing the study with Indian participants.

We found significant differences not only in feature use, privacy preferences, information sharing, and settings but also in the relationship between these observations and the demographics and privacy preferences of the populations. Similarities also emerged, for example, age was a stronger predictor of profile information sharing practices than cultural identity. Gender was a greater determinant of use of the blocking feature and profile information sharing than cultural identity. However, there was some subtlety in the comparisons in that the regression coefficients and levels of significance of age and gender varied within the populations.

We recommend specific features to address the privacy preferences of young and old, women as well as men, and Arabs as well as Indians. We also show that there is a need for further research on more tightly defined groups of participants by illustrating in analysis of the Arab populations that Saudi Arabs differ substantially from non-Saudi Arabs. We cannot conclude if this is a function of the ubiquity of WhatsApp in Saudi Arabia, cultural factors such as male guardianship, or differences in perceptions of social desirability. Without such research it remains possible that large numbers of people on the WhatsApp platform have technically straight-forward privacy requirements that are not implemented.

Our research questions address privacy concerns, perceptions, and What-

sApp information management practices. A cross-cultural analysis is especially important in messaging platforms as they become more pronounced in societies across the world. Our goal is to address the following three questions.

RQ1: To what extent do personal experiences on or perceptions of WhatsApp impinge feature usage?

We found that personal experiences like being contacted by strangers, as well as general perceptions about privacy impinge WhatsApp feature usage. We identified commonalities in privacy behaviors across both cultural groups. For instance, people who report sharing sensitive information also report using the blocking feature. Furthermore, those who would rather not be added to groups disable auto-download from these groups. We also explore some of the reasons that motivate such responses from both groups.

RQ2: To what extent does privacy behavior on WhatsApp vary between Arab and Indian participants?

To answer the above question, we analyze survey responses of 820 participants of self-identified Arabs and Indians. While these participants might or might not be currently residing in the respective countries, they identify themselves to have either Arab or Indian citizenship as appropriate. We qualitatively analyze the responses from these populations separately as well as comparatively. Our conclusions are focused on WhatsApp because of its ubiquity as the most popular messaging platform in Saudi Arabia and India; just as studies in western populations focus on Facebook Messenger [13]. Comparisons about the use of WhatsApp features and sharing settings offers insight into if and how individuals are influenced by their cultural, and demographic differences. The comparison between the two populations using a similar experimental approach may provide general insights on privacy decision-making, which can further inform design considerations on these platforms.

RQ3: How do demographics effect WhatsApp feature usage?

The WhatsApp questions included variables to address this question. Specifically the demographic questions were age, gender, educational background, and place of origin. The questions for both populations were further comprised of usage specific questions, WhatsApp privacy feature settings, and privacy behavior related questions. The privacy behavior questions were specific to WhatsApp (e.g., *Are you concerned about strangers contacting you with WhatsApp?*). The six features that we addressed across both populations were *Auto-Download*, *Blocking*, *Last Seen*, *Profile Photo*, *Location*, and *Status*. We noted some striking but subtle differences in the role of gender and age in blocking behavior, use of the Last Seen feature, and Profile Photo settings.

Our results reify previous findings that non-WEIRD (Western, Educated, Industrialized, Rich, and Democratic) populations differ from western popula-

tions [40]. While we explore some of the qualitative responses to better understand some of our quantitative findings, additional controlled laboratory investigations or qualitative research could illuminate the source of the differences more broadly. Our findings imply that we may not be able apply non-WEIRD results from different social media platforms to other WEIRD populations, as previously noted in [28]. However, this finding is impinged by the fact that the dominant social network platform is different in these populations, and thus the specific features vary. Thus location privacy on WhatsApp may not map directly to location privacy on Facebook, for instance.

It was not the case that geographical adjacency nor ethnic and religious similarities translated uniformly to privacy perceptions and behaviors. Specifically, there were results where non-Saudi Arabs were closer to Indian participants than the Saudi participants. From the observed nuances in privacy behavior, we can argue not only that research should embrace the diverse behavior in non-WEIRD populations for WhatsApp as experienced as a social media platform, but also that all non-WEIRD populations are not all the same. In fact, there can be nuanced behavior within subsets of these populations. We first found the demographics had different relationships to WhatsApp privacy behavior in the two populations: Arabs and Indians. Further analysis found that Saudi and non-Saudi Arabs had differences in privacy behavior as well, illustrating that there is a need for broader investigation and more targeted analysis of privacy preferences of global user populations. For example, our conclusions about the Indian participants may predict the behaviors and preferences of Bangladeshi or Burmese participants; but our findings argue against asserting this.

We believe that the findings from this study can benefit the way mobile messaging applications are designed in part by illuminating their role as social networking platforms. While we take a quantitative approach, we include qualitative validation of our recommendations. This is a first step crossing boundaries into understanding social media applications which have a more significant presence in non-WEIRD populations.

In the following section, we present an overview of other cross-cultural privacy studies. Section 3 details our recruitment, data compilation, and analysis procedures. Section 4 presents the results of the analyses of individual populations, and comparisons between them. We then discuss the implications of the results in Section 5, making recommendations for changes in WhatsApp and supporting these with both our numerical findings and comments from our participants.

2   Related Work

In December 2017, WhatsApp had 1.5 billion active users in contrast to Facebook's 600 million [37]. Other than the studies on Indian and Arab users (which are the subject of further analysis here) WhatsApp has been considered by western researchers in the security and privacy community as an instant messaging platform closer to Signal than an active social network closer to QZone [6]. De-

spite the ubiquity of WhatsApp in daily life in many non-WEIRD countries it is often not considered in social networking research.

A core motivation of our work is that privacy studies have primarily focused on WEIRD societies and social networking studies have focused on Facebook. Despite the role of WhatsApp in Arab and Indian daily life, these populations have largely remained understudied in academic literature. Here we address the privacy choices of these participants in a technology of daily use, and report how these interact with demographics.

In privacy as well as in other behavioral research, WEIRD populations are not necessarily representative of other populations [15]. It is reasonable to evaluate if privacy research on WEIRD populations predicts findings from South-East Asian and Middle-Eastern populations. Even within the WEIRD domain privacy perceptions and behaviors vary, where the dependant variables can vary based on the demographics, culture, occupation, interaction media, and other factors [10]. Cvrcek et al.'s study of privacy valuation across Europe found significant differences between Greek, Belgian, Czech, German, and Slovak populations in terms of location privacy indicating the importance of studying culturally varied populations [9]. Further afield, privacy risk perception of German participants were found to be higher than American participants, and both were higher than their Chinese counterparts [33]. The difference in the WEIRD populations were partly credited to the presence of data protection in Europe, but culture also played a role [32, 4, 8]. Similarly, privacy preferences were found to be distinct among American, Chinese, and Indian populations [40], with American users of Facebook expressing greater privacy concerns than their Indian or Chinese counterparts. This may reflect the fact that Facebook is dominant in the US, while not among the top social networks in China. In this paper, due to the sparse use of WhatsApp in WEIRD countries, we only compare Indians and Arabs' privacy perceptions.

A previous research on Indian populations' offline privacy attitudes and behavior surveyed 407 participants. The design was grounded in surveys that included only American participants [20]. We use a similar methodology where the same survey instrument was used in two populations.

Studies of offline risks have consistently found strong evidence that the tolerance for risk [33] and the cultural framing of risk vary significantly across nations [2]. These lower levels of concern about information disclosure reported by Indian participants may reasonably be expected to be reflected online. For example, a study of 201 Facebook users in the United Kingdom found that participants' perceived risk of sharing information on Facebook was a significant predictor of privacy concerns and precautionary behaviors [36]. Bellman et al. examined privacy concerns of internet users across different cultural and political settings as well as level of user expertise [3]. Indian participants were found to have higher levels of trust in information disclosure in the public and private sectors. Again there was evidence that privacy concerns vary significantly across cultures, and is an important factor organizations must consider before framing policies.

The differences between countries and lower privacy concerns in India were

both reified by cross cultural research on privacy by Wang in 2011. A study of 92 participants in three countries found that generally American respondents were the most privacy concerned, followed by the Chinese and with Indians showing the least concern [40]. However, research in risk perceptions on various other social media platforms (including Friendster, MySpace, and Facebook) has reported weak correlations between user's privacy choices and their online behavior [1]. Most of the users were unable to or uninterested in addressing privacy settings to control information sharing. The source of this 'privacy paradox' was investigated in a study of 232 Facebook users, where the perceived risk of sharing information was found to be the most important determinant of privacy behaviors. The usability of privacy controls was also statistically significant, but with a lower impact. Privacy preferences, measured using a standard Likert scale, were found to be significant but to have the least impact on behavior [14].

In contrast, King, Lampinen and Smolen report privacy attitudes to be a consequence of previous events rather than overall risk perception [18]. Supporting this result, Lewis, Kaufman and Christakis argue that privacy behaviors are a result of *'social influence'* and *'personal incentive'* [21] such as peer attitudes and cultural biases. If privacy attitudes are primarily a function of cultural attitudes, then examination of privacy is different cultures is needed to provide the support for different populations. Patil and Kobsa have similarly argued against risk perception being a primary determinant of privacy. Specifically their work has found that people are more privacy sensitive towards accessibility of information to strangers, content of the messages in communication, and reliability of the service. [27] In cases where privacy protecting behaviors are present, they suggest that this is a result of *impression* management, specifically in messaging apps. [19]

Privacy concerns were found to vary based on data type as well as data content. For example, perceptions and valuation of location sharing as a privacy risk vary across contexts and between individuals, and nations [7]. Saudi Arabia has the highest WhatsApp market penetration, with 78% of the population using WhatsApp [38]. Concurrently, it is not very surprising that WhatsApp in Saudi Arabia is subjected to targeted governmental regulations, including a previous ban on voice calls over the platform (This ban was lifted in 2017 [31]). WhatsApp is also treated as a credible source of information for law enforcement [23]. It is integrated into daily life, in educational institutions, and places of employment.

