Research Article

Cognitive reflection is associated with greater truth discernment for COVID-19 headlines, less trust but greater use of formal information sources, and greater willingness to pay for masks among social media users in Pakistan

We evaluated the relationship between individual differences in cognitive reflection and the ability to discern between true and false COVID-19 information, trust in information sources for receiving COVID-19 information and willingness to pay (WTP) for masks, using a sample of 621 low- and middle-income users in Pakistan. To capture WTP, participants play an incentive-compatible game involving real stakes. We find that higher cognitive reflection test (CRT) scores are associated with greater truth discernment for COVID-19 headlines, less trust but greater use of formal information sources, and greater WTP for KN95 masks. Thus, interventions to improve discernment of COVID-19 information should target users who rely on intuition and less on cognitive reflection.

Authors: Ayesha Ali (1), Ihsan Ayyub Qazi (2)
Affiliations: (1) Department of Economics, Lahore University of Management Sciences, Pakistan, (2) Department of Computer Science, Lahore University of Management Sciences, Pakistan

Research questions

- Is cognitive reflection associated with the ability to discern between true and false COVID-19 information among social media users in Pakistan?
- Is cognitive reflection related to trust in formal information sources and social media platforms for receiving credible COVID-19 information?
- Does cognitive reflection—as measured by cognitive reflection test (CRT) scores—correlate with real behavior such as actual willingness to pay for KN95 masks?

Essay summary

- In a phone survey of 621 social media users conducted in September 2020 in urban Pakistan,

---

participants were presented with a series of true and false headlines related to the COVID-19 pandemic. Respondents were asked to rate the accuracy of each headline. They also completed a two-question CRT for measuring cognitive reflection and provided information about basic demographics, social media use, and COVID-19-related behaviors. Finally, they participated in an incentive-compatible game designed to elicit actual willingness to pay (WTP) involving real stakes.

- We found that higher CRT scores are associated with greater truth discernment (i.e., the ability to distinguish between true and false headlines), greater use but less trust in formal information sources for receiving COVID-19 information (e.g., World Health Organization), and no discernible differences in trust in social media platforms (Facebook, Twitter, YouTube, and WhatsApp).
- We found that those with higher CRT scores are more likely to bid a higher price for KN95 masks in a real-stakes game, indicating a higher demand for masks among such users.
- Our results emphasize the relationship between cognitive reflection and susceptibility to COVID-19 misinformation in the context of a developing country and have implications for technology platforms, governments, and citizens. In particular, our work illustrates that low CRT users are more susceptible to misinformation and thus are important to target for interventions to improve discernment. Designing interventions that promote greater cognitive reflection may be useful for countering COVID-19 misinformation.

Implications

The uncertainty surrounding the COVID-19 pandemic has been fueled by the COVID-19 "infodemic." This has been particularly challenging for developing countries where lower literacy levels and lack of prior exposure to technology may make people more susceptible to COVID-19 misinformation (Guess et al., 2020). In this work, we ask whether *cognitive reflection* (Frederick, 2005; Evans & Stanovich, 2013)—the ability to use reflection to override intuitions—is associated with belief in COVID-19 misinformation, trust in different sources for receiving credible COVID-19 information, and COVID-19 preventative behaviors. More specifically, considering social media users in Pakistan, we *tested* whether higher cognitive reflection—as measured by CRT scores—predicts greater truth discernment, greater use of and trust in formal information sources, and higher WTP for masks.[2]

Examining these relationships in Pakistan, a developing country with the fifth highest population in the world and a fast-growing social media user base, contributes to our global understanding of the role cognitive reflection plays in explaining susceptibility to COVID-19 misinformation, since existing research has largely focused on the United States and Europe. We focus on cognitive reflection and the cognitive reflection test (CRT) (Frederick, 2005) as the latter has gained considerable popularity due to its power to predict a diverse range of beliefs, attitudes, and behavior. For example, higher CRT scores have been associated with fewer paranormal beliefs (Pennycook et al., 2012), lower susceptibility to bias in deductive reasoning (Toplak et al., 2011), greater rejection of conspiracy claims (Swami et al., 2014), lower perceived accuracy of fake news (Pennycook & Rand, 2019a), lower likelihood of sharing misinformation on social media (e.g., Pennycook et al., 2020), and greater sharing of reliable news sources on Twitter (Mosleh et al., 2021). However, there is a dearth of empirical evidence about CRT's power to predict belief in COVID-19 misinformation and preventative COVID-19 behaviors in the context of developing countries that generally have lower literacy rates and education levels. One notable exception is a study by Rosenzweig et al. (2021), which found CRT scores to be positively associated with discernment among a sample of Facebook users in Nigeria.

---

[2] Formal sources may be defined as those which are constituted in some regularized or legal manner in relation to the user (e.g., TV channels, newspapers, government sources, and international sources such as WHO) (Kaye, 1995).

In this paper, we aim to shed light on the relationship between cognitive reflection and vulnerability to COVID-19 misinformation using a sample of 621 social media users in Pakistan drawn randomly from low- and middle-income areas in the city of Lahore. We conducted a telephone survey due to COVID-19 restrictions (as opposed to a field survey) and to ensure representation of low digital literacy populations, who may not be well represented on online survey platforms (Hargittai & Shaw, 2020). We examine the association of cognitive reflection with truth discernment, trust in information sources, and WTP for KN95 masks via a real-stakes game.

To measure cognitive reflection, we used the verbal CRT proposed by Sirota et al. (2020), as the traditional CRT questions are known to confound reflection ability—or the disposition to think analytically—with mathematical ability (e.g., Campitelli & Gerrans, 2014; Liberali et al., 2012; Mastrogiorgio, 2015; Pennycook & Ross, 2016). As a result, they are less useful among populations with low numeracy skills or low literacy levels. In our study, we considered a sample of low- and middle-income social media users in Pakistan with a median education level, which lies between primary school and high school level education, and thus numeracy skills were likely to be low. More specifically, we used a 2-item CRT in our study adapted from the validated set of verbal CRT questions by Sirota et al. (2020).

To understand the relationship between cognitive reflection, beliefs about misinformation, and actual behavior, as well as to help address concerns that survey responses may not reflect real-world behavior, respondents participated in an incentive-compatible game—a variant of the Becker–DeGroot–Marschak (BDM) game (Becker et al., 1964)—designed to measure actual WTP involving relatively high-stakes. In this game, participants first chose a price at which they were willing to buy a KN95 mask in a given range of prices from PKR 0 to PKR 250 (USD 0 to USD 1.5 in 2020). A computer then randomly selected a price in the same range. If the computer's price was lower than the participant's bid, they get the KN95 mask, otherwise they receive money equal to the computer's price.[3] Thus, if a participant's WTP for the mask was greater (e.g., closer to PKR 250), they are more likely to win the mask. Moreover, a positive correlation between CRT scores and WTP would indicate that individuals with higher CRT scores were more likely to have accurate risk perceptions about COVID-19 and thus have a greater WTP for masks.

