

Less identifiable but no less helpful: Face masks do not go along with selfish behavior

Yashvin Seetahul (✉ yashvin.seetahul@uibk.ac.at)

Universität Innsbruck

Tobias Greitemeyer

Universität Innsbruck

Research Article

Keywords: Face Masks, Perceived Identifiability, Prosocial Behavior, Field Study

Posted Date: March 8th, 2022

DOI: <https://doi.org/10.21203/rs.3.rs-1314451/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Abstract

Covering the face with masks in public settings has been recommended since the start of the pandemic. Because faces provide information about identity, and that face masks hide a portion of the face, it is plausible to expect individuals who wear a mask to consider themselves less identifiable. Prior research suggests that perceived identifiability is positively related to prosocial behavior, and with two pre-registered field studies (total N = 5706) we provide a currently relevant and practical test of this relation. Our findings indicate that mask wearers and non-wearers display equivalent levels of helping behavior (Studies 1 and 2), although mask wearers have a lower level of perceived identifiability than those without a mask (Study 2). Overall, our findings suggest that claims that face masks are related to selfish behavior are not warranted, and that there is no practical link between perceived identifiability and prosocial behavior.

Statement Of Relevance

Although individuals who don't wear a face mask during the COVID-19 pandemic are accused of selfishness, prior research would suggest that wearing a face mask should decrease prosocial behavior. Several findings have linked reduced perceived identifiability to lower perceived risks of social reproach, lower incentives to help and thus a decrease in helping behavior. In the present work, we provide strong evidence for equivalence between individuals wearing a face mask and individuals not wearing one. Our findings are of practical relevance, given the ongoing pandemic, as they contribute to the overall understanding of the social and behavioral consequences of wearing face masks. Our findings are also of theoretical relevance as they provide counter evidence to prior research despite the nearly absolute power to detect small effects.

Introduction

Since the start of the COVID-19 pandemic, those who refuse to wear a sanitary face mask in the presence of others have often been perceived as selfish, whereas those who wear a face mask are considered altruistic (e.g., Christakis, 2021; Kelleher, 2021). Yet, throughout evolution, individuals have evolved to use faces as the primary cue to identify one another (Allen, 2006) and because perceived anonymity has been associated to reduced prosocial behaviors and increased selfishness (Zhong, Bohns and Gino, 2010), with two pre-registered field studies, we consider the possibility that wearing a face mask serves to avoid having to behave altruistically.

Costs and benefits of helping and refusing to help

Prosocial behaviors refer to a general category of voluntary behaviors that are carried out with the intent of providing benefits to others. They may benefit other individuals, entire groups, or society as a whole (Penner et al., 2005). It comes as no surprise that “[a]ssisting others, donating to charity, cooperating with others, and intervening to save another person’s life are all acts that societies generally value” (Dovidio,

1984, p. 364). However, while these behaviors will—by definition—provide benefits to the recipient, they come at a cost to the benefactor (e.g., time, effort, money). Although prosocial behaviors can also provide material, social, psychological, or even physical gains (Brown & Brown, 2015; Cabral, Ozbay, & Schotter, 2014; Poulin & Holmann, 2013; Wand, Deng, & Schen, 2019) to the benefactor, if perceived costs outweigh the benefits people will typically decide not to provide help (for reviews, Batson, Ahmad, & Stocks, 2011; Cialdini & Goldstein, 2004; Rahal & Fiedler, 2021).

Importantly, while behaving prosocially may come at personal costs, refusing to do so is not without its own toll. Not helping someone in need may lead to social reproach. Because selfishness is not valued by social norms (Schlenker & Leary, 1982), behaving selfishly may lead to public censure and ostracism (Huang, Zhu, & Zhang, 2021). Given that “human beings have a pervasive drive to form and maintain at least a minimum quantity of lasting, positive, and significant interpersonal relationships” (Baumeister & Leary, 2005, p. 497), they will be highly motivated to avoid the risk of being disliked or rejected (MacDonald & Leary, 2005). The risk of social reproach is all the more relevant when refusing to comply to an explicit and direct request for help, in which case “many people agree to things—even things they would prefer not to do—simply to avoid the considerable discomfort of saying ‘no.’” (Bohns, 2016, p. 120). In fact, individuals may prefer avoiding the social interaction altogether, rather than having to refuse helping someone (e.g., Andreoni, Rao, & Trachtman, 2017; Dana, Cain, & Dawes, 2006; DellaVigna, List, & Malmendier, 2012).

Hence, when deciding to help, individuals not only consider costs and benefits of helping but also the costs of refusing. If individuals perceive the costs of refusing as being higher, they may (unhappily) comply with the request. Here, we focus on a situation that reduces the risk of social reproach, and thus refusing to help is easier: when individuals are unidentifiable. In fact, individuals made 84% lower donations when their behavior was not observed and their names hidden, compared to when they were identifiable (Vesely & Klöckner, 2018), and the subjective belief that their behavior could be clearly identified increased individuals’ contribution behavior (Cress & Kimmerle, 2008; Kimmerle, Cress, & Hess, 2007). Similarly, Zhong et al. (2010) found that participants wearing sunglasses had an illusory sense of anonymity and behaved more selfishly than those wearing clear glasses. It is then plausible to expect that sanitary face masks would decrease compliance to helping requests.

