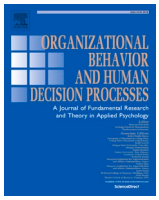
Organizational Behavior and Human Decision Processes 182 (2024) 104333

Contents lists available at ScienceDirect

Organizational Behavior and Human Decision Processes

journal homepage: www.elsevier.com/locate/obhdp



Different ally motivations lead to different outcomes: How self-transcendence and self-enhancement values predict effectiveness of self-identified allies☆

L. Taylor Phillips [a,*], Tamar A. Kreps [b], Dolly Chugh [a]

[a] Stern School of Business, New York University, New York, NY, United States
[b] Shidler College of Business, University of Hawai'i, Manoa, HI, United States

ARTICLE INFO

Keywords:
Allies/allyship
Schwartz values
Inequity/inequality
Privilege
Collective action

ABSTRACT

Dominant group members have different reasons for identifying themselves as allies to marginalized groups. How might these reasons relate to allies' effectiveness? We use Schwartz's values theory to integrate disparate work, focusing on two values that can underlie allyship: *self-transcendence*, or enhancing the welfare of others, and *self-enhancement*, or personal status and esteem. Across three yoked experiments ($N = 3016$), we tested how values relate to allies' intentions, behavior, and persuasiveness. Phase A of each study sampled self-identified allies (e.g., towards LGBTQ+ people, Black people, women). Both self-transcendence and, less consistently, self-enhancement predicted increased activism intentions; only self-transcendence predicted petition-signing behavior. Phase B sampled new participants, who viewed advocacy statements generated by allies in Phase A. We found that ally values affected audience reactions: ally self-transcendence was associated with greater persuasiveness, while self-enhancement was associated with lower persuasiveness. Although both values can generate ally engagement, self-transcendence may promote greater ally effectiveness.

1. Introduction

Organizations are increasing their efforts to become more inclusive, responding to recent social movements and business imperatives. These efforts have focused attention on the ways *dominant* group members—those who benefit from structural inequalities by virtue of their group membership on a given dimension—might exacerbate versus mitigate inequity (e.g., Cortland et al., 2017; Knowles, Lowery, Chow, & Unzueta, 2014; Phillips & Lowery, 2018; Wilkins & Kaiser, 2014). To mitigate inequity, dominant group members may become *allies*: that is, people who support the cause of a non-dominant, or marginalized, group they do not belong to (Ashburn-Nardo, 2018; Brown & Ostrove, 2013; Jun et al., 2023; Radke, Kutlaca, Siem, Wright, & Becker, 2020). For instance, men may ally with women and support #MeToo, White people may ally with Black people and support #BlackLivesMatter, and heterosexual-cisgender allies may support protections for the rights of lesbian, gay, bisexual, transgender, queer, and other gender and sexual minority people (LGBTQ+). Although which groups are "dominant" varies by context, dominant group members within a given context generally have increased access to the power to change minds and systems (Czopp & Monteith, 2003; Hekman et al., 2017; Jun et al., 2023; Rasinski & Czopp, 2010; Sherf, Tangirala, & Weber, 2017; Sidanius & Pratto, 1999), which can make them desirable coalition partners in social movements.

However, merely supporting a cause is not the same as helping it advance, and well-intentioned "allies" may not ultimately help the target group. In many contexts, being a self-declared ally is a socially desirable identity, and publicly amplifying others is a high-status behavior (Bain, Kreps, Meikle, & Tenney, 2021); some allies may thus join the cause in order to experience heightened status and admiration (e.g., "performative allyship"; Chugh, 2018; Crittle, 2017; Foster-Gimbel, Pillemer, & Phillips, 2022; Kristofferson, White, & Peloza, 2013). Such allies may not remain committed as time passes, or if the cause requires action rather than just good intentions, particularly if they are asked to bear personal risk or cost (Knowlton, Carton, & Grant, 2022; Sherf et al., 2017). Furthermore, some allies who do remain

☆ This article is part of the special issue "Allyship, Advocacy and Social Justice to Support Equality for Marginalized Groups in the Workplace" edited by Ellen Ernst Kossek, Jamie Ladge, Laura Little, Denise Loyd, Catherine Tinsley, and Alexis Washington.

* Corresponding author.
E-mail address: ltaylorphillips@nyu.edu (L.T. Phillips).

https://doi.org/10.1016/j.obhdp.2024.104333
Received 10 June 2022; Received in revised form 23 March 2024; Accepted 31 March 2024
Available online 24 April 2024

committed and active may nonetheless hurt the cause by expressing their message in a way that alienates, rather than persuading, potential supporters—for instance, with hostility (Knowlton et al., 2022). Thus, even self-professed and committed allies likely vary in their effectiveness. It is therefore important to understand what makes an ally not just well-intentioned, but also an active and persuasive advocate for the cause.

In this research, we examine how self-identified allies' personal values motivate their engagement with the cause (in both intentions and behavior), and their persuasiveness in advocating for it. Drawing on prior theoretical and qualitative work (Edwards, 2006; Russell, 2011; see also Schultz, 2001), we use Schwartz's values theory (Schwartz, 1992), which describes basic universal human values, as an organizing framework. In particular, we focus on *self-transcendence* values—valuing the welfare of others in one's society and the world at large—and *self-enhancement* values—valuing personal status, esteem, and success. Consistent with prior work (Edwards, 2006), we suggest that both of these values can provide routes toward ally engagement—but only self-transcendence is likely associated with more persuasive advocacy. Our work contributes to the field's understanding of what motivates allies by highlighting how distinct motivational routes can contribute to people's allyship. Moreover, it is also the first work to consider how values may be reflected in allies' communication with others, and ultimately affect how persuasive they are as allies. It thus has theoretical implications for the field's understanding of allyship, as well as practical implications for social movement leaders on how to motivate and deploy self-identified allies.

2. Schwartz values and collective action

We follow prior theoretical work in using the Schwartz values framework (Schwartz, 1992) to categorize ally motivations. *Values* are defined as generalized goals that motivate action and organize meaning, and are a stable individual difference (Fischer & Schwartz, 2011; Gouveia, Vione, Milfont, & Fischer, 2015; Schwartz, 2006; Uzefovsky, Döring, & Knafo-Noam, 2016; Vecchione et al., 2016). Values theory captures human motivations across contexts and cultures, in a single comprehensive structure. Over the past several decades, Schwartz's values framework has been validated repeatedly across hundreds of human populations, in over 60 countries (Sagiv & Schwartz, 2022; Schwartz & Boehnke, 2004; Schwartz, 1992, 1994, 2012); are robust across a larger range of societies than other frameworks (Smith, Peterson, & Schwartz, 2002); and unify many different motives or orientations developed across discrepant literatures (Smith et al., 2002). Because of this cross-cultural validity, Schwartz's values theory has been foundational in cross-cultural and social psychology, and is particularly well suited to describe motivation in allyship—a topic that by necessity includes multiple cultural sub-groups. In addition, Schwartz's values are theorized at the individual rather than national level (Hofstede, 2001, 2011; House et al., 2004; Smith et al., 2002; Schwartz, 2011), and have been used to describe and predict individual differences in behavior from voting to food consumption, in fields across the social and medical sciences (see reviews by Bardi & Schwartz, 2003; Sagiv & Roccas, 2021; Sagiv, Roccas, Cieciuch, & Schwartz, 2017; Schwartz, 1999; Sverdlik et al., 2012). By combining cross-cultural validity and universality with predictive validity and relevance at the individual level, Schwartz's values theory provides a particularly useful framework for assessing allies' motives and allyship behaviors.

