Emotion Regulation Contagion

Michael Pinus[1,2,3]

Eran Halperin[3]

Yajun Cao[1,2]

Alin Coman[4]

James J. Gross[5]

Amit Goldenberg[1,2,6]

[1] Harvard Business School

[2] Harvard University, Digital, Data and Design Institute

[3] Department of Psychology, Hebrew University of Jerusalem

[4] Department of Psychology, Princeton University

[5] Department of Psychology, Stanford University

[6] Department of Psychology, Harvard University

Abstract

In intergroup conflicts, emotion regulation interventions can decrease negative intergroup emotions and increase support for concessions. However, it is usually infeasible to provide emotion regulation interventions to everyone in a population of interest. This raises a "spill over" question about the relationship between the proportion of individuals who are treated with an emotion regulation intervention and its effectiveness at the whole-group level. To address this fundamental question, we had groups of six Israeli participants (N = 2,659) share real-time responses to anger-inducing, conflict-related stimuli. Before interacting with each other, we treated different proportions of each group with an emotion regulation intervention called cognitive reappraisal. This intervention involved teaching selected participants to change the interpretation of events to reduce negative emotions. Results indicated an exponential relationship between the proportion of treated participants and group reduction in negative emotions. Furthermore, targeting between 25%-40% of participants resulted in group emotional change. Using language analysis, we validated contagion in semantic content between treated and non-treated participants. These findings shed light on the potential for emotion regulation contagion to reduce groups' emotions, and more broadly, suggest the value of investigating the contagion of psychological interventions within groups.

Emotions are a central driver of human behavior, but they by no means always lead to desired outcomes. In many situations, achieving long-term individual and group goals requires controlling emotions via emotion regulation, defined by the activation of a goal to change the emotional trajectory [1]. To facilitate successful emotion regulation, researchers have developed a variety of emotion regulation interventions. The most prominent of these is a reappraisal intervention, which involves changing how one thinks about a situation with the goal of influencing one's emotional response[1,2]. In reappraisal interventions, participants are taught to generate alternative interpretations of emotional situations. The advantage of reappraisal is that it is cheap, quick, and easy to explain. It also is effective, and seems to consistently help individuals regulate their emotions in a variety of contexts [3−5].

One domain in which reappraisal interventions have proven helpful is intergroup conflicts. Intergroup conflicts are characterized by negative emotions that contribute to hostility and violence[6−8] and have a detrimental impact on the physical and mental wellbeing of millions around the globe [9−11]. Reappraisal interventions have been successfully employed in intergroup conflicts. For example, in the context of the Israeli-Palestinian conflict, teaching Jewish Israelis how to use reappraisal led to a reduction in negative emotions and an increase in support for conciliatory policies toward Palestinians [12,13] . These result were recently conceptually replicated in Columbia[14]. Porat and colleagues have further expanded the use of reappraisal by developing ReApp, an online application that gamifies reappraisal training for Israelis, showing that it reduced negative emotions towards the conflict[15].

But even in the examples above, in which reappraisal interventions were found successful in improving intergroup relations, they exclusively targeted small numbers of individuals. This is a limitation because intergroup emotional processes are group processes in nature, and treating

individuals without taking into consideration the group processes that will later unfold may not achieve the desired outcome. From an applied perspective, it is often impossible to conduct interventions on the whole group, both because of a lack of resources and an inability to access every group member. In light of the goal to influence the overall collective emotion[16], one central question related to employing emotion regulation interventions at scale is whether there is any "spill-over" of effects from treated to non-treated participants, and if so, what the relation might be between the proportion of treated participants and the whole-group impact via emotion regulation contagion.

When conducting an intervention on a proportion of a group, several outcomes are possible. First, it's possible that reappraisal would have a reduced effect on the treated participants, given that other people around them are not regulating. Second, it's possible that reappraisal would lead to a reduction in the treated participants' emotions, but that there would be no emotion regulation contagion to the non-treated group. Third, it's possible that non-treated participants would show increased responses, perhaps because they are increasing their emotions to balance out the reduction in the treated participants[17,18]. Finally, the most hopeful possibility is that reappraisal would spread from the treated to non-treated participants, leading to a reduction in negative emotions both among treated and among non-treated participants. Whether (and to what extent) one or more of these effects are evident might depend on the proportion of people within the group who are treated.

Using the Israeli-Palestinian conflict as a case study, we examined how changing the proportion of participants completing a reappraisal intervention within small groups would relate to negative emotion in both the treated and non-treated participants. This was done using a novel paradigm in which groups of six participants reacted to conflict-related images (by writing text and

rating) in real time, then saw each others' reactions. Prior to being exposed to the pictures and others'

emotional expressions, we manipulated the proportion of participants within the group who

completed a reappraisal intervention before conducting our interactive task to be between zero and six

participants. We then examined how our intervention impacted the group as a whole as well as both

the treated and non-treated participants' negative emotions.

Our first preregistered hypothesis (https://osf.io/d7u4h) was that reappraisal would be

effective for the treated individuals, decreasing their negative emotion. Our second preregistered

hypothesis was that reappraisal would spread from the treated individuals to the non-treated

individuals, but given the exploratory nature of the project, we did not predict the shape of the

relationship between the proportion of treated participants and negative emotion reduction. To assess

the mechanism for change in non-treated participants, we conducted Semantic Projection Analysis [19]

of the text produced by participants to test whether non-treated participants adopted the reappraisal

language used by treated participants.

Results

Israeli participants (N = 2,659) signed in to complete the study from their home computers.

After consenting, participants were assigned to a group of six participants (Figure 1). Before

interacting with each other as part of our emotional dynamics task, a portion of the group (from zero

to six participants) was assigned to a reappraisal intervention while the rest were assigned to a control

observing condition. Participants in the reappraisal condition received instructions that were adapted

from other reappraisal interventions[3] and fitted to the Israeli context (see SI for full text). Participants

were told that reappraisal is based on the insight that there are multiple interpretations for each

situation, and that our emotional responses depend on these interpretations. They were then asked to

practice reappraisal by looking at a picture of an amputee meeting with a doctor, and were then given examples of possible reappraisals to the situation. The observing condition was also used in previous tests of reappraisal interventions, and was found to lead to no significant changes in emotions compared to a passive empty control in which participants were not given any instructions about their emotions[3]. Participants in the observing control condition were instructed to observe their emotions as they naturally unfolded. Similar to the intervention, they were also asked to practice observing their emotions by looking at the same pictures as in the reappraisal intervention. They were also given examples of possible emotional reactions to the situation.

