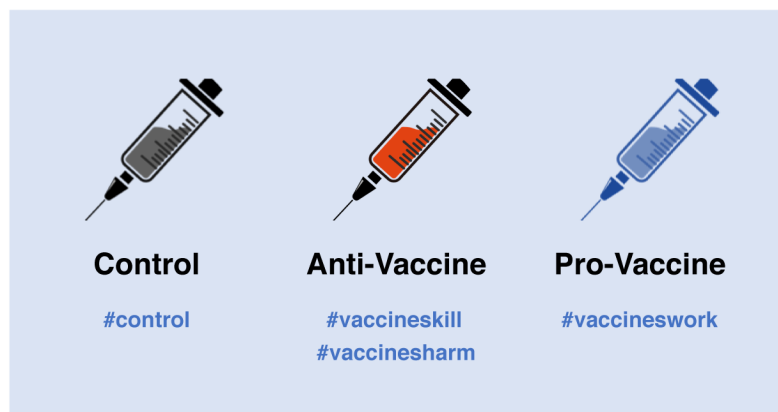


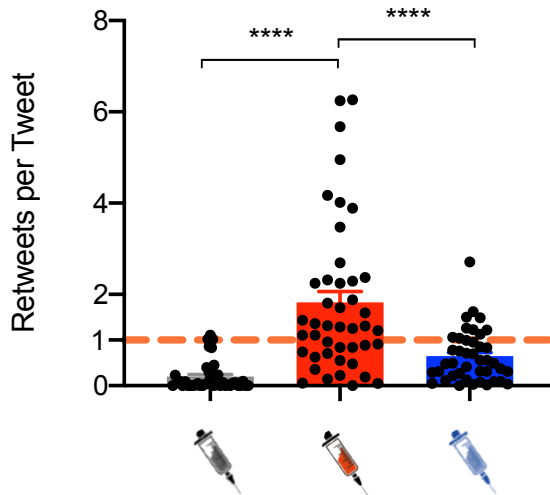
Supplementary Material

A revised behavioral analysis of the late 2020 anti-vaccination infodemic on Twitter

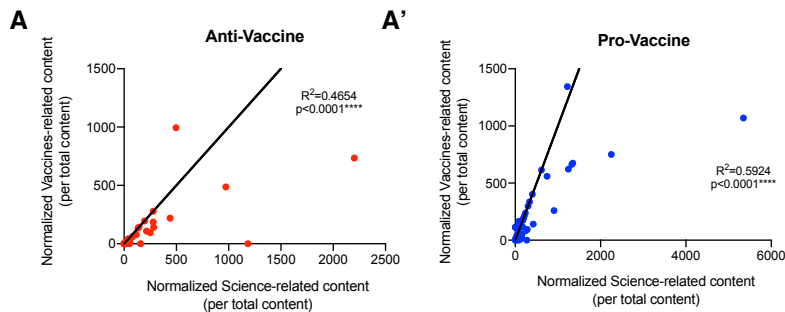
Supplementary Figures



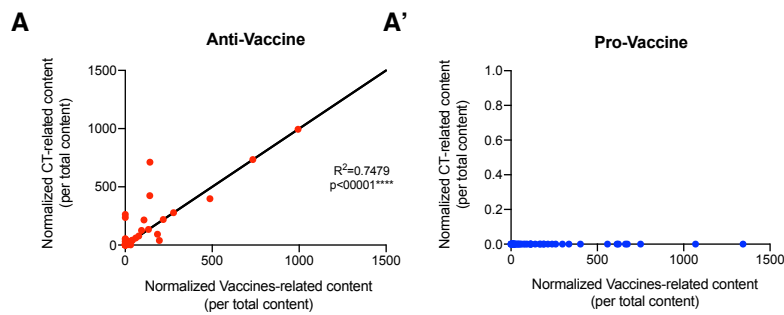
Supplementary Figure 1. Definition of the study groups and relative hashtags. We classified profiles in three groups: control (grey), anti-Vaccine (red) and pro-Vaccine (blue). Profiles (n=50 for each group) were identified automatically through the use of hashtags and the Twitter search function. Control profiles were selected for their use of randomly selected hashtags, anti-vaccination profiles for their use of widely chosen hashtags in the community (*#vaccineskill* and *#vaccinesharm*), whereas Pro-vaccine profiles were selected for their use of the *#vaccineswork* hashtag.