We found that cognitive reflection is associated with greater truth discernment (i.e., the ability to distinguish between true and false headlines),[4] greater use but less trust in formal information sources (e.g., WHO, newspapers, government sources such as https://covid.gov.pk/). In addition, we find that those with higher CRT scores are more likely to bid a higher WTP for a KN95 mask in a real-stakes game. The positive association we observed between CRT scores and truth discernment is consistent with existing research (Pennycook & Rand, 2019a). Our work extends insights from existing research in two aspects: (i) we considered a low- and middle-income sample from a developing country with lower literacy levels than U.S. and European samples considered in earlier works, and (ii) we studied behavior—rather than just self-reported beliefs—through a real-stakes game, which showed that CRT scores predict demand for KN95 masks and thus contribute positively to COVID-19 preventative behaviors.

Our results emphasize the relationship between cognitive reflection and susceptibility to misinformation in the context of social media users in developing countries. Our research supports the "reflectionist" account of cognitive reflection (Pennycook, 2018) which emphasizes the importance of analytic thinking as a positive force in judgment and decision-making. And more specifically, individuals who are more willing to think analytically when given reasoning problems are less likely to fall for false news. To uncover possible underlying mechanisms, our multivariate regression analysis involving different outcome variables, including truth discernment and WTP for masks, shows that our results are not

---

[3] We used *easyload* to transfer the received amount in airtime to the mobile phone of respondents. https://www.telenor.com.pk/easyload/
[4] This measure of truth discernment has been used in several earlier works (e.g., Pennycook et al., 2019a; Pennycook et al., 2019b; Guess et al., 2020; Sirlin et al. 2021; and Rosenzweig et al., 2021).

moderated by education, age, gender, or household expenditure;[5] see Appendix C. The association between CRT and truth discernment can be explained, at least in part, by differences in deliberative reasoning between high and low CRT users. High CRT users are more likely to engage in deliberative reasoning, which has been found to be causally associated with better truth discernment (Bago et al., 2020).[6] Thus, it seems that people fail to distinguish between truth and falsehood partly because they do not stop to reflect sufficiently on their prior knowledge (or have insufficient or inaccurate prior knowledge). One possible mechanism driving COVID-19 preventative behaviors (e.g., demand for KN95 masks) could be that since higher CRT users exhibit greater truth discernment, they have more accurate beliefs about the COVID-19 pandemic compared to low CRT users. This suggests that they are also likely to have more accurate risk perceptions about the expected cost of contracting COVID-19, thereby affecting demand for KN95 masks. Another possible channel could be differences in the risk preferences of high and low CRT users. Frederick (2005) showed greater risk aversion among high CRT scorers for choices involving losses. In case of COVID-19, this translates into willingness to pay more for a KN95 (and the possible discomfort of wearing a mask) compared to the expected loss of suffering from illness or even death due to the COVID-19 infection.

Our findings have implications for social media platforms, governments, and citizens interested in countering COVID-19 misinformation by providing insights into which users are most vulnerable to misinformation. Our results suggest that interventions designed to counter COVID-19 misinformation should target low CRT individuals, who are worse at discerning true from false COVID-19 information. Moreover, our study highlights the importance of designing interventions to prompt individuals to engage in more reflection (Milkman et al., 2009). One such intervention could be the use of metacognitive reflection prompts (Salovich & Rapp, 2021), which involve asking questions that encourage people to reflect on their own susceptibility to inaccurate information (e.g., "When was the last time you remember relying on inaccurate information while reading?"). Such prompts have been experimentally shown to increase resistance to inaccurate information. An important benefit of metacognitive reflection is that it may motivate a more careful consideration of the accuracy of information and may also increase awareness of potential consequences of inaccurate information. Similarly, asking participants to pause and explain why a news headline is true or false reduces their intentions to share false information (Fazio, 2020). Recent work shows that there are several light-weight implementations of such prompts[7]— designed to encourage people to reflect on the accuracy of presented information[8]—which are effective and can be delivered at scale (Epstein et al., 2021). There are several ways in which social media companies, governments, and civil society organizations could deliver such interventions intended to encourage information evaluation, for example through social media ads, by showing a reflection prompt that appears when one tries to share content on a social media platform,[9] via public service announcements, etc. Moreover, researchers have called for incorporating reflection-based approaches (e.g., those involving metacognition) into media literacy education for countering exposure to misinformation (Hodgin & Kahne, 2018; Vraga et al., 2020). Finally, reflection prompts also have the

---

[5] The correlation between CRT and truth discernment as well as CRT and WTP for KN95 masks is robust to adding controls for age, gender, and household expenditure (statistically significant at the 5% level); see Appendix C.

[6] In this study, the participants' level of deliberation was experimentally manipulated, which demonstrated a causal effect whereby deliberation reduced belief in false (but not true) news, regardless of partisan alignment.

[7] These prompts included (i) asking participants how important it was to them to share only accurate news, (ii) providing digital literacy tips and (iii) asking people to judge the accuracy of a news headline.

[8] Recent works show that individuals with higher CRT scores experience larger effects of light-weight accuracy prompt interventions, which also improved discernment even for participants who answered all CRT questions incorrectly (Epstein et al., 2021; Pennycook et al., 2021; Pennycook & Rand, 2022). Thus, Pennycook et al. (2022) hypothesize that accuracy prompts may in part work by inducing more reflection. Alternatively, they might just be changing factors that people reflect on without inducing more deliberation. However, further research is needed to distinguish between these cases.

[9] Social media companies already use prompting for different purposes, for example, Twitter prompts users when they try to share a tweet that was labeled for violating one of their policies. https://help.twitter.com/en/resources/addressing-misleading-info

normative benefit that they leverage an individual's own, often latent, ability to evaluate information rather than relying on external arbiters (e.g., social media organizations, fact-checking organizations) to discern between truth and falsehood.

We hope that our work will help encourage efforts on using CRT to identify populations who may be more vulnerable to misinformation as well as in designing interventions that promote greater cognitive reflection to reduce the spread of misinformation online.

Findings

Finding 1: Users with higher CRT scores show greater truth discernment.
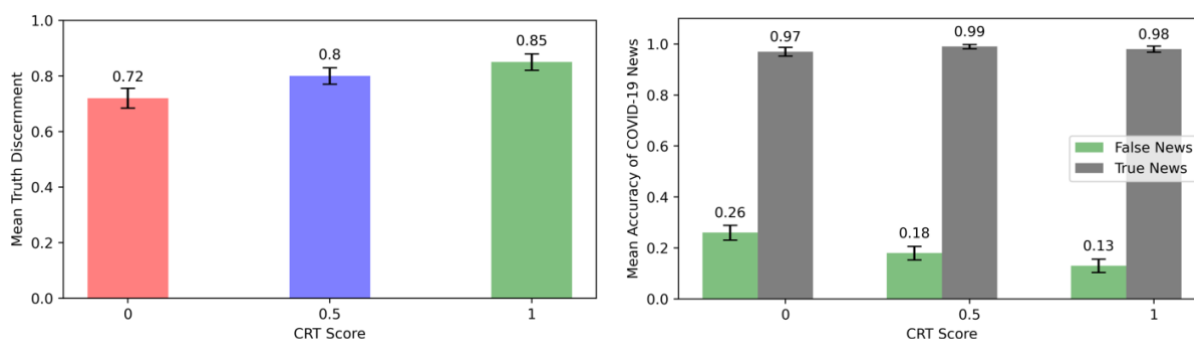