Building on work by Rokeach (1977), Schwartz's theory of human values identifies ten near-universal human values (e.g., "security," "benevolence") which cluster into four cardinal categories: self-transcending, self-enhancing, conservation, and openness ("cardinal values"; Schwartz, 1992, 1994; see recent review by Sagiv & Schwartz, 2022).[1] Based on previous work, which we review in the next section, we focus in the current investigation on self-transcendence and self-enhancement values. *Self-enhancement* values include orientations towards personal status, achievement, and power, that is, personal social superiority and esteem; and *self-transcendence* values include orientations towards universalism and benevolence, that is, the enhancement of others' welfare and transcending selfishness (Schwartz, 1992, 1994, 2012; Tamir et al., 2016). Thus, these two cardinal values differ at their core regarding concern for the self versus concern for others (Schwartz, 2012). Theoretically and empirically, these values are uncorrelated: one can have high or low self-enhancement values independently of having high or low self-transcendence values (Schwartz & Boehnke, 2004).

2.1. Schwartz values and allyship

The study of allies has increased in recent decades as social justice movements have intensified and, simultaneously, structural inequalities have remained steep. Dominant group members can play an important role in facilitating, or hindering, social movements. In some cases, dominant groups are numerical majorities (e.g., cisgender people); in all cases, dominant groups control or have increased access to important resources. And yet, those who identify as "allies" want to help those outside their group. What motivates this desire; what motivates not just desire, but action; and does motivation affect whether the action is actually helpful or not?

Previous researchers have applied the Schwartz taxonomy of values toward understanding what motivates different allies, or allies at different stages of their development as allies. In an especially impactful conceptual model of interracial ally development, Edwards (2006) asserts that allies can have self-enhancing orientations—aiming to enhance their own sense of status and achievement—or self-transcending orientations—aiming to achieve justice, lifting up others in general and the collective "us." In another foundational work for ally/activism scholars, Schultz (2001) uses a similar taxonomy to describe environmental activists (i.e. allies of the natural environment), distinguishing egoistic allies who value self-enhancement (protecting the planet for personal status, power, and use) from biospheric allies who value self-transcendence (protecting the planet for all beings). These models, spanning different allyship issues (racial justice and environment), suggest that both self-transcendence and self-enhancement values offer paths towards ally engagement. Following these two qualitative investigations, we measure self-transcendence values to capture allies' orientation to address injustice in the entire system, and self-enhancement values to capture their orientation to demonstrate personal status.

2.2. Other work on ally motives

A key insight from earlier theoretical work is that self-enhancement may motivate allyship—even if that allyship is performative—because being an ally can provide status, impact, and a sense of achievement. This research dovetails with other work, not using a values framework, that yields a similar conclusion about people's sometimes self-enhancing reasons for engaging with collective action and allyship (Leach, Snider, Iyer, & Smith, 2002; Radke et al., 2020). Whereas some work on allyship supports the importance of self-transcending motives, e.g. an orientation towards universal justice among men allying with women (Drury & Kaiser, 2014; see also Thomas, McGarty, Reese, Berndsen, & Bliuc, 2016; Van Zomeren, Postmes, Spears, & Bettache, 2011), other work finds support for more self-enhancing motives, e.g. among straight, cisgender LGBT-rights activists (Russell, 2011). In a recent piece, Radke et al.

[1] Because the cardinal values of conservation and openness are not of primary theoretical importance, we provide detailed definitions in the SOM.

(2020) describe morality motives for allyship as including interest in universalism and harm-reduction - aligning with Schwartz (1992) original definition of self-transcending values (universalism and benevolence). Radke et al. (2020) go on to contrast these motives with a personal focused motive including desire for reputation, money, or popularity, which likewise aligns with Schwartz (1992) original definition of self-enhancing values (personal status, achievement, and power). Finally, critical race and decolonial theories suggest that dominant group members might suppress prejudice or acknowledge privilege not only to change racist systems, but also for performative or instrumental reasons, to protect their sense of self or material interests (Chow, Lowery, & Hogan, 2013; Kendi, 2019; Phillips & Lowery, 2018; Salter & Adams, 2013; see also Bergsieker, Shelton, & Richeson, 2010; Plant & Devine, 1998). Although not using the Schwartz model, this work together highlights how both self-transcending and self-enhancing motives can underlie otherwise similar intergroup behavior (e.g., supporting affirmative action).

Work on moral cognition, without a specific focus on intergroup relations, further supports the idea that both self-transcending and self-enhancing motives can underlie prosocial action. Justice and care are common moral concerns (Graham et al., 2011; Rai & Fiske, 2011), which can motivate people to engage in advocacy on behalf of others even with no personal gain (Effron & Miller, 2012), providing support for a self-transcending motive. At the same time, people may only engage in prosocial behaviors deemed necessary for self-presentation (Barasch, Levine, Berman, & Small, 2014; Lin, Schaumberg, & Reich, 2016; Monin & Miller, 2001; Mullen & Monin, 2016), and narcissistic people are more likely to show token public support for moral causes (Konrath, Ho, & Zarins, 2016), providing support for a self-enhancing motive. In short, prior research from diverse theoretical perspectives provides support for the idea that not only self-transcending values but also self-enhancing values can drive allyship intentions and behavior.

To integrate this work in our consideration of ally engagement and persuasiveness, we turn to Schwartz for two reasons. First, Schwartz provides a framework that is cross-culturally relevant, and especially robust across decades of work (Sagiv & Schwartz, 2022). Second, Schwartz's emphasis on underlying values – simple enough to be broadly relevant across cultures and behaviors – ultimately map on well to existing intergroup and moral cognition literature considered above. That is, existing work that speaks more specifically to allyship contexts nevertheless relies, sometimes explicitly (Edwards, 2006; Schultz, 2001) and sometimes latently (Chow et al., 2013; Konrath et al., 2016; Radke et al., 2020), on the tension between Schwartz's self-transcending and self-enhancing values.