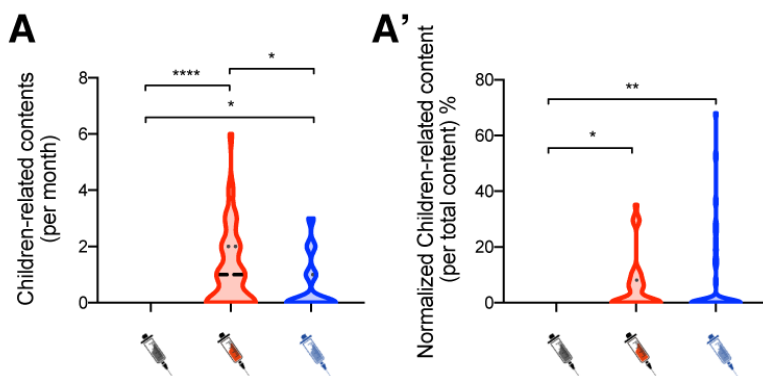
Supplementary Figure 2. Anti-vaccination profiles retweet more than they tweet. Profiles belonging to the control and pro-vaccination groups tweet more than they retweet (1 indicates an equal number of retweets and tweet on average in a month), whereas anti-vaccination profiles retweet more than they tweet. Ordinary one-way ANOVA; **** $p < 0.0001$; Outliers were excluded with ROUT, $Q = 0.1\%$; $n = 50$.



Supplementary Figure 3. The number of vaccines- and science-related contents shared by anti- and pro-vaccination profiles are correlated. For both the anti-vaccination group (red) (A) and the pro-vaccination group (blue) (A'), the higher the normalized number of science-related contents generated or shared (for the overall number of contents generated on any given topic), the larger the number of normalized vaccines-related tweets and retweets ($R^2 = 0.464$ and $R^2 = 0.5924$ respectively; **** $p < 0.0001$; Outliers were excluded with ROUT, $Q = 0.1\%$; $n = 50$).



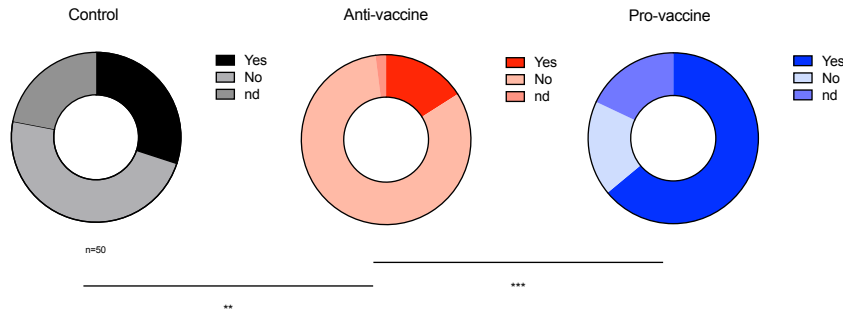
Supplementary Figure 4. Positive correlation between the number of normalized vaccines-related contents and conspiracy theories-associated for the anti-vaccination group. For both the anti-vaccination group (red), the higher the normalized number of vaccines-related contents generated or shared (for the overall number of contents generated on any given topic), the larger the number of normalized conspiracy theory (CT)-related tweets and retweets ($R^2=0.7479$; $****p<0.0001$) (A). For the pro-vaccination group, no correlation exists between the normalized number of vaccines-related tweets and the normalized number of tweets and retweets including CTs (B). Outliers were excluded with ROUT, $Q=0.1\%$; $n=50$.