After being assigned to the group of six and to their condition (reappraisal or control), participants completed our emotional dynamics task in their group of six. As mentioned, the proportion of people assigned to the reappraisal condition in each group varied between zero to six. During the task, participants saw pictures depicting either Palestinian resistance to the Israeli occupation or Palestinian violence against Israel. Pictures were mostly of terror attacks, or Palestinian demonstrations against Israel. These pictures were used in previous studies to elicit strong negative emotions among Jewish Israelis, mostly anger and sadness[20]. After viewing each picture, participants were asked to produce a brief text that expressed their emotions *("What comes up for you when you see the picture?"*, see further details in methods). Participants were then able to see the texts of all other participants in real time. They were then asked to rate their emotions in response to the picture on a 1-*neutral* to 10-*very negative* scale and again saw each others' real-time ratings (*"Please rate the degree of negative emotions you feel in response to the picture"*, see full description in Methods). Participants therefore had two points during each trial when they could impact each others' emotions, when observing each others' text and when observing each others' ratings (marked in red in Figure 1).

It is important to note that due to natural dropout in online studies, the size of the group, and the actual proportion of people that went through the reappraisal manipulation, sometimes changed during the task. To counter these variations during the task, we used the actual proportion of reappraisers within each group and in each trial rather than the assigned proportion. We also controlled for group size in all of the following models. After finishing the task, participants completed a survey that tested both manipulation checks such as intention to use reappraisal, and general sentiments towards Palestinians. These measures were designed to test whether there were changes in emotions could be seen when participants know that their ratings will be shown to others, in order to eliminate the possibility that compliance was driving participants' emotions. Using these more general ratings also allowed us to examine whether changes in emotional ratings throughout the task, extended to more general sentiments toward Palestinians.
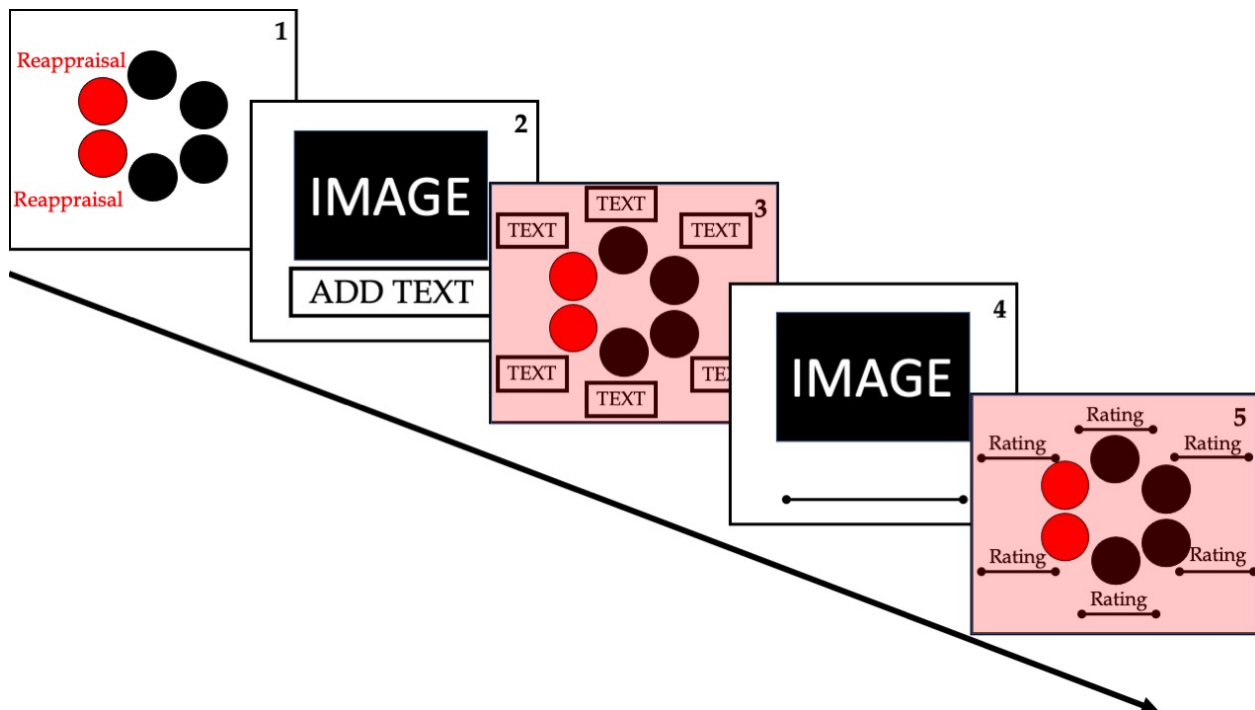
Figure 1. The structure of a trial in the emotion dynamics task (total of 20 trials). (1) Participants were assigned to groups of six and the proportion of participants who completed the reappraisal intervention was predetermined (in the example above, 2 out of 6 marked in red, see 1). (2) They then saw an image related to the conflict and were asked to provide their text to the picture. (3) Participants then saw all the texts produced by everyone in the group. (4) They were then asked to rate their negative emotions to the picture from 1-10, (5) and then saw each others' ratings. There are two steps within the task (steps 3 and 5, which are shaded) in which people see each others' responses in real time.

Before running the actual study, we conducted two pilot studies to validate key aspects for the analysis (see SI for full details). In the first pilot (N = 217), we tested a Hebrew version of a reappraisal intervention[3], examined among individuals and not in group contexts, and found the amount of people it required to show a reduction in negative emotions to conflict-related stimuli. In the second pilot (N = 379, see SI), we compared people's emotions in response to the stimuli either in groups of six – in a similar design to the one described above but with no people going through a reappraisal intervention – or when completing the task without being exposed to emotional responses of other group members. Results suggested that when participants were exposed to the stimuli in groups of six, but without having anyone assigned to the reappraisal intervention, they tended to express stronger emotions compared to when exposed to the stimuli separately, without seeing others' responses. Not only were participants' emotions stronger when seeing the stimuli in a group compared to separately, but their emotions also tended to intensify over trial numbers, suggesting a process of amplification over time. Finally, we examined emotion contagion by looking at changes in the variance of emotional ratings within the group over trial numbers. Results suggested that variance in emotions within the group decreased as the task progressed, providing evidence for contagion.