WhatsApp is also the dominant messaging application in India with over 200 million users [37]. Dramatically illustrating the reach of and trust in WhatsApp, a series of false reports about the kidnapping of children swept across India on WhatsApp in 2018. Information widely shared on the platform played a significant role in lynching and other mob violence in communities where the stories were widely believed [5]. Photos of Syrian children were distributed with claims that these were kidnapped and abused Indian children. Reacting to the brutality of the photos distributed over WhatsApp, five people were incorrectly identified as kidnappers and killed by the residents of isolated towns [42]. In

6

response, WhatsApp has implemented tagging to indicate when messages are forwarded, as opposed to having them appear to originate with the sender. WhatsApp has also limited the ability to forward a specific message to five people to prevent mass forwarding [44]. The goal was to enable recipients to distinguish first hand accounts from hearsay, and to explicitly curb spread of misinformation [45]. Though the latter event happened after the breadth of our study, it highlights how WhatsApp is an important focal point in behavior over social media, especially for these understudied populations.

However, to the best of our knowledge, there is limited literature that explores cross-cultural differences in how users in vulnerable populations interact with mobile messaging platforms. Previous research work indicates that the specific data shared by WhatsApp have privacy concerns in India [11] are different than concerns found in Saudi Arabian populations [30]. In the sections that follow, we delve more deeply into how culture affects the interactions among privacy perceptions, feature selection, and demographics.

3 Method

The study was conducted through a survey instrument targeted at WhatsApp users, above the age of 18, who identified themselves as either Arabs or Indians based on nationality. The instrument was initially developed as a bilingual self-reported survey for Saudis in Arabic and English. We adopted the English version for the Indian population, but did not translate into any of the major languages or dialects in India due to the fact that there are too many languages that could effect interpretation of the translated text. The research followed strict ethical guidelines and was approved by the Institutional Review Board (IRB). While the survey contains both quantitative and qualitative responses, we analyzed only the quantitative results in order to gain a broader view of how privacy preferences were affecting feature selection. We also occasionally report on some of the comments expressed by participants. In this section, we describe our survey instrument and in Section 5, we talk more about the potential limitations of our method.

3.1 Recruitment and Data Cleaning

Recall that we targeted nations with the most intensive use of WhatsApp (India) and the nation where WhatsApp has its largest user base (Saudi Arabia). WhatsApp is the most widely used instant messaging platform in both countries [37, 35]. The 186 non-Saudi Arabs were separated in the analysis to examine if we can observe if differences exist between Saudis and the greater Arab population. The Arab League defines its membership as 21 Arab countries of Algeria, Bahrain, Comoros, Djibouti, Egypt, Iraq, Jordan, Kuwait, Lebanon, Libya, Mauritania, Morocco, Oman, Qatar, Saudi Arabia, Somalia, Sudan, Syria, Tunisia, United Arab Emirates, and Yemen in addition to the Palestinian Territories. Of these countries, WhatsApp is banned in Syria and the

7

United Arab Emirates (which perhaps not coincidentally reported over ninety percent and eighty percent of the population as WhatsApp users, respectively, as recently as 2014[1]). As noted Saudi Arabia has the highest WhatsApp adoption of these countries.

In our study we separated Saudis based on previous analysis, and in doing so, continued to find significant differences between the Saudi and non-Saudi Arab populations. Our particular focus on Saudi users is based on the market penetration and widespread use of WhatsApp in the population.

Furthermore, both the Arab and Indian populations have been subject to fewer privacy studies than Americans and Europeans, and no comparative studies between the populations. While findings are often generalized for Western populations, we hypothesize that specific themes might come into focus when we study non-WEIRD populations[2].

This study was done in two phases over three years. The first phase was in Fall 2014 targeted at the Arab population. The participants were Arabs (548 being specifically from Saudi Arabia and 186 from other countries in the Arabian peninsula, after removing missing responses). We repeated the recruiting process for Indians in Fall 2017, resulting in 229 respondents who self-reported as Indians by nationality. Participants were recruited via email advertisements, forum postings, and snowball sampling through Facebook and WhatsApp. Both datasets have been combined for consistency across variables that can be compared for analysis. Questions that were present in one dataset but absent in the other were removed. We also removed responses which had missing values for any of the questions that remained in our dataset.

This reduced our dataset to 820 responses. The resulting data consisted of the responses from 146 non-Saudi Arabs, 222 Indians, and 452 Saudi Arabians. The demographic distribution of users is shown in Table 1. We assigned a variable to every question in our complete dataset, resulting in 24 variables, across 4 categories: general WhatsApp usage, feature settings, privacy preferences, and demographics.

1. Usage Variables: These variables measure the usage habits of users on WhatsApp. This includes information on the platform, version of the application, and length of usage. It also includes the type of data shared over WhatsApp and whom it is shared with. This serves to not only

---

[1]https://www.statista.com/statistics/509031/arab-countries-whatsapp-penetration/

[2]Consider North Dakota is closer to North Carolina than Riyadh to Baghdad (1640 miles versus 840). There may be more direct trade, and there is a common border between the two nations. One could argue that Iraqis and Saudis should have similar privacy preferences; that geographical adjacency, ethnic similarities, and a shared religious heritage would result in similar privacy preferences and behaviors.North Dakota is 89% white Christian and North Carolina is 70% white Christian compared with 75% Arab Islamic for Iraq and 85% Arab Islamic for Saudi Arabia. Thus in terms of sheer homogeneity, the comparisons are not prima facie illegitimate. Alternatively the sectarian religious differences and the radical Saudi male guardianship model of governance suggests these might be quite different, just strong differences between the Dakota's pure republican red and North Carolina's competitive purple may indicate differences. Research is the only way to know which of these is correct.

Table 1: The demographic distribution of the populations broken down into three populations.

Demographic Details		Saudis	non-Saudi Arabs	Indians
Gender	Male	159	34	140*
	Female	293	112	75*
Age (years)	18-24	99	74	80
	25-30	225	33	123
	31-40	103	27	17
	41-50	21	8	2
	over 50	4	4	0
Education	High School or Less	30	36	4
	Some College	38	58	2
	Bachelors	58	229	90
	Masters or Professional	18	99	114
	Doctoral	2	30	9
		452	146	222

*Responses from seven users who did not wish to specify their gender were excluded

understand the usage pattern but also gauge the technical familiarity of users.

2. Feature Setting Variables: Feature settings refer to the various options for the features available on WhatsApp. This refers to the use or non-use of a particular feature or visibility control settings when a feature is used. For example, settings indicate visibility to either *Everyone*, *My Contacts*, or *Nobody*.

3. Privacy Preference Variables: Users were queried about their privacy concerns specifically on WhatsApp. This removed concerns about generic assumptions about the nature of privacy. Participants were queried about their concerns and attitudes towards being contacted by strangers, receiving targeted advertisements, getting permission before being added to group conversations, interacting with professional contacts over WhatsApp, and sharing personal content which they believe contained sensitive information.

4. Demographic Variables: The demographic variables we considered obviously included region of origin. In addition participants were asked their age, gender, and level of education.

The list of variables and their description can be found in Appendix A.

As can be seen in Table 1, the demographic distribution is not representative of the individual populations under study. To address this imbalance in the participant pool we assigned post-stratification weights to our data based on gender distribution that is representative of the actual individual population. This would allow us to compare across populations to a certain extent, but must be considered with caution. We did not adjust weights for age and education because (i) age-distribution was fairly representative, and (ii) the age brackets used in our study is different from the age brackets that can be obtained from available population survey data. In order to adjust for gender ratio across populations, we have used population Census data made available through IPUMS International [3]. Table 2 shows the weights assigned after adjustment of the gender ratio. R1 is the gender ratio obtained from the IPUMS data, R2 is the ratio obtained from our dataset. The weight is calculated by dividing R1 by R2, and all the individual values among variables are adjusted according to the weight.

Table 2: Post-stratification weights of gender ratio to adjust for populations

Population		Census Ratio R1	Survey Ratio R2	Weight R1/R2
Saudis	Male	0.514	0.352	1.460
	Female	0.486	0.648	0.750
Arab (non-Saudis)	Male	0.509	0.233	2.184
	Female	0.491	0.767	0.640
Indian	Male	0.512*	0.651	0.786
	Female	0.488*	0.349	1.398

*Median taken from gender ratios of 22 Arab countries

3.2   Analysis

The analysis was done in three parts. First, we performed a correlation of preferences in feature choice against different privacy behavior across individual populations. This was done to understand how strongly privacy behavior was correlated with features choices across the three samples, in response to our first research question, RQ1. This reduced our number of variables under study to six feature choices, five privacy behaviors, and two demographic (age and gender) variables for each of Arab (non-Saudi), Saudi and Indian participants.

Second, we performed a binomial logistic regression for each of these feature choices against privacy behavior, platform usage and demographics. We further performed an ordered logistic regression for each of the profile information choices (profile photo, last seen, and status; considered under feature choices) to determine if they were a function of privacy concerns as well. We used the

---

[3]https://international.ipums.org/international/

ordered logistic regression model since the settings for each of these choices are ordered (*Everyone, My Contacts, and Nobody*, in terms of audience reach). The feature choices were:

1. **Blocked**, which allows users to block communication from individuals they do not want to interact with. This would also mean hiding all profile information.