The present research

The years 2020 and 2021 have been severely affected by the COVID-19 outbreak, and covering the face with masks is recommended and oftentimes mandatory in public settings to reduce the risk of transmission of the virus. With the possibility that face masks could remain highly present in social contexts in post-COVID-19 years, the health, social and psychological consequences of face masks have been topics of interest in public debates and in research (e.g., Barrick et al., 2020; Beltran et al., 2021; Crandall & Bahns, 2021; Fatfouta & Trope, 2021; Grahlow et al., 2021; Shehu et al., 2021; Stajduhar et al., 2021; Williams et al., 2021).

While individuals who refuse to wear a mask have been accused of being selfish, we investigate the possible increase in selfishness of those who do wear a mask as a byproduct of the masks themselves. Because these masks cover the mouth and nose area, they hide a large portion of the face and may reduce the actual and perceived identifiability, which then may decrease the extent to which they behave prosocially. With two field studies, we provide a currently relevant and practical test of the theoretical relation between identifiability and prosocial behaviors. We approached individuals in public areas who were either wearing a face mask or not wearing one. These individuals were provided the opportunity to behave prosocially to various degrees by answering questions in a long survey. Verbal consent was obtained from all individuals who agreed to answer the survey. The first page of the survey informed them that the participation was voluntary and that they could stop at any time (which was also reminded on all pages of the survey). We also counted how many individuals refused to answer the survey in each group (Mask groups vs. No Mask group). In both studies, we compared the number of questions answered by individuals wearing a face mask and individuals not wearing one. Study 2 also examined the impact of wearing a face mask on the participant's level of perceived identifiability. Both studies conform to the guidelines of the of the Austrian sanitary regulations in place during the data collection, the Declaration of Helsinki and the good scientific practice and recommendations of the European Union for Social Science and Humanities. According to the Ethics Committee of the University of Innsbruck, no ethical approval was required for this study as no identifiable human data was gathered and no intervention or manipulation took place in the context of this study. All data generated and analyzed during these studies as we followed the mentioned guidelines, as well as registrations, materials, and code are available on the OSF repository: <https://doi.org/10.17605/OSF.IO/CRZM3>.

Study 1

Results

Descriptive statistics. From our $N = 398$ participants, 186 wore a mask during the entire interaction (i.e., Mask group), 212 did not wear a mask during the entire interaction (i.e., No Mask group).

The overall mean for the prosocial behavior measure was 17.39 with a standard deviation of 35.2, a minimum of 0 and maximum of 102. The skewness of the distribution is 1.78 and the kurtosis is 4.44. Most participants in both groups answered 0 questions (see Figure 2). The prosocial measure for the Mask group was $M = 18.68$ ($SD = 35.89$) and for the No Mask group was $M = 16.25$ ($SD = 34.62$).

Hypothesis test. As per our pre-registration (<https://doi.org/10.17605/OSF.IO/UMY4F>), we conducted a one-sided Student's t-test on R to test whether the number of questions answered by participants in the Mask group was smaller than in the No Mask group. Given the continuous ratio nature of the dependent variable, and two-level categorical nature of the independent variable, the one-sided Student's t-test provides more power than alternative analytic strategies to observe whether the Mask group has a smaller mean than the No Mask group.

As indicated by the skewness and kurtosis of our measure in the descriptive statistics, the distribution of our measure violates the normal distribution resulting in a strong “floor effect” and a moderate “ceiling effect”. Transforming the distribution did not allow us to obtain a normal distribution, whether the transformation was Log (skewness = 1.31, kurtosis = 2.84), Square-Root (skewness= 1.53, kurtosis = 3.6) or Reciprocal (skewness = 1.78, kurtosis = 4.44). However, while normality of distribution of the dependent variable is necessary for small samples, non-normal distributions do not affect the validity of t-tests for large samples (Lumley, Diehr, Emerson, & Chen, 2002), hence the non-normal distribution of our prosocial measure is unlikely to impact the validity of our conclusions.

The result of our t-test indicates that the difference between the two groups was non-significant, $t(396) = -0.69, p = .754$ (one-tailed), $d = -0.07$, 95% CI [-Infinity, 8.267]. As the descriptive statistics indicate, participants in the Mask group actually answered more questions than participants in the No Mask group, which is the opposite of our hypothesis.

Additional test. We conducted an exploratory one-sided Student’s t-test to see if we could potentially make the inference that face masks increase prosocial behavior, but the difference was non-significant, $t(396) = -0.69, p = .246$ (one-tailed), $d = 0.07$, 95% CI [-3.4, Infinity].

Given the non-significant results of the t-tests, we conducted an exploratory equivalence test on R with the TOSTER package (Lakens, 2018). While non-significant t-tests by themselves provide inconclusive null results, equivalence testing allows for a more informative interpretation of the null. This method consists in specifying a “null interval” or an equivalence region, which is an interval within which the effect size would be too small to be of any interest and is therefore negligible, and then in testing whether the observed effect is within the bounds of this interval of negligible effect sizes (Dienes, 2021; Weber & Popova, 2012). Testing whether an effect is within the bounds of a specified interval is conducted by the combination of two one-sided t-tests (TOST, Lakens, 2017, 2018), a first one to test whether the effect is greater than the lower bound of the null region and a second one to test whether the effect is smaller than the upper bound. The combination of the two t-tests conducted with an α err probability of 0.05 provides a 90% CI (i.e., $1-2\alpha$), and the result of the equivalence test is significant when the 90% CI is within the null region and does not overlap with any of the two bounds. The result for the TOST suggests that we can reject the presence of the initially hypothesized difference of $d = 0.36$ or larger, $t(396) = 2.93, p = .002$, 90% CI [-0.097, 0.235] (see Figure 3). Further exploration indicates that the p value for the TOST remains below 0.05 ($p = .0496$) for an effect as small as $d = 0.24$, 90% CI [-0.097, 0.235].