Was reappraisal effective, even when some people in the group were not treated with reappraisal? We first tested the effect of reappraisal by comparing negative emotion ratings as a function of whether participants were assigned to the reappraisal intervention or control condition, across trials and regulators' proportions. We preregistered three ways to examine the effect of the intervention, and all of our tests were significant (See SI). Here we report the simplest way to examine whether the reappraisal treatment produced the hypothesized effect: a linear mixed model which predicted rating with the regulation/control condition, controlling for the actual proportion and the actual number of participants in the group in each trial. The model included random intercepts for the stimuli, the group, and the individual participants (nested within groups). The effect of the manipulation was statistically significant and negative (β = -.18, 95% Confidence Intervals [-.23, -.12], $t_{(2337.12)}$ = -6.71, p < .001) indicating that participants who were treated with the reappraisal intervention reported less negative emotions than the non-treated participants. We also tested the effect of reappraisal in a group setting by comparing participants' self-reported use of emotion regulation (see methods, $t_{(2335.51)}$= 44.09, p < 0.001, d = 1.88, 95% Confidence Intervals [1.78, 2.13]), and the effort exerted on emotion regulation, (see methods, $t_{(2551.57)}$= 39.26, p < 0.001, d = 1.51, 95% Confidence Intervals [1.45, 1.63]). These effects were statistically significant, strong, and in the expected direction.

To supplement our primary statistical analysis, we also qualitatively evaluated whether participants were producing different content in the two conditions. We generated a word cloud based on the only-control (no reappraisal in the group of six) and only-reappraisal (six participants treated with reappraisal) conditions to qualitatively examine differences in the produced text between the two conditions (Figure 4A). Results show that expressions in the control condition – where words like

"anger", "fear" and "hatred" were the most frequent – differed from expressions in the reappraisal condition – where words like "luck", "hope", "maybe" were more frequent. While these word clouds only provide qualitative results of the produced content, they illuminate the substantial difference in content between the two conditions. We provide further analysis of the text below.

What is the relationship between the proportion of participants who went through the intervention and its effectiveness? We hypothesized that higher proportions of treated participants in the group would lead to greater reduction in negative emotion, both within the treated and control participants in each group. However, we had no specific prediction as to the shape of the reduction effect (i.e., linear or non-linear), so we compared alternative models to find the best approximation of the regulation agents' proportion dosage effect.

Our first preregistered model was a repeated measures model in which we examined whether the proportion of reappraisal within the group (0%-100%) predicted negative emotions. The model included random intercepts for the stimuli, the group, and the individual participants (nested within groups). We fitted models representing different dosage levels in terms of the proportion of treated participants: linear, quadratic, cubic, logarithmic, and exponential. Results suggested that the strongest model was the exponential model (β = -.15, 95% Confidence Intervals [-.18, -.12], $t_{(885.01)}$ = -9.27, p < .001) such that the reduction in emotions became increasingly larger with the increase in reappraisal proportion. We had a similar model with an interaction between proportion and condition which is reported in SI.

An important limitation of the model described above is that it assumes the same dosage effect for both the control and the reappraisal participants. However, looking at the group as a whole may

10

miss important information, as it seems possible that there might be different dynamics for each condition, and that the model would average over these differences. To account for this limitation, our pre-registered analysis plan included performing the model comparison procedure described above for the treated and non-treated participants separately.

For the participants treated with reappraisal, all the models performed with similar AICs, but the model with the best fit was the cubic model, AIC = 83,465.37, which suggested that reduction in emotion was stronger within low and high proportions of reappraisal (see Figure 2A). Within the cubic model, the effect of the proportion of the regulators was significant, (β = -.08, 95% Confidence Intervals [-.12, -.04], $t_{(621.17)}$ = -3.88, p < .001). In a similar manner, in the non-treated subset of participants (i.e., those that were not presented with the reappraisal instructions), the model with the best fit was the quadratic model, AIC =11,082.4 (see Figure 2A). which suggested that there is relatively little change in the emotions of the non-treated participants until a certain proportion, at which point changes become much greater with every increase in the proportion of treated participants. Within the quadratic model, the effect of the proportion of the regulators was again significant, (β = -.06, 95% Confidence Intervals [-.09, -.02], $t_{(1369.4)}$ = -3.46, p < .001).

Overall, our results indicate that the dosage effect of the emotion regulation intervention is exponential at the entire group level. In addition, our results indicate that while the dosage effect might be different for the treated and non-treated participants, both subsamples demonstrated a degree of emotion regulation contagion, and showed non-linear reduction in conflict-related negative emotions.
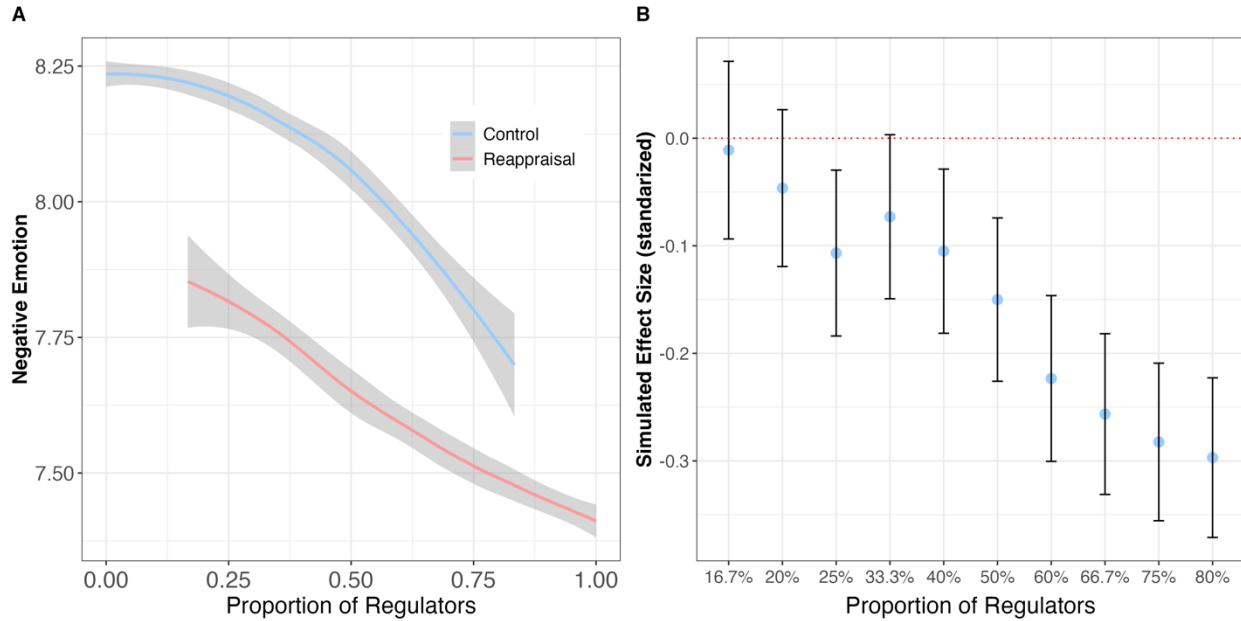
Figure 2. Results from the main Study. Panel A captures the impact of proportion of regulators on negative emotions. The red line represents the participants that were treated with reappraisal within each group of six. The blue line represents the control participants within each group. Grey areas represent standard errors. Results suggest that for the reappraisal condition, the best fitting model was a cubic model, although this model was very similar to others. On the control condition, the best fitting model was the quadratic model. Panel B captures the results of the simulation testing the proportion of regulators needed to reach a significant change within the control conditions. The blue dots and error bars represent the average standardized effects and their 95% critical intervals in 1,000 iterations of simulation. Using simulated data, we made sure that proportion bins included the same number of control users. We then compared the emotions of the groups with only control participants to participants in the control condition in each of the proportions. Results suggest a significant difference, with a reduction of 0.1 sd already at 25% proportion of regulators.

