Cosmopolitan morality trades off in-group for the world, separating benefits and protection

Xuechunzi Bai[a,b], Varun Gauri[a], and Susan T. Fiske[a,b,1]

[a]Princeton School of Public and International Affairs, Princeton University, Princeton, NJ 08540; and [b]Department of Psychology, Princeton University, Princeton, NJ 08540

Contributed by Susan T. Fiske, July 1, 2021 (sent for review January 17, 2021; reviewed by Ronnie Janoff-Bulman, David G. Rand, and John F. Dovidio)

Global cooperation rests on popular endorsement of cosmopolitan values—putting all humanity equal to or ahead of conationals. Despite being comparative judgments that may trade off, even sacrifice, the in-group's interests for the rest of the world, moral cosmopolitanism finds support in large, nationally representative surveys from Spain, the United Kingdom, Germany, China, Japan, the United States, Colombia, and Guatemala. A series of studies probe this trading off of the in-group's interests against the world's interests. Respondents everywhere distinguish preventing harm to foreign citizens, which almost all support, from redistributing resources, which only about half support. These two dimensions of moral cosmopolitanism, equitable security (preventing harm) and equitable benefits (redistributing resources), predict attitudes toward contested international policies, actual charitable donations, and preferences for mask and vaccine allocations in the COVID-19 response. The dimensions do not reflect several demographic variables and only weakly reflect political ideology. Moral cosmopolitanism also differs from related psychological constructs such as group identity. Finally, to understand the underlying thought structures, natural language processing reveals cognitive associations underlying moral cosmopolitanism (e.g., *world*, *both*) versus the alternative, parochial moral mindset (e.g., *USA*, *first*). Making these global or local terms accessible introduces an effective intervention that at least temporarily leads more people to behave like moral cosmopolitans.

cosmopolitanism | intergroup morality | cognitive processes | public policy

Morality beyond borders seems to challenge basic human impulses. In the midst of the pandemic, each country had to decide how much to trade off their own nation's interests against foreign nations' interests, under conditions of limited resources. Similarly, other contemporary global challenges—climate change, migration, global poverty, and trade policy—require explicit comparative moral judgments. This paper investigates moral cosmopolitanism: the psychological puzzle of overcoming a comparative preference for one's own nation, relative to the world as a whole, in the context of an explicit tradeoff. This paper examines which people are willing to sacrifice the in-group's interests to help the world.

Tribal Loyalty versus Universality

Studies of intergroup relations and moral decisions generate an obvious prediction, but closer examination reveals an alternative. In-group favoritism and parochial morality suggest people would support helping their nation, an easy case to make. Nevertheless, out-group neutrality and universal morality hint that at least some people prioritize helping the world.

On one hand, the human tendency toward tribal loyalty seems utterly obvious and makes moral cosmopolitanism seem impossible. Much morality research emphasizes just one's own group, construing morality mostly as an in-group problem. Theories focus on morality as an intragroup phenomenon whose primary evolutionary purpose is to reconcile the individual's self- interest with the group's well-being (1, 2). When other groups do join the consideration set, they often take collateral damage from strong in-group morality. For instance, parochial altruism— benefiting in-group at the cost of depriving out-group—prevails in human behavior, helping the in-group to survive intergroup competition (3). Likewise, people display parochial cooperation, parochial trust, parochial empathy, and parochial harm sensitivity in intergroup judgments (4–6). When situations are direct, zero-sum, and forced-choice, people are much more parochial than cosmopolitan (7).

On the other hand, cosmopolitan morality is feasible: consideration beyond in-group welfare has precedents. Theories of intergroup relations contend that in-group love and out-group hate are not reciprocal; attachment to one's in-group does not require hostility toward out-groups (8–9). Such psychological asymmetry suggests room for cosmopolitan morality. For example, in some intergroup economic games, participants most often choose the option that benefits the in-group without affecting the out-group (10). Likewise, cosmopolitan philosophers argue that one's fellow citizens' interests should matter no more than those of nonnationals, other things being equal (11, 12). Corroborating this ideal, evolutionary theory and laboratory research converge on altruism as standing alongside self-interest to drive interpersonal behavior (13, 14). For example, in other intergroup economic games, participants who give more in general also affiliate more with all participants regardless of team, suggesting morality could be universal (15).

What Is Missing

Serving in-group interest is understandable. The question here is who wants to help the world, even at the cost of the nation? When facing conflictual world events that demand explicit and direct

Significance

Contemporary world challenges—pandemic response, climate change, migration, global poverty, and trade policy—require comparative moral judgments (tradeoffs) between one's own and other nations. Such choices distribute benefits and protect security for just the national in-group or the whole world. In representative surveys from eight countries, most respondents are willing to sacrifice national well-being to aid foreigners, more to protect from harms than to redistribute benefits. This disaggregates two dimensions of moral cosmopolitanism (benefiting versus protecting), distinguishes it from adjacent constructs (such as group identity), and demonstrates links to attitudes and behaviors. Cognitive associations (global, inclusive words) reflect cosmopolitan thinking; making the world salient can elicit cosmopolitan behavior, at least temporarily. Understanding this global perspective can guide its enhancement.

Author contributions: X.B., V.G., and S.T.F. designed research; X.B. performed research; X.B., V.G., and S.T.F. contributed new reagents/analytic tools; X.B. analyzed data; and X.B., V.G., and S.T.F. wrote the paper.

Reviewers: R.J.-B., University of Massachusetts Amherst; D.G.R., Massachusetts Institute of Technology; and J.F.D., Yale University.

The authors declare no competing interest.

[1]To whom correspondence may be addressed. Email: sfiske@princeton.edu.

This article contains supporting information online at https://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2100991118/-/DCSupplemental.

Published September 27, 2021.

tradeoffs (e.g., allocating international aid budgets or COVID-19 resources), who are these moral cosmopolitans? Understanding these individuals allows a methodological entry point into the investigation of the social psychological processes of interest (cf. refs. 16, 17). Theoretically, understanding moral cosmopolitan individuals could potentially resolve the opposing predictions in the literature. Our analyses go beyond polarized alternatives and identify richer psychological processes. Pragmatically, policy makers increasingly need to know who is more likely to vote for world-oriented bills and how to promote more world-beneficial policies. Our data provide initial clues.

Moral Cosmopolitans

Existing psychological measures provide valuable insights but miss our target. Not all elements of moral cosmopolitanism appear together in any one previous work; per our opening definition, tradeoffs, morality, context, and group components need simultaneous measurement, but different approaches have each excluded some part of the concept, respectively. In some cases, tradeoffs are missing. Specifically, identity-based measures such as identification with all humanity (18) and expansion of the moral circle (19) require moral inclusion of other groups but they do not require sacrificing one's in-group. In other cases, morality is missing. Hierarchy-based measures such as social dominance orientation (20) capture group comparison, but they differ from intergroup morality. In still other cases, context sensitivity is missing. Morality-facet measures such as moral foundations (21) classify moral judgments into clear-cut dimensions, but intergroup morality involves an interaction among harm, fairness, and loyalty. Finally, sometimes groups are missing. Utilitarianism measures (22) propose multiple dimensions in moral judgments but are not group-based. Likewise, emotions such as empathic concern (23) measure feelings of compassion for others but are not necessarily group-based. To understand the all-in-one moral cosmopolitanism, we first created a psychometrically valid scale.

Psychological Processes

Following our concept, moral cosmopolitanism should be a multidimensional cognitive construct. Being a moral cosmopolitan can either help the world or not harm the world. Moreover, being a moral cosmopolitan requires moral reasoning but not world identity. The combination of these features distinguishes us from past work.

First, moral cosmopolitanism has two psychologically distinct dimensions: benefit redistribution and harm protection. Other work differentiates benefit and harm but not at an intergroup level. Specifically, at the individual level, recent work in utilitarian moral judgments differentiates permissive attitudes toward instrumental harm from impartial concern for the greater good (22). At the intragroup level, separate moral motives also differentiate a prescriptive morality that benefits others from a proscriptive morality that protects others from harm (24, 25). Changing to the intergroup level, the closest to separating benefit from harm is the attitude asymmetry in intergroup dynamics, which finds that in-group love can occur without out-group hate (9). This suggests helping an in-group does not imply harming out-groups. In moral cosmopolitanism, however, individuals sacrifice in-group interests to help the world. Hence, in the domain of comparative moral decisions (tradeoffs) between nation states, we predict people will protect the world more than benefit it since the latter involves resource reallocation and necessarily comes at a cost to the in-group.