Discussion

The data from this first study did not support our predicted difference of $d = 0.36$ (12.8 questions) between the Mask and No Mask groups. There are two potential explanations for the null result. First, it may be that the lack of identifiability does not truly reduce prosocial behaviors. This would imply that conclusions in prior research in the literature were erroneous (i.e., false positives). Second, it may be that face masks do not actually reduce individuals’ perception of their own identifiability. One untested auxiliary assumption of our study was that individuals wearing sanitary face masks during the COVID-19

pandemic felt more anonymous and less identifiable, which is why, in turn, they would have behaved less prosocially. It is possible that, given that individuals focus on the eye area for face recognition (Barton et al., 2006), individuals wearing face masks would expect others to be able to identify them even with the masks. We addressed these caveats in a second field study.

Study 2

Having observed no significant difference in the first study, this time we hypothesized that participants wearing a mask and participants not wearing one would conduct equivalent levels of prosocial behaviors. We also measured perceived identifiability and hypothesized that individuals with and without a mask will have equivalent scores. Finally, we addressed the potential inflation of effect sizes in the literature by considering a much smaller effect size of interest. A recent study by Lovakov and Agadullina (2021) estimated that the median effect size for non-experimental studies in social psychology was $d = 0.19$. We rounded this value and considered an effect size of $d = 0.2$.

Results

Descriptive statistics. From our $N = 5308$ participants, 2250 wore a mask during the entire interaction (i.e., Mask group) and 3058 did not (i.e., No Mask group).

The overall mean for the prosocial behavior measure was 8.38 with a standard deviation of 26.14, a minimum of 0 and a maximum of 106. The skewness of the distribution is 3.17 and the kurtosis is 8.58. As with the first study, in both groups, most participants answered 0 questions. The prosocial measure for the Mask group was $M = 9.05$ ($SD = 27.06$) and for the No Mask group was $M = 7.88$ ($SD = 25.44$).

The overall mean of the perceived identifiability measure was 3.12 ($SD = 0.72$), a minimum of 1.25 and a maximum of 5. The skewness of the distribution is 0.03 and the kurtosis is -0.48. The perceived identifiability score for the Mask group was $M = 3.05$ ($SD = 0.75$) and for the No Mask group was $M = 3.19$ ($SD = 0.68$).

Hypothesis test. As per our pre-registration (<https://doi.org/10.17605/OSF.IO/7642E>), we conducted an equivalence test in the form of a TOST on R to test whether the number of questions answered by participants was equivalent in both groups. We defined equivalence by an effect size smaller than $d = 0.2$ (which corresponds to 5.22 questions answered out of 106). The result of the TOST was significant and indicates that we can reject the presence of a difference of $d = 0.2$ or larger, $t(5306) = -5.59$, $p < 0.001$, 90% CI [-0.0009, 0.09].

As per our pre-registration, we conducted a TOST to test whether the level of perceived identifiability was equivalent in both groups. Again, we defined equivalence by an effect size smaller than $d = 0.2$ (which corresponds to 0.143 point estimates). The result of the TOST was non-significant and indicates that we cannot reject the presence of a difference of $d = 0.2$ or larger, $t(655) = 0.0524$, $p = 0.479$, 90% CI [-0.324, -0.064]. (see Figure 5).

Given our deviation from the pre-registration, we had a power of 64.16% to reject the presence of an effect of $d = 0.2$, but we had a power of 0.8 to reject the presence of an effect of $d = 0.23$, however, the result of the TOST was non-significant for a difference of $d = 0.23$ or larger either, $t(655) = 0.437$, $p = 0.331$, 90% CI [-0.324, -0.064].

Additional test. A non-significant TOST for the perceived identifiability measure only allows us to state that we cannot reject the presence of an effect of a pre-specified size. In order to test for the presence of an effect different from 0, we conducted an exploratory one-sided t-test to test whether perceived identifiability significantly decreased in the Mask group compared to the No Mask group. The difference between the group was significant, $t(655) = -2.414$, $p = .008$, $d = .19$, 95% CI [-Infinity, -0.04].

While the significant TOST for the prosocial behavior measure allows us to reject the presence of an effect of a pre-specified size, it does not allow the inference that the effect is non-zero. Thus, we conducted an exploratory one-sided t-test to test if the slight increase in prosocial behavior in the Mask group compared to the No Mask group was significantly different from zero, but this difference was not significant, $t(5306) = -1.62$, $p = .053$, $d = .04$, 95% CI [-0.01, Infinity].

We further explored to what degree we could narrow the null region of the equivalence test for the prosocial behavior and maintain a p-value below 0.05. The TOST remains significant ($p = .023$) for an effect as small as $d = .01$.

Finally, among the 653 participants who answered the perceived identifiability measure, the level of perceived identifiability was not correlated to the measure of prosocial behavior, $r = -0.008$, $p = .83$, 95% CI [-0.08, 0.07]. We provide a wide range of alternative analyses in the supplementary materials.

Discussion

Again, individuals wearing a face mask and individuals not wearing one had equivalent levels of prosocial behavior. Importantly, individuals wearing face masks do, in fact, have lower levels of perceived identifiability. This lower level of perceived identifiability, however, was not correlated to the level of prosocial behavior.