Based on these bodies of prior work, we suggest that both self-enhancement and self-transcendence should drive allies to care about the allyship cause and to engage in activism, such as signing petitions and intending to protest. Activism behavior likely provides a sense of connection to others, focuses individuals beyond the self, and provides a means for achieving collective and justice goals (Bai et al., 2017; Piff, Dietze, Feinberg, Stancato, & Keltner, 2015). The more an ally values self-transcendence, the more they should seek these outcomes and thus engage in allyship. Simultaneously, engaging in activism for a cause one believes in can provide achievement and self-expansion (Besta & Zawadzka, 2019; see also Schultz, 2001), as well as status, impact, and power (Bain et al., 2021; Drury & Kaiser, 2014; Klar & Kasser, 2009; Konrath et al., 2016). The more allies value self-enhancement, the more they should seek these outcomes and thus engage in allyship behaviors. Thus, both self-enhancement and self-transcendence values should positively relate to intended and actual activism behavior.

H1a: Self-transcendence values will positively relate to engagement in allyship activism intentions.
H1b: Self-enhancement values will positively relate to engagement in allyship activism intentions.
H2a: Self-transcendence values will positively relate to engagement in allyship activism behavior.
H2b: Self-enhancement values will positively relate to engagement in allyship activism behavior.

3. Different values, differing persuasiveness

Even if both self-enhancement and self-transcendence values can motivate people to engage in allyship behavior, not all allies are effective. Might allies oriented towards self-enhancement engage in allyship behavior differently than allies oriented towards self-transcendence? Supporting this possibility, prior work by Kristofferson et al. (2013) has found that people who engage in public, low-cost, token advocacy—who perhaps value self-enhancement (see Konrath et al., 2016; Lee & Hsieh, 2013)—are less likely to stay engaged with the cause over time, compared to those who express support privately. This finding lends credence to the idea that different underlying values for allyship may lead to different patterns of behavior. Here, we specifically theorize that allies' underlying values may relate to their persuasiveness to others when promoting their cause.

We expect that the underlying values that motivate allies may express themselves in the style in which those allies communicate to others when writing advocacy messages. Prior research provides multiple reasons to make this prediction. One possibility is that allies may egocentrically try to frame messages in ways that center their own values, believing that such messages will be more persuasive (see Feinberg & Willer, 2013). Another intriguing possibility is that allies' values could show up in more unintended ways, as if those values were unconsciously "leaking" into the message and leaving a mark. Supporting this possibility, prior work has shown that a person's personality (Boyd & Pennebaker, 2017) and their values in particular (Boyd, Wilson, Pennebaker, Kosinski, Stillwell, & Mihalcea, 2015; Chen, Hsieh, Mahmud, & Nichols, 2014) affect subtle aspects of their language use such as verb tenses and pronouns (e.g., self-transcendence values are positively associated with using "we" or "us" pronouns, and self-enhancement values are negatively associated with using these pronouns; Chen et al., 2014). Prior work has also demonstrated that social advocacy messages can include subtle and implicit markers of an author's actual prejudice (such as expressing a belief that people are personally responsible for their outcomes: Jacoby-Senghor, Rosenblum, & Brown, 2021; see also Correll, Weisshaar, Wynn, & Wehner, 2020). Furthermore, other research has found that different kinds of advocacy messages can be more or less persuasive to audiences. For example, normative appeals that emphasize collective action and beliefs can be powerful motivators (Cialdini & Goldstein, 2004; Nolan, Schultz, Cialdini, Goldstein, & Griskevicius, 2008); allies canvassing door-to-door are more persuasive when they emphasize the perspective of others (rather than self; Broockman & Kalla, 2016); and personally-focused, angry language can be unpersuasive (DeSteno, Dasgupta, Bartlett, & Cajdric, 2004; Nabi, 2002).

Prior work has suggested the role of liking as a mechanism for these differences in persuasiveness. When audiences like messages more, they find those messages more persuasive (Crano & Prislin, 2006; Kruglanski & Thompson, 1999; McGinnies & Ward, 1980; Petty & Cacioppo, 1986). Moreover, audiences are often persuaded when they like either the message or the messenger, not always distinguishing between the two (Crano & Prislin, 2006). Although the vast literature on attitudes and persuasion continues to debate whether liking affects persuasion due to automatic/non-deliberative heuristics (e.g., the "likability heuristic" in dual-process models; Petty & Cacioppo, 1986) or for other reasons (see Kruglanski & Thompson, 1999), these different theories of persuasion converge in their agreement that liking is a critical antecedent to persuasion. As such, we measure audience liking as a key mechanism connecting allies' underlying values to the ultimate persuasiveness of their advocacy messages.

We hypothesize that audiences will be more persuaded by statements

from allies with self-transcendence values, and *less* persuaded by statements from allies with self-enhancement values. We expect this will happen because allies' values will affect the statements they write—whether through intentional strategy or unintentional "leakage" of their values into their language—and that self-transcending values will increase, and self-enhancing values will decrease, the likability of the statement and author. In particular, the Brunswik (1956) lens model has been used to show how underlying traits (such as values) can be detected by observers through implicit and latent cues (e.g., verbal, nonverbal; see Gosling, Ko, Mannarelli, & Morris, 2002). For instance, Black targets are able to detect underlying prejudicial and/or stereotyping beliefs of the messengers (Jacoby-Senghor et al., 2021; see also Bergsieker, Leslie, Constantine, & Fiske, 2012; Dupree & Fiske, 2019). Language likewise carries the residue of gender bias, as evidenced in professional performance reviews (Correll et al., 2020) and academic recommendations (Madera, Hebl, & Martin, 2009). We suggest author values may thus "leak" not only into an author's statement, but also, via audience liking, into audiences' own propensity to take action for the ally cause (e.g., Jacoby-Senghor et al., 2021). As reviewed above, self-transcending values include benevolence and universalism, which are likely associated with warmth and thus positive in perceivers' eyes. Self-enhancing values, on the other hand, are associated with self-focus and status seeking, which are likely perceived more coldly, especially in the context of allyship. Therefore, we suggest:

H3a: Self-transcendence values will positively relate to audience liking and thus persuasiveness.
H3b: Self-enhancement values will negatively relate to audience liking and thus persuasiveness.

4. Current research

Our work builds on the prior work reviewed above to provide the first quantitative, large-sample test of the prediction that both self-transcendence and self-enhancement values can spur ally behavior. In addition, we build on prior work examining ally beliefs and behavior as outcome variables, and add a focus on allies' persuasiveness. To test our predictions, we conducted three, two-phase yoked experiments (see Fig. 1 for method diagram and theoretical model). In Phase A of each study, we examined whether self-transcendence and self-enhancement values relate to engagement in activism behavior. In Studies 1 and 3, participants self-nominated a cause with which they were allied (i.e. advocated for a marginalized group they themselves did not belong to; e.g. Black Lives Matter, LGBTQ+ rights). In Study 2, we sampled self-identified straight (heterosexual), cisgender allies of lesbian, gay, bisexual, and transgender (LGBT) people, which allowed us to measure not only intentions to engage in activism but also real activism behavior (petition signing). We chose to sample self-proclaimed allies even though some of our participants were likely "allies" in name only, because this was precisely part of the variance we hoped to model; at the same time, understanding which dominant group members identify as "allies" at all was outside the scope of our investigation. In short, our aim was to understand predictors of how engaged and effective allies were once at least minimally aligned with the cause.