Second, moral cosmopolitanism, as the philosophers' abstract principle (11, 12), is more about cognitive associations (what concepts to consider when making decisions); it is less a function of group identity (which affinity group to consider). To the question of what makes people morally cosmopolitan, current explanations largely center on social identity theory (26). At the intergroup level, given people in general favor their in-groups, one way to make out-groups more favorable is to recategorize the out-groups as a larger common in-group (27). In moral judgments, identification with the world (18) and expansive moral boundaries (19) effectively apply this principle. In contrast, we propose a cognitive association perspective, for several reasons. First, the concept of moral cosmopolitanism, in philosophy, is an ideal that rises above personal attachment. Second, the global identity view includes the in-group in the larger group, which can be a sufficient mechanism, but moral cosmopolitanism requires an even more minimal process—a change at the cognitive level. We predict that as long as the most relevant set of cognitive concepts are invoked in people's minds, making people behave like moral cosmopolitans should be plausible, even in the absence of an inclusive world identity.

Overview of Results

Here large-scale empirical evidence documents the prevalence and provenance of moral cosmopolitanism, filling gaps from previous arguments often based on American samples. Our findings are based on nationally representative samples from eight countries in North America, Latin America, Europe, and Asia (total $n = 9{,}871$). After validating a moral cosmopolitanism scale with two factors, across eight countries, the equitable benefit (EB) dimension predicts preferences for national or global redistributive choices, with about half of the samples scoring as moral cosmopolitans. In contrast, the second dimension, equitable security (ES), approaches ceiling, with most people avoiding differential harm beyond borders. Moral cosmopolitanism correlates with attitudes toward salient world policies (e.g., immigration) and predicts charitable and COVID-19 resource decisions. For these purposes, moral cosmopolitanism differs from related psychological constructs (e.g., world identity, moral circle, and political ideology). Using linguistic analysis to identify psychological processes (e.g., reasoning globally rather than identifying with the world), EB and ES are associated with concrete social cognitive concepts. From that comes an intervention that temporarily leads people to behave like moral cosmopolitans in benefit redistributions. Without invoking social identity or deliberate calculations, making the cosmopolitan-related thinking cognitively accessible was sufficient to cause a new sample of American participants ($n = 1{,}162$) to endorse cosmopolitan actions in benefit-redistributive decisions. Summary results appear in Table 1.

Study 1: EB and ES Confirm Distinct Psychological Factors in Moral Cosmopolitanism. Existing concepts of cosmopolitanism mostly come from sociology (e.g., ref. 28), emphasizing openness to new cultures or enjoyment of travel but typically not moral judgments. Existing concepts of moral judgments mostly come from psychology (e.g., ref. 21), discussing general moral values but not considering dilemmas that explicitly pit one's group against others. We developed a scale that simultaneously captures cosmopolitanism and moral judgments.

We collected 81 initial items (*SI Appendix*, SI1.1, *Item generation*) based on existing literature and contemporary concerns, consulted with leading researchers to refine the items, ran pilot studies with online participants from Western and Eastern cultures ($n = 1{,}575$). This process reduced the most effective candidate items to 21 (*SI Appendix*, SI1.1, *Exploratory factor analysis* and *Confirmatory factor analysis*). We then ran nationally representative samples in eight countries (United States, Guatemala, Columbia, Spain, United Kingdom, Germany, China, and Japan; $N = 700+$ as the prespecified sample size per country; $n = 5{,}772$; *SI Appendix*, SI1.2), translating and back-translating questionnaires that asked for (dis) agreement on six-point scales for each item. To extract latent dimensions and test structural (in)variance across national samples, we randomly split responses into exploratory ($n = 200$) and confirmatory ($n = 500+$) datasets for each country and performed multiple-group factor analyses using maximum likelihood estimation with robust SEs (*SI Appendix*, SI1.1, *Multination factor analysis*) (29). The multinational factor analysis yielded nine items which appear in Table 2.

Table 1. Summary results for moral cosmopolitanism

Hypotheses	Participants	Analyses	EB subscale	ES subscale
1. Some people are moral cosmopolitans	n = 1,068 US; n = 507 CN; n = 5,772 replication from eight countries: CN, CO, DE, ES, GT, JP, UK, and US	Item generation; US and China validation; exploratory factor analysis; multigroup confirmatory factor analysis	About half above midpoint = 3.5	Majority above midpoint = 3.5
2.1. Moral cosmopolitans behave in two ways: (a) benefit the world and (b) protect the world	n = 5,772 replication from eight countries: CN, CO, DE, ES, GT, JP, UK, and US	Dictator game: Teach for [own country] vs. **Teach for All**	***Predictive***	Weaker
		Trolley dilemma: **[not] sacrifice immigrants** vs. sacrifice citizens	No effects	***Predictive***
	n = 200 from US	Sign petitions via mock Twitter click: Mask the USA vs. **Mask the World**	***Predictive***	No effects
		Sign petitions via mock Twitter click: Vaccinate the USA vs. **Vaccinate the World**	Weaker	***Predictive***
2.2. Moral cosmopolitans have distinct opinions: (a) benefit the world and (b) protect the world	n = 5,772 replication from eight countries: CN, CO, DE, ES, GT, JP, UK, and US	**Benefit:**		
		Allocate budget to international affairs	***Supportive***	No effects
		Increase immigration	***Supportive***	No effects
		Value international organizations	Supportive	Supportive
		Protect:		
		Criticize own leaders	No effects	***Supportive***
		Avoid domination	No effects	***Supportive***
		Prevent harm to foreign workers	No effects	***Supportive***
		Reduce country CO_2	No effects	***Supportive***
3.1. Moral cosmopolitans differ from other people in distinct ways: divergent validity	n = 5,772 replication from eight countries: CN, CO, DE, ES, GT, JP, UK, and US	Demographics: gender, race, socioeconomic, etc.	No correlations	No correlations
		Political orientation: liberal	Small	Small
		Trait empathy	Small	Medium
		Parochial empathy	Small	No correlations
		Moral circle including foreigners	Small	Small
		Identification with the world	Small	Small
	n = 504 from US	Moral foundations: in-group, authority	Small	Small
		Moral foundations: harm, fairness, purity	No correlations	Small
3.2. Moral cosmopolitans differ from other people in distinct ways: discriminant validity	n = 5,772 replication from eight countries: CN, CO, DE, ES, GT, JP, UK, and US	Utilitarianism: **impartial benevolence**	***Medium***	Small
		Utilitarianism: **instrumental harm**	No correlations	***Small***
		Social dominance orientation: **dominance**	Small	***Medium***
		Social dominance orientation: **egalitarianism**	***Small***	Medium
4. Moral cosmopolitans have distinct thoughts	n = 5,772 replication from eight countries: CN, CO, DE, ES, GT, JP, UK, and US	Natural language processing on **free responses** after donation	***Correlate with EB:*** world, both, education	Similar to EB but weaker
5. Experiments can produce more moral cosmopolitans	n = 1,162 from US	Prime with sentences from moral cosmopolitanism vs. parochialism before donation: Donate to the USA vs. **Donate to the World**	***Predictive***	Weaker

Bold-italics in last two columns indicate distinct predictions for EB and ES as hypothesized. Bold text in third column indicates direction of the effect. Predictive, weaker, and no effects are shorthand for statistically significant and larger effect size, statistically significant but smaller effect size, and no statistically significant effects, respectively. Divergent and convergent validity show semipartial correlations, medium refers to $r < 0.5$, small refers $r < 0.3$, and no correlation refers to statistically nonsignificant effects. All regression analyses (thus reported effects in this table) controlled for political orientation and other convergent/divergent validity measures. Country names are abbreviated: CN, China; CO, Colombia; DE, Germany; ES, Spain; GT, Guatemala; JP, Japan; UK, United Kingdom; and US, United States, in alphabetical order.

A stable two-dimensional construct showed both country differences and individual differences in all sampled countries (see scale in Table 2 and response distributions in Fig. 1). The first dimension, EB (five items), captures the tendency to maximize well-being for broader humanity, even at the expense of one's nation or conationals. For example, "Citizens in our country should accept a reduction in their standard of living if doing so increases global equality" and "It is wrong to always prioritize our citizens over people in other countries." Mean responses centered on 3.39, where 6 indicates complete agreement. Empirically, around half (54%) of the participants showed parochial tendencies (below 3.5 on the scale), and 46% showed cosmopolitan tendencies (above 3.5 on the scale) in cross-national benefit redistributions.

The second dimension, ES (four items), captures the tendency to protect all people from harm, not just fellow citizens: "It is as bad when terrorist attacks happen elsewhere as when they happen in our nation" and "It is as bad when our military accidentally kills foreign civilians as when a foreign military accidentally kills our

civilians." Mean responses approach ceiling, with an average of 5.27. Empirically, 95% participants showed cosmopolitan tendencies in cross-national harm allocations (above 3.5 on the scale), and only a very small proportion (5%) of them showed parochial tendencies (*SI Appendix*, SI1.3).