General Discussion

The present research showed that individuals wearing a face mask did not differ in terms of their level of prosocial behavior from individuals without a mask, even though they had lower levels of perceived identifiability. The large sample size allows us to claim that face masks do not decrease prosocial behavior with almost no risk of type II error. These findings are inconsistent with previous research that found that reduced perceived identifiability results in reduced prosocial behavior. Although our study differs in many ways from Zhong et al. (2010), it is noteworthy that the two studies they conducted were considerably underpowered to detect typical social psychological effects. Recent meta-scientific work (Lovakov & Agadullina, 2021) indicates that the median effect size in the field of social psychology

overall is $d = .36$. The $N = 50$ and $N = 83$ for the study 2 and study 3 respectively by Zhong et al. (2010) would have provided only 23.87% and 36.73% power to detect an effect of $d = .36$. Underpowered studies may lead to largely overestimated effect sizes (Fraleley & Vazire, 2014) which would explain the effect sizes of $d = .57$ and $d = 0.6$ detected by Zhong et al. (2010). Given that our findings are based on high powered studies, we feel it more realistic to conclude at this point that perceived identifiability is unlikely to have an impact on prosocial behavior.

When considering the implications of our results, one should keep in mind that the design of our two studies does not allow us to test causal claims and there could be confounding variables that had an impact on the relationship between face masks and helping behavior. For example, participants that were wearing a mask might have been more altruistic in nature than those not wearing a mask, which then canceled out the effect of the first group being less identifiable (i.e., a potential decreased helping behavior). A second limitation was that our research assistants were always wearing face masks themselves. It is plausible that they were perceived as being part of an in-group with mask wearing participants which would, in turn, increase the level of prosocial behavior (Hackel, Zaki, & Van Bavel, 2017) and thus undermined the potential negative effect of perceived identifiability on prosocial behavior. Note, however, that the measure of perceived identifiability was not related to the amount of helping in Study 2, suggesting it to be unlikely that wearing a face mask has a diminishing impact on prosocial behavior via decreased identifiability.

A third limitation of our conclusions concerns our measure of perceived identifiability itself. The explicit nature of the items may have induced our participants to focus their attention on the degree to which they are identifiable resulting in the observation that face masks decrease the level of perceived identifiability. It's possible that if they are not explicitly asked about it, individuals wearing face masks would not necessarily feel less identifiable. However, given that neither our first study (without the perceived identifiability questions) nor our second study (with the perceived identifiability questions) provide evidence for an effect on prosocial behavior, our data provide no reason to believe that a potential question-behavior effect on perceived identifiability would impact prosocial behavior.

Regardless of these limitations, our two studies have strong significance from theoretical and practical perspectives. Taken together, our observations contribute to the cumulative knowledge in the prosocial literature as they provide strong evidence against the idea that perceived identifiability is related to prosocial behavior. Although potential confounds could have influenced our findings, the null result despite our nearly absolute statistical power points to the colossal sample sizes that would be required to conduct randomized controlled laboratory tests of these third variables. Nonetheless, our findings provide two positive implications. First, while our results show no decrease in self-interest behavior when wearing a face mask, we also provide strong evidence that individuals not wearing one do not behave more selfishly either. This, indirectly, points to the need to focus on alternative reasons why individuals don't wear face masks. The second positive implication of our study should bring comfort to researchers in the field of psychological science. Notwithstanding the arduousness of finding participants for studies in

ordinary times, researchers need not worry about additional potential hardship caused by face masks during the pandemic. As our results indicate, individuals agree to take part in studies all the same.

Study 1 Method

Data collection and sampling strategy. We used a correlational observational design. Seven research assistants (five female and two male assistants) individually approached participants in public areas outside in Austria from the 19th of January 2021 to the 5th of February 2021 (average temperature = 1.59°C, SD = 4.34). The research assistants presented themselves as university students who needed data to complete a course and earn credits. They were instructed to only approach solitary adults and to approach approximately the same number of individuals wearing a mask and not wearing a mask. The research assistants were always wearing a mask.

The sample size was determined a priori by a power analysis conducted on R Studio (R Studio Team, 2020) and the pwr package (Champely, Ekstrom, Dalgaard et al., 2018) for a one-sided t-test, an α error probability of 0.05, a Power ($1-\beta$ error probability) of 0.95 to observe an effect size of $d = 0.363$. We expected to observe an effect of this size by averaging the findings by Zhong et al. (2010), Kimmerle et al. (2007) and Cress and Kimmerle (2008). The study by Vesely and Klöckner (2018) was not considered for the power analysis because they confounded observability and identifiability. The minimum sample size, based on this power analysis was $N = 330$.

A total number of 415 participants were approached. Our first two questions assessed age and occupation. All individuals who were approached and refused to answer any question were coded as having conducted the lowest level of prosocial behavior (see measures section), however, given that they did not answer any questions (including the first two) we could not assess the age and occupation for these participants (see Figure 1). Age and occupation were only assessed for those who answered at least the first two questions, of which, four participants were below 18 years old and were excluded from the analysis. Thirteen participants changed their mask status during the interaction. In our main analysis, we also excluded these participants, but we also conducted alternative analyses by including these observations (see supplementary materials).

The final sample size consisted of $N = 398$ participants (206 females), of which $n = 96$ participants answered at least the first question ($M_{\text{age}} = 36.79$, $SD_{\text{age}} = 16.50$; 58 females). A sensitivity analysis showed that our actual sample size of $N = 398$ would have had a Power of 0.975 to detect the predicted effect of $d = 0.363$, and that the smallest effect size detectable with a Power of 0.95 would have been of $d = 0.33$.

Measures. At the time of data collection, wearing a mask was mandatory inside closed public places (e.g., malls, shops, offices) but not outside, and participants were only approached outside. Therefore, all participants who wore a mask chose to do so.