Figure 1. Relationship of CRT scores with truth discernment (left plot), defined as average accuracy ratings for true news minus average accurate ratings for false news) and accuracy of true news and false news (right plot). *For example, if a participant rated 60% of the true headlines as accurate and 20% of false headlines as accurate, their level of truth discernment would be 0.6 - 0.2 = 0.4. Thus, discernment values can range from -1 to 1. The differences in estimates are statistically significant at the 5% level between every pair of CRT groups for truth discernment and accuracy of false news but not for accuracy of true news. Error bars indicate 95% confidence intervals.*

High CRT scores are positively associated with truth discernment—which is the ability of respondents to distinguish between true news and false news stories (*defined as average accuracy ratings for true news minus average accurate ratings for false news*)—for COVID-19 news stories. We found a significant positive relationship of CRT scores with truth discernment (ß = 0.139, 95% CI = [0.092, 0.185], $p < 0.001$); see Figure 1. These results are robust to adding controls for age, gender, education level (primary, up to high school, college education or above), and household expenditure (which we use as a proxy for income); see Appendix C. These findings are consistent with existing research, which found that higher CRT scores are correlated with lower perceived accuracy of non-COVID fake news on a sample of U.S. adults recruited online (e.g., Lazer et al., 2018; Pennycook & Rand, 2019a) and COVID-19 misinformation among a sample of social media users in the United States (e.g., Pennycook et al., 2020) and Nigeria (Rosenzweig et al., 2021).

Next, we unpack whether truth discernment is driven by greater perceived accuracy of true news or lower perceived accuracy of false news by considering accuracy for false news and true news separately. We find that truth discernment is primarily driven by lower perceived accuracy of false news by respondents with higher CRT scores (ß = -0.132, 95% CI = [-0.172, -0.091], $p < 0.001$); see Figure 1. We find no significant relationship of CRT scores with the perceived accuracy of true news (ß = 0.0071, 95% CI = [-0.172, -0.091], $p = 0.453$); see Figure 1.[10]

---

[10] One possible reason for the high accuracy of true news could be that when we conducted the phone survey (i.e., in September 2020), the infection rates had declined considerably in Pakistan from the peak daily rate of around 7000 cases to 500 daily cases (see https://covid.gov.pk/stats/pakistan) and that people had formed accurate beliefs about the true news we asked them, which

Finding 2: Higher CRT scores are associated with less trust but greater use of formal information sources.

Examining trust of respondents on different sources, we found that high CRT scores are negatively associated with trust and confidence in formal information sources, which includes TV news channels, Urdu and English newspapers, radio, World Health Organization (WHO), and the government (ß = -0.343, 95% CI = [-0.508, -0.178], *p* < 0.001); see Figure 2.[11] However, high CRT users more frequently obtain COVID-19 information from formal information sources compared to low CRT users; see Appendix E, Figure D. These results suggest that even though high CRT users more frequently obtain COVID-19 information, cognitive reflection might make users less trusting of content consumed from these sources.[12] In comparison, for social media platforms (Facebook, WhatsApp, YouTube, and Twitter), we did not find a statistically significant association (at the 5% significance level) between CRT scores and trust (ß = 0.124, 95% CI = [-0.0094, 0.256], *p* = 0.067); see Figure 2. Moreover, this association is not robust to adding frequency of use of social media sources as a control (ß = -0.0139, 95% CI = [-0.143, 0.115], *p* = 0.832).
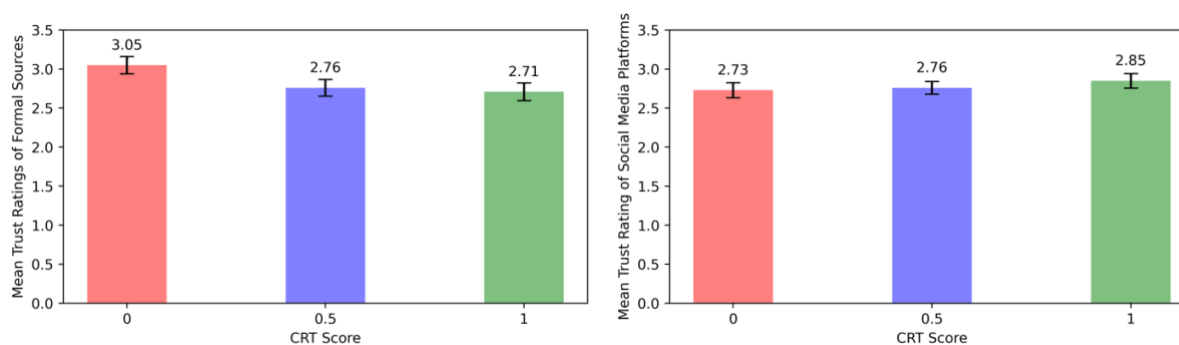

Figure 2. Relationship of CRT scores with trust in formal sources (left plot) and social media platforms (right plot). We measure participants' trust in formal information sources by their responses to the question, "How much trust and confidence do you have in each of the following sources when it comes to reporting about the coronavirus? TV News channels, Urdu & English Newspapers, Radio, World Health Organization, and the Government (left plot). For each platform, the response options were, Very low trust (1), Low trust (2), High trust (3), Very high trust and (4) Do not use this source (N/A). Similarly, we report the mean trust for social media platforms, which includes Facebook, Twitter, YouTube, and WhatsApp (right plot). We report the average trust rating of respondents with different CRT scores. The differences in estimates are statistically significant at the 5% level between CRT groups 0 and 0.5 and groups 0 and 1 for trust ratings of formal sources but not for trust ratings for social media platforms. Error bars indicate 95% confidence intervals.

Is greater frequency of use of formal sources correlated with truth discernment? We found that greater use of formal sources is associated with better truth discernment (ß = 0.37, 95% CI = [0.158, 0.589], *p* = 0.001). We also found a positive association between frequency of use of social media platforms and truth discernment (ß = 0.32, 95% CI = [0.231, 0.409], *p* < 0.001). Thus, a potential channel through which higher CRT scores might be affecting discernment is through greater exposure and use of several sources to access information, thereby enabling users to form more accurate beliefs.

---

were related to handwashing, wearing a mask, and physical distancing. Thus, it is likely that the correlation between CRT scores and truth discernment is underestimated due to ceiling effects for true news even though we expect mean discernment to be higher for CRT groups without ceiling effects.

[11] These results are robust to the inclusion of controls for age, education, gender, and household expenditure; see Appendix Figure A.

[12] This is plausible given the COVID-19 "infodemic" and the uncertainty surrounding the COVID-19 pandemic especially in 2020.

Finding 3: Outcomes of an incentive-compatible game show that users with higher CRT scores have greater willingness to pay for KN95 masks.