Countries showed both similarities and differences. The two dimensions consistently emerge in all sampled countries with satisfactory model fits, indicating the psychological constructs' generality regardless of cultural or economic contexts (*SI Appendix*, SI1.1, *Multination factor analysis*). However, we also observe meaningful mean differences: one-way ANOVA revealed significant country differences in EB [$F(7,5764) = 49.894$, $P < 0.001$] and ES [$F(7,5764) = 147.08$, $P < 0.001$]. Post hoc Tukey HSD (Honestly Significant Difference) tests revealed that the United States ($M = 2.93$, SD = 1.37) scored lower on EB than all other countries, and Spain ($M = 4.01$, SD = 1.22) scored higher on EB than all other countries. Speculating, these ranks may reflect the countries' respective roles in global redistribution (US citizens see their country as donors; Spanish citizens see themselves as a potential recipient of European Union [EU] aid), or they may reflect differing US–EU attitudes toward the role of public support more generally.

For ES, Japan ($M = 4.55$, SD = 0.95) scored lower than all other countries, perhaps because of wariness toward military action arising from the atomic bombings in World War II, or generally being more protective of conationals rather than other nations. At the other extreme, the Spanish-speaking countries, Guatemala ($M = 5.58$, SD = 0.74), Colombia ($M = 5.55$, SD = 0.84), and Spain ($M = 5.66$, SD = 0.68), scored higher on ES than the remaining countries.

To summarize, moral cosmopolitanism captures two distinct but related psychological dimensions, EB (resource redistribution) and ES (harm protection). The presence of the two dimensions is consistent with our theoretical priors that morality draws on two distinct forms of regulation: approach or promotion (benefit), and prevention or avoidance (respecting security) (24, 30). Crucially, the two factors function differently within the intergroup context. Consistent with prior work on asymmetric intergroup evaluations (8, 9), people tend to exhibit moral parochialism (in-group favoritism) in benefit redistribution and moral cosmopolitanism in harm prevention (protect all groups). The fact that the two factors emerge naturally suggests people may think in different ways about benefits versus harms to other groups. The two-factor structure is stable and general across populations but specific enough to identify differences across countries.

Study 2: EB Predicts Resource Redistribution and ES Predicts Avoiding Harm Allocation. Given that moral cosmopolitanism incorporates two distinct factors of benefiting and not harming out-groups, one would expect that those dimensions predict distinct comparative moral choices. EB should correlate positively with redistributing resources for a global cause relative to a national one, such as charitable donation or COVID-19 mask distribution.

ES should be correlated with refusal to harm a nonnational to help a conational, such as sacrificing an immigrant to save citizens in the trolley dilemma (31, 32) or sacrificing others to save citizens by monopolizing COVID-19 vaccines.

To assess a standard beneficial moral behavior, charitable donation, the same online participants from eight countries completed an economic dictator game at the beginning of the study. In addition to their fee, they received $1, which they could donate to one or divide between two organizations—Teach for All vs. Teach for America (or their country's equivalent). The outcome was donations to Teach for All divided by the total amount available to donate. A higher value indicates a morally cosmopolitan preference, while a lower value indicates a morally parochial preference. As a robustness check, we randomly assigned half of the participants to donate on behalf of the research team, while the remaining half had a third option to keep the money for themselves. We ran multilevel regression analyses with individual responses nested within countries allowing for by-country random intercepts and random slopes (*SI Appendix*, SI2.1).

As expected, EB significantly predicted cosmopolitan donations, such that for each unit increase in EB, participants donated 4% (SE = 0.01, $P < 0.001$) more to Teach for All (Fig. 2*A*). ES also predicted cosmopolitan donations but not as strongly. For each unit increase in ES, participants donated 1% (SE = 0.00, $P = 0.026$) more to Teach for All. As expected, game-condition main effects emerged: participants donated 19% (SE = 0.03, $P < 0.001$) less to Teach for All if they had the option to keep the money (i.e., self-interest). We did not observe significant interaction effects, meaning self-interest did not moderate the morally patriotic or cosmopolitan donation.

Table 2. Final scale items for moral cosmopolitanism

Subscale	Item
Equitable Benefits	Citizens in our country should accept a reduction in their standard of living if doing so increases global equality.
	Citizens in our country are obliged to lower their standard of living if doing so substantially contributes to global justice.
	Citizens in our country must lower their standard of living if that is necessary to achieve a higher standard of living for every person in the world.
	It is a problem if most people in our country generally agree on the principle of putting our own citizens first.
	It is wrong to always prioritize citizens over people in other countries.
Equitable Security	It is as bad when terrorist attacks happen elsewhere as when they happen in our nation.
	When a hostage is being held by terrorists, it is equally bad if the hostage is from other nations and from our nation.
	It is as bad when our military accidentally kills foreign civilians as when a foreign military accidentally kills our civilians.
	When our friend, who is a fellow citizen, makes a distasteful joke about others, it is equally inappropriate when the joke concerns foreigners as when it concerns fellow citizens.

Items are shown in randomized order. For each item, participants rate on a scale from 1 to 6, where 1 indicates strongly disagree and 6 indicates strongly agree.

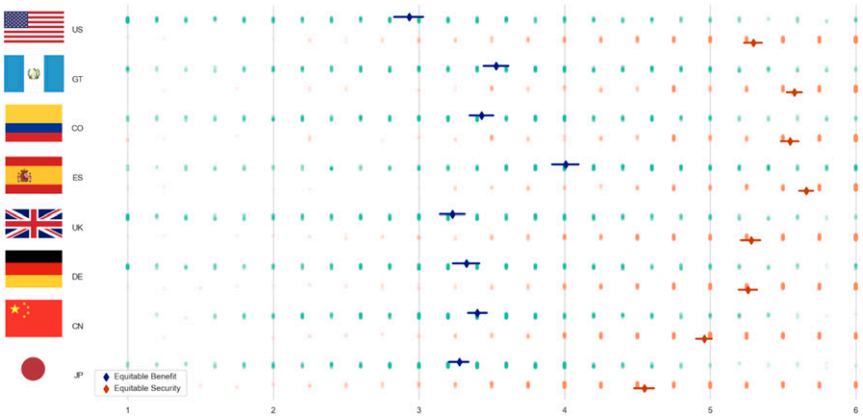

Fig. 1. Self-reported moral cosmopolitanism, EB (green) and ES (red), by country means and 95% confidence interval, with raw scores scattered in the background. Country names are given on the *y* axis, with geographical locations in the world map: United States (US), Guatemala (GT), Colombia (CO), Spain (ES), United Kingdom (UK), Germany (DE), China (CN), and Japan (JP). On the *x* axis, scores approaching 6 indicate relatively moral cosmopolitan judgment, while scores approaching 1 indicate relatively moral patriotic judgments.

To assess a standard harmful moral behavior, participants also completed an international trolley dilemma, making moral judgments on whether it is acceptable to push an immigrant to save five middle-class fellow citizens from a runaway trolley (*Methods* and *SI Appendix*, SI2.2). As expected, participants who scored higher on ES were less likely to push the immigrant ($b = -0.37$, SE = 0.07, $P = 0.001$); EB was unrelated ($b = 0.024$, SE = 0.03, $P = 0.433$) (Fig. 2*B*). Given that one central correlate in the trolley dilemma is utilitarianism, we put ES and instrumental harm, taken from the Oxford Utilitarian Scale (22), into the same model. We observed a significant interaction between the two ($b = -0.12$, SE = 0.03, $P < 0.001$): For each unit increase in instrumental harm, participants were 0.49 more likely to condone pushing pushing the immigrant ($b = -0.49$, SE = 0.04, $P < 0.001$). However, if participants were also one unit higher on ES, they were only 0.37 more likely to condone pushing the immigrant. In other words, the association between utilitarianism and the international trolley decision attenuated among moral cosmopolitans (*SI Appendix*, SI5.1, documents a replication and extension of this experiment).

To explore COVID-19–related moral behavior, a preliminary survey addressed commonly discussed choices. Wearing a mask and being vaccinated both provide benefits and protection against harm, but they have different social connotations: "Masks are marketed as tools to help protect others, while the coronavirus vaccine, which may not prevent transmission to others, is perceived as something that helps oneself" (33). Admittedly, when we planned a survey (*SI Appendix*, SI2.3), we did not have these predictions. However, our findings showed that ES (endorsing harm prevention for all) was correlated with vaccine distribution. EB (sharing resources that help members of all communities) was correlated with mask distribution. With these priors in mind, we then conducted a Twitter-click study (reported below).

First, COVID-19 beneficial moral behaviors plausibly concern mask redistribution. Since the COVID-19 pandemic began, there have been debates surrounding international vs. domestically focused policies. New online participants in the United States ($n = 100$) could sign petition A (Mask the World) or petition B (Mask America) via a mocked Twitter-click interface (*Methods* and *SI Appendix*, SI2.4). Respondents indicated endorsement by clicking an interactive retweet button on the screen as many times as they wanted. The petition proposed that the US government and companies should make, sell, and donate masks to (A) everyone in the world vs. (B) our own citizens. The proportion of clicks to the world, which approximates a costly moral behavior (34), was our behavioral outcome. We ran a linear regression with moral cosmopolitanism rating as the independent variable, the world click ratio as the dependent variable, and political ideology as the control variable.