2. **Auto-Download**, which lets users choose whether or not to download media from other users in a group.

3. **Last Seen**, where users can see the time stamp of when another user they are communicating with was last active on WhatsApp.

4. **Profile Photo**, which is a display picture for users to identify other users.

5. **Status**, WhatsApp allows users to share an update about themselves that is visible to people who have their contact number and are on WhatsApp for up to 24 hours. As a privacy choice, users are able to limit who sees their status update, much like a 'story' on Instagram.

6. **Location**, which allows users to share their present co-ordinates over WhatsApp so that others are able to locate them easily. Please note that if a user shares location in a group that consists of individuals they have blocked, the blocked individuals would still be able to see their location. Thus, in terms of location-sharing, group sharing gets preference over individual sharing.

These feature choices were evaluated against privacy concerns. The exogenous variables were if users were concerned about being contacted by strangers (*scc or stranger contact concern*), being added to a group with consent (*groupadd*), sharing of information over WhatsApp which they consider sensitive (*sensitive*), targeted advertisements through surveillance on WhatsApp (*ads*), or contacting co-workers over an informal platform like WhatsApp (*proff*).

Third, we wanted to examine if these feature choices were dependent on the cultural influences of the populations under study as an independent variable, (*origin*). We divided our participants responses into two groups, Arab and Indian users. As noted earlier, we also evaluated Saudi and non-Saudi Arabs separately to examine if privacy behavior of a larger cultural group (Arab) could be represented by participants from a specific place of origin (Saudi Arabia). We have considered Saudi Arabia in this case because it has the largest number of WhatsApp users in the world [35]. In order to study privacy choices for features (blocked, auto-download, last seen, profile photo, status, and location) against origin, we perform a Pearson's Chi-squared Test on the following null hypothesis:

Privacy choices for features in each case is NOT dependent on origin.

This is in response to our second research question, where we want to see if privacy behavior varies across the populations and if origin is a factor driving

Table 3: Usage characteristics for WhatsApp across populations (%)

		Arab (non-Saudi)	Saudi	Indian
Operating System	Android	65.0	41.4	77.9
	iOS	27.4	55.3	19.8
	Other	7.5	3.3	2.2
Length of Use	Less than a year	9.5	0.4	0
	1-2 years	28.8	14.8	1.3
	2-3 years	0	0	9.0
	3-4 years	50.6	63.7	28.4
	4-5 years	0	0	28.8
	5 years or more	10.9	21.0	32.4
Updated Version	Yes	81.5	90.0	96.8
	No	6.8	6.2	0.4
	Don't Know	11.6	3.8	2.7

the feature choices enumerated above. This gave us six hypothesis to check for each of the features. For example, *"Privacy choices for blocking is NOT dependent on origin"*. We then performed a Chi-Squared test on each of these null hypotheses as a test of significance.

Finally, in order to address our third research question, RQ3, we measure the effect of age and gender on individual feature choices. While we acknowledge that gender and age are not distributed adequately to be representative, it is a reasonable initial measure of the how the population decides on privacy features. Due to the uneven distribution of educational backgrounds across the populations, we excluded it from analysis. This is further discussed under limitations.

4   Results

4.1   Usage Statistics

Table 3 shows the mobile platform used, the length of usage and whether participants used the latest update of WhatsApp or not. Most users across the three populations have been using WhatsApp for over three years (61.5% for Arabs, 84.7% for Saudis and 89.6% for Indians), which ascertains their familiarity with the platform. There is also a consistent percentage of users over the different lengths of usage for the Indian population, which indicates a steady growth of WhatsApp as a messaging platform in the population. A vast majority of each population also uses the updated version as of 2016-17, It is unclear if they knew about it and consciously downloaded it or if it was automatically updated. Android is the dominant platform for the Indian and Arab populations at large, but Saudis are fairly divided between Android (41.4%) and iOS (55.3%) users, with iOS having the upper hand.

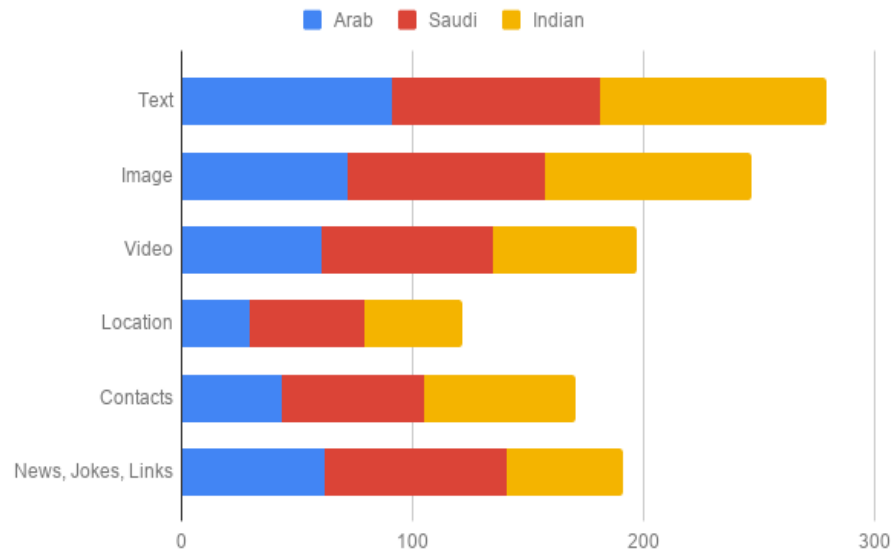Figure 1 shows the types of data (text, image, video, contacts, location,

Figure 1: Types of Media Shared Across Populations (expressed as % of populations for all three groups)

news, jokes and links) shared by users across the three populations, expressed in terms of percentages of responses. As can be seen, text is the most widely used medium of communication, followed by images and videos. WhatsApp has recently introduced video and voice call features, which were not available during most of the breadth of this study, and hence could not be included in the types of data, and remains a subject of investigation in future. The use of different media of communication is fairly consistent across the three populations, indicated by the equal length of the bars. There is a slight difference however, for location sharing, with more Indians willing to share their location compared to Arabs.

WhatsApp users broadly communicated with the following groups over WhatsApp: friends, family, people living far away (*Far Away*), co-workers, and sometimes, even people they did not know very well (*"Not Well Known"*). Over 85% of users from all three populations said that they used WhatsApp either *Nearly Always* or *Fairly Often* to communicate with a smaller social circle, comprising of friends and family. More than half of the users across the three groups also used WhatsApp to communicate with people living far away from them, not excluding near ones. These observations were consistent across the three groups.

However, a significant difference was observed in case of contacting recipients who were not well known. To put it in perspective, 29.73% Indian participants said that they used WhatsApp to contact people they did not know well, in contrast with 11.64% for Arabs and 11.73% for Saudis. Conversely, 60.27%

Table 4: Usage by message recipient (%)

		Arab (non-Saudi)	Saudi	Indian
Not Well Known	Nearly Always	4.11	4.65	12.61
	Fairly Often	7.53	7.08	17.12
	Rarely	28.08	33.18	39.64
	Never	60.27	55.09	30.63
Co-workers	Nearly Always	17.12	27.21	27.48
	Fairly Often	30.14	30.75	35.13
	Rarely	34.93	31.86	27.48
	Never	17.81	10.18	9.91
Far Away	Nearly Always	24.66	28.98	49.55
	Fairly Often	38.36	37.39	34.23
	Rarely	33.56	29.87	15.76
	Never	3.42	3.76	0.45
Family	Nearly Always	51.37	62.83	49.10
	Fairly Often	39.04	28.54	36.49
	Rarely	9.59	8.18	13.06
	Never	0	0.44	1.35
Friends	Nearly Always	54.11	65.04	75.67
	Fairly Often	30.82	28.54	18.92
	Rarely	13.01	5.97	4.95
	Never	2.05	0.44	0.45

Arabs and 55.09% Saudis said that they would NEVER contact anyone they did not know well over WhatsApp (probably because it involves exchanging phone numbers with such users). In comparison less than a third of Indian users expressed this sentiment (30.63% for Indian users). The probability of using WhatsApp to contact co-workers was almost same across all three populations, with nearly half of each population using WhatsApp in a professional setting. Surprisingly, this had little effect on Saudis, whose privacy behavior was not significantly dependent on whether they contacted users professionally over WhatsApp. In contrast, the choice of features as will be elaborated in Section 4.4 for Indian users were significantly correlated with contacting professionals over the platform.

4.2   Effect of Origin on Feature Settings

As described in the previous section, we observed a number of privacy behavior factors effecting privacy feature choices across the three populations. Since we were measuring similar parameters for the same application in different cultural settings, we presumed it would be interesting to study if cultural background itself was a predictor of privacy behavior, in response to RQ2. In order to study privacy choices for features (blocked, auto-download, last seen, profile photo, status, and location) against origin, we perform a Pearson's Chi-squared Test

Table 5: Pearson's Chi-sq test performed for each of the feature-specific hypothesis, showing the X-squared coefficients. The p-values are shown in parenthesis.