Next, in the second phase of each yoked study, we examined whether self-transcendence increases, but self-enhancement decreases, allies' persuasiveness. We presented advocacy statements written by author participants in Phase A to diverse audience participant samples. This approach allowed us to test how ally authors' values affected audience reactions to their advocacy statements. See Supplemental Online Material (SOM) for full materials and additional analyses. Survey materials and deidentified data are available here: https://researchbox.org/2728.

5. Study 1

Our first study examined how allies' values related to their engagement with a cause with which they claimed to be allied, as well as their advocacy persuasiveness as rated by a separate audience in the second

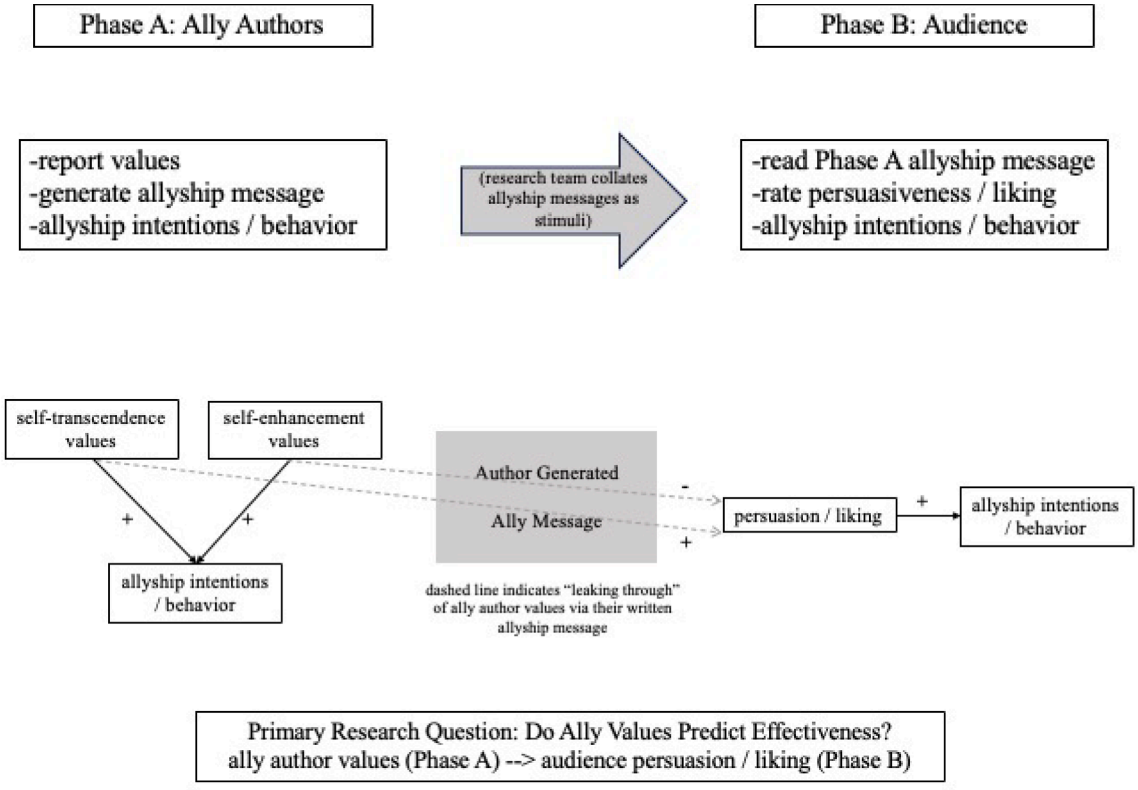


Fig. 1. Method Diagram & Theorized Model.

phase of the yoked design. In Phase A, we asked participants to identify a cause with which they were allied, and then generate statements about this cause, similar to ones they might share via social media. We then measured their personal values and their allyship engagement (activism intentions). We expected both self-transcendence and self-enhancement values to be positively associated with engagement (H1a-b). Then, in Phase B, we showed new participants a random sample of six advocacy statements written by participants in Phase A, and measured these new audience participants' responses—how much they liked the statements, and their own activism intentions after reading each one (H3a-b). We also measured audience petition signing for each issue they read about, to see whether participants who liked the messages more would actually engage in more activism.

5.1. Method

Participants. Table 1 presents participant demographics across all studies. We did not complete power analyses prior to data collection, particularly given unknown effect size. However, across studies we aimed to recruit $N = 100$ (and more when funding allowed) in Phase A, following current best practices for correlational designs of small-to-medium effect size (Simmons, Nelson, & Simonsohn, 2018). We aimed to recruit $N = 10$ per stimulus (and more when funding allowed) in Phase B, following best practices for stimulus sampling designs of unknown effect size (Judd, Westfall, & Kenny, 2012). All sample sizes were determined a priori.

In Phase A, we recruited 100 U.S. adult volunteers from Amazon's Mechanical Turk (MTurk). We received 116 responses (including incompletes, which we downloaded solely to check for duplicates), then removed 11 duplicate responses and 12 incompletes, for a final sample of $N = 93$. For readability, we refer to Phase A participants as "authors."

In Phase B, we recruited 300 U.S. adult volunteers from MTurk. We received 338 responses, then removed 43 duplicates and 25 incompletes, yielding $N = 270$. Each participant rated liking and activism intentions for six statements, yielding $N = 1620$ observations for these measures. They were offered the opportunity to sign three petitions, yielding $N = 810$ observations for this measure. For readability, we refer to Phase B participants as "audience" participants.

Table 1
Participant demographics, Studies 1–3.

Variable	*Study*					
	Study 1 (Phs A)	Study 1 (Phs B)	Study 2 (Phs A)	Study 2 (Phs B)	Study 3 (Phs A)	Study 3 (Phs B)
1. Age:						
Mean (*SD*)	32.30 (9.44)	24.66 (9.81)	35.14 (12.21)	35.27 (11.42)	38.96 (13.78)	39.92 (13.63)
2. Gender Identity:						
Female	48 %	44 %	48 %	52 %	47 %	49 %
Male	52 %	54 %	52 %	47 %	51 %	50 %
Non-binary or Unknown	–	2 %	–	<1%	1 %	1 %
3. Race:						
Asian/Asian-American	9 %	9 %	9 %	3 %	7 %	7 %
Black/African-American	8 %	9 %	8 %	7 %	7 %	7 %
Native-American	1 %	<1%	–	–	1 %	<1%
Latino/Latino-American	5 %	4 %	5 %	8 %	7 %	5 %
White/European-American	73 %	74 %	75 %	75 %	74 %	75 %
Multiracial/Other/Unknown	4 %	5 %	2 %	6 %	4 %	6 %

Procedure. In Phase A, participants ("authors") first identified a group-based cause that they supported, but for which they were not a member of the relevant group (e.g., a White participant supporting Black Lives Matter). Next, authors wrote social media statements about the cause, of the sort they would share online, into a Facebook interface. Finally, they completed measures of activism intentions, values, and demographic information. We had participants write statements before completing these quantitative measures, because we wanted to get a sample that best approximated the statements that participants might spontaneously post on real social media sites, with minimal priming or spillover from first completing closed-ended measures. However, to address the possibility that writing the statements might contaminate the later measures, we later used a time-lagged design in Study 3.