Masks were supposed to be a resource benefiting other people in the community, especially in the US context. Indeed, EB significantly predicted more support for Mask the World ($b = 0.16$, SE = 0.03, $P < 0.001$), controlling for political ideology ($b = -0.04$, SE = 0.02, $P = 0.08$), and ES ($b = 0.02$, SE = 0.04, $P = 0.60$; neither being predictive). That is, for each unit higher on the EB dimension, participants on average clicked Mask the World petition 16% more times.

COVID-19 harmful moral behaviors plausibly concern vaccine monopoly. Another set of online American participants ($n = 100$) had a choice to sign a different petition: (A) Vaccinate the World or (B) Vaccinate America. The petition proposed that the US government and companies should make, sell, and donate COVID-19 vaccines to (A) everyone in the world vs. (B) our own citizens. Again, we used the proportion of clicks to the world as the behavioral outcome.

We hypothesized that vaccines prevent harm because people are less likely to get infected when vaccinated. As hypothesized, ES significantly predicted more support for Vaccinate the World ($b = 0.13$, SE = 0.04, $P = 0.003$), controlling for political ideology ($b = 0.04$, SE = 0.02, $P = 0.08$, not predictive) and EB ($b = 0.07$, SE = 0.03, $P = 0.04$; weaker but predictive). That is, for each unit increase in the ES dimension, participants on average clicked Vaccinate the World petition 13% more times.

Turning to general policy attitudes: EB correlated with attitudes approving of globally redistributive policies including, for instance, expansive immigration policies (e.g., when jobs are scarce, employers should not give priority to people of this country over immigrants, $b = -0.37$, SE = 0.05, $P < 0.001$, and not reducing the number of immigrants to our country nowadays, $b = -0.21$, SE = 0.04, $P = 0.001$). EB also predicted positive attitudes toward international organizations (e.g., in general, international institutions do valuable work, $b = 0.17$, SE = 0.03, $P < 0.001$) and support for international agreements (e.g., increasing budget support to the UN, $b = 0.23$, SE = 0.03, $P < 0.001$, and more government budget to assist other countries, $b = 0.33$, SE = 0.04, $P < 0.001$).

ES predicted attitudes toward avoiding active harm including, for instance, criticizing own government and private companies (e.g., boycott companies who make products in other countries just to take advantage of lower safety and health standards, $b = 0.26$, SE = 0.07, $P = 0.006$); resisting own country's immoral behaviors, as shown by freely criticizing leaders ($b = 0.22$, SE = 0.07, $P = 0.009$);

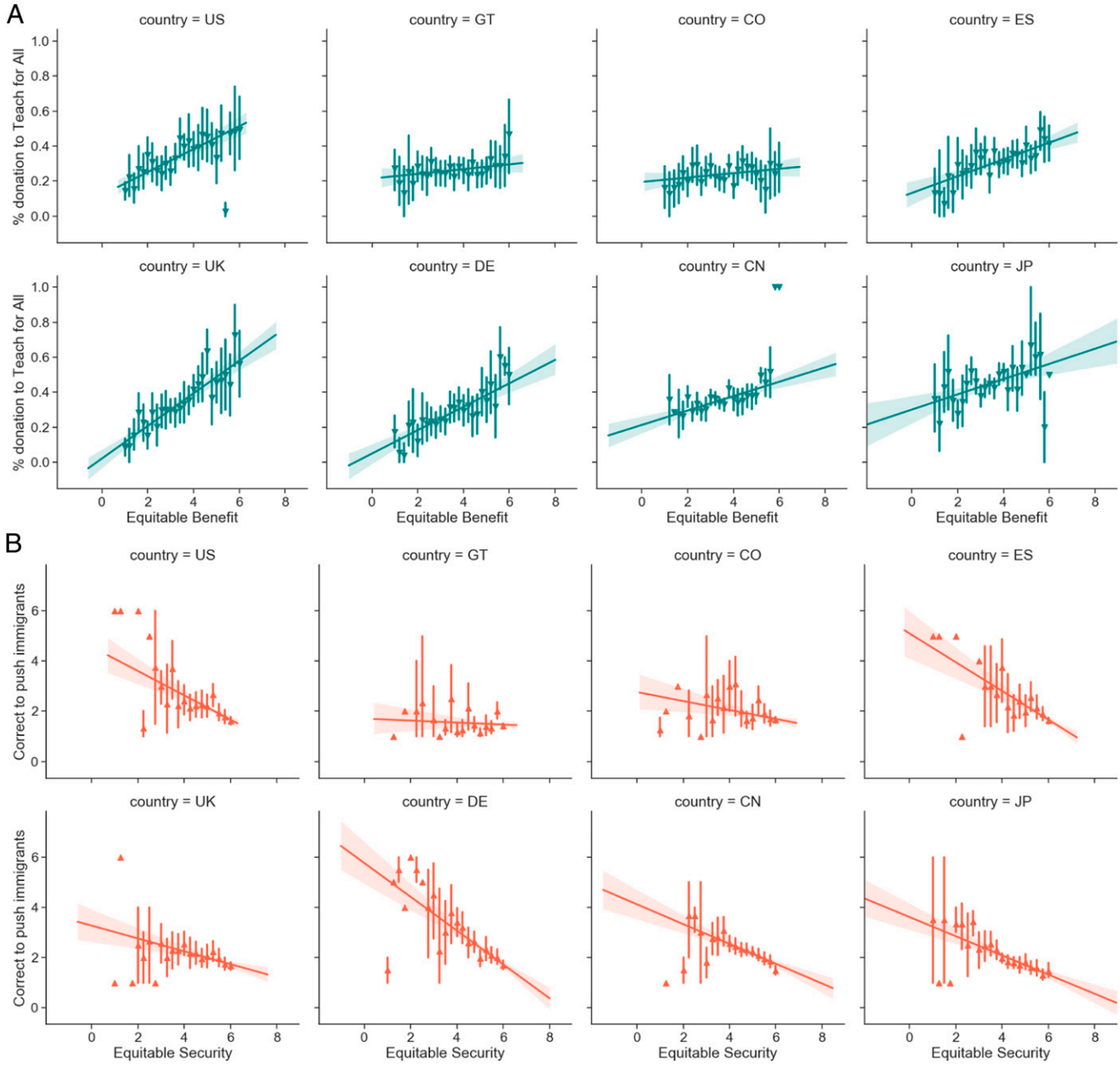

Fig. 2. Relations of cosmopolitan moral judgments with (*A*) comparative charitable donations and (*B*) international trolley dilemma, displaying country central tendency and 95% confidence intervals. Full model estimates individual-level linear effects while controlling for within-country dependencies with random slopes and random intercepts estimation. *x* axis denotes self-reported cosmopolitan moral judgments. *y* axis denotes donation proportion (*A*), where a higher value indicates more to Teach for All, and trolley dilemma decisions (*B*), where a lower value indicates reluctance to push an immigrant to save five citizens.

and not advocating for the biggest say in deciding world policies ($b = -0.14$, SE = 0.05, $P = 0.025$). It also predicted country-level CO_2 reduction ($b = 0.35$, SE = 0.05, $P < 0.001$) but less so personal CO_2 reduction ($b = 0.09$, SE = 0.03, $P = 0.005$), perhaps because country-level reduction can reduce active harm to the world, whereas individual-level reduction might be less effective (more items and results in *SI Appendix*, SI2.5).

To summarize, the EB factor of moral cosmopolitanism correlated with comparative moral judgments regarding real-world charitable donations, mask redistribution, and redistributive international policies. The ES factor of moral cosmopolitanism correlated with not sacrificing an immigrant to save fellow citizens, preferences to avoid vaccine monopolies, and opposing policies that actively harm citizens in other countries.

Study 3: Moral Cosmopolitanism Correlates with, but Differs from World Identity, Inclusive Moral Values, and Political Ideology. To understand the mechanisms of moral cosmopolitanism, we examined its potential correlates. Analyses considered plausible demographic associations, as well as other constructs from social psychology regarding group identity and moral psychology that relate to the greater good but are not explicitly intergroup tradeoffs or do not distinguish benefits and harms.

In relation to demographics, one might suspect that life experiences, including travel, socioeconomic status, age, gender, and

political beliefs, were drivers. However, tests generally found slight to null correlations with cosmopolitan donations (*SI Appendix*, SI3.1), except for political beliefs. Being liberal was slightly and positively correlated with both factors and the actual donations (EB, $b = 0.27$, SE = 0.05, $P = 0.001$; ES, $b = 0.13$, SE = 0.02, $P = 0.001$; and donation to Teach for All, $b = 0.03$, SE = 0.01, $P = 0.016$).

