# Supplemental Material
# A: Data Management

OVERALL

In total, after accounting for the respondents who started, but did not complete the survey, there was 3.69% missing data (there was overlap between missingness on demographics, contact, norms, intergroup anxiety, and social distance questions).

After using listwise deletion, there were 787 respondents in the final sample.

Incompletion

At Time 1, 3,403 participants began the survey. Of those, 16.37% did not complete the survey – meaning they started the survey, but skipped one or more major sections of the survey.

About 6% of respondents did not complete any demographic information.
Just over 10% didn't complete any demographic or social distance information.

After removing these participants, we were left with 2,846 participants at Time 1.

At Time 2, 1,913 participants began the survey. Of those, 20.18% did not complete the survey – meaning they started the survey, but skipped one or more major sections of the survey.

About 10% didn't complete any demographic, social distance, or social norms information.

About 10% of respondents did not complete any demographic information or norms information.

Since these participants did not complete the focal variables, it was difficult to examine potential missing data mechanisms. Because using imputation or full-information maximum likelihood when data is MNAR or MCAR can lead to biased estimates, we chose to be conservative and use listwise deletion (Allison, 2002; Rubin, 1976)

After removing these participants, we were left with 1,527 participants at Time 2.

Attrition

We examined the factors that predicted participants' likelihood of completing the study at Time 1, but not at Time 2 (i.e., attrition). First, we assessed if any of the independent variables, as reported at Time 1, were predictive of completing the follow-up survey. Indeed, women were more likely than men, LGBTQ individuals were more likely than non-LGBTQ individuals, those with a stable childhood socioeconomic status (SES) were more likely to complete than those who did not, and first-generation students were more likely than second generation+ students.

Importantly, some of the independent variables that were predictive of survey completion at Time 2, are also predictive of the levels of the dependent and mediating variables at Time 1. For example, gender, LGBTQ status, and to a lesser extent, childhood SES are predictive of desired social distance at Time 1. Similarly, gender, LGBTQ status, childhood SES, and parental education are predictive of intergroup anxiety. Finally, gender and LGBTQ status are predictive of contact at Time 1.

Therefore, when examining if these dependent and mediating variables (contact, intergroup anxiety, and desired social distance) are predictive of responding to both surveys, we include interaction terms for those independent variables that are predictive of both the mediating/dependent variables and missingness at Time 2. With these interaction terms included, we examine if completing the follow-up survey is predictive of the dependent and mediating variables' values at Time 1.

Importantly, the completion of the follow-up study is not predictive of desired social distance at Time 1. Said otherwise, there is no difference in T1 desired social distance, intergroup anxiety, or contact between those who did and did not complete the survey (when including interaction terms for confounding variables). These findings large satisfy the missing at random assumption (Allison, 2002; Rubin, 1976).

We then removed all people who responded at Time 2, but not at Time 1, this was about 47% of respondents. Therefore, we were left with a total of 814 participants.

Other Missing Data

Below, we discuss the way we addressed missing data using listwise deletion. We chose to use listwise deletion and proration for simplicity; however, because proration can result in biased estimates (Mazza et al., 2015), we include the regression table that includes results with full-information maximum likelihood, below (FIML).

*Social Distance*

Time 1: Of those who completed both surveys, ~4% were missing one or more items on the 10-point social distance scale item measures. We ended up using a composite measure for social distance – averaging individuals scores across items. Therefore, if they were missing less than 3 items, they were still included in the sample. Respondents missing more than 3 items were excluded. This was a total of 4 respondents, or less than 1% of the data (respondents who completed the study).

Time 2: Of those who completed both surveys 3.94% were missing one or more items on the 10-point social distance scale item measures. We ended up using a composite measure for social distance – averaging individuals scores across items. Therefore, if respondents were missing less than 3 items, they were still included in the sample. Respondents missing more than 3 items were excluded. This was a total of 7 respondents, or less than 1% of the data (out of respondents who completed the study).

*In-group/Out-group*

When T1 and T2 data was combined, we used complete data to address missing data on past history of mental illness.

If a respondent said they had never had a mental illness at Time 1 & Time 2, they were categorized as not having had a mental illness.

If they did not respond at Time 1, but at Time 2 said they had never had a mental illness, they were categorized as not ever having had a mental illness at Time 1 as well.