The outcomes of our real-stakes game designed to measure WTP show that higher CRT scores are positively associated with WTP for KN95 masks (ß = 21.7, 95% CI = [4.7, 38.8], *p* = 0.013); Figure 3. As a robustness check, we also report the proportion of KN95 masks won by respondents with different CRT scores. We find that these results are consistent with the WTP results. Taken together these findings suggest that respondents with higher CRT scores are more likely to have accurate beliefs about COVID-19 and therefore more accurate perception of the risks posed by COVID-19, resulting in greater willingness to pay for KN95 masks.
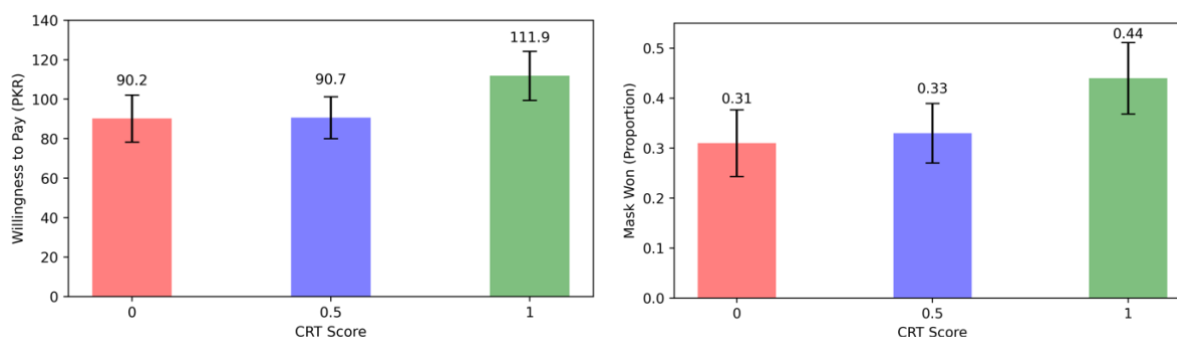

Figure 3. Relationship of CRT scores with demand for masks as measured via an incentive-compatible game. *Shown are the average WTP values for respondents with different CRT scores (left plot) and proportion of KN95 masks won across CRT scores (right plot). The differences in estimates are statistically significant at the 5% level between CRT groups 0 and 1 and groups 0.5 and 1 for WTP and the proportion of KN95 masks won. Error bars indicate 95% confidence intervals.*

These results are also consistent with existing evidence relating cognitive reflection with actual behavior. For example, considering a sample of mostly UK- and US-based individuals, Mosleh et al. (2021) found that Twitter users with higher CRT scores are more discerning in their social media use, as evidenced by the types and number of accounts followed, and by the reliability of the news sources they shared. An intervention designed to nudge people to think about accuracy increased *actual* sharing discernment in a Twitter field experiment involving American individuals (Pennycook et al., 2021).

Methods

From September 1, 2020, to September 21, 2020, we conducted a phone survey of social media users chosen randomly based on population density in the city of Lahore,[13] the capital of the Punjab province in Pakistan, to understand the relationship between cognitive reflection and truth discernment of COVID-19 headlines, trust in social media platforms for receiving credible COVID-19 information, and COVID-19 preventative behaviors. Our final dataset included a total of 621 individuals: mean age = 29.3, 50.4% female, mean monthly household expenditure = PKR 35,575 (USD 214.5 in 2020), and a median education between primary school and high school.

Our survey included questions about demographic characteristics, social media use, and COVID-19 information sources and attitudes, and CRT questions. In addition, we asked a series of questions about 10 COVID-19 news stories (seven false and three true); see Table 1 in Appendix A for the list of headlines used. We obtained the true stories from major news outlets in Pakistan including DAWN and The Express

---

[13] We use population density as a proxy for income levels. We obtained contact information and consent to participate in follow-up surveys of a field sample recruited in an earlier study. See Appendix E for details about sample selection.

Tribune as well as the World Health Organization (WHO). We used false COVID-19 news stories that had appeared online but were subsequently fact-checked and verified as untrue by various organizations including AFP Fact Check, PolitiFact, WHO, and UNICEF. Finally, the respondents participated in a real-stakes mask game. While all participants were presented with the same set of news stories, they were presented in a random order to each participant. After hearing each news story, we asked if they thought the news was true or not ("To what an extent do you agree with the statement?") to which the participants responded using a slider scale 0 (*disagree*), 0.5 (*not sure*), and 1 (*agree*).[14] We then found respondent-level truth discernment, which is defined as average accuracy ratings for true news minus average accurate ratings for false news and is a measure of respondents' ability to distinguish between true news and false news stories. The mean discernment across all respondents was 0.79.

Traditional CRT questions (Frederick, 2005) are known to confound reflection ability with mathematical ability and thus they are likely to be less useful for populations with low numeracy skills. Given the relatively low education level in our sample, we used a 2-item CRT from the validated set of verbal CRT questions by Sirota et al. (2020).[15] Each question had an obvious but incorrect response to assess individuals' ability to suppress a default (intuitive) incorrect response in favor of a correct reflective response. We asked both questions in Urdu, the national language of Pakistan. The two CRT questions taken from Sirota et al. (2020) were as follows:

- "If you were running a race, and you passed the person in second place, what place would you be in now?". The correct answer is "second" whereas the intuitive answer is "first."
- "Fatima's father has 4 daughters but no sons—Tina, Mina, Rina. What is the fourth daughter's name?"[16] The correct answer is "Fatima" whereas the intuitive answer could be any name ending with the "ina" sound (e.g., Bina, Sina).

The CRT score was coded as follows: 0 (*neither correct*), 0.5 (*one correct*), 1 (*both correct*). One criticism of the traditional CRT questions has been their high familiarity (Sirota et al., 2020). In our sample, 88% of the respondents had not heard of the verbal CRT questions we asked. We found that our results are robust after controlling for a binary variable capturing whether a respondent was already aware of the CRT questions (1 if aware of both the questions, 0 otherwise).

We measured trust in social media platforms using the question, "How much trust and confidence do you have in news about the coronavirus received through the following social media platforms: WhatsApp, Facebook, Twitter, and YouTube." For each platform, the response options were (1) *Very low trust*, (2*) Low trust,* (3) *High trust,* (4) *Very high trust,* (N/A) *Do not use this source*. We constructed a respondent-level trust variable by taking the average across all four platforms. The mean trust across all respondents was 2.78. We did not assign any value to the response *Do not use this source*. Similarly, we measured trust in formal information sources using the question, "How much trust and confidence do you have in each of the following sources when it comes to reporting about the coronavirus?: News Channels, Urdu Newspapers, English Newspapers, Government Sources (e.g., covid.gov.pk, press briefings, etc), Radio, the World Health Organization (WHO)" and reported the average trust across all these sources (the response options were the same as above) across CRT scores. The mean trust for formal sources across all respondents was 2.83.