We used 87 of the 93 statements generated in Phase A as stimuli for Phase B. All of these statements were about one of the three most common issues participants identified: LGBTQ+ (49 statements), race (24), or gender (14). The remaining statements were about 6 various other issues and, for simplicity, were not used in Phase B.

In Phase B, participants ("audience") first completed demographics, including those relevant to this study's allyship issues (LGBTQ+ identity, race, gender). Then, they read a total of six social media statements generated by Phase A "authors" (see Table 5 for examples). Based on the approximate proportions of statements on each issue, we showed each "audience" participant three randomly selected LGBTQ+ statements, two randomly selected race statements, and one randomly selected gender statement. Statements were selected randomly for each audience participant from the full set, so across participants, all 87 statements were represented. Participants read statements grouped by issues (e.g. all three LGBTQ+ statements one after another), with issues in counterbalanced order. After reading each post, participants reported their liking for the post (as a measure of self-reported persuasion) and their activism intentions. Finally, after seeing all six statements about the three different issues, participants were offered the chance to sign one petition supporting equal rights for each group.[2]

Measures.[3]

Author Values were measured using the ten-item Short Schwartz Value Survey (e.g., "BENEVOLENCE (helpfulness, honesty, forgiveness, loyalty, responsibility)"; 1 = Opposed, to 8 = Extremely important; Lindeman & Verkasalo, 2005; see also Schwartz, 1992). Following previous research, we collapsed these into four cardinal values: self-transcendence (benevolence & universalism; $r = 0.50$), self-enhancement (power & achievement; $r = 0.51$), openness (self-direction, hedonism, & stimulation; $\alpha = 0.67$), and conservation (tradition, security, & conformity: $\alpha = 0.77$).

Author Activism Intentions were measured with ten items from the Activism Orientation Scale (Corning & Myers, 2002; e.g., "How likely is it that you will engage in each of the following activities in the future? participate in a protest supporting the cause"; 1 = Not at all, to 7 = Extremely; $\alpha = 0.95$).

Audience Liking was measured following each statement with ten

[2] We included items in Phase B to measure participants' group membership (e.g., LGBTQ+), their position on the issue (political support and allyship), and values. Audience individual differences were not our theoretical focus, but in exploratory analyses we were surprised to find no effects of these measures. We return to this issue in Study 2, and present full results in SOM.

[3] In Studies 1 and 2, we also included exploratory measures of correlates of values, including participants' sense of self (self-construal, contingencies of self-worth, identification with humanity, ally identity), and belief in a just world. In Study 2, we also measured moral foundations. Finally, in all studies we ran each allyship statement through LIWC (Pennebaker, Booth, & Francis, 2007) to explore some linguistic differences that have been found to correlate with different values (see Chen et al., 2014; Fredrickson, 2004; Magee & Smith, 2013; Tiedens, 2001), and we tested whether these linguistic differences may relate to audience perceptions. We found no consistent effects. These are not the focus of the current theory, and materials and results are in SOM.

items (from 1 = Not at all, to 7 = Extremely): "I liked this post", "I like the person who wrote the post", "I felt offended or annoyed by the post" (reversed), "The person who wrote the post has good intentions", "The person who wrote the post should write more posts on the topic", "This post would make [target group] members feel supported", "This post was persuasive", "This post will help the cause that the post was about", "I would share this post on Facebook", "I would 'Like' this post on Facebook". We combined all items (α = 0.93).

Audience Activism Intentions was measured following each statement with the same items as Phase A (α = 0.95).

Audience Petition Signing was measured by offering each participant the chance to sign petitions relevant to LGBTQ+ rights, racial minority rights, and women's rights (one petition per issue; Signed = 1; Did Not Sign = 0). Importantly, this is our key indicator of effective persuasion, following the focus on behavioral measures of persuasion in the existing literature (see Crano & Prislin, 2006).

5.2. Results

Analytic Approach. Empirical approaches to values theory have demonstrated that the best way to assess individual values is to consider the full values circumplex: the relative prioritization of target values among the full set of values is understood to be key to predictive validity. As such, the standard practice in this literature is to measure all ten constituent values, and include these simultaneously as controls, in order to isolate the effects of the specific values of interest (Lindeman & Verkasalo, 2005; Sagiv & Schwartz, 2022; Sandy, Gosling, Schwartz, & Koelkebeck, 2017; Schwartz, 2007, 2012). Therefore, across studies, we entered all values as simultaneous predictors and standardized all predictor measures, to: (1) reduce multicollinearity and rating bias; and, (2) more readily compare the unique variance (coefficients) contributed by the values of interest (self-transcendence and self-enhancement). However, given the additional values in the circumplex (e.g., self-conservation) are not the focus of our theorizing, we report these details in SOM.

Table 2
Correlations among variables, Studies 1–3.

Variable	Mean (SD)	1	2	3	4
1. Self-Transcendence	S1 5.94 (1.49) S2 5.90 (1.49) S3 6.08 (1.16)	–	–	–	–
2. Self-Enhancement	S1 4.45 (1.59) S2 4.58 (1.49) S3 4.12 (1.28)	S1 0.07 S2 0.10 S3 0.04	–	–	–
3. Openness	S1 4.97 (1.40) S2 5.00 (1.41) S3 5.08 (1.16)	S1 0.32* S2 0.27* S3 0.51*	S1 0.54* S2 0.48* S3 0.36*	–	–
4. Conservation	S1 5.07 (1.65) S2 4.72 (1.58) S3 4.80 (1.42)	S1 0.19 S2 0.25* S3 0.18*	S1 0.46* S2 0.36* S3 0.36*	S1 0.22* S2 0.10 S3 0.02	–
5. Activism Intentions	S1 4.36 (1.53) S2 3.68 (1.59) S3 4.01 (1.50)	S1 0.46* S2 0.35* S3 0.36*	S1 0.17 S2 0.13 S3 0.05	S1 0.11 S2 0.24* S3 0.25*	S1 0.05 S2 -0.06 S3 -0.20*

Note. *$p < 0.05$.

Phase A: Authors. Table 2 presents descriptive statistics for Studies 1–3. Our main analyses used the four cardinal values as predictors (Table 3a), but we also conducted analyses using the ten constituent values (Table 3b). All results were robust to political ideology controls (see SOM).

As expected, both author self-transcendence ($b = 0.81$, $SE = 0.15$, $t(88) = 5.33$, $p < .001$) and author self-enhancement ($b = 0.46$, $SE = 0.19$, $t(88) = 2.45$, $p = .016$) positively predicted author activism intentions.