Feature Setting	Chi-squared Test
Blocked	**24.714 (0.0003858)**
Auto-Download	**101.34 ($<$2.2e-16)**
Last Seen	**70.498 (3.231e-13)**
Profile Photo	**87.513 ($<$2.2e-16)**
Status	**123.43 ($<$2.2e-16)**
Location	**25.089 (4.828e-05)**

on null hypothesis H0, elaborated in Section 3. This gives us the following six hypothesis to test:

- H0(a): Privacy choices for blocking strangers is NOT dependent on origin

- H0(b): Privacy choices for preventing auto-download is NOT dependent on origin

- H0(c): Privacy choices for hiding last seen is NOT dependent on origin

- H0(d): Privacy choices for hiding profile photo is NOT dependent on origin

- H0(e): Privacy choices for hiding status is NOT dependent on origin

- H0(f): Privacy choices for location sharing is NOT dependent on origin

We then performed a Chi-Squared Test on each of these null hypotheses as a test of significance and obtained the p-values for each of the above hypotheses as summarized in Table 5. As observed, none of the p-values obtained were of significance, which means that we cannot reject any of our hypotheses H0(a) through H0(f). Thus, origin plays a key factor in influencing privacy choices for the different features in consideration. In the following sections, we will examine individual populations to see the underlying demographic characteristics and privacy choices that drive these differences across cultures.

4.3 Privacy Behavior

Figure 2 shows the privacy concern of Arab and Indian users across three categories: asking before being added to a group, sharing of sensitive information, and concern over receiving targeted advertisements that uses data from their conversations. When asked if they would liked to be asked before being added to a group, more than three-fourths of all participants across the three population groups agreed that their consent was important. Currently, WhatsApp allows anyone to be added to a group conversation without their consent, and this control over being able to choose whether they wanted to be a part of a group
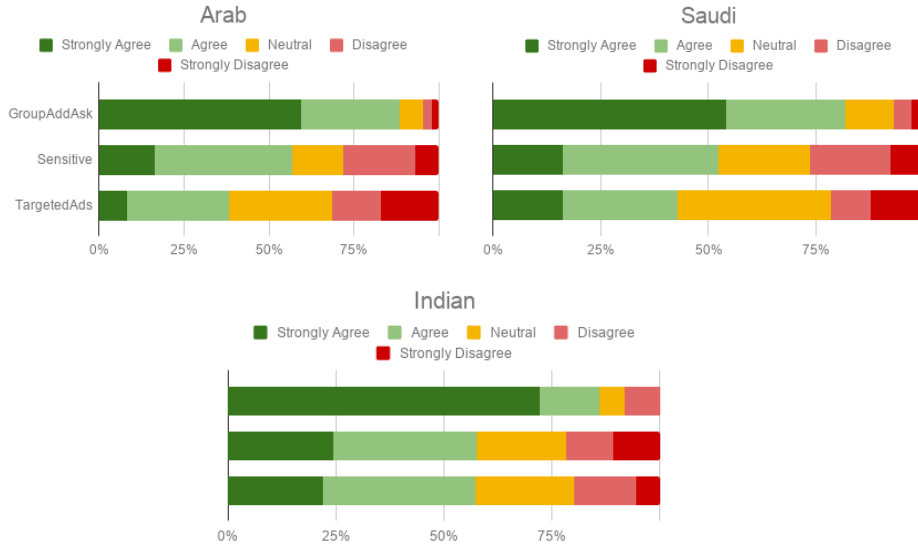
Figure 2: Privacy concerns by population

or not was expressed by most users. More than half of the participants across the populations also admitted to sharing sensitive information over WhatsApp.

When asked if participants were concerned that their messages could be read and targeted advertisements could appear based on their conversations over WhatsApp, Indians agreed that it was more likely, compared to their Arab counterparts. There is also a lack of trust in the end-to-end encryption mechanism implemented on WhatsApp, which theoretically makes it impossible for WhatsApp as a platform to read users' texts [41]. This is a surprising finding, because Arab participants believed that there were *less* chances of targeted advertisements based on their conversation even before end-to-end encryption was implemented in 2016. In contrast, over 50% of Indian participants were concerned about the appearance of targeted advertisements even after the implementation of end-to-end encryption.

Participants also expressed concern over being contacted by strangers over WhatsApp, with Saudis (54.6%) expressing the most concern and Indians having the least (33.3%). Over a quarter of Indians also said that they were not at all concerned about being contacted by strangers (contact information can be easily accessed through groups). This was not a sentiment reflected by Saudis (16.1%) or Arabs (12.3%) in general.

4.3.1 Gender was an important factor in privacy choices

Figure 3 shows privacy-preserving feature use (or non-use) based on gender for each population. While extent of individual feature choices vary across pop-
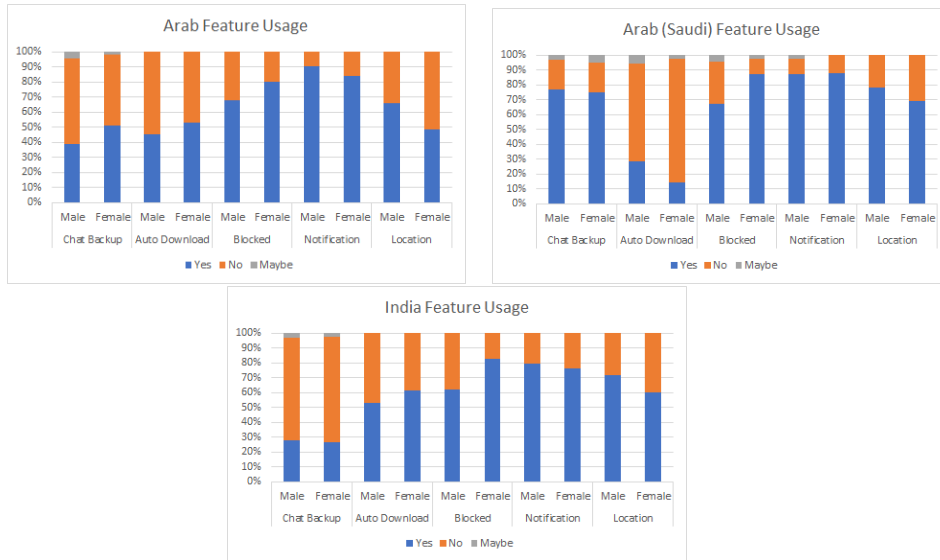
Figure 3: Feature usage by gender for three populations

ulations, There are sharp differences between female and male participants, especially for blocking and location-sharing. For example, location sharing was heavily reliant on gender across all populations. Women were less likely to share their location with other WhatsApp users. A positive correlation was found between the being female and using the blocking feature. Users who identified themselves as women were also more concerned about being contacted by strangers. This is a probability that it might be due to mobile phone sharing practices elaborated in [34]. These findings are consistent across Arab, Saudi Arabian, and Indian cultures. This is further explained according to the population in consideration.

1. **Arab Users**
   Women desired greater control over what they share and receive over WhatsApp. They used the blocked feature to prevent being contacted by strangers more than men, chose to receive fewer notifications from WhatsApp chats, and restricted access to their location more often. This was true both for Arabs at large, and for Saudis in particular. The use of auto-download (which allows users to download content automatically from other users as long as they are connected to the internet) and chat backup (to backup their chats to their local device) were the two features distinguishing the two groups of users. While the use of chat backup was more prevalent (over 70%) for Saudis, it was limited in case of Arabs (less than 50%). Similarly, content control was more used by Saudis, who seldom used auto-download. This can be seen in Figure 3.

2. **Indian Users**

17

Table 6: Regression Values for Saudi Population

Concerns/ Demographics	Feature Setting				Profile Information		
	Blocking	Auto-Download	Location	Notification	Last Seen	Profile Photo	Status
Gender	**1.03971**	0.277012	**-0.57406**	-0.21121	0.18509	-0.06686	0.02339
	(8.57e-06)	(0.183)	**(0.0079)**	(0.37620)	(0.3232)	(0.7516)	(0.9132)
SCC	**0.32476**	-0.009273	-0.15552	-0.05960	**-0.44136**	**-0.51587**	**-0.52927**
	(0.0257)	(0.946)	(0.2860)	(0.71382)	**(6.00e-04)**	**(7.22e-04)**	**(6.70e-04)**
GroupAdd	0.12775	**0.431583**	0.12182	-0.25389	-0.16205	0.04462	0.04782
	(0.2198)	**(2.43e-05)**	(0.2354)	(0.05792)	(0.07887)	(0.6635)	(0.6425)
Sensitive	0.09946	-0.115293	0.03915	0.17036	0.10324	0.04153	-0.02203
	(0.3164)	(0.200)	(0.6671)	(0.09422)	(0.2006)	(0.6462)	(0.8102)
Ads	-0.06327	0.060491	0.01253	**-0.29728**	-0.090499	-0.08074	-0.02598
	(0.4959)	(0.478)	(0.8870)	**(0.00397)**	(0.2337)	(0.3618)	(0.7706)
Proff	0.07153	0.049144	0.00178	-0.05007	-0.02447	-0.02189	-0.03139
	(0.4040)	(0.533)	(0.9826)	(0.57664)	(0.7303)	(0.7845)	(0.6985)
Age	-0.08904	0.047405	-0.18938	**-0.27760**	0.00461	-0.06623	**-0.18777**
	(0.5083)	(0.698)	(0.1339)	**(0.04631)**	(0.9678)	(0.6021)	**(0.1424)**
Platform	0.04935	**0.612170**	0.24874	-0.01367	**-0.46577**	**-0.47230**	**-0.31209**
	(0.7424)	**(5.62e-05)**	(0.0672)	(0.92983)	**(9.27e-04)**	**(6.97e-03)**	**(0.0582)**

The extent of feature usage (chat backup, auto-download, blocked, notification, and location) was fairly similar for both Indians and Arabs. Gender again was a dominant factor in the use of blocking feature, with women using the same more frequently. Like their Saudi counterparts, Indian women also chose to restrict location sharing and limited notifications from chats. However, the use of chat backup and auto-download among Indians were more similar to Arabs at large than the specific Saudi population. Chat backup was rarely used by men who were not Saudis, and auto-download was more used by women than men.