In relation to adjacent psychological constructs, one might also suspect that basic moral values or other psychological constructs explain moral cosmopolitanism. Semipartial correlation tests found the two factors correlate with distinct sets of psychological constructs (*SI Appendix*, SI1.4). Specifically, EB does not relate to moral foundations of care/harm ($r = 0.07$, $P = 0.136$), fairness ($r = 0.05$, $P = 0.311$), or purity ($r = -0.02$, $P = 0.738$) but does relate slightly to in-group ($r = -0.16$, $P < 0.01$) and authority ($r = -0.21$, $P < 0.01$; from ref. 21). In contrast, ES relates to all fundamental moral motives (harm $r = 0.29$; fair $r = 0.28$; in-group $r = -0.15$; authority $r = -0.19$; purity $r = -0.15$, $P < 0.01$). Benefiting versus not harming operates differently.

EB correlates more with beneficial and comparative measures, such as utilitarianism for good, $r = 0.40$, $P < 0.01$, and comparatively more empathy to out-groups than in-groups, $r = 0.18$, $P < 0.01$ (from ref. 4). ES correlates more with harmful and generic measures, such as utilitarianism for bad, $r = -0.20$, $P < 0.01$ (from ref. 22); not including foreigners in moral concern, $r = -0.16$, $P < 0.01$ (from ref. 19); general empathic concern ($r = 0.36$, $P < 0.01$; from ref. 23); and lower social dominance orientation (dominance dimension, $r = -0.34$ and equity dimension, $r = -0.31$, $P < 0.01$; from ref. 35). Both factors positively, although not strongly, correlated with world identification (EB $r = 0.25$, ES $r = 0.25$; from ref. 18; $P < 0.001$ unless otherwise denoted). World identity may not be necessary.

In summary, we did not find strong demographic differences in moral cosmopolitanism, with the exceptions of political beliefs, which had small effects. The benefit redistributive factor correlates with in-group loyalty, consistent with the finding that national boundaries do play a role in benefit allocation. The harm protective factor correlates with general moral concerns, consistent with our finding that national boundaries are less relevant in harm protection. Both dimensions show small effects with an inclusive world identity indicating a different mechanism. These findings support the notion that people, without identifying with the world, may want to benefit their in-groups, but they do not necessarily want to harm out-groups (8, 9).

Study 4: EB and ES Highlight Common and Distinct Social Cognitive Concepts from Linguistic Analysis. As foreshadowed in the opening discussion and the discriminant analyses, inclusive identity may

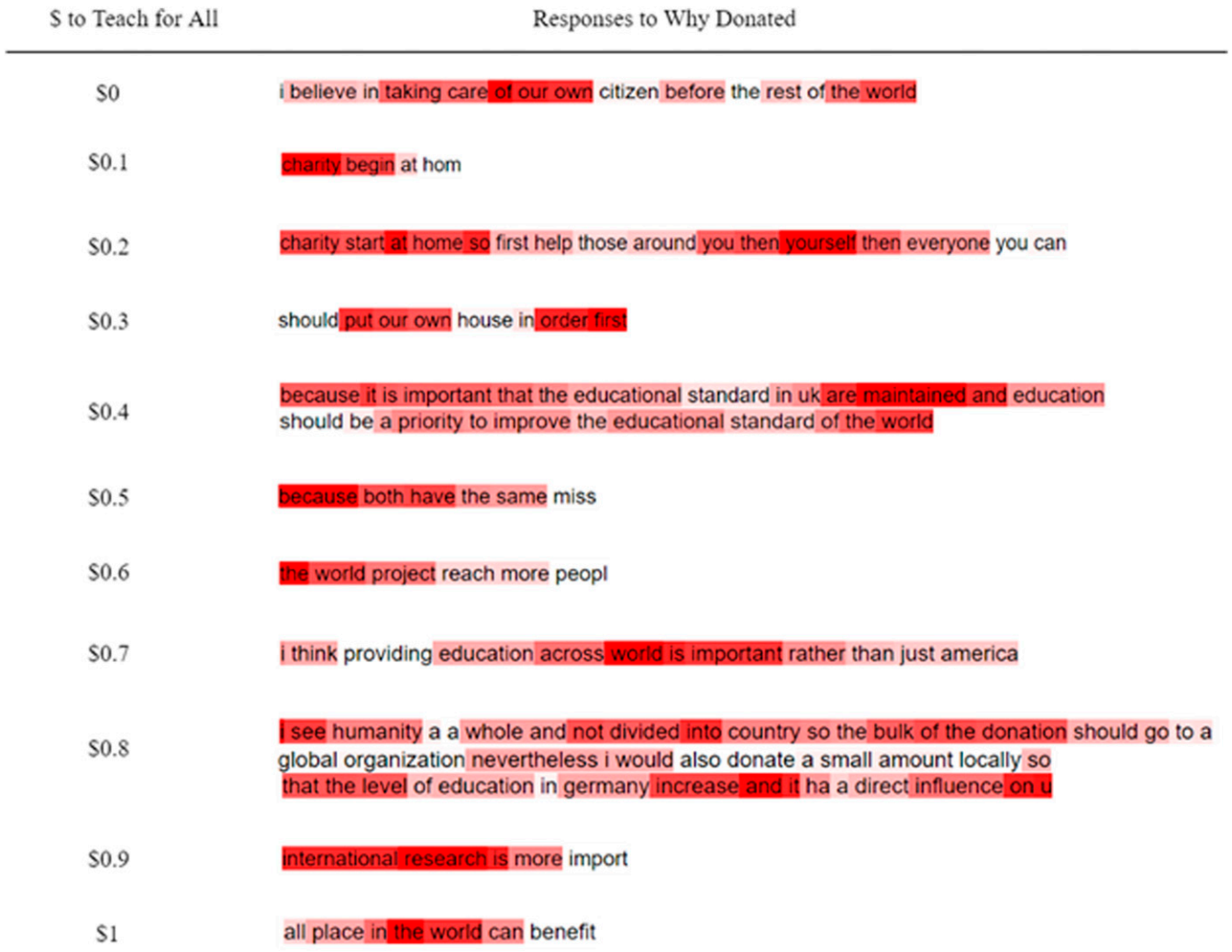

Fig. 3. Attention heat map learned from a linguistic model using responses translated by Google Translate from eight countries. See *SI Appendix*, SI3.2, for model architecture details. Each line is an example response from patriotic ($0 to Teach for All) to cosmopolitan ($1) donations. Red denotes model estimated word weights; the darker the color, the more important a word is.

be one sufficient mechanism but is not necessary. Independently, we probed participants' thoughts by content analyzing free responses. This analysis helps to understand the cognitive mechanism underlying moral cosmopolitanism which could inform interventions to encourage it.

To extract the most salient concepts, in the format of words, from sentences, in the eight- country study, after making their comparative donations, participants read: "Please tell us, briefly, why you decided as you did." This netted 5,772 free responses from people's thoughts about their donations. We developed computational linguistic models that predict which words are most important in predicting their donations (long short-term memory recurrent neural networks with self-attention) (36, 37) (*SI Appendix*, SI3.2; examples are given in Fig. 3). We cross-validated the relevant word importance by regressing the attention weights on participants' donations and moral cosmopolitanism scale scores. In multilevel modeling for each keyword, the independent variable was the attention weight for that keyword for each participant, and the dependent variable was that participant's scale scores and actual donations; errors were clustered at the country level. The regression coefficients thus show which words and concepts matter to moral cosmopolitans, relative to moral patriots (*SI Appendix*, SI3.2).

Three sets of concepts in the free-response answers were particularly predictive of participants' actual donations. These predictive concepts were group (world vs. our own), (im)partiality (both vs. first), and concrete construal of the cause (education vs. charity). Although cosmopolitans and patriots both think about the benefited group (not surprising since they were making group-level decisions), the default group differs. Moral cosmopolitans thought more about the world (attention to "world" predicts the scale, $b = 0.68$, SE = 0.31, and predicts donation, $b = 0.77$, SE = 0.07). In contrast, moral patriots thought more about their own country (attention to "our," $b = -2.68$, SE = 0.60 for scale; $b = -0.83$, SE = 0.14 for donation; attention to "own," $b = -2.73$, SE = 0.97 for scale; $b = -0.73$, SE = 0.23 for donation). Second, concepts of impartiality differed. Moral cosmopolitans paid more attention to equal treatment ("both," scale $b = 1.23$, SE = 0.30; donation $b = 0.50$, SE = 0.07), while moral patriots paid more attention to prioritizing or ranking ("first," $b = -1.55$, SE = 0.35; $b = -0.55$, SE = 0.08). Third, the levels of other-directed concreteness were different. Moral cosmopolitans focused on improving education for recipients ("education," scale $b = 18.46$, SE = 6.96; donation $b = 5.17$, SE = 1.63), whereas moral patriots focused on the action of the person giving away money ("charity," $b = -1.97$, SE =0.60; $b = -0.44$, SE = 0.14; all $P < 0.05$).

Asking directly for five salient words supplemented the ecologically unconstrained sentence analysis, in a laboratory-oriented, simpler design with American participants, which commenced after the pandemic began. After making their click-for-petition decisions, they read, "Please list briefly some thoughts you had while deciding and while retweeting." We counted the most frequently mentioned words in the Mask the World, Mask America, Vaccinate the World, and Vaccinate America groups.