Phase B: Audience. To account for our multiple observations per participant, we used mixed model analyses, with random intercepts by topic, order, and participant. The fixed effects were the statement *author's* four cardinal values (standardized). We used one model to predict audience liking, and a separate model to predict audience activism (Table 6).

As expected, audience liking was positively predicted by author self-transcendence ($b = 0.06$, $SE = 0.03$, $t(625) = 2.12$, $p = .035$) and *negatively* predicted by author self-enhancement ($b = -0.13$, $SE = 0.04$, $t(1137) = -3.41$, $p < .001$). Unexpectedly, audience activism intentions were not related to either author value ($ps > 0.220$).

Audience Petition Signing and Mediation by Liking. We next tested whether this relationship between author values and audience ratings would translate into actual differences in audience behavior. Recall that each participant had an opportunity to sign only one petition for each topic, even when they had read multiple statements on that topic. In our preliminary analyses predicting petition signing, we therefore used issue as the level of analysis. We created composites of author values and audience liking for each issue by averaging across each participant's statements for that issue. We then used generalized mixed models with random-intercepts by topic and participant. Author values had no total effects on audience petition signing ($ps > 0.555$). However, in a separate model, audience liking did predict petition signing ($b = 1.56$, $SE = 0.33$, $z = 4.69$, $p < .001$).

Hence, we tested for an indirect effect of author values on audience petition signing via liking. There may be indirect effects even when total effects are nonsignificant (Rucker, Preacher, Tormala, & Petty, 2011). We returned to advocacy statement as the level of analysis, and for each one we coded whether participants signed the relevant petition (yes or no). We then used the piecewiseSEM package in R (Lefcheck, 2016) to perform a path analysis: this allowed us to test for indirect effects while

Table 3a
Effect of values on activism intentions (Studies 1–3; Phase A Ally Authors).

Study 1 (Phase A)	Activism intentions			
	b	95 % C.I.	*SE*	*t*
Intercept	4.36	[4.08,4.64]	0.14	31.35***
Self-enhancement	0.46	[0.09, 0.83]	0.19	2.45*
Self-transcendence	0.81	[0.52, 1.10]	0.15	5.33***
Openness	−0.30	[−0.65, 0.05]	0.18	−1.67
Self-conservation	−0.22	[−0.53, 0.09]	0.16	−1.35
Study 2 (Phase A)	**Activism intentions**			
	b	95% C.I.	*SE*	*t*
Intercept	3.68	[3.45, 3.92]	0.12	31.92***
Self-enhancement	0.17	[−0.10, 0.44]	0.14	1.19
Self-transcendence	0.56	[0.32, 0.80]	0.12	4.52***
Openness	0.19	[−0.08, 0.46]	0.14	1.34
Conservation	−0.32	[−0.57, −0.07]	0.13	−2.44*
Study 3 (Phase A)	**Activism intentions**			
	b	95% C.I.	*SE*	*t*
Intercept	4.00	[3.80, 4.20]	0.10	39.81***
Self-enhancement	0.25	[0.01, 0.49]	0.12	2.05*
Self-transcendence	0.64	[0.40, 0.88]	0.12	5.14***
Openness	−0.03	[−0.28, 0.22]	0.13	−0.22
Conservation	−0.50	[−0.72, −0.28]	0.11	−4.43***

Note. *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$. All values standardized.

Table 3b
Effect of values on activism intentions (Studies 1–3; Phase A Ally authors).

	Activism intentions (Study 1 Phase A)			Activism intentions (Study 2 Phase A)			Activism intentions (Study 3 Phase A)		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*
Intercept	4.36	0.14	30.52***	3.70	0.12	31.51***	4.01	0.10	40.16***
Power	0.13	0.20	0.66	−0.00	0.14	−0.01	−0.04	0.12	−0.32
Achievement	0.40	0.21	1.96^{+}	0.12	0.15	0.43	0.25	0.12	2.13*
Universalism	0.50	0.19	2.62*	0.41	0.14	2.84**	0.51	0.13	3.85**
Benevolence	0.40	0.19	2.11*	0.33	0.15	2.22*	0.29	0.12	2.35*
Hedonism	−0.16	0.19	−0.82	−0.04	0.16	−0.24	−0.13	0.11	−1.14
Self-Direction	−0.18	0.21	−0.86	−0.20	0.15	−1.32	−0.12	0.13	−0.94
Stimulation	−0.12	0.21	−0.58	0.41	0.18	2.32*	0.25	0.12	2.05*
Security	−0.02	0.20	−0.10	−0.12	0.15	−0.82	−0.25	0.12	−2.06*
Conformity	−0.22	0.21	−1.03	−0.15	0.17	−0.85	−0.04	0.13	−0.31
Tradition	0.01	0.23	0.03	−0.05	0.17	−0.28	−0.25	0.14	−1.76

Note. *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$. All values standardized.

Table 4
Effect of values on petition interest and petition signing (Study 2, Phase A Ally Authors).

Study 2 (P1)	Petition Interest			Petition Signing		
	b	*SE*	*z*	*b*	*SE*	*z*
Intercept	−1.05	0.19	−5.46***	−0.51	0.17	−3.00**
Self-enhancement	0.05	0.22	0.25	−0.16	0.21	−0.78
Self-transcendence	0.75	0.22	3.34**	0.59	0.20	3.04**
Openness	−0.15	0.22	−0.67	−0.14	0.20	−0.72
Conservation	−0.23	0.20	−1.16	0.04	0.19	0.22

Notes. $^{+}$ $p < .10$; *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$. All values standardized.

Table 5
Examples of advocacy statements in Study 1 Phase B (written by participants in Study 1 Phase A).

LGBTQ+ Rights Focus	Women's Rights Focus	Racial Justice Focus
I believe anyone should be able to love whoever they want to. As long as they respect people and are fair in life you can be whatever you want to be I don't think it's anyone else's business to judge.	I think it is only reasonable that women should receive equal pay for doing the same work and job as anyone else. There is no logical or fair reason that any company or employer should expect to pay a woman less for the same responsibility, production and hours that any other employee puts in.	It continually blows my mind how people treat others differently based on the color of skin. It is shallow and so unfair, and especially un-Christian.
Every human deserves equal rights. Regardless if you are gay, straight, or transgender, we are all humans.	No woman should get paid less for doing the same work as a man!	Crazy thought here, but maybe…, people that aren't white deserve to be treated like humans too.
Marriage equality is a right, not a privilege.	Yeah, don't really want to get preachy but I just think it should be up to the woman is all. I don't think it's any other person's place to decide what they do.	If black lives matter then why do the cops keep firing at them?

still accounting for the multi-level structure of our data. We modeled paths for the effect of each author cardinal value on audience liking, including random-intercepts by topic, participant, and order; and a path for the effect of liking on petition signing, including random-intercepts by topic and participant. This model showed high goodness-of-fit (Fisher's $C(8) = 5.69$, $p = .682$; BIC = 101.45), indicating a significant indirect effect of author values on petition signing through liking (Fig. 2). Specific path coefficients revealed significant effects of each author value on audience liking, including a positive path for self-transcendence ($b = 0.06$, $SE = 0.03$, $p = .035$) and a negative path for self-enhancement ($b = -0.13$, $SE = 0.04$, $p < .001$). In turn, audience liking affected petition signing ($b = 0.92$, $SE = 0.17$, $p < .001$). Thus, allies' values can influence their effectiveness in persuading others, which can then influence others' actual petition signing behavior. Most relevant to our theorizing, the more an author valued self-transcendence, the more audiences liked their statements, and hence the more they signed the petition; and the more an author valued self-enhancement, the *less* audiences liked their statements, and hence the less they signed the petition.