---

[14] An important distinction between our phone survey and earlier studies that used online media formats to present false and true news is that the latter includes contextual information, such as the source of news stories whereas in our phone survey, no source information was provided. Thus, our survey measures the accuracy of already held beliefs about COVID-19 rather than their ability to use contextual clues (e.g., source) to determine the veracity of information. Therefore, our findings may generalize to situations where exposure to (mis)information might be high, and people have formed some beliefs about the accuracy of various statements.

[15] Earlier works have used CRTs comprising a different number of questions including a 3-item (Frederick 2005), a 4-item (Thomson et al., 2016), and a 7-item (Pennycook et al., 2019a) CRT.

[16] We changed the names in the original question from (Mary, Nana, Nene, Nini, Nono) to (Fatima, Tina, Mina, Rina) but retained the structure to reflect greater awareness of latter names in the local context.

To understand the relationship between cognitive reflection and actual behavior, respondents participated in an incentive-compatible game designed to measure WTP involving real stakes. We measured WTP for KN95 masks at the start of the game using the question, "Please choose a price at which you are willing to buy a KN95 mask in the range of PKR 0 - PKR 250."[17] After choosing a price, a computer randomly selected a price in the same range. For ease of delivering money, the computer chose a price from a discrete but uniformly spaced set (0, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250). If the computer's price was lower than the participant's bid, they received a KN95 mask,[18] otherwise, they received an amount equal to the computer's price as mobile airtime. Thus, if a participant's WTP for the mask was greater (e.g., closer to PKR 250) then they were more likely to win the mask. In the text, we have reported the regression coefficient, *p*-value, and 95% confidence interval of the estimated coefficient from a regression of the outcome on the CRT score.

Bibliography

Ali, A., & Qazi, I. A. (2021). Countering misinformation on social media through educational Interventions: Evidence from a randomized experiment in Pakistan. *arXiv*. https://arxiv.org/abs/2107.02775

Bago, B., Rand, D. G., & Pennycook, G. (2020). Fake news, fast and slow: Deliberation reduces belief in false (but not true) news headlines. *Journal of Experimental Psychology: General, 149*(8), 1608–1613. https://doi.org/10.1037/xge0000729

Becker, G. M., Degroot, M. H., & Marschak, J. (1964). Measuring utility by a single-response sequential method. *Journal of the Society of Systems Research*, *9*(3), 226–232. https://doi.org/10.1002/bs.3830090304

Campitelli, G., & Gerrans, P. (2014). Does the cognitive reflection test measure cognitive reflection? A mathematical modeling approach. *Memory & Cognition, 42*, 434–447. https://doi.org/10.3758/s13421-013-0367-9

Ellis, P., Roberts, M. (2016). *Leveraging urbanization in South Asia*. The World Bank Group. https://openknowledge.worldbank.org/bitstream/handle/10986/22549/9781464806629.pdf?sequence=17&isAllowed=y

Epstein, Z, Berinsky, B., Cole, R., Gully, A., Pennycook, G., & Rand, D. G. (2021). Developing an accuracy-prompt toolkit to reduce COVID-19 misinformation online. *Harvard Kennedy School (HKS) Misinformation Review, 2*(3). https://doi.org/10.37016/mr-2020-71

Evans, J. S., & Stanovich, K. E. (2013). Dual-process theories of higher cognition: Advancing the debate. *Perspectives on Psychological Science, 8*(3), 223–241. https://doi.org/10.1177/1745691612460685

Fazio, L. (2020). Pausing to consider why a headline is true or false can help reduce the sharing of false news. *Harvard Kennedy School (HKS) Misinformation Review, 1*(2). https://doi.org/10.37016/mr-2020-009

Frederick, S. (2005). Cognitive reflection and decision making. *Journal of Economic Perspectives, 19*(4), 25–42. https://doi.org/10.1257/089533005775196732

Guess, A., Lerner, M., Lyons, B., Montgomery, J. M., Nyhan, B., Reifler, J., & Sircar, N. (2020). A digital media literacy intervention increases discernment between mainstream and false news in the United States and India. *Proceedings of the National Academy of Sciences, 117*(27), 15536–15545. https://doi.org/10.1073/pnas.1920498117

---

[17] In 2020, PKR 250 was equal to USD 1.5. The survey participation compensation was PKR 100 (USD 0.6), which was separate from the money involved in the game.
[18] The KN95 masks were delivered to the homes of the survey takers by our implementation partner.

Hargittai, E., & Shaw, A. (2020). Comparing internet experiences and prosociality in Amazon Mechanical Turk and population-based survey samples. *Socius*. https://doi.org/10.1177%2F2378023119889834

Harriri, M. (2020). Cities in bad shape: Urban geometry in India. *American Economic Review, 10*(8), 2377–2421. https://www.aeaweb.org/articles?id=10.1257/aer.20171673

Hodgin, E., & Kahne, J. (2018). Misinformation in the information age: What teachers can do to support students. *Social Education, 82*(4), 208–212. https://www.socialstudies.org/system/files/publications/articles/se_8204208.pdf

Kaye, D. (1995). Sources of information, formal and informal. *Management Decision*, *33*(5), 13–15. https://doi.org/10.1108/EUM0000000003898

Lazer, D. M. J., Baum, M., Benkler, Y., Berinsky, A. J., Greenhill, K. M., Menczer, F. Metzger, M., Nyhan, B., Pennycook, G., Rothschild, D., Schudson, M., Sloman, S. A., Sunstein, C., Thorson, E. A., Watts, D. J., & Zittrain, J. L. (2018). The science of fake news. *Science*, 359(6380), 1094–1096. https://doi.org/10.1126/science.aao2998

Liberali, J. M., Reyna, V. F., Furlan, S., Stein, L. M., & Pardo, S. T. (2012). Individual differences in numeracy and cognitive reflection, with implications for biases and fallacies in probability judgment. *Journal of Behavioral Decision Making, 25*(4), 361–381. https://doi.org/10.1002/bdm.752

Mastrogiorgio, A. (2015). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 6*(532), 936. https://doi.org/10.3389/fpsyg.2015.00936

Milkman, K. L., Chugh, D. & Bazerman, M. H. (2009). How can decision making be improved? *Perspectives on Psychological Science, 4*(4), 379–383. https://doi.org/10.1111/j.1745-6924.2009.01142.x

Mosleh, M., Pennycook, G., Arechar, A., & Rand, D. G. (2021). Cognitive reflection correlates with behavior on Twitter. *Nature Communications, 12*(921). https://doi.org/10.1038/s41467-020-20043-0

Pennycook, G. (2018). *The new reflectionism in cognitive psychology: Why reason matters*. Routledge.