4.4 Effect of Privacy Attitude on Feature Choices

4.4.1 Saudi Arabia

Table 6 shows the regression values of feature use using independent privacy and demographic characteristics for the Saudi population. The co-efficients for which the p-value is significantly less have been written in bold (the p-values themselves are indicated within brackts) for Tables 6, 7, and 8. For the Saudi population, concern about being contacted by strangers (SCC) is a major criteria that determines how people use (or not use) privacy preserving features like hiding their last seen, profile photo, status update or blocking another user.

There is an effect of platform choice for these settings. Android users for the Saudi population tend NOT to hide their profile information like last seen, profile photo and status. They also tend to use the auto-download feature more.

WhatsApp on the Android platform has these settings on by default, which suggests that users might not have changed their privacy settings for such profile information. This indicates a need for better and more transparent privacy settings for Android users, which is nearly half of all WhatsApp users we surveyed. We did not ask about income, so there may be a hidden variable underlying the importance of platform choice. However, education did not appear to be highly significant in platform choice and education is often a good proxy for income.

Some of our qualitative responses show how personal interpretations address feature choices among Saudis. For example, one of the Saudi participants' mentioned, *"I used Whatsapp app because I believe[d] it will protect my privacy however now I feel it is becoming less priva[te]!"*. Similarly, there was a reflected feeling of loss of privacy when one of the user also said, *"..Worst thing that may have happened to Whatsapp is that they got bought over, its not about Facebook, it could have been Google/Microsoft or the likes, my concern is that now we're sure that our data is in the hands of people who can do anything with it"*.

Gender plays a key role by bearing a strong correlation with both blocking and location hiding in these groups. Women have been found to be the most active in blocking strangers who contact them via WhatsApp and also choose to not share their location. One of our female Saudi participants commented, *"I once had to delete my phone number because of a st[al]ker who kept sending me messages from different phone number[s]. Blocking his number was not effective. That was a breach of my privacy"*.

For Arabs in general, however, there are very few factors which effect privacy choices, as a marked departure from our Saudi users. Table 7 shows the regression values of feature use on independent privacy and demographic characteristics for the greater Arab population. There were no significant factors that effected the sharing of profile information. Blocking other users was unsurprisingly strongly dependent on being contacted by people they did not know (SCC) and was done more by users who shared a lot of sensitive information over WhatsApp. Location sharing on the other hand, was strongly gender correlated, with women hiding their location more often. The only privacy setting that was consistent among Saudis and the greater Arab population was the dependence of use (or non-use) of the auto-download feature on the platform, with more Android users downloading media by default compared to their iOS counterparts.

4.4.2 India

The privacy behavior among Indians were more computationally complex and feature preferences were a result of a number of different inter-dependent factors, as seen in Table 8. That is, while gender dominated Saudi choices, gender interacted with other demographics and variables. For example, the use of blocking feature was strongly dependent on a number of factors like preference of being added to a group without permission (groupadd) [*"Want [WhatsApp users to] request permission before adding to a group"*], sharing sensitive information, and contacting co-workers over WhatsApp. It seemed like there is a notion of

Table 7: Regression Values for Greater Arab Population

Concerns/ Demographics	Feature Setting				Profile Information		
	Blocking	Auto-Download	Location	Notification	Last Seen	Profile Photo	Status
Gender	0.28273 (0.5170)	-0.00951 (0.9794)	**-0.93527** **(0.0161)**	-1.36175 (0.1026)	-0.13628 (0.6718)	-0.13628 (0.6718)	0.00729 (0.9850)
SCC	**0.88818** **(0.0056)**	0.29130 (0.2782)	0.34419 (0.2163)	**-1.51886** **(0.0238)**	-0.02776 (0.9079)	-0.02776 (0.9079)	-0.21210 (0.4648)
GroupAdd	-0.02646 (0.91014)	0.21581 (0.2784)	**-0.43717** **(0.0586)**	0.13650 (0.7149)	0.00358 (0.9836)	0.00358 (0.9836)	-0.16971 (0.4452)
Sensitive	**0.47200** **(0.00930)**	0.04985 (0.7508)	-0.00251 (0.9877)	0.32079 (0.2576)	-0.07628 (0.5903)	-0.07628 (0.5903)	-0.25066 (0.1384)
Ads	-0.16895 (0.38508)	-0.24602 (0.1176)	0.11100 (0.4869)	0.09067 (0.7519)	0.02063 (0.8780)	0.02063 (0.8780)	0.15030 (0.3428)
Proff	-0.07959 (0.66768)	-0.03216 (0.8362)	-0.20761 (0.2148)	-0.19440 (0.4836)	-0.20641 (0.1243)	-0.20641 (0.1243)	0.06905 (0.6764)
Age	0.07499 (0.7143)	0.07242 (0.6772)	0.05074 (0.7800)	0.60724 (0.1699)	-0.14745 (0.3592)	-0.14745 (0.3592)	-0.06612 (0.7305)
Platform	-0.25015 (0.33988)	**0.74211** **(0.0061)**	**0.64044** **(0.0094)**	0.32429 (0.4247)	0.03891 (0.8339)	0.03891 (0.8339)	0.21114 (0.3337)

close-knit community since privacy behavior in a group was a significant factor in privacy choices among Indians. It was also perceived to be affecting interpersonal relationships, when a user mentioned, *"....Block[ing] systems should not be kept for lovers as [it makes] the prob[lem] more messy"*.

It was also related with demographic variables like gender and age. For instance, participants who wanted to be asked before being added to a group without consent, were more likely to use the blocking feature. This was mentioned by one of our female participants, who said *"WhatsApp......should add permanently images/ video delete feature to let a person safe from being harassed"*. WhatsApp has since then, introduce a delete feature for users to delete their conversations. Younger adults and women were also more likely to block other users more.

Table 8 shows the regression values of feature use on independent privacy and demographic characteristics for the Indian population.

Indians were also found to be more sensitive about sharing their profile information possibly because of extensive use of WhatsApp in a professional setting (71.2%). Profile information like last seen, profile photo, and status were found to have limited sharing preferences, among users who contacted professionals over WhatsApp. Such users also tended not to share their location with other users.

Even though gender was related to limiting content and users, it played a more significant role in profile information sharing. Women, more than men, in case of Indian users, were less likely to share their location, profile photo, or

Table 8: Regression Values for Indian Population

Concerns/ Demographics	Feature Setting				Profile Information		
	Blocking	Auto-Download	Location	Notification	Last Seen	Profile Photo	Status
Gender	**0.7900** **(0.0390)**	**-0.79770** **(0.0248)**	**-0.50682** **(0.0914)**	-0.08195 (0.8560)	**-0.95812** **(0.98e-04)**	**-1.08534** **(0.0002)**	**-0.53954** **(0.0455)**
SCC	0.2266 (0.3405)	-0.07372 (0.7407)	-0.21241 (0.3089)	-0.30686 (0.3107)	-0.07381 (0.6485)	-0.27881 (0.1461)	-0.14989 (0.4105)
GroupAdd	**0.3937** **(0.0348)**	**-0.28301** **(0.1134)**	0.22672 (0.1989)	0.02874 (0.9136)	-0.05016 (0.7287)	**-0.29727** **(0.0786)**	0.03808 (0.8169)
Sensitive	**0.3772** **(0.0088)**	-0.03378 (0.8106)	-0.01369 (0.9173)	0.02367 (0.9003)	0.08538 (0.4177)	-0.16140 (0.1829)	-0.06538 (0.5755)
Ads	0.1400 (0.3936)	0.12255 (0.4464)	0.03597 (0.8072)	0.04650 (0.8290)	0.00649 (0.9536)	-0.17083 (0.2007)	-0.01785 (0.8876)
Proff	**0.2784** **(0.0193)**	-0.13380 (0.2581)	**0.25231** **(0.0239)**	**0.38039** **(0.0136)**	**0.18789** **(0.0402)**	**0.17690** **(0.1001)**	**0.17869** **(0.0862)**
Age	**-0.6264** **(0.05013)**	-0.09446 (0.7460)	**-0.49116** **(0.0666)**	-0.34307 (0.3808)	**0.49159** **(0.0202)**	**0.67316** **(0.0087)**	0.36171 (0.1211)
Platform	0.1798 (0.62490)	-0.41813 (0.2299)	0.46511 (0.1646)	0.46084 (0.2612)	-0.13551 (0.5762)	-0.27400 (0.3658)	-0.04563 (0.8709)

last seen.

Additionally, unlike Saudi users, age played a major role in the way privacy features were selected, across most categories. In India, the users who are older tend to contact coworkers through WhatsApp more than Arab and Saudi Arabian users. Younger adults were less likely to share profile information like last seen and profile photo. Younger adults were also more likely to use the blocking feature. On the other hand, older adults were more willing to share their location.

Taken together, gender, age and professional status were more likely to effect privacy choices for Indians. Thus, young women with professional contacts on WhatsApp were more likely to limit sharing their profile information, compared to older adults, men or users using WhatsApp for informal purposes.