To measure prosocial behavior, the research assistants asked participants to take part in a survey to help them earn credits for a university assignment. We used the German long version of the HEXACO survey (Lee & Ashton, 2018). The survey was selected for its length and because of the repetitive and monotonous aspect of the questions. Any other similar survey could have been used, but an existing standardized survey such as the HEXACO increases the realistic aspect of the study. The participants were given the survey in a 14-pages paper format and a pen. The research assistants emphasized that answering the questions was completely voluntary and that participants were free to answer the number of questions they wanted, however participants were informed that the more information they could provide the better. Participants were also reminded at the bottom of each page that they could stop if they wanted. Answering the questions was a costly behavior due to its length, monotonous aspect, and the discomfort of answering to 14-pages paper format survey while standing outside in the cold weather. But importantly, answering resulted in no material gain of any sort. Agreeing to do so was for the benefit of the research assistant, and the degree to which it would benefit the research assistant was determined by the number of questions. Prosocial behavior was then measured by the total number of questions the participants answered.

The original survey had 100 questions and we included two questions about the age and occupation on the first page. Consequently, our measure of prosocial behavior ranged between 0 (answering no question) to 102 (answering all the questions). We chose to include participants who did not answer any question in order to assess the full range of the behavioral response, as not answering any question may be interpreted as not being prosocial at all in that situation and answering all the questions may be interpreted as being completely prosocial.

Given the predicted standardized effect size of $d = 0.363$, we predicted that participants in the Mask group would answer, on average, at least 12.8 fewer questions than participants in the No Mask group (i.e., 12.54% difference).

Study 2 Method

Data collection and sampling strategy. The sampling procedure was the same as with the first study, with a few differences. This time we had four research assistants (2 male and 2 female assistants). The study was conducted from the 24th of March 2021 to the 18th of July 2021 (average temperature = 10.31, SD = 5.22).

As with our first study, we considered participants who refused to answer any question and they were coded as having conducted the lowest level of prosocial behavior. Consequently, and given that our measure of perceived identifiability was in the form of questions (see measures section), we had fewer participants for this latter measure than for the prosocial behavior measure (i.e., participants who did not answer the perceived identifiability questions were still counted in the prosocial behavior measure). For this reason, and in order to have enough power for both measures, we conducted an a priori power analysis on R Studio (R Studio Team, 2020) with the TOSTER package (Lakens, 2017) to estimate the

minimum sample size to detect equivalence for the perceived identifiability measure only, and aimed to meet this number. With this approach, the sample size for the prosocial behavior would naturally be the same size or larger. As indicated in our pre-registration (<https://doi.org/10.17605/OSF.IO/7642E>), the minimum sample size we wished to obtain for the perceived identifiability measure was $N=858$ which would have provided a Power ($1-\beta$ err probability) of 0.80 to detect an effect of $d = 0.2$ should there be one. However, finding enough participants proved to be more complicated than we expected.

For feasibility reasons, we had to deviate from the pre-registration and terminate the data collection with fewer participants after 133 days of data collection. After excluding 17 participants who were younger than 18 years and 85 participants who changed their mask status during the interaction, the sample size for the perceived identifiability measure consisted of $N = 657$ ($n_{\text{Females}} = 385$; $M_{\text{age}} = 32.08$, $SD_{\text{age}}=14.26$), see Figure 4. This sample provided a power of 80% to reject the presence of an effect size smaller than $d = 0.23$ (the power to reject the presence of an effect size smaller than $d = 0.2$ is then of 64.16%). The sample size for the prosocial behavior measure was $N = 5308$ (Females = 2939). This sample provided a power of nearly 100% to reject the presence of an effect size larger than $d = 0.2$, and the smallest detectable effect with a power of 0.80 was $d = 0.076$. It follows that we deviated from the pre-registration in regard to the sample size because the costs of collecting more data outweighed the benefits. Given that the effect size of $d = 0.23$ for the perceived identifiability measure is very close to our desired $d = 0.2$, and given the nearly absolute power for the prosocial behavior measure, we consider our study to provide informative observations nonetheless.

Measures. Our addition to the method used in the first study was the measure of perceived identifiability. We included four items at the beginning of the survey that came after the questions regarding age and occupation. The questions were the German translation of the following:

1. "I feel anonymous right now."
2. "I feel identifiable by others right now."
3. "At this very moment I feel like nobody around me knows who I am."
4. "If someone who knows me was walking near me at this very moment, they would be able to identify me."

These four items were at the beginning to increase the likelihood that they would be answered by participants. In order to have a similar format as the HEXACO items, participants had to report on a 5-level scale to what level they agreed with each of the previous sentences (i.e., 1 = strongly disagree, 2 = disagree, 3 = neutral, 4 = agree, 5 = strongly agree). The items 2 and 4 and the reversed items 1 and 3 were averaged into an overall score of perceived identifiability. The four items had a Cronbach's α of .61 and a McDonald's ω of .66. Given the four additional questions, our measure of prosocial behavior ranged between 0 (answering no question) to 106 (answering all the questions).

Declarations

Data Availability Statement.

The datasets generated and analyzed during the current studies are available in the OSF repository: <https://doi.org/10.17605/OSF.IO/CRZM3>.