The top 10 words (with frequencies; distinctive words are italicized) in Mask the World are *world* (13), safety (12), *health* (11), *global* (10), helpful (9), help (9), caring (8), masks (6), *everyone* (6), and *love* (5). The top 10 words in Mask America are *America* (24), safety (19), mask (19), *American* (15), help (11), *USA* (10), *protect*/*protection* (10), *home* (9), *first* (8), and helpful (7). The top 10 words in Vaccinate the World are help (23), world (22), *everyone* (18), *global* (18), health (16), *vaccine* (14), *compassion* (11), *equality* (10), safety (9), and *hope* (9). The top 10 words in Vaccinate America are help (14), America (13), health (9), care (6), world (6), safety (6), *cost* (5), *first* (5), *give* (5), and *citizens* (4). An immediate observation is that again, we see that group (world, global, everyone vs. America/n, USA, and citizens) and (im)partiality (everyone, equality vs. first, and home) are salient constructs for people to make their retweeting decisions. New here is a distinction between other-focused moral cosmopolitan sentiments (compassion, hope, and love) vs. self-interest (cost and give). "Safety" appears in all conditions. However, "protection" appears more frequently in Mask America condition (*SI Appendix*, SI3.3).

In summary, linguistic reports of the activated cognitive concepts were robust predictors of people's donations. Concepts such as "world, both, education," as compared to "own country, first, charity," are especially predictive if a person is going to donate more to an international cause instead of a national one. We found similar sets of linguistic markers among mask and vaccine distribution petition supporters. Concepts of "world vs. USA" and "both vs. first" appeared frequently. "Protection" was salient for mask distribution.

Study 5: Making the Cosmopolitan-Related Thinking Cognitively Accessible Caused Participants to Endorse Cosmopolitan Actions in Benefit Redistributive Decisions. Having identified an apparent cognitive mechanism, we wanted to use it to experimentally encourage people to engage in cosmopolitan actions. This experiment examined the causal effects of the three classes of cosmopolitan concepts, identified above, using semantic priming. Following the eight-country studies, we hypothesized that if concepts like "world, both, education" are the active concepts of cosmopolitan-related thinking, making them cognitively accessible would lead people to engage in cosmopolitan action. Conversely, we thought that concepts like "our own, first, charity" would activate moral patriotism and parochial generosity. We did not intervene in harm protective decisions because they are already endorsed almost at ceiling; instead we focused on the hard decisions—providing benefits to the world at in-group expense.

For a semantic intervention, new American online participants ($n = 1{,}162$) were randomly assigned to read either moral cosmopolitan or moral patriotic (parochial) concepts. We selected sentences from previous American participants that included at least two of the key concepts. Moral cosmopolitan treatment has sentences such as "There is an immediate need for education worldwide especially for women and children living in impoverished countries."

Moral patriotic treatment has sentences such as "I think we should help our own first then broaden the scope" (see full stimuli in *Experiment*). Note that these sentences simply supplied individual reasoning, not behavioral norms or overt persuasion. Crucially, to make sure participants actively engaged in the concepts and thought about them, rather than receiving them passively, we deployed methods from the elaboration likelihood model (38) and narrative paradigm (39) to assist the semantic priming task. Participants typed the pros and cons or told a personal story regarding each treatment sentence.

The outcome measure was real-world charity donation (share of donations, out of $1, to international organizations rather than national ones). To reduce demand effects, we listed six nongovernmental organizations (NGOs; national and international efforts in education, poverty, and emergency medicine) (*SI Appendix*, SI4.1).

The baseline cosmopolitan donation rate from previous studies was 35.7%, meaning that on average, participants donated more to national NGOs than to international NGOs. Our experiment revealed a significant average accessibility effect (Fig. 4). Participants treated with moral cosmopolitanism concepts donated 50% to international NGOs, 11.8% (SE = 0.02, $P < 0.001$); those treated with moral patriotic concepts donated 38%. We did not find engagement- style differences: participants who wrote pros and cons donated similarly to those who wrote personal stories. Although we tentatively hypothesized that extreme believers (in either direction) might be less influenced by our manipulation than moderate believers, we did not observe such treatment heterogeneity. To rule out the potential confounding effect of support for education NGOs, we tested out-of-domain outcomes by excluding Teach for All/America and found a similar effect ($b = 0.10$, SE = 0.02, $P < 0.001$). This test also mitigates concerns related to demand effects (*SI Appendix*, SI4.2).

In summary, confirming the linguistic evidence of correlated accessibility, this experiment demonstrated the causal impact of moral cosmopolitan concepts on international donations. Manipulating participants to think like moral cosmopolitans changed their donations in a resource allocation context, moving them to help the world even at the expense of their own nation and fellow citizens. Resource allocations are influenced by tribal morality (2, 3), but the current experiment shows that interventions could plausibly break down tribal (parochial) boundaries, at least temporarily.

Discussion

This paper investigated what people do when facing direct, comparative moral judgments between their own and other nations. Our conceptualization of moral cosmopolitanism as combining tradeoff, morality, context, and group elements mimics the hard political tradeoffs between the in-group and others in contemporary world challenges. For example, should people prioritize giving masks and vaccines to their own citizens or make the resources equally available to people around the world? Under the broad umbrella of moral cosmopolitanism, defined as overcoming a comparative preference for one's own nation, relative to the world as a whole, in the context of an explicit tradeoff, we found significant individual differences along two psychological dimensions: benefit redistribution and harm protection.

EB, on one hand, relates to redistributing resources to nonnationals. Consistent with in-group biases in the moral judgment literature, people in such comparative choices tend to remain parochial. ES, on the other hand, entails not harming nonnationals. Most people seem to overcome comparative choices on that dimension, endorsing safety for all. Benefiting and harming out-groups are independent. On the basis of sizable samples from eight countries in North America, Latin America, Europe, and Asia, we provide evidence for this distinction.

This distinction has theoretical implications. Past research shows mixed results: some find people are mostly parochial, others find people do possess a global morality. Joining the conversation, our results suggest focusing on whether a judgment concerns benefit redistribution or harm protection. They are not symmetrical, especially in intergroup judgments. When people are ready to help the in-group, we should observe more tribal or parochial morality; when people are unwilling to harm out-groups, we should observe more cosmopolitan morality. Future research might draw on this differentiation when assessing intergroup moral judgments and global thinking, more generally.

This differentiation can also be practical because the two dimensions predict distinct sets of policy attitudes and real-world moral behaviors. EB predicts redistributive policies such as immigrant aid, charitable donations, and mask distributions. ES predicts harm protective policies such as human rights violations, civilian casualties in war, and vaccine hoarding. This differentiation can inform expectations about how people will react toward policies that either imply benefitting an in-group or harming an out-group. When the policy is (or is framed as) in-group beneficial, the EB dimension is likely to be operative in people's minds. When the policy is (or is framed as) out-group harmful, we should expect ES psychology to be at work. However, due to ethical reasons, we did not include real-world harmful decisions other than thought experiments on the trolley dilemma and demonstrations of vaccine policy via mouse-click behaviors. Future work should investigate how ES operates in real harm decisions, in an ethical manner.

Throughout these analyses, political ideology only slightly correlates with cosmopolitan judgments; its predictive power diminishes after controlling for moral cosmopolitanism. This indicates that many controversial world policies may transcend partisanship and may entail deeper psychological underpinnings.

The mechanism underlying moral cosmopolitanism seems to suggest distinct sets of cognitive concepts when people are making comparative decisions. Our finding is consistent with philosophical accounts that emphasize the role of reason in generating global concern (12, 40). Linguistic analysis and experiments found that moral behavior can change by manipulating the framing of global thinking. In the experiment, we did not invoke a larger moral circle or a common global identity or encourage utilitarian consideration. Linguistic cues given to moral patriots sufficed to increase cosmopolitan donations regardless of participants' original political ideology, age, gender, education level, or place on the social ladder. This result speaks to prior research that found altruistic behavior unresponsive to framing (41). To be sure, our manipulation changed only temporary behaviors; we do not have data on long-term effects.

People probably require prolonged exposure to cosmopolitan concepts to overcome their everyday exposure to parochialism. Future research could explore longitudinal experiments.