Pennycook, G., Cheyne, J. A., Seli, P., Koehler, D. J., & Fugelsang, J. A. (2012). Analytic cognitive style predicts religious and paranormal belief. *Cognition, 12*3(3), 335–346. https://doi.org/10.1016/j.cognition.2012.03.003

Pennycook, G., Epstein, Z., Mosleh, M., Arechar, A. A., Eckles, D., & Rand, D. G. (2021). Shifting attention to accuracy can reduce misinformation online. *Nature*, *592*, 590–595. https://doi.org/10.1038/s41586-021-03344-2

Pennycook, G., McPhetres, J., & Zhang, Y. (2020). Fighting COVID-19 misinformation on social media: Experimental evidence for a scalable accuracy-nudge intervention. *Psychological Science, 31*(7), 770–780. https://doi.org/10.1177%2F0956797620939054

Pennycook, G., & Rand, D. G. (2019a). Lazy, not biased: Susceptibility to partisan fake news is better explained by lack of reasoning than by motivated reasoning. *Cognition*, *188*, 39–50. https://doi.org/10.1016/j.cognition.2018.06.011

Pennycook, G., & Rand, D. G. (2019b). Fighting misinformation on social media using crowdsourced judgments of news source quality. *Proceedings of the National Academy of Sciences of the United States of America, 116*(7), 2521–2526. https://doi.org/10.1073/pnas.1806781116

Pennycook, G., & Rand, D. G. (2019c). Cognitive reflection and the 2016 U.S. Presidential election. *Personality and Social Psychology Bulletin, 45*(2), 224–239. https://doi.org/10.1177/0146167218783192

Pennycook, G., & Rand, D. G. (2022). Accuracy prompts are a replicable and generalizable approach for reducing the spread of misinformation. *Nature Communications, 13*(1), 2333. https://doi.org/10.1038/s41467-022-30073-5

Pennycook, G., & Ross, R. M. (2016). Commentary: Cognitive reflection vs. calculation in decision making. *Frontiers in Psychology, 7*(9). https://doi.org/10.3389/fpsyg.2016.00009

Rosenzweig, L. R., Bago B., Berinsky, A. J., & Rand, D. G. (2021). Happiness and surprise are associated with worse truth discernment of COVID-19 headlines among social media users in Nigeria. *Harvard Kennedy School (HKS) Misinformation Review*, *2*(4). https://doi.org/10.37016/mr-2020-75

Salovich, N. A., & Rapp, D. N. (2021). Misinformed and unaware? Metacognition and the influence of inaccurate information. *Journal of Experimental Psychology: Learning, Memory, and Cognition*, *47*(4), 608–624. https://doi.org/10.1037/xlm0000977

Sirlin, N., Epstein, Z., Arechar, A. A., & Rand, D. G. (2021). Digital literacy is associated with more discerning accuracy judgments but not sharing intentions. *Harvard Kennedy School (HKS) Misinformation Review*, *2*(6). https://doi.org/10.37016/mr-2020-83

Sirota, M., Dewberry, C., Juanchich, M., Valus, L., & Marshall, A. C. (2020). Measuring cognitive reflection without maths: Development and validation of the verbal cognitive reflection test. *Journal of Behavioral Decision Making, 34*(3), 322–343. https://doi.org/10.1002/bdm.2213

Swami, V., Voracek, M., Stieger, S., Tran, U. S., & Furnham, A. (2014). Analytic thinking reduces belief in conspiracy theories. *Cognition, 133*(3), 572–585. https://doi.org/10.1016/j.cognition.2014.08.006

Thomson, K. S., & Oppenheimer, D. M. (2016). Investigating an alternate form of the cognitive reflection test. *Judgment and Decision Making, 11*(1), 99–113. http://dx.doi.org/10.1037/t49856-000

Toplak, M. E., West, R. F., & Stanovich, K. E. (2011). The cognitive reflection test as a predictor of performance on heuristics-and-biases tasks. *Memory & Cognition, 39*, 1275–1289. https://doi.org/10.3758/s13421-011-0104-1

Vraga, E. K., Tully, M., Maksl, A., Craft, S., & Ashley, S. (2020). Theorizing news literacy behaviors. *Communication Theory, 31*(1), 1–21. https://doi.org/10.1093/ct/qtaa005

Competing interests
The authors do not have any competing interests.

Ethics
The data collection was approved by the institutional review board (IRB). Participants agreed on participation via voluntary opt-in and informed consent procedures. Participants could withdraw participation or fill out other/don't know categories for questions involving i.e., gender. All data was processed anonymously.

Data availability
All materials needed to replicate this study are available via the Harvard Dataverse:
https://doi.org/10.7910/DVN/ZOAKLW

Appendix A: COVID-19 headlines

Table 1. COVID-19 headlines used in the study.

	COVID-19 Headline	Veracity
1	5G mobile networks spread coronavirus	False
2	There is no coronavirus. These are all lies	False
3	Coronovirus is a conspiracy against Islam and Muslims	False
4	Senna leaves can treat coronavirus	False
5	Only old people are at risk of contracting or dying from coronavirus	False
6	Hair dryers are effective in killing the coronavirus	False
7	Pneumonia vaccine provides protection against coronavirus	False
8	Social (physical) distancing can reduce the spread of coronavirus	True
9	Washing your hands with soap and water kills viruses that may be on your hands	True
10	Wearing a mask can reduce the spread of coronavirus	True

Appendix B: Mask game

(a) Protocol
Participants are provided the following information about the benefits of KN95 masks:
- KN95 masks provide 95% protection against viruses including coronavirus, bacteria, dust, etc.
- KN95 masks are easy to wear, can be worn for a long time, and do not cause a rash or skin irritation
- KN95 masks can be reused
- KN95 masks are approved by the World Health Organization as a safety measure from COVID-19

(b) Now we will play a game in which you will have a chance to win a KN95 mask or money.

(c) Game instructions:
- Please choose a price at which you are willing to buy a KN95 mask in the range of Rs. 0 - Rs. 250
- The computer will randomly select a price in the same range, i.e., Rs. 0 - Rs. 250 (for ease of delivering money, computer will choose price from a discrete set 0, 25, 50, 75, 100, 125, 150, 175, 200, 225, 250)
- If computer's price <= your price, you will get the KN95 mask
- If computer's price > your price, you will get money = computer's price
- Some examples below can help participants understand the game better:
  - If you select Rs. 100 and the computer picks Rs. 50 you will get the mask
  - If you select Rs. 100 and the computer picks Rs. 100 you will get the mask
  - If you select Rs. 100 and the computer picks Rs. 150 you will get Rs. 150
- Participants are asked to think carefully before choosing their price.
- Participant is informed of the outcome.