References

1. Adams, R. B., Ambady, N., Macrae, C. N., & Kleck, R. E. Emotional expressions forecast approach-avoidance behavior. *Motivation and Emotion*, 30(2), 177–186. <https://doi.org/10.1007/s11031-006-9020-2> (2006).
2. Allen, G. L., Peterson, M. A., & Rhodes, G. Seeking a Common Gestalt Approach to the Perception of Faces, Objects, and Scenes. *The American Journal of Psychology*, 119(2), 311. <https://doi.org/10.2307/20445341>(2006).
3. Andreoni, J., Rao, J. M., & Trachtman, H. Avoiding the Ask: A Field Experiment on Altruism, Empathy, and Charitable Giving. *Journal of Political Economy*, 125(3), 625–653. <https://doi.org/10.1086/691703> (2017).
4. Barrick, E. M., Thornton, M. A., & Tamir, D. I. Mask exposure during COVID-19 changes emotional face processing. *PLOS ONE*, 16(10), e0258470. <https://doi.org/10.1371/journal.pone.0258470> (2021).
5. Barton, J. J. S., Radcliffe, N., Cherkasova, M. V., Edelman, J., & Intriligator, J. M. Information Processing during Face Recognition: The Effects of Familiarity, Inversion, and Morphing on Scanning Fixations. *Perception*, 35(8), 1089–1105. <https://doi.org/10.1068/p5547> (2006).
6. Batson, C. D., Ahmad, N., & Stocks, E. L. Four forms of prosocial motivation: Egoism, altruism, collectivism, and principlism. In *Social motivation* (pp. 103–126). Psychology Press. (2011).
7. Batson, C. D., & Powell, A. A. Altruism and Prosocial Behavior. In *Handbook of Psychology* (pp. 463–484). American Cancer Society. <https://doi.org/10.1002/0471264385.wei0519> (2003).
8. Beltran, D. G., Isch, C., Ayers, J. D., Alcock, J., Brinkworth, J. F., Cronk, L., Hurmuz-Sklias, H., Tidball, K. G., Horn, A. V., Todd, P. M., & Aktipis, A. Mask wearing behavior across routine and leisure activities during COVID-19. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/2qya8> (2021).
9. Berman, J. Z., & Small, D. A. Self-Interest Without Selfishness: The Hedonic Benefit of Imposed Self-Interest. *Psychological Science*, 23(10), 1193–1199. <https://doi.org/10.1177/0956797612441222> (2012).
10. Bohns, V. K. (Mis)Understanding Our Influence Over Others: A Review of the Underestimation-of-Compliance Effect. *Current Directions in Psychological Science*, 25(2), 119–123. <https://doi.org/10.1177/0963721415628011> (2016).
11. Brown, S. L., & Brown, R. M. Connecting prosocial behavior to improved physical health: Contributions from the neurobiology of parenting. *Neuroscience & Biobehavioral Reviews*, 55, 1–17. <https://doi.org/10.1016/j.neubiorev.2015.04.004> (2015).

12. Cabral, L., Ozbay, E. Y., & Schotter, A. Intrinsic and instrumental reciprocity: An experimental study. *Games and Economic Behavior*, 87, 100–121. (2014). <https://doi.org/10.1016/j.geb.2014.05.001>
13. Champely, S., Ekstrom, C., Dalgaard, P., Gill, J., Weibelzahl, S., Anandkumar, A., Ford, C., Volcic, R., De Rosario, H., & De Rosario, M. H. Package 'pwr.' R Package Version, 1(2). (2018).
14. Christakis, N. A. (2021, October 20). *Sometimes Altruism Needs to Be Enforced*. The Atlantic. <https://www.theatlantic.com/ideas/archive/2021/10/pandemic-altruism-selflessness-punishment/620427/> (2021).
15. Cialdini, R. B., & Goldstein, N. J. Social Influence: Compliance and Conformity. *Annual Review of Psychology*, 55(1), 591–621. <https://doi.org/10.1146/annurev.psych.55.090902.142015> (2004).
16. Crandall, C., & Bahns, A. How Much Do Masks Affect Social Interaction? Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/q3jus> (2021).
17. Cress, U., & Kimmerle, J. Endowment heterogeneity and identifiability in the information-exchange dilemma. *Computers in Human Behavior*, 24(3), 862–874. <https://doi.org/10.1016/j.chb.2007.02.022> (2008).
18. Dana, J., Cain, D. M., & Dawes, R. M. What you don't know won't hurt me: Costly (but quiet) exit in dictator games. *Organizational Behavior and Human Decision Processes*, 100(2), 193–201. <https://doi.org/10.1016/j.obhdp.2005.10.001> (2006).
19. Dienes, Z. Obtaining Evidence for No Effect. *Collabra: Psychology*, 7(1). <https://doi.org/10.1525/collabra.28202> (2021).
20. Emmons, R. A. The Psychology of Gratitude: An Introduction. In *The psychology of gratitude* (pp. 3–16). Oxford University Press. <https://doi.org/10.1093/acprof:oso/9780195150100.003.0001> (2004).
21. Fatfouta, R., & Trope, Y. Keeping one's distance: Mask wearing is implicitly associated with psychological distance. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/jpsw5> (2021).
22. Grahlow, M., Rupp, C., & Derntl, B. The impact of face masks on emotion recognition performance and perception of threat. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/6msz8> (2021).
23. Hackel, L. M., Zaki, J., & Van Bavel, J. J. Social identity shapes social valuation: Evidence from prosocial behavior and vicarious reward. *Social Cognitive and Affective Neuroscience*, 12(8), 1219–1228. <https://doi.org/10.1093/scan/nsx045> (2017).
24. Haxby, J. V., Hoffman, E. A., & Gobbini, M. I. The distributed human neural system for face perception. *Trends in Cognitive Sciences*, 4(6), 223–233. [https://doi.org/10.1016/S1364-6613\(00\)01482-0](https://doi.org/10.1016/S1364-6613(00)01482-0) (2000).
25. Huang, L., Zhu, Y., & Zhang, D. Low Altruism as a Cause of Ostracism. *Frontiers in Psychology*, 12, 2438. <https://doi.org/10.3389/fpsyg.2021.653217> (2021).
26. Kelleher, O. (2021, December 16). *Judge criticises 'absurd selfishness' of woman with eight convictions for not wearing mask*. The Irish Times. <https://www.irishtimes.com/news/crime-and-law/courts/district-court/judge-criticises-absurd-selfishness-of-woman-with-eight-convictions-for-not-wearing-mask-1.4757365> (2021)