Estimating when the impact becomes significant for the control condition. One important question is what proportion of treated participants is required in a group for the non-treated participants in that group to be influenced by the reappraisal intervention. It is impossible to answer this question without data imputation with simulations, because in the raw data there are different numbers of reappraisal and non-treated participants in each proportion. The unbalanced number of reappraisal and non-treated participants in each proportion create unequal variances which make it

difficult to compare effects. To equalize sample sizes for each proportion, we binned the data based on actual proportions of treated participants (see methods for detailed description). We then kept proportion bins that had more than 20 non-treated participants, and simulated the missing data to have 243 control participants, which was the largest number of controls in each proportion. Imputation was done first by creating new groups for each proportion. Group size was determined based on the group size distribution within the task (see methods). We then populated these groups by simulating data based on the ratings of the participants in each proportion. The result of each simulation was groups of 243-245 controls that were equal in size for each proportion (numbers varied because of differences in group sizes). In order to test whether increasing the proportion of reappraisers led to a significant reduction in the ratings of the non-treated participants, we compared the ratings of the non-treated in the baseline condition (0% reappraisal) to those of the non-treated in the different proportions of reappraisal. To make sure that our results were not driven by a specific simulation, we repeated the process 1000 times, each time comparing the control only to all other conditions. In figure 4b we reported the standardized differences for these 1000 comparisons, with 95% confidence intervals. Results showed a reduction in controls' rating which became significant with a reduction of 0.1 SD in ratings already at 25% (Figure 3B), and that treating between 25-40% of the group with a reappraisal intervention results in significant and reliable emotion contagion at the group level. With the largest proportion (80%) in the sample, non-treated participants' negative emotions were reduced by nearly 0.3 *SD*. While this very much depends on the size of the bins, it provides a sense of comparison for future studies.

Impact on general sentiments towards Palestinians. After completing the task, participants were

asked to complete a survey of more general sentiments towards Palestinians. We use the term

sentiments, inspired by Frijda's conceptualization[21], because they do not reflect emotional responses to

a specific situation but rather more general feelings towards Palestinians. Unlike the ratings during the

task, participants knew that their ratings to these sentiment questions were not going to be shown to

others. This served as a good opportunity to examine changes in emotion without the peer pressure of

having others view their ratings. We reasoned that finding differences in the expression of sentiments

towards Palestinians would be another indication that genuine contagion occurred.

Participants were asked to rate nine negative sentiments towards Palestinians (e.g., *"Generally*

speaking, when you think about the Palestinians in the Palestinian territories, to what extent do you

feel fear towards them?"). Each item was rated on a 1-6 scale (1 - not at all, 6 – very much). We created

a negative emotional attitudes scale using the relevant emotional items (fear, anger, hatred, disgust,

α=.81). We report results on positive emotions and guilt in supplementary information.

To examine changes in sentiments towards Palestinians following the task we averaged both

the proportion of treated participants and the group size across all trials. We then conducted a mixed

model interaction between the proportion of participants treated with reappraisal in the same group

and the condition assigned to the specific participant predicting the different sentiments. For the

proportion of treated participants, we used an exponential model, as this was suggested to be the best

fitting model for the interaction, but results were similar with a cubic or linear model (see SI). Our

model also included a random intercept of group as well as condition nested within group, as

participants were nested within different groups. Looking first at negative emotions and exploring the

main effects, results suggested that increasing the proportion of participants treated with reappraisal

led to reduction in negative sentiments for both treated and non-treated participants (β =-.13, 95%

Confidence Intervals [-.24, -.04], $t_{(1772.34)}$ = -2.50, p = .01). There was not a significant main effect

between the treated and non-treated participants when ignoring the proportion of treated participants

(β =-.02, 95% Confidence Intervals [-.08, .06], $t_{(1767.21)}$ = -.80, p = .42). However, we did find an

interaction between the proportion of treated participants and condition, such that the relationship

between proportion of treated participants and reduction in negative sentiment was stronger for the

control participants (β =.10, 95% Confidence Intervals [.01, .21], $t_{(1227.51)}$ = 1.97, p = .04).

Overall, these results suggest that the effect of proportion of treated participants on emotional

ratings was also extended to negative general sentiments towards Palestinians. These results are

encouraging because participants provided these ratings knowing that no other participants would see

them. It is therefore another support that the manipulation led to real changes in emotion. It's worth

mentioning that we also measured general attitudes as well as dehumanization towards Palestinians.

Results pointed to significant reduction in dehumanization (which are closely related to negative

sentiments of anger, hate, and contempt) and marginally significant reduction in negative attitudes

towards Palestinians, but as expected these results were weaker than the emotional results (see SI).

Providing evidence for the spread of reappraisal in semantic content. Our results provide

evidence for reduction in emotion as a result of the increased proportion of participants treated with

reappraisal, but we have not yet provided evidence that reduction in ratings within the control

participants is driven by changes in their interpretation of the situations. One way to address this

limitation is to examine changes in the text that participants produced as a function of the proportion

of participants treated with reappraisal. To do this, we utilized a method called Semantic Projection

Analysis [19], which is based on the idea that semantic meanings can be estimated by subtracting one linguistic representation from another. For example, to create a semantic representation of the term Queen, one can take a semantic representation of the term King and subtract the difference between the semantic representation of the term man and woman. Using the same idea, to generate a linguistic representation of a "pure reappraisal" content we can take the content produced by participants who were assigned to exclusively reappraisal groups (i.e., all 6 participants were treated with the reappraisal intervention) and subtract the content produced by participants who were assigned to exclusively control groups (i.e., none of the 6 participants were taught to reappraise). The result is a semantic representation of "pure reappraisal". We can then compare this pure reappraisal representation to the texts that participants produced throughout the task. At this point it is important to acknowledge that some control participants may spontaneously reappraise, and some participants treated with reappraisal may not reappraise (despite their instructions to do so). This means that any findings from this analysis must rise above this noise.