4.5 Summary of Results

4.5.1 Privacy Concerns are not Culturally Independent

Different populations had different privacy behaviors that resulted in specific privacy choices. We performed the Pearsonś Chi-Squared test on privacy choices for blocked ($p = 0.0003858$), auto-download (p ¡ 2.2e-16), last seen ($p = 3.231$e-13), profile photo (p ¡ 2.2e-16), status (p ¡ 2.2e-16) and location ($p = 4.828$e-05) based on origin, which was an evidence against our null hypothesis H0. This confirms the importance of cultural context in privacy choices.

Arabs had a stronger relationship between gender with blocking of other users and location sharing, compared to their Indian counterparts. Saudi Arabian women reported more usage of the blocking feature to prevent unknown individuals from contacting them (p=8.57e-06) compared to Indian women (p=0.0390). Similarly, Arab women (p(Saudi)=0.0079, p(non-Saudi)=0.0161) limited sharing of their location data with other WhatsApp users compared to their Indian counterparts (p=0.0914).

Gender effected the way in which profile information was shared among Indians. Features like limiting the audience of their last online log (last seen, p=0.98e-04) and profile photo (p=0.0002) from Everyone to Contacts Only or Nobody was done more by Indian women than men. Conversely, for the Arabs, gender was not a significant factor in how they chose to share their profile information.

In fact, stranger contact concern was found to have a more significant effect on how Saudis shared their last seen (p=6.00e-04) or profile photo (p=7.22e-04), which was not a significant factor among Indians. The stress of blocking and stranger contact and the corresponding need to explore possible design solutions are implied by our results.

4.5.2 Privacy Behavior can be Different in Sub-Cultures

Cross cultural studies, like this one, have focused on large populations. We also initially focused on nations in this study, specifically Saudi Arabia and India. However, we had a large number of participants who identified as Arab but not Saudis take the survey (as these were separate questions, as recommended by the Saudi researcher). Yet Privacy behavior that effected certain privacy choices among Arabs were not necessarily same as the behavior that effected privacy choices among Saudis. For example, concern over being contacted by strangers had a strong relationship with how Saudis chose to hide their profile information like last seen (p=6.00e-04), profile photo (p=7.22e-04) and status (p=6.70e-04).

However, this was not the case among Arabs at large, who did not have any significant specific privacy behavior effecting their profile information sharing habits. The platform on which Saudis installed WhatsApp also had a strong relationship with the sharing of profile information. Thus, WhatsApp Android users did not hide their profile information as much as WhatsApp users on iOS platforms did.

4.5.3 Some Privacy Behavior effected Choices Similarly Across Cultures

Age, to a certain extent, effected the way users chose to share profile information across all populations in our study. While younger Saudi users chose to share their status updates (p=6.70e-04) more widely, younger Indian users controlled visibility of their profile photo (p=0.0087).

Disabling automatic downloads was correlated with a preference not to be added to a group without prior permission. This was highly significant for

Saudi participants (p=2.43e-05) and marginally significant for Indian participants (p=0.1134).

Participants who preferred to be asked before being added to a group chose not to automatically download images, videos and documents. However, it is not clear if this is a preference due to memory and data usage limitations or an unwillingness to receive a lot of information from groups they did not wish to be included in. Additionally, both Arabs and Indians who self-identified as users who share sensitive information over WhatsApp tended to use the blocking feature more to limit interaction with people they did not know.

In both populations, concern over receiving behavioral, targeted advertisements was not perceived as a significant privacy threat. We found no evidence that concerns about surveillance had an impact on privacy behavior.

5   Discussion and Implications

WhatsApp is a widely adopted mobile instant messaging application across Arab and Indian subcontinents. It is designed to accept any Unicode character, making the use of regional languages and dialects simple. The central focus on voice and video using a mobile device enables wide participation. Additionally, WhatsApp enables individual control of privacy by enabling opt-outs for various features. The choice of opt-out means the privacy settings, by default, offer unrestricted access unless specified otherwise. We inquired about the behaviors and choices on WhatsApp features in populations from regions of high WhatsApp adoption. In our study, we look at both how people control information about themselves (i.e., profile photo, last seen, status updates, and location) as well as information they receive from other users (i.e, Blocking and Auto-download features).

We found that a simple comparison of feature usage analysis does not illustrate the spectrum of privacy preferences of WhatsApp users. Delving more deeply, we found that the demographics that impinge privacy choices are quite different. Cultural influences on privacy behavior and feature selection, as our work highlights, is inherently complex. As in case of Arab and Indian population, privacy behavior impinges feature choices in very different ways for different populations. This was particularly true for stranger contact concern and use of WhatsApp in professional settings. From qualitative inputs, Indians who used WhatsApp professionally (i.e. Work Groups) were more privacy sensitive than Arabs who limited their WhatsApp usage for personal interactions. Similarly, choosing to use the blocking feature and not sharing location was more common among Saudis.

Gender and age also played significant roles in how users interacted with others over WhatsApp. Due to the variability in privacy preferences among cultures, it is difficult to make an overarching default recommendation for all three populations. Instead, it encouraged us to rethink privacy controls that WhatsApp provides. While the following design implications are rudimentary and yet to be tested, these are provided as possibilities to enable WhatsApp to be

more privacy-sensitive about its users, across different cultural backgrounds and to cater to complex privacy expectations. Our recommendations are actionable with privacy sensitive modifications and additions to the existing interface for better risk communication and improved trust.

5.0.1   Privacy by Default

Privacy concerns have always been generated negative experience and mistrust among social media users due to lack of data transparency [25, 26]. Thus, we recommend to produce more privacy-preserving interfaces that set higher barriers for information disclosure by opt-in sharing for the populations in consideration. For example, location sharing should be set as opt-in at the outset, instead of users going back and restricting their location settings on their phone for WhatsApp. WhatsApp has already implemented the feature of a two-way communication about 'Read Receipts' where a user can opt-out of the feature but cannot see that their message has been read if they opt out. Such privacy preserving options can be provided to the users by default.

5.0.2   Offer an Option for Permissions-Based Contact

Stranger contact can be a risky endeavor for any person and anyone can contact a person in WhatsApp. Participants across all populations for example, indicated that they did not want to be added to a group without their consent. Our research argues strongly for the option of permission-based contacts. This would mean that only those people who are in the person's contact list could connect with her. Alternatively, contact information could be masked so that users are only able to contact each other via WhatsApp (explained further in Section 5.0.3 below). This recommendation would be effective not only for the stranger-averse but also for individuals that are well-known, are being bullied, or other situations such as sudden Internet infamy.

5.0.3   Protect Contact Information

WhatsApp users have their contact information shared automatically in a group. This is made more problematic by the fact that anyone can be added to a group without their permission. Permissions-based contact (discussed above) and permissions to be added to a group (noted below) resolve some of these problems. However, the widespread adoption of WhatsApp not only in social situations but employment and educational settings means that people must join groups to perform the activities of daily life.

Contact information of users is easily accessible over WhatsApp to other users if they are in the same group, even if they do not know each other. It is not possible to block this sharing other than leaving the group. Sharing of phone numbers in this case can be perilous, because allows strangers to contact these users *outside* WhatsApp (over direct messaging or phone call). Strangers can also access an individual's WhatsApp information from different phone numbers

even if they are blocked on one number. In comments on the survey, one of the participants mentioned that WhatsApp should allow "Adding people by some username instead of [a] phone number". Currently there is no setting that allows a person not to share phone numbers in a WhatsApp group, and this should change. Choosing to share a phone number in a group should an explicit choice per group, and is a core component of privacy by design in WhatsApp.

5.0.4 Choice and Consent in Joining Groups

As mentioned earlier, over 80% users across the three population groups wanted to be asked before being added to a group. Better group controls, especially being asked before being added to a group, is a widely desired function. Secondly, people would like to mute their departure from a group. For example, one of the participants said, *"[WhatsApp should] [a]sk permission before adding to groups and [allow users to] leave groups silently without alerting everyone"*. While we suspect that this is a more social construct, such preferences should exist and our participants requested these. As mentioned earlier, WhatsApp is frequently used in India for contacting colleagues in workplaces and other professional contacts. In our study, this was found to effect the way Indians shared information about themselves or interacted with other WhatsApp users. Users who mentioned that communicate with professional contacts or co-workers over WhatsApp were found to use the blocked feature more, hide their last seen, profile photo, status, and limit location sharing. This was not a significant factor among Arabs.

Enabling easy *segregation of users into high level groups such as work and family* would make it simple to share more with family without exposing information to the workplace. At this writing status can be shared selectively by adding each contact separately, one by one. The ability to form large-scale groups could empower people where WhatsApp is part of the culture.

5.0.5 Reporting and Age Appropriateness

In case of Indian users, and Arab users to a lessor extent, age played an important role in how profile information was shared. Young people shared less information compared to older adults. In case of WhatsApp however, there is no method to assert your own age on the application; nor a way for guardian consent during enrollment. As long as the phone number is validated, anyone can communicate over WhatsApp. One way to address this is to have stronger defaults if a user asserts a young age, addressing in spirit the legislation implemented in law in the US with the Children's Online Privacy Protection Act (15 U.S.C. §§6501–6506). WhatsApp does have an age criteria in its policy, but enforcing that in the regions we considered is explicitly *not* a recommendation. Denying access to people under 18 is not a recommendation that we would consider ethical, as this could result in denying social, educational, and employment opportunities to young people. Conversely, young users may be particularly vulnerable to stranger contact. Further, in every culture there is

information considered to be inappropriate content for children and younger adults. Thus, one of our core recommendations is querying new users about their age and increasing the stringency of defaults based on that response.