27. Kimmerle, J., Cress, U., & Hesse, F. W. An interactional perspective on group awareness: Alleviating the information-exchange dilemma (for everybody?). *International Journal of Human-Computer Studies*, 65(11), 899–910. <https://doi.org/10.1016/j.ijhcs.2007.06.002> (2007).
28. Lakens, D. Equivalence Tests: A Practical Primer for t Tests, Correlations, and Meta-Analyses. *Social Psychological and Personality Science*, 8(4), 355–362. <https://doi.org/10.1177/1948550617697177> (2017).
29. Lakens, D., Scheel, A. M., & Isager, P. M. Equivalence Testing for Psychological Research: A Tutorial. *Advances in Methods and Practices in Psychological Science*, 1(2), 259–269. <https://doi.org/10.1177/2515245918770963> (2018).
30. Lee, K., & Ashton, M. C. Psychometric Properties of the HEXACO-100. *Assessment*, 25(5), 543–556. <https://doi.org/10.1177/1073191116659134> (2018).
31. Lin, S. C., Schaumberg, R. L., & Reich, T. Sidestepping the rock and the hard place: The private avoidance of prosocial requests. *Journal of Experimental Social Psychology*, 64, 35–40. <https://doi.org/10.1016/j.jesp.2016.01.011> (2016).
32. Lovakov, A., & Agadullina, E. R. Empirically derived guidelines for effect size interpretation in social psychology. *European Journal of Social Psychology*, 51(3), 485–504. <https://doi.org/10.1002/ejsp.2752> (2021).
33. Lumley, T., Diehr, P., Emerson, S., & Chen, L. The Importance of the Normality Assumption in Large Public Health Data Sets. *Annual Review of Public Health*, 23(1), 151–169. <https://doi.org/10.1146/annurev.publhealth.23.100901.140546> (2002).
34. Open Science Collaboration. Estimating the reproducibility of psychological science. *Science*, 349(6251), aac4716–aac4716. <https://doi.org/10.1126/science.aac4716> (2015).
35. Penner, L. A., Dovidio, J. F., Piliavin, J. A., & Schroeder, D. A. Prosocial Behavior: Multilevel Perspectives. *Annual Review of Psychology*, 56(1), 365–392. <https://doi.org/10.1146/annurev.psych.56.091103.070141> (2005).
36. Poulin, M. J., & Holman, E. A. Helping hands, healthy body? Oxytocin receptor gene and prosocial behavior interact to buffer the association between stress and physical health. *Hormones and Behavior*, 63(3), 510–517. <https://doi.org/10.1016/j.yhbeh.2013.01.004> (2013).
37. Rahal, R.-M., & Fiedler, S. Cognitive and Affective Processes of Prosociality. *Current Opinion in Psychology*. <https://doi.org/10.1016/j.copsyc.2021.10.007> (2021).
38. RStudio Team. RStudio | Open source & professional software for data science teams. RStudio: Integrated Development for R. RStudio, PBC, Boston, MA URL <http://www.rstudio.com/> (2020).
39. Schlenker, B. R., & Leary, M. R. Audiences' reactions to self-enhancing, self-denigrating, and accurate self-presentations. *Journal of Experimental Social Psychology*, 18(1), 89–104. [https://doi.org/10.1016/0022-1031\(82\)90083-X](https://doi.org/10.1016/0022-1031(82)90083-X) (1982).
40. Shehu, H. A., Browne, W. N., & Eisenbarth, H. WHO CAN READ YOUR FACIAL EXPRESSION? A Comparison of Humans and Machine Learning Classifiers Detecting Emotion from Faces of People

with Sunglasses and Facemasks. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/mnsyx> (2021).

41. Stajduhar, A., Ganel, T., Avidan, G., Rosenbaum, R. S., & Freud, E. Face Masks Disrupt Holistic Processing and Face Perception in School-Age Children. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/fygjq> (2021).
42. Vesely, S., & Klöckner, C. A. How anonymity and norms influence costly support for environmental causes. *Journal of Environmental Psychology*, 58, 27–30. <https://doi.org/10.1016/j.jenvp.2018.07.013> (2018).
43. Wang, M., Wang, J., Deng, X., & Chen, W. Why are empathic children more liked by peers? The mediating roles of prosocial and aggressive behaviors. *Personality and Individual Differences*, 144, 19–23. <https://doi.org/10.1016/j.paid.2019.02.029> (2019).
44. Weber, R., & Popova, L. Testing Equivalence in Communication Research: Theory and Application. *Communication Methods and Measures*, 6(3), 190–213. <https://doi.org/10.1080/19312458.2012.703834> (2012).
45. Williams, W. C., Haque, E., Mai, B., & Venkatraman, V. Face masks influence how facial expressions are perceived: A drift-diffusion model of emotion judgments. Preprint at PsyArXiv. <https://doi.org/10.31234/osf.io/a8yxf> (2021).
46. World Health Organization. When and how to use masks. <https://www.who.int/emergencies/diseases/novel-coronavirus-2019/advice-for-public/when-and-how-to-use-masks> (2020).
47. Zhao, X., & Knobel, P. Face mask wearing during the COVID-19 pandemic: Comparing perceptions in China and three European countries. *Translational Behavioral Medicine*, ibab043. <https://doi.org/10.1093/tbm/ibab043> (2021).
48. Zhong, C.-B., Bohns, V. K., & Gino, F. Good lamps are the best police: Darkness increases dishonesty and self-interested behavior. *Psychological Science*, 21(3), 311–314. <https://doi.org/10.1177/0956797609360754> (2010).