To conduct our Semantic Projection Analysis, we processed the text to derive a 768-dimension embedding vector for each text response, using AlephBERT, a large pre-trained language model for modern Hebrew [22]. Next, we created aggregated baseline vectors for the regulated responses and the control response, by selecting only the responses that were provided by participants who were in groups that were pre-allocated to either 0% or 100% regulators. We then subtracted the control baseline vector from the regulation baseline vector, to derive the "pure reappraisal" vector. We then computed the cosine distance of each individual's text responses in our dataset to the "pure reappraisal" vector, to estimate the usage of reappraisal language in it. Lastly, we fitted a mixed-linear model to predict the usage of reappraisal language as a function of the proportion of reappraisers in

16

each group. More specifically, we conducted a three-way interaction between condition (reappraisal or control), the proportion of reappraisers in the group (0%-100), and the trial number (see Figure 5). Similar to previous models, we used random intercepts for stimuli, group, and individual participant (nested within groups).

Focusing on the main effects, results showed that regardless of the proportion of treated participants, those who were assigned to the reappraisal condition (and thus exposed to the reappraisal intervention) within each group of 6 were more similar to the "pure reappraisal" content than the control condition, (β =.35, 95% Confidence Intervals [.30, .40], $t_{(1220.79)}$ = 13.76, p < .001). This was expected given the assigned conditions and served as a sanity check. Results also indicated a main effect of proportion: increase in the proportion of participants in the group who were taught to reappraise led to an increase in similarity to the "pure reappraisal" semantic representation (Figure 4B, β =.08, 95% Confidence Intervals [.03, .12], $t_{(2060.14)}$ = 3.46, p < .001). Finally, we also found a significant effect of time, such that an increase in trial number led to an increase in similarity to the "pure reappraisal" semantic representation, (β =.07, 95% Confidence Intervals [.06, .08], $t_{(28097.55)}$ = 10.57, p < .001).

Having established these three main effects, we then examined interactions. The only significant interaction was that between condition (control or reappraisal) and trial number (β =-.08, 95% Confidence Intervals [-.10, -.06], $t_{(28098.55)}$ = -7.65, p < .001,), suggesting that the association between trial number and similarity to "pure reappraisal" was stronger for the control condition than the reappraisal condition (Figure 4C). This is a very important finding, as it emphasizes that controls were much more influenced by the reappraisal language as a function of the proportion of reappraisers than the treated participants.

To further investigate the relationship between proportion of treated participants and semantic similarity to "pure reappraisal," we examined the simple effect of each condition separately. Results suggested that increasing the proportion of reappraisal led to a significant increase in similarity to pure reappraisal for the control condition (β =.08, 95% Confidence Intervals [.03, .10], $t_{(951.21)}$ = 3.72, p < .001) and a marginally significant increase in the reappraisal condition (β =.05, 95% Confidence Intervals [.03, .10], $t_{(824.57)}$ = 1.98, p = .05). These results provide evidence for the spread of reappraisal language as a function of increases in the proportion of reappraisers within each group.
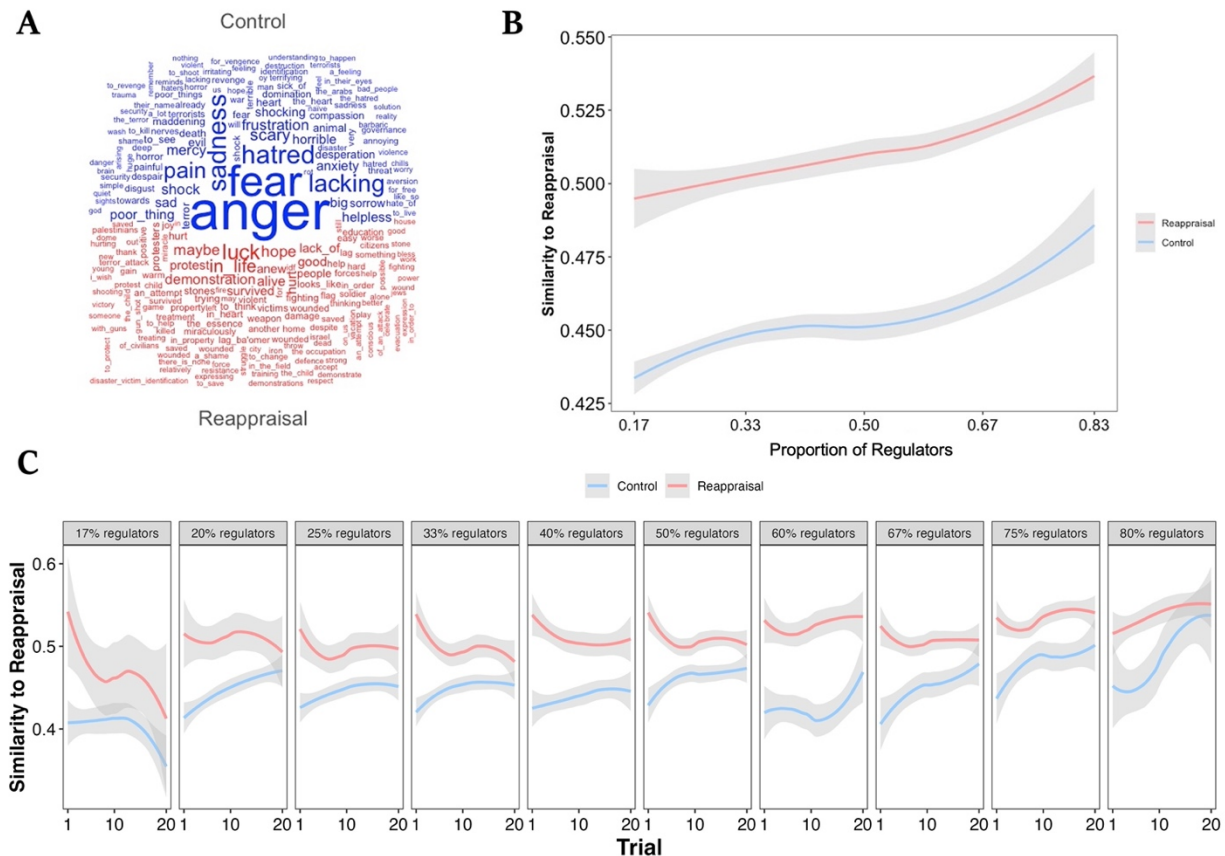
Figure 4. Semantic analysis of the text responses. (A) Word cloud of 1-grams and 2-grams from either the only control condition (no reappraisal in the group of six) and only reappraisal condition (six participants trained in reappraisal). Size of the word reflects frequency of use. (B) Similarity to reappraisal, evaluated using semantic projection analysis, for both the reappraisal and control conditions. The x axis represents the proportion of participants treated with reappraisal in each group