5.0.6 Allow cooperative blocking

Participants, especially from Saudi Arabia, have indicated being stalked (*"I once had to delete my phone number because of a st[al]ker who kept sending me messages from different phone number[s]. Blocking his number was not effective."* outside WhatsApp. While it remains to be investigated if there have been repeated offenders, WhatsApp should have a way for users to report another user, to prevent such behavior. Due to the nature of WhatsApp as a messaging platform, there is not a centralized control where one user is able to report other usersábuse of the platform. Blocking is limited to individuals. As a result when one person abuses the platform many people have to block that individual. When one group of people serially target people, as with GamerGate, 4chan, or other Internet mobs, collective blocking can reduce the effort on the target. Allowing communities to self-organize and block individuals collectively would enhance the usability of the platform and the autonomy of users, as has been done by Twitter which enables importing of block lists [4]. On WhatsApp the current implementation of group membership exacerbate these problems.

5.0.7 Information pedigree indicators

Users in our population group have already demonstrated a lack of trust (receiving targeted advertisements because their WhatsApp messages can be read) and caution over received content (lack of use of auto-download). With the inability to distinguish between trustworthy and untrustworthy information, some people expressed a lack of trust in WhatApp as a whole. Indicators are needed to distinguish familiar from unfamiliar, as is being implemented in forwarding information. Specifically, WhatsApp recently introduced the a label for text messages indicating if they have been forwarded from other users or groups. This was done in response to various acts of violence, some deadly, that were driven by information disseminated on WhatsApp where recipients assumed the forwarding (trusted) parties were the initiating (unknown, untrusted) parties. While this is an admirable first step, it might not be enough to inform users. More information on *why* forwarding labels are used could help more people understand the implications of trusting messages, especially ones which are politically or financially situated. Warning were added without cross-cultural analysis, and we could find no documentation or report of user studies.

In warning science it is the case that the users with the lowest expertise are targeted by design, because aiming for average means missing half the users. This principle could be used in developing indicators for integration into WhatsApp, and could further be subject to research for efficacy in different cultural contexts.

---

[4]https://help.twitter.com/en/using-twitter/advanced-twitter-block-options

5.0.8 Privacy feedback

In the same vein, and also potentially informed by warning science, privacy nudges could inform users of risk. Since more than half of the population in each group we surveyed said that they share sensitive information to some degree, reminders about sharing preferences before posting video content or joining groups could assist users.

5.0.9 Enable settings for contact groups instead of individual lists

Participants in our study indicated significantly different usage behavior for different audience. For example, WhatsApp was more frequently used in personal settings (within family or friends) rather than to contact people they did not know very well. Furthermore, feature selection and profile information sharing as shown in Section 4.4.2 was strongly correlated with whether or not users communicated with their professional contacts over WhatsApp, especially for Indian users. This is possibly because users have different self-presentation for their personal and professional lives [28]. Thus, contact groups (friends, family, acquaintances or co-workers) is a suggested design modification that is also culturally situated. Currently, WhatsApp allows users to individually add black list or white list other users who receive the information that they share. However, audience boundary in the form of contact groups would ensure that shared information is contained within certain groups. For example, users might choose to share an image with their friends, but hide it from their co-workers.

5.0.10 More transparent policies

WhatsApp has a number of well-defined privacy policies [43] that are all merged together under Terms of Service, and not presented to the clearly to the user during account creation. It is well known that the cognitive cost of reading long terms of service and privacy policies is high [24, 39]. Furthermore, due to the differences in privacy perceptions across populations, these policies should also be situated in a local context similar as suggested by Bellman [3]. For example, Indian users might have distinct clauses that outline the use of WhatsApp in political context. A a quick highlight of important terms of service could be productive in making users more informed.

6 Limitations and Future Work

WhatsApp is widely adopted across Arab and Indian cultures. We were able to evaluate differences in the effect of privacy attitudes and behaviors, but the source of the privacy concerns were not always clear. For example, users who block someone may do so because they want to avoid a marginally known acquaintance, or based on a concern about a friend of a friend with excessive use of groups, or to avoid contact by strangers. The privacy behavior in all of these

cases was collectively assessed as users having concerns over being contacted by strangers.

Additionally, in order to reduce further variation between populations, we have not included technical expertise questions [29], or standard measures of general privacy sensitivity [22, 20]. We specifically choose not to include general privacy sensitivity questions because these instruments were developed with WEIRD populations. A better cross-cultural privacy instrument is a significant challenge, one that will hopefully be met in the future.

We recognize there might be a social desirability bias since we only had an 80-question long survey instrument, which could have been supplemented with additional interviews. Similarly, the survey is also slightly biased towards English-speaking Indians since we have used snowball sampling, and hence our results, even though balanced across the populations, are not necessarily representative. However, our findings and recommendations are meant to be an initial attempt at studying how these understudied populations compare with respect to their privacy behavior.

We considered our populations distinct from non-western countries not only because privacy is understudied among these populations, but also because of the fact that they have significant representation in WhatsApp usage. Not only do privacy preferences vary between WEIRD populations and other populations, privacy preferences vary within large cultural groups. The context and social expectation of use varies across groups and cultures. The concentration of technology design in small areas and the homogeneity of users can be addressed through testing with different populations to understand their privacy risks and concerns. While a diverse set of users would be ideal, there is a need for more immediate results. A comparative study across western and non-western populations could highlight the contrasts in privacy behavior and the underlying decisions. Furthermore, a study looking at educational and income demographics could reveal additional information about privacy practices.

7 Conclusions

Our findings indicate individual perceptions, region of origin, and demographics impinge feature usage to a high degree. We found that these were not independent predictors, but rather interacted differently with privacy behaviors on WhatsApp. In other words, privacy choices were dependent not only on cultural background, but were a result of complex, inter-related behavior situated in WhatsApp usage.

Particularly, in response to our first research question, RQ1, personal experiences like being contacted by strangers and perceived lack of trust were strongly correlated with the use or non-use of certain features. Note that privacy behavior was also socially and culturally situated, with Indian participants most likely to hide their profile information and location from professional contacts rather than friends or family. Most participants in both populations wanted to be able to control the content and recipient of their shared information.

As described in Section 4.2 feature choices were dependent on privacy behavior, which in turn was dependent on place of origin of participants. Feature choices for different populations were based on different privacy behaviors, some of which were common across the three groups, but most were very culturally situated. This responds to our second question, RQ2, of whether privacy behavior varies across the populations.

Gender and age influenced the use of features and sharing behavior, both within and across the three population groups. Gender played a key role in blocking and location sharing preferences across populations, and specifically influenced how users shared their profile information among Indian participants. Age was also a key factor for feature choices for Indian users. Demographics effected each population differently in response to RQ3, but it was common that they were an important factors in how users interacted with the application. We identified privacy concerns in all the groups of participants, albeit socially motivated.

A core observation, and one which calls for more research rather than conclusions, is that WhatsApp is experienced as a social network application not as a messaging app. The embedded use of groups, the multimedia interactions, and the intensity of use are very different in the high-adoption regions we examined than what we see described in the US. The difference in actual use (i.e., social network) from perception of use (i.e., encrypted SMS) could underlie the recommendations.

In any use case there is an argument for more nuanced reexamination of how privacy controls are implemented. Of course our results must be treated with the same caution as with other kinds of generalizations from limited groups of participants. If nothing else, the results serve to inform the importance of cultural inclusiveness on the design of privacy research, and design choices for privacy–impinging technologies that reach across the globe.

8   Acknowledgments

References

[1] Alessandro Acquisti and Ralph Gross. "Imagined communities: Awareness, information sharing, and privacy on the Facebook". In: *International workshop on privacy enhancing technologies*. Springer. 2006, pp. 36–58.

[2] Ulrich Beck and Cordula Kropp. "Environmental risks and public perceptions". In: *Handbook on Environment and Society*. Ed. by Jules Pretty and Julia S. Guivant. Sage, 2007. Chap. 41, pp. 1205–1216.

[3] Steven Bellman et al. "International differences in information privacy concerns: A global survey of consumers". In: *The Information Society* 20.5 (2004), pp. 313–324.

[4] Susan Bennett et al. "Gdpr: Change to European privacy laws and its impact on Australian businesses". In: *Governance Directions* 70.2 (2018), p. 85.

[5] Soutik Biswas. *On the frontline of India's WhatsApp fake news war*. 2018. URL: https://www.bbc.com/news/world-asia-india-45140158.

[6] Karen Church and Rodrigo De Oliveira. "What's up with whatsapp?: comparing mobile instant messaging behaviors with traditional SMS". In: *Proceedings of the 15th international conference on Human-computer interaction with mobile devices and services*. ACM. 2013, pp. 352–361.

[7] Sunny Consolvo et al. "Location disclosure to social relations: why, when, & what people want to share". In: *Proceedings of the SIGCHI conference on Human factors in computing systems*. ACM. 2005, pp. 81–90.

[8] Marc Cornock. *General Data Protection Regulation (GDPR) and implications for research*. 2018.

[9] Dan Cvrcek et al. "A study on the value of location privacy". In: *Proceedings of the 5th ACM Workshop on Privacy in Electronic Society*. ACM. 2006, pp. 109–118.

[10] Sanchari Das, Javon Goard, and Dakota Murray. "How Celebrities Feed Tweeples with Personal and Promotional Tweets: Celebrity Twitter Use and Audience Engagement". In: *Proceedings of the 8th International Conference on Social Media & Society*. ACM. 2017, p. 30.