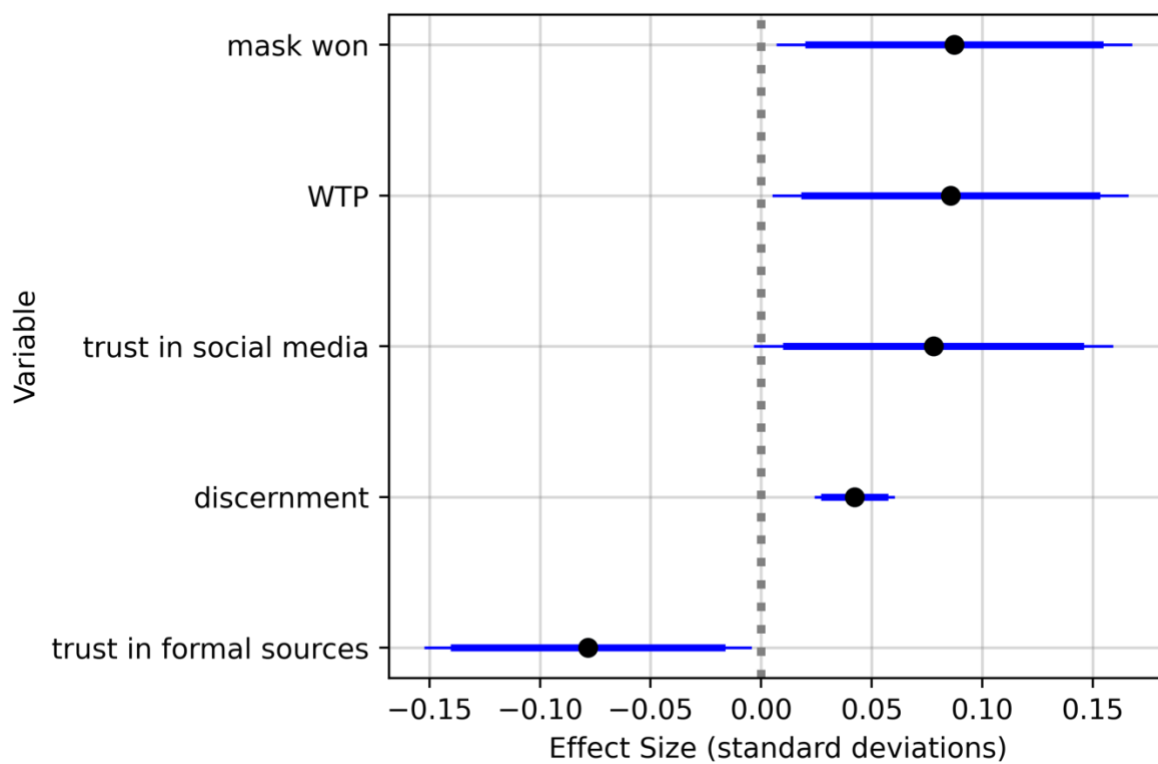## Appendix C: Robustness checks

Figure A. Regression coefficients of CRT scores for different dependent variables (shown on the y-axis) along with confidence intervals after controlling for age, gender, education, and income. *All variables have been standardized, mean-centered and scaled by standard deviation for comparability. Thin lines indicate 95% confidence intervals (CIs), thick lines indicate 90% CIs. Right of the dotted zero line (i.e., positive values) indicates a positive correlation.*
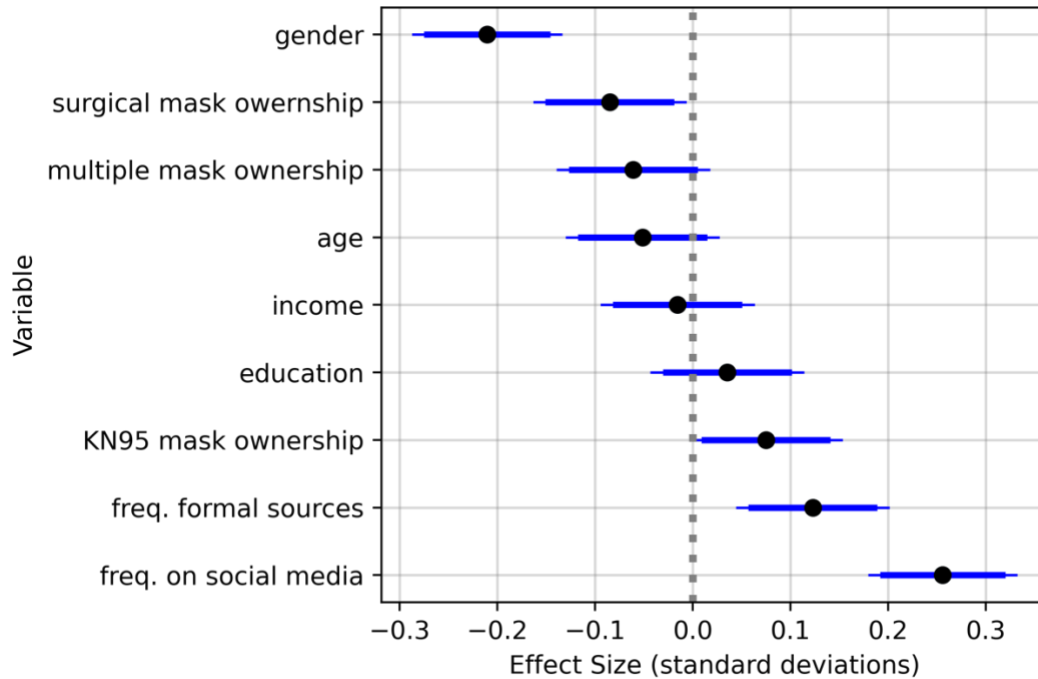
Appendix D: Additional results

Figure B. Correlates of CRT scores. *All variables have been standardized, mean-centered and scaled by standard deviation for comparability. The variable "freq. on social media" corresponds to the question, "How frequently do you use the following social media platforms for receiving/sharing news about the coronavirus? WhatsApp, Facebook, Twitter, YouTube" with response options: Often (1), Sometimes (0.67), Rarely (0.33), Never (0), Do not use this source (0). The variable "freq. formal sources" corresponds to, "How frequently are you getting information from each of the following sources about the coronavirus through any medium (including reading online, watching on TV, etc.)? News Channels, Urdu Newspapers, English Newspapers, Government Sources, Radio, World Health Organization" with the same response options as in the previous question. Gender=1 for females and 0 otherwise. Thin lines indicate 95% confidence intervals (CIs), thick lines indicate 90% CIs. Right of the dotted zero line (i.e., positive values) indicates a positive correlation.*
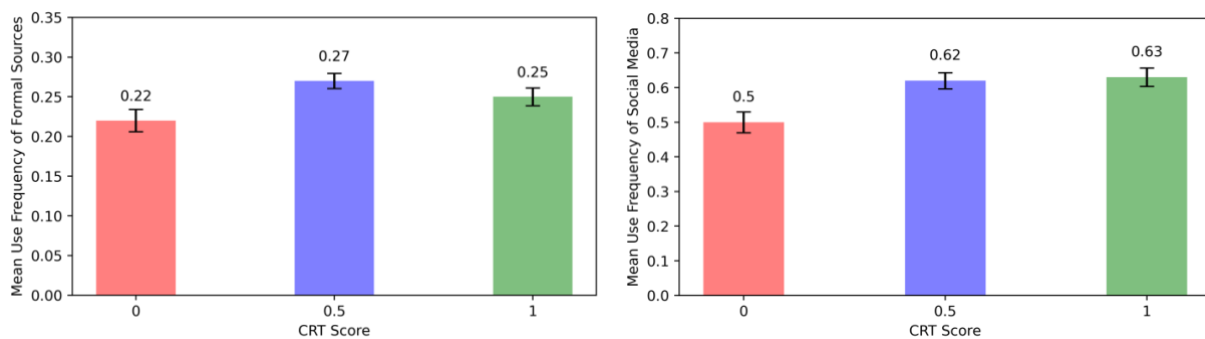