of six. The y axis represents the semantic similarity to the "pure reappraisal" semantic representations. Results suggest that increase in the proportion of participants who were treated within each group of six led to a marginally significant increase in similarity in the reappraisal condition (red line) and to a significant increase in similarity to reappraisal in the control condition. (C) Similarity to reappraisal over time as binned by the proportion of reappraisers in each group, where the x axis represents the trial number (1-20), and the y axis represents the semantic similarity to the "pure reappraisal" semantic representations. Results suggest that when the number of regulators is low (17%) participants become more distant from reappraisal over time. However, as the number of reappraisers increases within the group, we see an increase in similarity to reappraisal language over time.

Discussion

In the current study, we examined how the proportion of participants treated with a reappraisal intervention impacted reduction in negative emotions within the group. More specifically, using intergroup conflict as the context for the investigation, we designed a paradigm in which groups of six participants responded emotionally to conflict-related stimuli. We then tested how increasing the proportion of participants treated with reappraisal impacted the emotions of the control participants.

We found that reappraisal reduced participants' emotions, even if the non-treated participants themselves were not instructed to use reappraisal. We also found that the relationship between the number of treated participants and the reduction in the control participants' emotion was exponential, and that when the proportion of treated participants was between 25% and 40%, there was a reliable difference in emotion ratings (compared to ratings of control participants in groups with only control participants). Analyzing participants' text using Semantic Projection Analysis provided evidence for change in language produced by the control participants as a function of the proportion of treated participants, providing support for linguistic contagion between the treated and non-treated participants.

Emotion regulation interventions designed to improve intergroup relations aim to impact the collective as a whole. Given limited resources and the fact that its rarely possible to access all members of a certain group, influencing individual emotions will not impact the group without processes of emotion regulation contagion. The current project turns the focus towards collective-level outcomes. This new focus on the collective leads to a completely different set of questions. For example: How many people need to be treated with an intervention in order to achieve an overall outcome? Who are the right targets for such interventions? We believe that pondering these questions could improve the utility of emotion regulation intervention but also other, more general psychological interventions. We hope that this project is a significant step in a broader examination of the spread of psychological interventions within groups and collectives.

Limitations and Future Directions

This project has limitations related to the interpretation of the findings and the translation of these findings to useful applications. The first limitation is the concern that changes in ratings throughout the task may have been driven by compliance to other group members rather than real changes in emotion. To help address this concern, we measured participants general sentiments towards Palestinians after the task, when participants knew that their ratings would not be observed by other participants. Finding differences in these reported sentiments as a function of treated participants reduced the chance that results were driven by compliance. Furthermore, even if results were all driven by compliance to other group members, we still think such change in expression could have substantial effect on groups' behaviors. If reappraisals continue to spread in the network regardless the authenticity of whoever created them, they may contribute to an overall reduction in

emotion. Nevertheless, future studies should use other non-self-report methods to examine changes in emotion as a result of regulation contagion.

A second limitation relates to the fact that the size and nature of the groups – as well as the way in which they communicated – are different from typical groups, making it hard to generalize to natural interactions. Groups often vary in terms of size, hierarchy, and network typology, and all of these aspects might influence how emotion regulation spreads. For example, it is likely that targeting central nodes in a network, or people with high power, may change the relationship between the number of treated participants and their impact on those who were not treated. We chose to simplify all of these aspects in order to be able to reduce the noise and examine these contagion effects in a simple and controllable design. However, future studies should not only vary these group features, but should also try to examine the spread of interventions in natural field experiments.

Finally, a third limitation of the current project is that it focuses on reappraisal but did not compare its impact to another emotion regulation strategy. Furthermore, participants in the study were instructed to use the strategy on their own emotions, whereas in many situations people have an explicit goal to influence the emotions of others. Future studies should examine how different emotion regulation strategies and goals influence the relationship between proportion of treated participants and impact on the whole group.

Despite these limitations, we believe that the current project represents an exciting new step for research on emotion regulation interventions and for psychological interventions in general.

Methods

Participants. Our power analysis for the main study was based on a pilot study in which we examined the number of participants required to achieve a significant difference between the reappraisal and

control condition. We conducted a pilot study in which 217 participants were assigned to either a reappraisal intervention or a control condition (see SI for full description). Results suggested that the effect size of the reappraisal intervention was d =.20, which suggested that 200 participants would be needed to detect a difference between the two conditions in terms of negative emotions. Because we did not have a good estimate of the expected size for the effect of the proportion of treated participants, and because we realized that the study may require more sensitivity, we decided to double the estimate and set the planned sample to 400 participants per condition, with total of 2,800 participants (preregistration - https://osf.io/d7u4h).

The study was approved by both the Harvard IRB and Hebrew University. Participants were Jewish Israelis who were recruited through iPanel, an Israeli survey company in exchange for 25 NIS (~$7). The study was conducted in multiple runs during April 2022. In each run, participants were randomly assigned to a group of 6 participants and to one of the 7 conditions, corresponding to the proportion of participants trained with reappraisal. 2,830 participants completed the task and the following survey (see full breakdown from raw data to this number in SI). We applied a few preregistered exclusion criteria to remove participants during the study. First, because the study was conducted synchronously, participants could go through the task without providing any ratings or text responses. We therefore concluded that participants who timed out for more than 2 consecutive trials or for a total of six trials would be eliminated from the data. Second, we removed participants who failed an attention check and who provided nonsensical responses when either asked to describe the instructions or in providing text to the pictures. Third, we removed groups who had technical issues or that ended up with less than 2 participants at the end of the task (See SI for full description). Our

final sample was 2,659 participants (Gender: 1,129 males, 1438 females, 92 other or refused to say;

Age: *M = 42.07, SD =*14.618).