[11] Jayati Dev, Sanchari Das, and L Jean Camp. "Privacy Practices, Preferences, and Compunctions: WhatsApp Users in India". In: *Proceedings of the Twelfth International Symposium on Human Aspects of Information Security & Assurance (HAISA 2018)*. Lulu. com, p. 135.

[12] Jayati Dev, Sanchari Das, and L Jean Camp. "Understanding privacy concerns of whatsapp users in India: poster". In: *Proceedings of the 5th Annual Symposium and Bootcamp on Hot Topics in the Science of Security*. ACM. 2018, p. 28.

[13] Catherine Dwyer, Starr Hiltz, and Katia Passerini. "Trust and privacy concern within social networking sites: A comparison of Facebook and MySpace". In: *AMCIS 2007 proceedings* (2007), p. 339.

[14] Vaibhav Garg and L Camp. "Ex ante vs. ex post: Economically efficient sanctioning regimes for online risks". In: (2013).

[15] Joseph Henrich, Steven J Heine, and Ara Norenzayan. "Most people are not WEIRD". In: *Nature* 466.7302 (2010), p. 29.

[16] *How WhatsApp helped turn an Indian village into a lynch mob - BBC News.* `https://www.bbc.com/news/world-asia-india-44856910`. (Accessed on 02/15/2019).

[17] *Indian political parties abuse WhatsApp service ahead of election: Reuters.* `https://www.reuters.com/article/us-india-whatsapp/indian-political-parties-abuse-whatsapp-service-ahead-of-election-executive-idUSKCN1PV1E3`. (Accessed on 02/15/2019).

[18] Jennifer King, Airi Lampinen, and Alex Smolen. "Privacy: Is there an app for that?" In: *Proceedings of the Seventh Symposium on Usable Privacy and Security.* ACM. 2011, p. 12.

[19] Alfred Kobsa, Sameer Patil, and Bertolt Meyer. "Privacy in instant messaging: An impression management model". In: *Behaviour & Information Technology* 31.4 (2012), pp. 355–370.

[20] Ponnurangam Kumaraguru and Lorrie Faith Cranor. "Privacy indexes: a survey of Westin's studies". In: (2005).

[21] Kevin Lewis, Jason Kaufman, and Nicholas Christakis. "The taste for privacy: An analysis of college student privacy settings in an online social network". In: *Journal of Computer-Mediated Communication* 14.1 (2008), pp. 79–100.

[22] Naresh K Malhotra, Sung S Kim, and James Agarwal. "Internet users' information privacy concerns (IUIPC): The construct, the scale, and a causal model". In: *Information systems research* 15.4 (2004), pp. 336–355.

[23] Sara Malm. *Saudi Arabian man gets 40 lashes for insulting ex-wife on WhatsApp.* 2018. URL: `https://www.dailymail.co.uk/news/article-6388353/Saudi-Arabian-man-gets-40-lashes-insulting-ex-wife-WhatsApp.html`.

[24] Aleecia M McDonald and Lorrie Faith Cranor. "The cost of reading privacy policies". In: *ISJLP* 4 (2008), p. 543.

[25] Abu Saleh Md Noman, Sanchari Das, and Sameer Patil. "Techies Against Facebook: Understanding Negative Sentiment Toward Facebook via User Generated Content". In: *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems.* ACM. 2019.

[26] Greg Norcie, and L Jean Camp. "The Price Of Privacy: An Examination of the Economic Costs of Abstention from Social Networks". In: *APC 2015: Amsterdam Privacy Conference.* Oct. 2015.

[27] Sameer Patil and Alfred Kobsa. "Instant messaging and privacy". In: *Proceedings of HCI.* Vol. 4. 2004, pp. 85–88.

[28] Sameer Patil et al. "Comparing privacy attitudes of knowledge workers in the US and India". In: *Proceedings of the 3rd international conference on Intercultural collaboration.* ACM. 2010, pp. 141–150.

[29] Prashanth Rajivan et al. "Factors in an end user security expertise instrument". In: *Information & Computer Security* 25.2 (2017), pp. 190–205.

[30] Yasmeen Rashidi, Kami Vaniea, and L Jean Camp. "Understanding Saudis' privacy concerns when using WhatsApp". In: *Proceedings of the Workshop on Usable Security (USEC'16)*. 2016.

[31] Reuters. *Saudi is lifting Skype, WhatsApp ban, but will censor calls.* 2017. URL: `https://in.reuters.com/article/saudi-telecoms-ban/saudi-is-lifting-skype-whatsapp-ban-but-will-censor-calls-idINL5N1M214R`.

[32] Family Educational Rights. *Privacy Act of 1974, Pub.* 1974.

[33] Towards Perceived Risk. "Cross-cultural differences in risk perception". In: *Management Science* 44.9 (1998), p. 1205.

[34] Nithya Sambasivan et al. ""Privacy is not for me, it's for those rich women": Performative Privacy Practices on Mobile Phones by Women in South Asia". In: *Fourteenth Symposium on Usable Privacy and Security ({SOUPS} 2018)*. 2018, pp. 127–142.

[35] *Saudi Arabia Social Media Statistics 2018 - Official GMI Blog.* `https://www.globalmediainsight.com/blog/saudi-arabia-social-media-statistics/`. (Accessed on 02/15/2019).

[36] Paul van Schaik et al. "Security and privacy in online social networking: Risk perceptions and precautionary behaviour". In: *Computers in Human Behavior* 78 (2018), pp. 283–297.

[37] Manish Singh. *WhatsApp hits 200 million active users in India.* 2017. URL: `http://mashable.com/2017/02/24/whatsapp-india-200-million-active-users/#Dka5Ao6c5sqW`.

[38] Statistica. *Share of population in selected countries who are active WhatsApp users as of 3rd quarter 2017.* 2018. URL: `https://www.statista.com/statistics/291540/mobile-internet-user-whatsapp/` (visited on 04/15/2018).

[39] Tony Vila, Rachel Greenstadt, and David Molnar. "Why we can't be bothered to read privacy policies models of privacy economics as a lemons market". In: *Proceedings of the 5th international conference on Electronic commerce*. ACM. 2003, pp. 403–407.

[40] Yang Wang, Gregory Norice, and Lorrie Faith Cranor. "Who is concerned about what? A study of American, Chinese and Indian users' privacy concerns on social network sites". In: *International Conference on Trust and Trustworthy Computing*. Springer. 2011, pp. 146–153.

[41] *WhatsApp FAQ - End-to-end encryption.* `https://faq.whatsapp.com/en/android/28030015/`. (Accessed on 02/15/2019).

[42] *WhatsApp India: Five lynched after online child kidnap rumors - CNN.*
      `https://www.cnn.com/2018/07/02/asia/india-lynching-whatsapp-intl/index.html`. (Accessed on 02/15/2019).

[43] *WhatsApp Legal Info.* `https://www.whatsapp.com/legal/#terms-of-service`. (Accessed on 02/14/2019).

[44] *WhatsApp to cap message forwarding to 5 chats globally - Times of India.*
      `https://timesofindia.indiatimes.com/business/india-business/whatsapp-to-cap-message-forwarding-to-5-chats-globally/articleshow/67623233.cms?utm_source=contentofinterest&utm_medium=text&utm_campaign=cppst`. (Accessed on 01/25/2019).

[45] *WhatsApp Tries To Curb Misinformation, And Annoying Aunts And Uncles : NPR.* `https://www.npr.org/2019/01/22/687439015/whatsapp-tries-to-curb-spread-of-misinformation-by-limiting-message-forwarding`. (Accessed on 02/15/2019).

A   Appendix: Variable Description in Dataset

A.1   Demographic Variables

1. Origin - Classifies participants into Arab (non-Saudi), Arab (Saudi) and Indian users
2. Gender - Gender (Male, Female). Additional 'Do not wish to specify option' for 2017-18 study
3. Age - All participants were required to be above 18 years, so the age categories were 18-24, 25-30, 31-40, 41-50 and above 50
4. Education - Educational qualification of participants

A.2   Usage Variables

1. Platform - Mobile operating system used (Android/ iOS / Other)
2. Length - Amount of time (in years) of WhatsApp usage
3. Updated - Whether user uses the current and updated version of WhatsApp
4. Frequency - How frequently users use WhatsApp in their life
5. Speak. Friends, Speak. Family, Speak. FarAway, Speak. CoWorkers, Speak. NotWellKnown - Frequency of contacting specific user groups
6. Use. SendText, Use. Image, Use. Video, Use. Location, Use. Contact, Use. NJL - The different use cases for WhatsApp - sending texts, images, videos, sharing location, contact information and content (news, jokes, links)

A.3   Feature Setting Variables

Features available on WhatsApp (Chat Backup, Auto Download media content, Block users, setup Notifications, share Location), including,
1. Chat Backup (backup)
2. Auto-Download (ad)

3. Blocked

4. Notification (notif)

5. Location

6. Last Seen (lseen), Profile Photo (pp) , Status - Privacy setting for each of Last Seen, Profile Photo visibility and Status (Everyone, My Contacts, Nobody, Dont Know)

A.4   Privacy Preference Variables

1. Stranger Contact Concern (SCC) - Frequency of being contacted by people not in user's contact list

2. Targeted Ads (ads) - Concern about targeted ads using WhatsApp data

3. GroupAddAsk- Ask users before adding them to a group

4. Professional Contact (proff) - Frequency of interacting with professional contacts via WhatsApp

5. Sensitive Data (sensitive) - Likelihood of sharing sensitive data over WhatsApp