If they said they had a mental illness at Time 1 but not at Time 2 (14 participants), we replaced their Time 2 data with their Time 1 response. We made this decision because, the question asks if respondents ever had a mental illness, so if they had it at the first time point, we would assume they also had it at the second point. Importantly, this did not change results.

If they responded at Time 1 but not at Time 2, we replaced their missing Time 2 responses with their Time 1 values (2 participants).

Less than 1% of respondents did not answer at both Time 1 and Time 2 and were dropped from the sample.

*Contact*

There were several indicators of contact with mental illness.

Of those who completed Time 1 and Time 2, 2.46% were missing on the question "how many people have you known with mental illness", however, we decided not to use this measure because there was no option for "0".

Instead, to mirror close contact, we used a network-based measure. For this, respondents listed friends and then were asked if the friends had a mental illness or not. Respondents who said "don't know" were counted as not having that particular friend with mental illness. After accounting for participants who didn't complete the study (see Incomplete Data), there was no remaining missing data.

*Norms*

Norms questions were only asked at Time 2. We use three items to assess perceived descriptive norms on campus. About 4.5% were missing on at least one question.

Any respondent who was missing more than one question was removed from the sample, this was 1.72% of respondents.

*Intergroup Anxiety*

Intergroup anxiety was asked at Time 1 and 2. About were missing and 0.86% and 1.47% were missing on intergroup anxiety questions at Time 1 and Time 2, respectively.

Participants were removed if they were missing more than one of four questions, this was 2 participants, less than 1% of the data.

*Demographics*

To account for missing data on demographics at T1 or T2, we used the T1 data to replace missing demographic data at T2 and vice-versa. After replacing missing data from T1 with data at T2 and vice-versa, only 0.49% of respondents had remaining missing demographic data. Therefore, we removed these participants from the final sample.

Between T1 and T2, some respondents changed their responses to characteristics that would be seemingly stable, e.g., racial/ethnic identification and childhood socioeconomic status. We examined these respondents' other demographic data and the responses were consistent across time points. Additionally, their other responses seemed high quality. Therefore, we kept them in the sample.

*Errata:*

One person listed their T1 age as 1717. We assume this individual meant 17, since, at Time 2, they responded 19. We replaced their T1 value to be 17.

Another person listed their T1 age as 12 but their T2 age as 20. We assume that there was a typo at T1 and imputed age as 18 for this respondent.

Table SA1**. Structural Equation Model using Full Information Maximum Likelihood (N=814)**

| | Direct | | Indirect | | Total | |
|---|---|---|---|---|---|---|
| Δ Contact | | | | | | |
| Perception of Improved Norms | 0.378 | ** | | | 0.378 | ** |
| Δ Intergroup Anxiety | | | | | | |
| Δ Contact | -0.055 | *** | | | -0.055 | *** |
| Perception of Improved Norms | 0.000 | | -.021 | ** | -0.021 | ** |
| Δ Desired Social Distance | | | | | | |
| Δ Contact | -0.010 | * | -.006 | *** | -0.017 | *** |
| Δ Intergroup Anxiety | 0.115 | *** | | | 0.115 | *** |
| Perception of Improved Norms | -0.137 | *** | -.006 | ** | -0.143 | *** |
| T1: Desired Social Distance | -0.658 | *** | | | -0.658 | *** |
| *Demographic Covariates*[1] | | | | | | |
| Δ Mental Health Diagnosis | -0.051 | ** | | | -0.051 | ** |
| Gender: Woman | -0.053 | ** | | | -0.053 | ** |
| Age in Years | 0.003 | | | | 0.003 | |
| White | -0.043 | | | | -0.043 | |
| International Student | -0.004 | | | | -0.004 | |
| First Generation Student | 0.004 | | | | 0.004 | |
| Childhood Financial Stability | 0.008 | | | | 0.008 | |
| Δ Gender or Sexual Orientation | -0.050 | | | | -0.050 | |

Notes—.Fit Statistics: $X^2$ (19)=69.057 p<0.001; CFI=0.933; RMSEA=0.057 pclose=0.194; TLI=0.884; SRMR=0.033. †p<0.1, *p<0.05, **p<0.01, ***p<0.001, two-tailed tests; [1]Covariates represent overall effect on social distance, measured at Time 1 with the exception for change in mental health diagnosis and change in LGBTQ+ identification, which represent differences between T1 and T2

For the FIML model, all scales were estimated by first using confirmatory factor analysis and FIML, then the difference measures were created, then the final model was estimated using structural equation model with FIML.