Figures

Figure 1

Alluvial diagram of the structure of our sample for study 1 **Note**. This diagram represents the data collection process for the final sample we used in our main analysis. This diagram excludes participants who changed their mask status during the interaction with the research assistant. RA = Research Assistant.

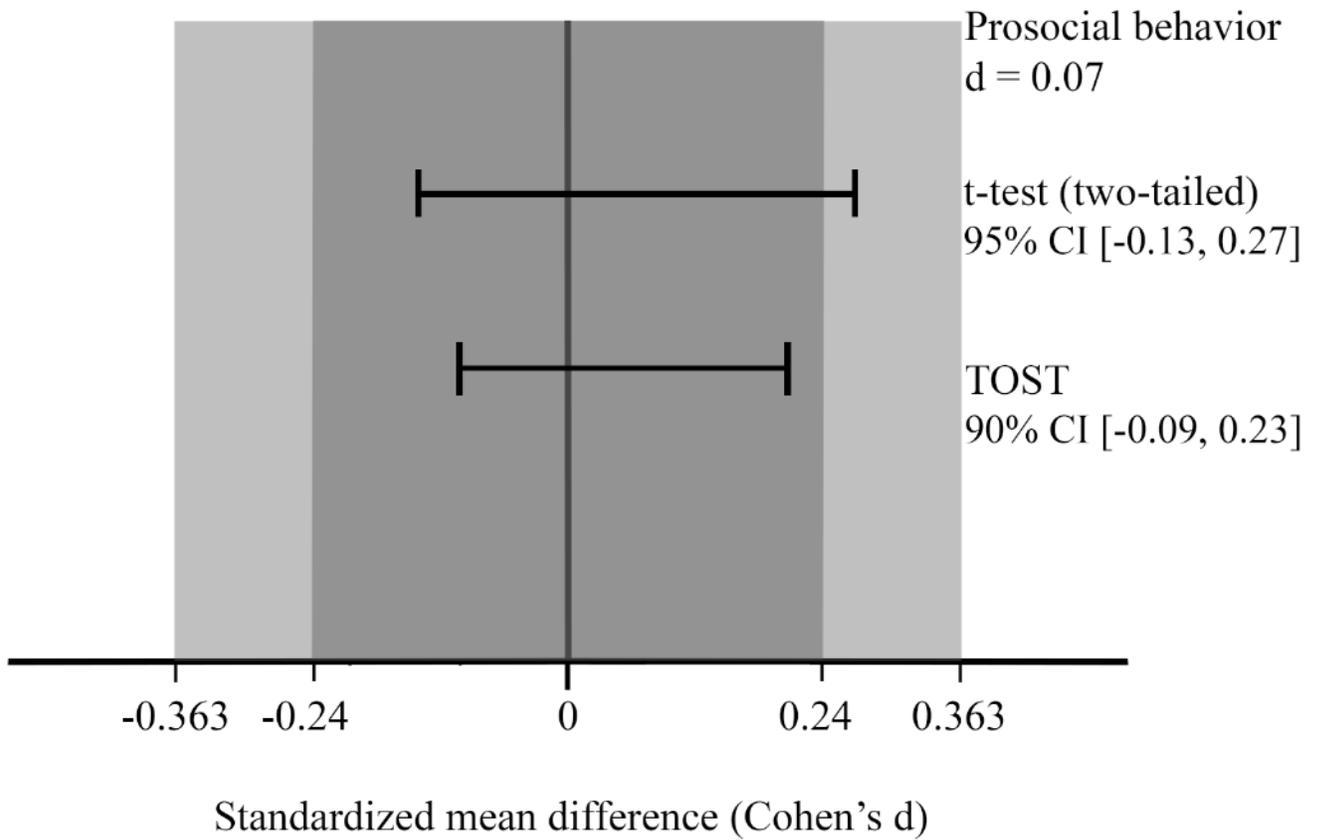


Figure 2

Raincloud plots of prosocial measure scores for the two groups.

Figure 3

Standardized mean differences for the 'Mask' group and the 'No Mask' group in Study 1 **Note.** The t-test indicates that the difference between the two groups is non-significantly different from 0. The TOST indicates that the difference between the two groups is significantly smaller than $d = 0.363$ and smaller than $d = 0.24$.

Figure 4

Alluvial diagram of the structure of our sample for study 2 **Note.** This diagram represents the data collection process for the final sample we used in our main analysis. This diagram excludes participants who changed their mask status during the interaction with the research assistant. RA = Research Assistant.

Figure 5

Standardized mean differences for the 'Mask' group and the 'No Mask' group in Study 2 **Note.** The t-test for perceived identifiability indicates that the difference between the two groups is significantly different from 0. The TOST for perceived identifiability indicates that the difference between the two groups is not significantly smaller than $d = 0.2$, nor is it smaller than $d = 0.23$. The t-test for prosocial behavior indicates that the difference between the two groups is non-significantly different from 0. The TOST indicates that the difference between the two groups is significantly smaller than $d = 0.2$.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [Supplementaryfilerobustnesschecks.pdf](#)