Task. When logging in to the task, participants were told that they were going to do a study in real

time with 5 other participants. Participants were asked to choose a name and were told that other

participants would see that name during the task. Following this stage, participants were forwarded to

a waiting room where there were assigned to a group of six participants. Once six people logged in,

the group was assigned a condition (0-6 people reappraising) and each person in the group was

assigned to either the control or reappraisal condition. Both control and reappraisal conditions were

based on a recent reappraisal intervention that was validated in a large global sample[3], and in a pilot

study as a preparation for the current project (see SI). In the control condition, participants were told

that they would be asked to implement a strategy called *observing*, which involves paying attention to

emotions as they unfold. We chose to use this active control condition – in which participants were

asked to engage with their emotions – because it has been used in previous studies exploring

reappraisal. It is, however, worth mentioning that a recent study that compared observing to a more

passive control condition in which participants were merely instructed to respond did not find

consistent differences between the two conditions [3]. Participants in the reappraisal condition received

instructions that were similar to Wang et al., but were slightly modified to the Israeli context (see SI

for full text). Participants were told that reappraisal is based on the insight that there are multiple

interpretations for each situation, and that our emotional responses depend on these interpretations.

Participants in both conditions then observed a practice image an amputee meeting with a doctor

who is holding a prosthetic limb. They were then asked to respond to the picture and saw example

responses based on the condition. Participants then completed an open ended question in which they

were asked to describe the instructions of the task as well as answered a multiple choice question in which they were asked to select the description of their condition. We removed participants who failed to properly answer both of these questions. Notice that the example given to participants was very different from the pictures in the actual study. This was done to avoid linguistic copying as much as possible from the practice stage to the actual task.

Finally, before the start of the task, participants were shown two pictures. For one picture (a picture of a truck) participants were told that the average negative emotion that the picture elicited was one, for the second (a picture of a child corpse) participants were told that the average rating was 10. The reason we added these descriptions is because in initial piloting we found that participants' responses to the pictures were almost at ceiling and we wanted to reduce the average emotional rating.

The actual task was 20 trials long. In each trial, participants were synchronously presented with a picture related to the Israeli-Palestinian conflict. Pictures were chosen from a sample of pictures used in a previous study because they elicited strong negative emotions, especially anger, among Jewish Israelis [23]. Each picture contained a location, to clarify where they were taken. While observing the picture, participants were told: "*Try to use the method you learned, observing/reappraisal, and to express your emotional response to this picture. The response should be short (one sentence). What comes up for you when you see the picture?*" Participants were asked to enter a text to the picture and had up to 35 seconds to do so. Following this stage, participants were forwarded to a window in which they saw the name of the person responding (names were selected by participants in the beginning of the task) and their text in response to the picture. Participants were able to observe each others' responses for 15 seconds. Following this stage, participants were asked to rate their negative emotions

in response to the picture on a scale of 1-no negative emotion to 10- very strong negative emotion.

Participants had 25 seconds to rate their emotional responses and following this stage they again saw

each user's name and their rating to the picture for another 15 seconds. Participants completed 20

trials of the task. After 10 trials, participants received a reminder of their instructions for the task

(either reappraisal or control). Following the task, participants completed a few questions about

demographics as well as a few exploratory surveys (see SI) that were mainly designed to examine

potential mediators and were not preregistered in the analysis.

Measures. In addition to the measures taken in the task which are described above, participants also

completed a survey following the task. The survey included three parts. The first part measures

aspects related to emotion regulation. The second measured participants' sentiments towards

Palestinians. The third measured attitudes towards Palestinians (described and reported in SI).

Emotion Regulation. Participants were first asked about their emotion regulation attempts using

a three item scale that was adopted from questions used in a previous reappraisal intervention[24](α =

.87): *"to what extent (if any) did you try to control your emotions while watching the pictures", "To*

what extent did you try to reduce negative emotions that came up while watching the pictures", "while

watching the pictures, how much effort did you put to regulate your emotions?"). Responses were

rated on a scale of 1-not at all, to 6-very much so. Participants also rated the degree to which they used

reappraisal in the task using a four item scale that was adapted from the same intervention reported

above[24] (α = .84): *"While watching the pictures I tried to change their meaning.", "While watching the*

pictures I tried to give them a more positive meaning", "While watching the pictures I tried to

understand why people do what they do.", "While watching the pictures I tried to give them a new

meaning." Responses were rated on a scale of 1-not at all, to 6-very much so. In addition to these two

measures participants also completed the Emotion Regulation Questionnaire (ERQ)[25]

General Nagative Sentiments Towards Palestinians. In addition to the measures in the task, we

examined participants' general sentiments towards Palestinians. We measured four negative emotions

(fear, anger, hatred, disgust, α=.81), and additional 4 positive emotions and guilt (reported fully in SI).

For each emotion, participants were asked: *"Generally speaking, when you think about the*

Palestinians in the Palestinian territories, to what extent do you feel [emotion] towards them?". Each

emotion was rated on a 1-6 scale (1-not at all, 6 – very much). We created a negative emotions scale

(fear, anger, hatred, disgust) by averaging all of the negative emotion items.

Simulation. We conducted a simulation analysis to estimate when the impact of reappraisal becomes

significant for the control participants. The goal of the simulation is to mitigate the statistical bias

caused by the unequal amount of control participants under different conditions. In our experiment

sample, conditions with lower control proportions had fewer observations of control participants'

ratings. This would lead to higher variances in statistical estimates when we compared the reappraisal

effect for control participants across different proportion conditions. To make the reappraisal effects in

all conditions comparable, it is necessary to make sure that these conditions have an equal number of

observations for control participants.

Data generation. We populated the control participant sample by creating new groups and

simulated control participants' ratings in each group. We first needed to decide on the group size of

each simulated group as the group sizes in the actual experiment varied across groups due to

dropouts. We sampled the group size data from a normal distribution with the mean and standard deviation of the group sizes of all existing groups in the original sample (M = 3.72, SD = 1.10). We randomly generated numbers from this distribution and rounded them to the closest integers as the new group sizes. For numbers smaller than one (larger than six), we forced them to be one (six). After the group size was decided, we calculated the number of control individuals in each new group by multiplying the group size by the proportion of control participants.

Next, for each control participant, we simulated 20 trials of negative emotion ratings which was consistent with the experiment trial number. The simulated ratings in each proportion condition were generated from a distribution of ratings of the corresponding proportion in the experiment. This rating distribution was estimated first as a normal distribution based on the mean and standard deviation of ratings in each proportion condition. We then squeezed the range of the distribution to [1, 10] and rounded each number drawn from the distribution to the closest integer as the rating.