Figure C. Relationship of CRT scores with frequency of use of formal information sources (left plot) and social media platforms (right plot). *The outcome variable (left plot) corresponds to the question, "How frequently are you getting information from each of the following sources about the coronavirus through any medium (including reading online, watching on TV, etc.)? News Channels, Urdu Newspapers, English Newspapers, Government Sources, Radio, World Health Organization" with response options: Often (1), Sometimes (0.67), Rarely (0.33), Never (0), Do not use this source (0). The outcome variable (right plot) corresponds to, "How frequently do you use the following social media platforms for receiving/sharing news about the coronavirus? WhatsApp, Facebook, Twitter, YouTube" with the same response options as in the previous question. We report the average values for use frequency for respondents with different CRT scores. Error bars indicate 95% confidence intervals.*

Appendix E: Sample selection

From September 1, 2020, to September 21, 2020, we conducted a phone survey of 621 social media users who were recruited as part of an earlier field study on misinformation, conducted in the city of Lahore, Pakistan (Ali & Qazi, 2021). The original field study involved a random sample of 750 social media users drawn from low- and middle-income areas of the city. During the baseline survey conducted in May 2019, we obtained contact information and consent to participate in follow-up surveys. We obtained consent and phone numbers of 695 users out of which we were successfully able to contact and complete the phone survey of 621 users (i.e., response rate of 89.3%). Out of the 695 phone numbers, we found that 3.5% were wrong numbers, 5.2% were switched off, and 1.4% did not pick up the phone. Our final dataset included a total of 621 individuals: mean age = 29.3, 50.4% female, mean monthly household expenditure = PKR 35,575 (USD 214.5 in 2020), and a median education between primary school and high school. The phone survey was conducted by our implementation partner, the Survey Wing of the Institute of Development Alternatives (IDEAS).

Field sample recruitment. The original field sample of 750 social media users was randomly drawn from low- and middle-income areas of the city of Lahore. We used inverse population density as a proxy for income. Congestion is common in older and poorer areas of Lahore, due to the non-compact and horizontal pattern of urban development observed commonly in South Asian cities (Ellis & Roberts, 2016; Harriri, 2020). Therefore, we used AsiaPop (2013) satellite data to provide population counts at a spatial resolution of 100 m by 100 m to identify congested low- and middle-income neighborhoods. The areas selected for drawing the sample accounted for nearly 35% of the city's total population and covered seven of the fourteen National Assembly constituencies in the city. The median population density in the selected areas of our study was 109 persons per 100 m by 100 m grid whereas the median density in areas not covered by our sample was 28 persons per 100 m by 100 m grid. Figure D shows the inverse relationship between population count on a 100 m x 100 m grid and the average household's monthly expenditures reported in our survey. A random sample of 200 grids was drawn from the selected areas and five households per grid were surveyed, covering a total of 1000 households in the baseline survey. To initiate the data collection within the chosen grids, a random point (x and y coordinate) was dropped within the grid. The enumerators arrived at the point and used the left-hand rule to survey, within each grid, five households where at least one social media user was present. The definition of social media user for our survey was that the respondent must be at least 18 years of age and be a social media user.
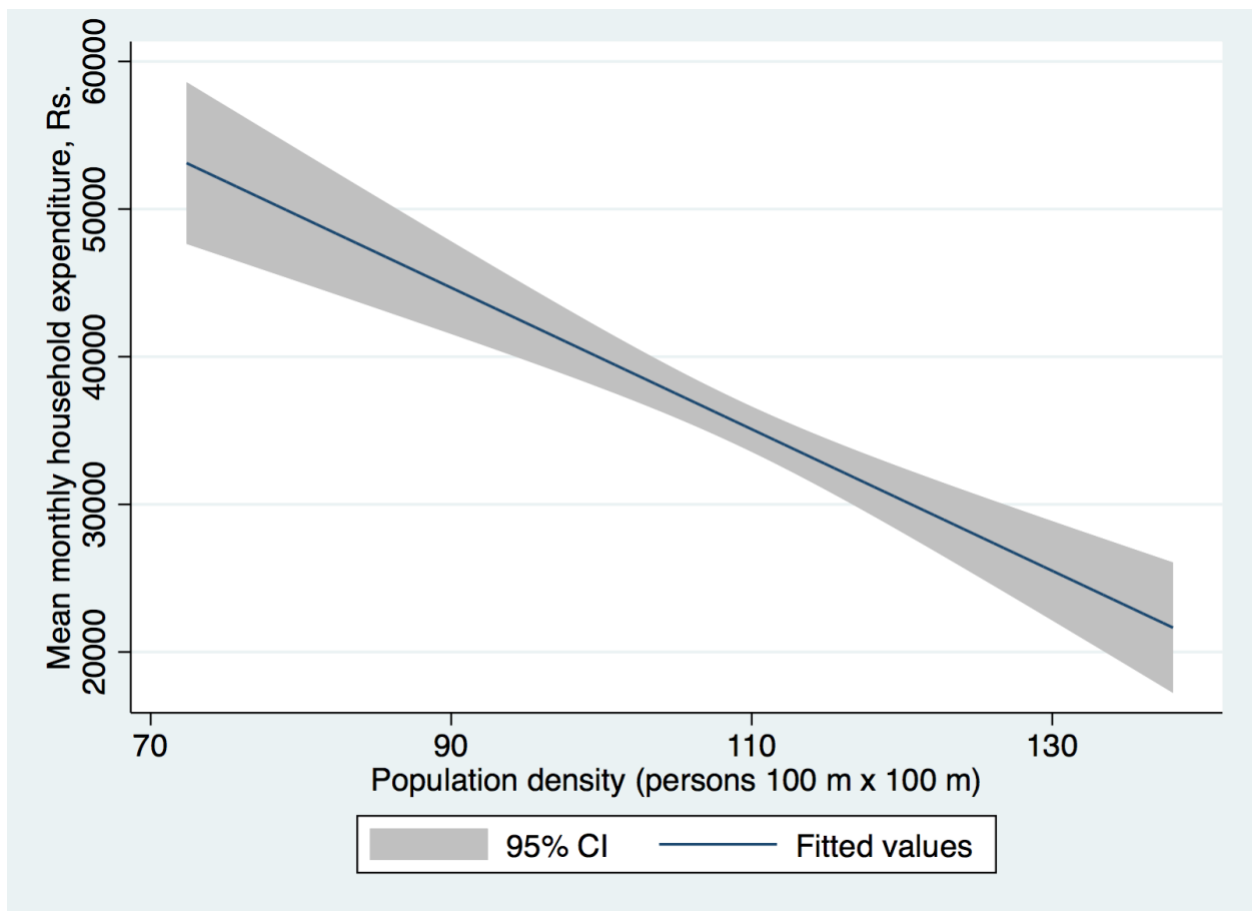
Figure D. Relationship between mean monthly household expenditure and population density (persons per 100m x 100m) in survey data collected in Lahore, Pakistan.