We focused on the nation as the group boundary of interest because it is psychologically understudied, yet relevant to important political, economic, and social conflicts. Furthermore, moral patriotism is philosophically defensible, at least in comparison to other morally arbitrary group categories, such as race and gender (42). For instance, when someone says "Whites first" or "men first," that person is rightly criticized. However, "our nation first" is more widely acceptable, whether because national cooperation creates moral obligations or because living under common laws,

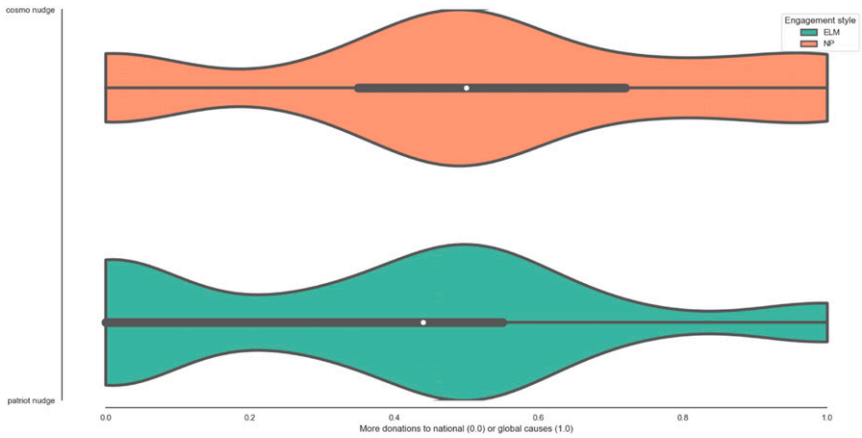

Fig. 4. Average treatment effects on international donation by American online participants. The null hypothesis of no difference was rejected, by an 11.8% increase. Plot features central tendency and kernel density estimations for experimental conditions.

moving toward a common destiny, or sharing a common history establishes a morally justified preference for conationals. At the same time, unlike the family or the clan, a nation is an imagined and abstract community (43, 44), and members are strangers to each other (45), in the same way that "citizens of the world" are strangers to each other. As a result, out-group preference in the world–nation comparison may be more plausibly compared than the family–stranger comparison that some utilitarian thinkers examine.

Converging with recent evidence on the flexibility of local versus global moral judgments (46–48), our work offers global evidence of moral cosmopolitan judgment, as well as evidence of its dual factors, psychological sources, cognitive mechanisms, and associated interventions. In the context of the many intergroup moral problems the world faces—restrictions on refugees, climate change, pandemics, regulation of global trade, and the public good of global scientific collaboration—our research may offer some hope that the human capacity for reasoned judgment may be up to the challenge.

Methods

All studies are approved under the Institutional Review Board at Princeton University, protocol 10027. Participants completed a consent form before proceeding to the study. International participants received a translated consent form in their native language.

Scale. To develop and validate the new scale, we conducted six studies with 7,388 total participants (*SI Appendix*, SI1.1–SI1.4). In this iterative process, we designed and refined online surveys asking participants to indicate the degree to which they agree or disagree with each statement (from 1, strongly disagree, to 6, strongly agree). Studies 1.1 to 1.3 were exploratory tests of face validity and incorporated expert opinions. Studies 1.4 and 1.5 were confirmatory tests on both samples of American and Chinese participants, as well as adding relevant psychological scales to test convergent and divergent validity, such as moral foundation scales and moral expansiveness scales. Study 1.6 was a full launch of our refined scale with international participants, which included behavioral measures of real-world charitable donations and free responses. These data appeared in our main analysis in the main text.

Predictive Validity. Regarding charitable donation (*SI Appendix*, SI2.1), participants read the following:

"We invite you to make a donation to a charitable cause. You can donate up to 10 of your Shells to one or both of the following organizations: Teach for All, an international NGO, or Teach for America [Teach First UK, and other country equivalent], a national NGO. The mission of both organizations is to improve education and expand opportunities for children in disadvantaged communities. Teach for All works worldwide, while Teach for America works in the US. Your donation decision will have a real-world effect, as we will donate as the way you suggested. Please indicate how much would you like to donate to: Teach for All (slider from 0 to 10) and Teach for America (slider from 0 to 10)."

Their total must sum to 10. In the self-interest condition, participants read "You could either keep the bonus to yourself, or donate up to 10 of the bonus to one or both of the following organizations."

Regarding the international trolley dilemma (*SI Appendix*, SI2.2), participants read the following:

"Suppose a runaway trolley is speeding down a track toward five middle-class citizens of our country. You are on a bridge over the tracks, and you can stop the trolley by putting something very heavy in front of it. As it happens, there is an immigrant next to you, and your only way to stop the trolley is to push him over the bridge and onto the track, killing him to save five. Is pushing him the right or wrong thing to do? 1 is completely wrong to 6 is completely right."

Regarding the COVID-19 petition (conducted in August 2020; *SI Appendix*, SI2.3 and SI2.4), participants read the following:

"In this section, we invite you to have your voice heard by the US Congress. On the next page, you will have a chance to influence a COVID-19 related petition we will send to Congressional Representatives on behalf of this study's participants. We will not mention your name, just the total number of participants and your collective choice. If you think that a petition should receive more support, please indicate by clicking on the retweet button. Retweet, broadly, means sharing the petition with other people. Just as in real life, not every retweet will get a response. You will choose one of the two petitions. We will send only one of these two petitions to members of Congress. The petition that receives most retweets will be sent out. So the more times you retweet, the more likely your desired petition, and the position you advocate, will be heard."

On the next page, they read the following:

"People all over the world are interested in receiving vaccines against COVID-19. But people advocate for different strategies on how those vaccines should be made and distributed. Petition A supports Vaccinate [Mask] the World. It proposes that the US government and companies should make, sell, and donate COVID-19 vaccines [masks] to everyone in the world. Petition B supports Vaccinate [Mask] America. It proposes that the US government and companies should make, sell, and donate COVID-19 vaccines [masks] to our own citizens. On the next page, you will have a chance to retweet for the petition you prefer. They then look at two retweet buttons, under each petition. After clicking the retweet, a number will show up reminding participants how many times they have already clicked."

Note this is a hypothetical but incentivized Twitter interface: participants get to experience the retweet button on the interface by clicking on it as many times as possible, but nothing gets retweeted. It is incentivized though because the strongly endorsed petitions were sent out to representatives of Congress.

Regarding personal opinions on world policies, we prepared patriotism (e.g., "I love my country"), immigration (e.g., "I would not like to have immigrants or foreign workers as my neighbors"), loyal resistance (e.g., "Sometimes a true patriot has to criticize our leaders"), evaluations of international organizations (e.g., "In general, international institutions do valuable work"), and CO_2 opinions (e.g., "Our country should reduce its carbon footprint because climate change is causing some islands nations to disappear"). (See details in *SI Appendix*, SI2.5.)

Mechanism from Linguistic Data. For linguistic modeling (*SI Appendix*, SI3.1 and SI3.2), with international samples, we used Google Translate (application programming interface) to translate all participants' responses into English and trained our computational linguistic models to classify moral cosmopolitans from moral patriots. Our model is based on long short-term memory recurrent neural networks with self-attention (36, 37) to predict donations from the text inputs. Just as humans do not process a full picture to understand what is going on but selectively attend to useful information, we trained our model to selectively detect influential linguistic cues by explicitly requiring the computation of attention weights for each input word when making predictions. For example, our linguistic model predicts the sentence "Both are important" to be moral cosmopolitan, while also yielding the attention weights for each word (0.42, 0.14, and 0.23), which means that "both" receives the most weight (0.42) and contributes significantly to this prediction.

For frequency counts (*SI Appendix*, SI3.3), in American post–COVID-19 samples, we gathered participants' responses in terms of five single words. After making retweet decisions, participants read the following:

"Please list briefly some thoughts you had while deciding and while retweeting. Single words are the best, one in each box. Please list at least one word, up to five."

We then classified these responses into Mask the World, Mask America, Vaccinate the World, and Vaccinate America groups. After basic data cleaning (e.g., remove stop words, lowercase, and lemmatization), we counted top words in each category (word clouds in *SI Appendix*, SI3.3).

Experiment. See preregistration report for design and model details (*SI Appendix*, SI4.1 and SI4.2). We present experimental stimuli here. Cosmopolitan treatment condition is as follows:

(1) I think that anyone in the world has the right to education, it is something similar to food, so it must be donated in equal parts to both organizations. (2) I believe everyone should have the same opportunity to earn a quality education, both in the US and all over the world. I would like to see other countries flourish and improve. (3) There is an immediate need for education worldwide especially for women and children living in impoverished countries. (4) Worldwide many students especially girls cannot get an education any other way. (5) Education is of paramount importance, both here and abroad. (6) I chose to donate half of the money to Teach for All because I think international education needs more help. Though I am keenly aware of the educational issues in America, I think that the world problem is much larger and needs more help. (7) I think providing education across world is important rather than just America. (8) Because the first [organization] acts globally and that includes America.

Patriot treatment condition is as follows:

(1) I think we should help our own first then broaden the scope. (2) [We] need to help our own country first. (3) We need to take care of our own

people and stop help[ing] the world—our country needs help for our children. (4) I believe we should help Americans first. (5) While I would like to support children in other countries, domestic issues are a priority for me. (6) While education is needed throughout the world, we need to remember that our USA children are in need also. (7) I think education opportunities should begin at home in America. However, I do recognize the need for assistance in other countries as well. (8) Education is good for everyone but more should be put at home.