We kept generating new groups for every proportion condition until the total number of control participants (both simulated and original numbers) reached a target number (243 participants per proportion condition). The target number was determined by the largest number of control participants among all proportion conditions in the original experiment. An iteration was completed when the control participant numbers in all proportion conditions were equal to or larger than the target number. Note that we excluded one proportion condition (83.3% treatment) from all simulation processes because it only had 4 control participants in the original data. We were not able to estimate its distribution of ratings due to the very limited sample size. The result of each iteration was groups of 243-245 control participants.

Text processing. We processed text in the task using AlephBert, which is a large pre-trained language model in Hebrew[22]. Prior processing the text via AlephBert, we removed punctuation and capital letters from the text, replaced symbols with words, removed double spaces and made sure that there were spaces after commas and periods. We then used AlephBert to generate embeddings for each text that was produced by each participant in each trial. To generate the "pure reappraisal" vector we averaged all vectors of the all reappraisal condition (all six participants in the reappraisal condition). Similar process was done to the control condition: we averaged all the responses in the all control condition (all six participants in the control condition). We then subtracted the control ratings from the reappraisal, which produced a vector representation of "pure reappraisal". We then compared the pure reappraisal vector to each of the participants' responses using cosine similarity.

References

1. Gross, J. J. Emotion regulation: Current status and future prospects. *Psychological Inquiry* **26**, 1–26 (2015).

2. Gross, J. J. Antecedent- and response-focused emotion regulation: Divergent consequences for experience, expression, and physiology. *Journal of Personality and Social Psychology* **74**, 224–237 (1998).

3. Wang, K. *et al.* A multi-country test of brief reappraisal interventions on emotions during the COVID-19 pandemic. *Nat Hum Behav* **5**, 1089–1110 (2021).

4. Liu, J. J. W., Ein, N., Gervasio, J. & Vickers, K. The efficacy of stress reappraisal interventions on stress responsivity: A meta-analysis and systematic review of existing evidence. *PLoS One* **14**, e0212854 (2019).

5. Preuss, H., Capito, K., van Eickels, R. L., Zemp, M. & Kolar, D. R. Cognitive reappraisal and self-compassion as emotion regulation strategies for parents during COVID-19: An online randomized controlled trial. *Internet Interventions* **24**, 100388 (2021).

6. Bar-Tal, D., Halperin, E. & De Rivera, J. Collective emotions in conflict situations: Societal implications. *Journal of Social Issues* **63**, 441–460 (2007).

7. Halperin, E. *Emotions in conflict : Inhibitors and facilitators of peace making*. (Routledge, 2016).

8. Goldenberg, A., Halperin, E., van Zomeren, M. & Gross, J. J. The process model of group-based emotion: Integrating intergroupe emotion and emotion regulation perspectives. *Personality and Social Psychology Review* **20**, 118–141 (2016).

9. Davies, S., Pettersson, T. & Öberg, M. Organized violence 1989–2021 and drone warfare. *Journal of Peace Research* **59**, 593–610 (2022).

10. Institute for Economics & Peace. *Economic Value of Peace 2021: Measuring the global economic impact of violence and conflict*. http://visionofhumanity.org/resources (2021).

11. Wenzel, T., Schouler-Ocak, M. & Stompe, T. Editorial: Long Term Impact of War, Civil War and Persecution in Civilian Populations. *Front. Psychiatry* **12**, 733493 (2021).

12. Halperin, E., Porat, R., Tamir, M. & Gross, J. J. Can Emotion Regulation Change Political Attitudes in Intractable Conflicts? From the Laboratory to the Field. *Psychol Sci* **24**, 106–111 (2013).

13. Halperin, E., Pliskin, R., Saguy, T., Liberman, V. & Gross, J. J. Emotion regulation and the cultivation of political tolerance: Searching for a new track for intervention. *Journal of Conflict Resolution* **58**, 1110–1138 (2013).

14. Hurtado-Parrado, C. *et al.* Emotion Regulation and Attitudes Toward Conflict in Colombia: Effects of Reappraisal Training on Negative Emotions and Support for Conciliatory and Aggressive Statements. *Frontiers in Psychology* **10**, (2019).

15. Porat, R., Erel, L., Pnueli, V. & Halperin, E. Developing ReApp: an emotion regulation mobile intervention for intergroup conflict. *Cognition and Emotion* **34**, 1326–1342 (2020).

16. Goldenberg, A., Garcia, D., Halperin, E. & Gross, J. J. Collective emotions. *Current Directions in Psychological Science* **29**, 154–160 (2020).

17. Goldenberg, A., Saguy, T. & Halperin, E. How group-based emotions are shaped by collective emotions: evidence for emotional transfer and emotional burden. *J Pers Soc Psychol* **107**, 581–596 (2014).

18. Goldenberg, A., Enav, Y., Halperin, E., Saguy, T. & Gross, J. J. Emotional compensation in parents. *Journal of Experimental Social Psychology* **69**, 6–11 (2017).

19. Grand, G., Blank, I. A., Pereira, F. & Fedorenko, E. Semantic projection recovers rich human knowledge of multiple object features from word embeddings. *Nat Hum Behav* **6**, 975–987 (2022).

20. Pliskin, R., Halperin, E., Bar-Tal, D. & Sheppes, G. When Ideology Meets Conflict-Related Content: Influences on Emotion Generation and Regulation. *Emotion* **18**, 159–170 (2018).

21. Frijda, N. H. *The emotions*. (Cambridge University Press, 1986).

22. Seker, A. *et al.* AlephBERT:A Hebrew Large Pre-Trained Language Model to Start-off your

    Hebrew NLP Application With. Preprint at https://doi.org/10.48550/arXiv.2104.04052

    (2021).

23. Pliskin, R., Bar-Tal, D., Sheppes, G. & Halperin, E. Are leftists more emotion-driven than

    rightists? The interactive influence of ideology and emotions on support for policies.

    Personality & social psychology bulletin **40**, 1681–97 (2014).

24. Wang, K. *et al.* A global test of brief reappraisal interventions on emotions during the

    COVID-19 pandemic. *Nature Human Behaviour* (2020) doi:10.31234/osf.io/m4gpq.

25. Gross, J. J. & John, O. P. Individual differences in two emotion regulation processes:

    Implications for affect, relationships, and well-being. *Journal of personality and social*

    psychology **85**, 348–362 (2003).