5.3. Discussion

In Study 1, we found that both author self-transcendence and author self-enhancement predicted increased activism intentions, supporting H1a and H1b. In turn, author self-transcendence also predicted more positive audience reactions to advocacy messages, supporting H3a. In contrast, allies' self-enhancement values led to more negative audience reactions to advocacy messages. In turn, audiences' liking ultimately related to their willingness to take action by signing petitions supporting the allyship cause, supporting H3b. Thus, allies' values were not only associated with their own activism intentions as allies, but also—by leaking through into the types of advocacy statements they made—the responses and ultimately the engagement of others who heard them.

Although we did find an indirect effect of ally author values on audience behavior via audience liking, we did not find a significant total effect where people were altogether more likely to sign the petition after reading statements from self-transcending authors. This may be because our design allowed us to collect just one binomial measure of audience behavior per multiple author statements. In other words, our indirect effects held even though audience members read several other statements: the effect of an author's self-transcendence was strong enough to indirectly make an audience member more likely to sign a petition, even when that person also read other statements before having the opportunity to sign. We later addressed this issue in Study 3, in which we were able to more precisely measure how a single statement affected audience behavior.

6. Study 2

Study 2 was similar to Study 1 but used just a single issue—LGBT[4] rights—instead of allowing participants to choose their own allyship cause. This change had several advantages. First, we expected that this

[4] Although other groups (e.g., Queer, Intersex, Asexual) are often included in more expansive versions of this acronym, we specifically used the acronym LGBT in our Study 2, and therefore use "LGBT" when referring to procedures, materials, data, or results for this study.

Table 6
Effect of author values on audience reactions (Studies 1–3; Phase B Audience).

Study 1 (Phase B)	Audience Liking			Audience Activism intentions			Audience Petition Signing		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	z
Intercept	3.93	0.09	43.92***	3.16	0.11	28.94***	−9.13	0.81	−11.31***
Self-enhancement (author)	−0.13	0.04	−3.41***	−0.04	0.03	−1.23	−0.13	0.40	−0.33
Self-transcendence (author)	0.06	0.03	2.12*	0.01	0.03	0.58	−0.01	0.29	−0.04
Openness (author)	0.19	0.03	5.49***	0.07	0.03	2.31*	0.19	0.31	0.62
Conservation (author)	−0.10	0.03	−3.00**	−0.10	0.03	−3.50***	−0.10	0.30	−0.33
Study 2 (Phase B)	**Audience Liking**			**Audience Activism intentions**			**Audience Petition Signing**		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*
Intercept	4.21	0.06	64.85***	3.39	0.10	35.31***	−0.34	0.12	−2.95**
Self-enhancement (author)	−0.09	0.03	−3.56***	−0.05	0.02	−2.13*	−0.39	0.44	−0.90
Self-transcendence (author)	0.07	0.02	3.10**	0.01	0.02	0.54	0.46	0.36	1.27
Openness (author)	0.06	0.02	2.57**	0.01	0.02	0.65	0.58	0.42	1.38
Conservation (author)	−0.02	0.02	−0.89	−0.02	0.02	−0.91	−0.42	0.40	−1.05
Study 3 (Phase B)	**Audience Persuasion**			**Audience Liking**			**Audience Activism intentions**			**Audience Petition Signing**		
	b	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*t*	*b*	*SE*	*z*
Intercept	3.53	0.04	85.79***	4.06	0.05	86.79***	2.87	0.04	65.99***	−0.48	0.05	−9.69
Self-enhancement (author)	0.06	0.05	1.20	0.10	0.06	1.78	0.09	0.05	1.65	0.03	0.06	0.49
Self-transcendence (author)	0.12	0.05	2.39*	0.17	0.06	2.91**	0.05	0.05	0.91	0.10	0.06	1.52
Openness (author)	−0.14	0.05	−2.59*	−0.20	0.06	−3.16**	−0.05	0.06	−0.95	−0.01	0.07	−0.15
Conservation (author)	−0.03	0.05	−0.58	−0.07	0.05	−1.23	−0.07	0.05	−1.30	−0.08	0.06	−1.42

Notes. $^{+}p < .10$; $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. All values standardized.

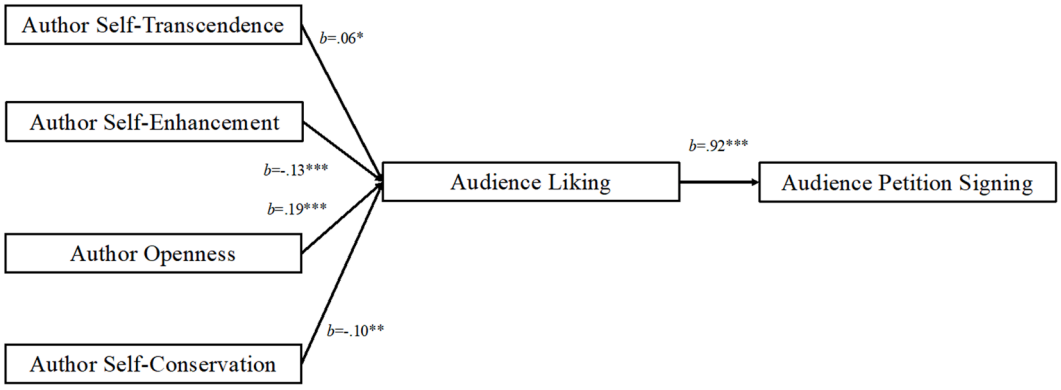


Fig. 2. Structural Equation Model for Study 1. Notes. $^{*}p < 0.05$; $^{**}p < 0.01$; $^{***}p < 0.001$. All values standardized.

would increase variance in ally engagement. In Study 1, when authors self-nominated their own cause, they probably chose causes they were relatively engaged with. We expected that choosing a cause for participants would increase the number of participants who were only minimally identified or engaged (while still sampling many who were highly engaged). Second, using a single issue allowed us to offer statement authors, not just audience members, an opportunity to sign a petition and thus measure authors' real activism behavior in addition to intentions. Third, by using a single issue we could make sure to recruit audience participants who were part of the target group as well as those who were part of the dominant group (whether allied or not). This allowed us to test how audience members' own identities might affect the way they responded to advocacy messages.