Data Availability. Materials, data, code, preregistration, tables, and figures have been deposited in Open Science Framework (OSF), https://osf.io/63k9t/?view_only=5477b1adb0b24e05bb66698412d8c946. All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. For valuable comments, the authors thank the reviewers, Alexander Cappelen, Peter Singer, Bertil Tungodden, FiskeLab, and participants of the 2021 Society for Philosophy and Psychology.

1. N. Ellemers, *Morality and the Regulation of Social Behavior: Groups as Moral Anchors* (Routledge, 2017).
2. J. D. Greene, *Moral Tribes: Emotion, Reason, and the Gap between Us and Them* (Penguin, 2013).
3. J. K. Choi, S. Bowles, The coevolution of parochial altruism and war. *Science* **318**, 636–640 (2007).
4. E. G. Bruneau, M. Cikara, R. Saxe, Parochial empathy predicts reduced altruism and the endorsement of passive harm. *Soc. Psychol. Personal. Sci.* **8**, 934–942 (2017).
5. P. Molenberghs, J. Gapp, B. Wang, W. R. Louis, J. Decety, Increased moral sensitivity for outgroup perpetrators harming ingroup members. *Cereb. Cortex* **26**, 225–233 (2016).
6. C. K. De Dreu, D. Balliet, N. Halevy, "Parochial cooperation in humans: Forms and functions of self-sacrifice in intergroup conflict" in *Advances in Motivation Science*, A. J. Elliot, Ed. (Elsevier Science, 2014), 1, pp. 1–47.
7. S. Davidai, M. Ongis, The politics of zero-sum thinking: The relationship between political ideology and the belief that life is a zero-sum game. *Sci. Adv.* **5**, eaay3761 (2019).
8. M. B. Brewer, The psychology of prejudice: Ingroup love and outgroup hate? *J. Soc. Issues* **55**, 429–444 (1999).
9. A. Mummendey, S. Otten, Positive–negative asymmetry in social discrimination. *Eur. Rev. Soc. Psychol.* **9**, 107–143 (1998).
10. N. Halevy, G. Bornstein, L. Sagiv, "In-group love" and "out-group hate" as motives for individual participation in intergroup conflict: A new game paradigm. *Psychol. Sci.* **19**, 405–411 (2008).
11. T. Pogge, World poverty and human rights. *Ethics Int. Aff.* **19**, 1–7 (2005).
12. P. Singer, *The Expanding Circle* (Clarendon Press, Oxford, 1981).
13. D. G. Rand, M. A. Nowak, Human cooperation. *Trends Cogn. Sci.* **17**, 413–425 (2013).
14. Susan T. Fiske, "Helping: Pro-social behavior" in *Social Beings* (Wiley, New York, 4th Ed., 2018).
15. M. Mosleh, A. J. Stewart, J. B. Plotkin, D. G. Rand, Prosociality in the economic Dictator Game is associated with less parochialism and greater willingness to vote for intergroup compromise. *Judgm. Decis. Mak.* **15**, 1–6 (2020).
16. M. Snyder, W. Ickes, "Personality and social behavior" *Handbook of Social Psychology*, G. Lindzey, E. Aronson, Eds. (Random House, 1985), vol. 2, pp. 883–947.
17. P. Glick, S. T. Fiske, The ambivalent sexism inventory: Differentiating hostile and benevolent sexism. *J. Pers. Soc. Psychol.* **70**, 491 (1996).
18. S. McFarland, D. Brown, M. Webb, Identification with all humanity as a moral concept and psychological construct. *Curr. Dir. Psychol. Sci.* **22**, 194–198 (2013).
19. D. Crimston, P. G. Bain, M. J. Hornsey, B. Bastian, Moral expansiveness: Examining variability in the extension of the moral world. *J. Pers. Soc. Psychol.* **111**, 636–653 (2016).
20. F. Pratto, J. Sidanius, L. M. Stallworth, B. F. Malle, Social dominance orientation: A personality variable predicting social and political attitudes. *J. Pers. Soc. Psychol.* **67**, 741 (1994).
21. J. Graham *et al*., Mapping the moral domain. *J. Pers. Soc. Psychol.* **101**, 366–385 (2011).
22. G. Kahane *et al*., Beyond sacrificial harm: A two-dimensional model of utilitarian psychology. *Psychol. Rev.* **125**, 131–164 (2017).
23. M. H. Davis, Measuring individual differences in empathy: Evidence for a multidimensional approach. *J. Pers. Soc. Psychol.* **44**, 113 (1983).
24. R. Janoff-Bulman, S. Sheikh, S. Hepp, Proscriptive versus prescriptive morality: Two faces of moral regulation. *J. Pers. Soc. Psychol.* **96**, 521–537 (2009).
25. R. Janoff-Bulman, N. C. Carnes, Surveying the moral landscape: Moral motives and group-based moralities. *Pers. Soc. Psychol. Rev.* **17**, 219–236 (2013).
26. H. Tajfel, J. C. Turner, "An integrative theory of intergroup conflict" in *The Social Psychology of Intergroup Relations*, W. G. Austin, S. Worchel, Eds. (Brooks/Cole, 1979), pp. 33–37.
27. S. L. Gaertner, J. F. Dovidio, P. A. Anastasio, B. A. Bachman, M. C. Rust, The common ingroup identity model: Recategorization and the reduction of intergroup bias. *Eur. Rev. Soc. Psychol.* **4**, 1–26 (1993).
28. M. Cleveland, M. Laroche, Acculturaton to the global consumer culture: Scale development and research paradigm. *J. Bus. Res.* **60**, 249–259 (2007).
29. Y. Rosseel, Lavaan: An R package for structural equation modeling and more. Version 0.5–12 (BETA). *J. Stat. Softw.* **48**, 1–36 (2012).
30. E. T. Higgins, "Promotion and prevention: Regulatory focus as a motivational principle" in *Advances in Experimental Social Psychology*, M. P. Zanna, Ed. (Academic Press, New York, NY, 1998), 30, pp. 1–46.
31. M. Cikara, R. A. Farnsworth, L. T. Harris, S. T. Fiske, On the wrong side of the trolley track: Neural correlates of relative social valuation. *Soc. Cogn. Affect. Neurosci.* **5**, 404–413 (2010).
32. P. Foot, The problem of abortion and the doctrine of double effect. *Oxford Rev.* **5**, 5–15 (1967).
33. M. Hesse, Why might women be less eager to get the coronavirus vaccine? An investigation. *Washington Post*. https://www.washingtonpost.com/lifestyle/style/women-covid-vaccine-skeptical/2020/12/15/63551cac-3a61-11eb-9276-ae0ca72729be_story.html. Accessed 15 December 2020.
34. D. L. Ames, S. T. Fiske, Perceived intent motivates people to magnify observed harms. *Proc. Natl. Acad. Sci. U.S.A.* **112**, 3599–3605 (2015).
35. A. K. Ho *et al*., The nature of social dominance orientation: Theorizing and measuring preferences for intergroup inequality using the new SDO_7 scale. *J. Pers. Soc. Psychol.* **109**, 1003–1028 (2015).
36. S. Hochreiter, J. Schmidhuber, Long short-term memory. *Neural Comput.* **9**, 1735–1780 (1997).
37. Z. Lin *et al*., A structured self-attentive sentence embedding. arXiv [Preprint] (2017). https://arxiv.org/abs/1703.03130. Accessed 10 December 2018.
38. R. E. Petty, J. T. Cacioppo, "The elaboration likelihood model of persuasion" in *Communication and Persuasion* (Springer, New York, NY, 1986), pp. 1–24.
39. W. R. Fisher, The narrative paradigm: An elaboration. *Commun. Monogr.* **52**, 347–367 (1985).
40. S. Pinker, *The Better Angels of Our Nature. Why Violence Has Declined* (Viking, 2011).
41. J. Hoover, K. Johnson, R. Boghrati, J. Graham, M. Dehghani, Moral framing and charitable donation: Integrating exploratory social media analyses and confirmatory experimentation. *Collabra Psychol.* **4** (9) (2018).
42. J. Rawls, The law of peoples. *Crit. Inq.* **20**, 36–68 (1993).
43. B. Anderson, *Imagined Communities: Reflections on the Origin and Spread of Nationalism* (Verso Books, 2006).
44. P. James, Forms of abstract "community" from tribe and kingdom to nation and state. *Philos. Soc. Sci.* **22**, 313–336 (1992).
45. V. O. Packard, *A Nation of Strangers* (McKay, 1972).
46. L. Bian, S. Sloane, R. Baillargeon, Infants expect ingroup support to override fairness when resources are limited. *Proc. Natl. Acad. Sci. U.S.A.* **115**, 2705–2710 (2018).
47. B. Gaesser, Y. Shimura, M. Cikara, Episodic simulation reduces intergroup bias in prosocial intentions and behavior. *J. Pers. Soc. Psychol.* **118**, 683–705 (2020).
48. M. Kleiman-Weiner, R. Saxe, J. B. Tenenbaum, Learning a commonsense moral theory. *Cognition* **167**, 107–123 (2017).