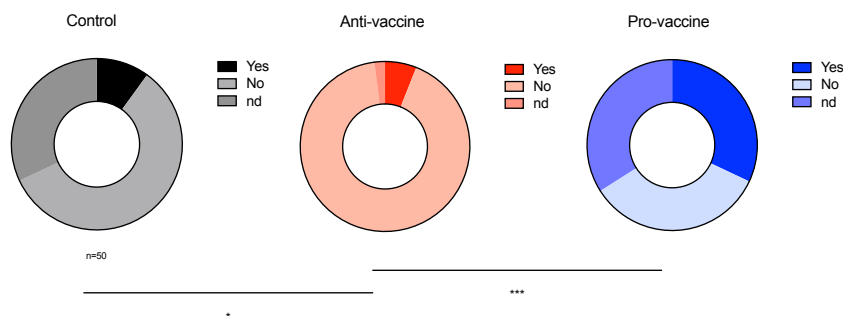
Supplementary Figure 5. Both anti- and pro-vaccination groups share contents associated to children. Anti- (red) and pro-vaccination (blue) profiles share children-related contents, with the anti-vaccination group being the largest net producer of children-related contents on Twitter (A). We calculated the number of children-related content (tweets and retweets) published in the 24 hours before data analysis and normalized it for the total number of tweets published on average during a single day. 100 percent indicates that all generated contents are estimated to be children-related. Natural fluctuations above 100 percent are due to the variation

between the activity on Twitter during the 24 hours prior to data analysis compared with an average day (A'). Ordinary one-way ANOVA; * $p < 0.05$; ** $p < 0.01$; **** $p < 0.0001$; Outliers were excluded with ROUT, $Q = 0.1\%$; $n = 50$.

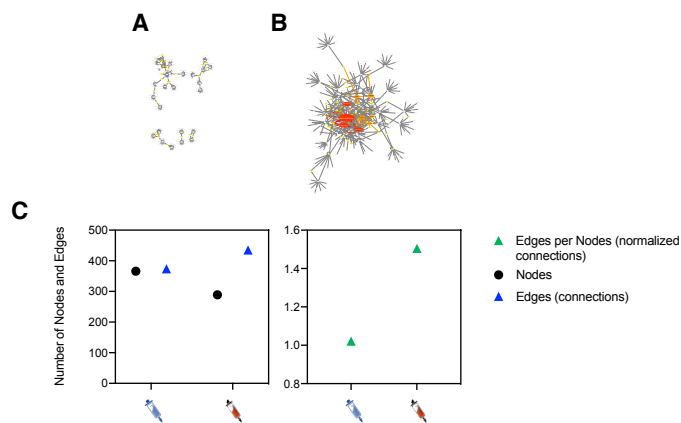
A Can the profile be tracked?



B Is the education and/or profession declared?



Supplementary Figure 6. Anti-vaccination profiles are less prone to declare their identity, education or profession when compared with control and pro-vaccination profiles. 30% of control profiles (grey shades) declare their identity (name and surname, and a profile picture depicting a real person). In comparison, pro-vaccination profiles (blue shades) are more likely to declare their identity (64%) and only 16% of anti-vaccination profiles (red shades) declare their identity (A). 10% of control profiles declare either their education level or current profession. This percentage increases substantially for the pro-vaccination group (32%) and drops further for the anti-vaccination group (6%) (B). Profiles are defined as trackable when users publicly release their name, surname and a valid profile picture. Profiles are not defined as trackable when they fail to meet one of the aforementioned parameters. nd (not defined) indicates the above-mentioned criteria are not applicable (for instance, in the case of institutions without a verified badge on Twitter). This approach was also used for determining whether users declare their education level or profession. Chi-square test; * $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$; $n = 50$.



Supplementary Figure 7. Anti-vaccination profiles are better connected with each other and establish a community, when compared with the pro-vaccination group. The pro-vaccination (**A**) and anti-vaccination (**B**) Twitter webs, scaled 1:1. Yellow colour represents Twitter profiles (nodes) with 2 to 4 anti-vaccination profiles preferentially retweeting their contents within the top 10 most retweeted users (edges; $2 \leq E \leq 4$; $n=42$). Orange nodes represent profiles with 5 to 9 edges ($5 \leq E \leq 9$; $n=42$), whereas red nodes indicate profiles with more than 10 connecting edges ($E \geq 10$; $n=42$). Size of the nodes is linearly scaled depending on the number of edges connecting the node (**A**, **B**). Number of nodes and edges for anti- (blue syringe) and pro-vaccination groups (red syringe). The anti-vaccination group has more edges than nodes, when compared with the pro-vaccination group. The number of edges per node is higher in the anti-vaccination web, when compared with the pro-vaccination web (**C**). Graphical representation and web parameters were generated with Cytoscape.