We chose to use LGBT rights as the issue, because the largest number of Study 1 participants identified themselves as allies to lesbian, gay, bisexual, and/or transgender people. Thus, in Phase A, our sample comprised participants who identified as allies for this cause (77–78 % of those initially recruited). In Phase B, we over-sampled LGBT participants so that we could explore how they might respond to different ally statements (see Jacoby-Senghor et al., 2021). Other than the addition of author petition signing in Phase A, we measured the same things as in Study 1.

6.1. Method

Participants. *Phase A (Authors):* We recruited 200 U.S. adults from MTurk. We received 216 responses. We then removed 5 duplicates and 4 incompletes; 18 who identified as LGBT; and 28 who answered "Not at

all" to the question "to what extent do you identify as an ally toward LGBT individuals.[5] Thus, our final sample was $N = 161$ allies.[6]

Phase B (Audience): As discussed above, we wanted to recruit an audience sample that included both members of the target group (LGBT) and dominant group members (cisgender/heterosexual). All participants, who were U.S. adults recruited on MTurk, had previously completed a separate pre-screening survey in which they indicated their sexual orientation and gender identities. We recruited 130 who had pre-identified as LGBT, and 200 who had pre-identified as cisgender and heterosexual. We received 500 responses, then removed 55 duplicates and 128 incompletes, yielding $N = 317$. Each participant rated ten posts, yielding $N = 3170$ observations.

Procedure. In Phase A, participants ("authors") wrote social media statements about LGBT rights, then completed the same measures of activism intentions and values as in Study 1. Next, authors were given the opportunity to sign a pro-LGBT-rights petition: we asked whether they were interested in signing a "petition to support equal rights for LGBT individuals" (Yes = 1; No = 0), and then gave all authors the opportunity to actually sign by inputting their initials and email address (Signed = 1; Did Not Sign = 0). Finally, they completed demographic information.[7] We used 165 posts from Phase A authors as stimuli for Phase B, omitting 28 posts that research staff judged to be harmful or irrelevant (e.g., supporting conversion therapy for LGBT people).[8]

In Phase B, "audience" participants first completed demographics to confirm pre-screened sexual orientation and gender identity, then they read ten statements randomly selected from among the 165 statements from Phase A. After reading each statement, participants rated the post (liking) and reported their own activism intentions. At the end of the study, after reading and rating all ten statements, participants were offered one opportunity to sign a petition for LGBT equal rights.

Measures. All measures were the same as in Study 1, with three exceptions. First, Phase A participants were offered the chance to sign a single petition for LGBT rights at the end of their survey. Second, and likewise, the petition at the end of the Phase B study was a single petition for LGBT rights. Third, we added six more items to the audience liking measure, for a total of sixteen items: "This post would be persuasive to non-LGBT individuals", "This post was a good allyship post", "Allies should not write posts like this (reversed)", "The person who wrote this post would be supportive toward LGBT individuals", "The person who wrote this post wants to help LGBT causes", "The person who wrote this post is a good ally." We originally intended these new items to form a separate composite capturing liking for the author rather than the post. However, they were highly correlated with the other ten items, so we combined all sixteen ($\alpha = 0.95$).

6.2. Results

Phase A: Authors. Replicating Study 1, author self-transcendence positively predicted activism intentions ($b = 0.56$, $SE = 0.13$, $t(156) = 4.52$, $p < .001$; Table 3a). Unexpectedly, author self-enhancement was not a significant predictor ($b = 0.17$, $SE = 0.14$, $t(156) = 1.19$, $p = .236$).

We used binomial regressions to test how authors' values related to their self-reported interest in signing the petition, and actual signing behavior. As predicted, author self-transcendence positively predicted interest in signing ($b = 0.75$, $SE = 0.22$, $z = 3.34$, $p < .001$) and actual signing ($b = 0.59$, $SE = 0.20$, $z = 3.04$, $p = .002$; Table 4). In contrast, self-enhancement had no effect on petition interest ($b = 0.05$, $SE = 0.22$, $z = 0.25$, $p = .806$) or signing ($b = -0.16$, $SE = 0.21$, $z = -0.78$, $p = .436$).

Phase B: Audience.[9] As expected, and replicating Study 1, author self-transcendence was positively related to audience liking ($b = 0.07$, $SE = 0.02$, $t(2694) = 3.10$, $p = .002$; Table 6), and again unrelated to audience activism intentions ($p = .589$). Also as expected, author self-enhancement was *negatively* related to audiences' liking ($b = -0.09$, $SE = 0.03$, $t(2676) = -3.56$, $p < .001$) and activism intentions ($b = -0.05$, $SE = 0.02$, $t(2635) = -2.13$, $p = .033$).

Audience Petition Signing and Mediation by Liking. As in Study 1, we averaged across the values of the authors of each participant's ten statements, then performed binomial logistic regressions predicting audience petition signing. We again found no total effect of any author values on audience petition signing ($ps > 0.195$). However, replicating Study 1, we did find that audience liking predicted audience petition signing ($b = 0.94$, $SE = 0.14$, $z = 6.85$, $p < .001$). We therefore performed the same piecewiseSEM path analysis, with paths specified as in Study 1. Replicating Study 1, this model showed high goodness-of-fit (Fisher's $C(8) = 3.29$, $p = .915$; BIC = 83.14), indicating a significant indirect effect of author values on audience petition signing, through audience liking (Fig. 3). Specific path coefficients revealed significant effects of three author cardinal values on audience liking, including a positive path for self-transcendence ($b = 0.07$, $SE = 0.02$, $p = .002$) and a negative path for self-enhancement ($b = -0.09$, $SE = 0.03$, $p < .001$). In turn, audience liking significantly affected audience petition signing ($b = 0.45$, $SE = 0.03$, $p < .001$).

6.3. Discussion

Replicating Study 1, Study 2 found that self-transcendence values among heterosexual, cisgender allies predicted increased activism intentions, supporting H1a. Furthermore, self-transcendence values predicted increased likelihood of a real activism behavior—signing a petition for LGBT rights—supporting H2a. These results replicate our earlier findings in a group of allies who were assigned their cause (rather self-assigning a cause), and thus support a positive relationship between self-transcendence motives and ally engagement.

We found mixed results for self-enhancement. We did find that author self-enhancement was associated with decreased audience liking, echoing the finding from Study 1 that self-enhancement values in allies may be associated with lower persuasiveness. However, we did not find that authors' self-enhancement values predicted their own activism intentions; in other words, we did not find evidence in this study that allies with high self-enhancement values were more committed allies. However, in a later replication study (see SOM), authors' self-enhancement values *did* predict their own activism intentions, replicating Study 1. These mixed findings partially support H1b.

[5] Results are robust to using different lower-bound cutoffs for being an "ally" (see SOM).

[6] In this context, individuals who report "allyship" could in theory support policies that are actually opposed to LGBT interests, such as conversion therapy. Therefore, we measured seven additional items: allyship toward "LGBT marriage rights" and "LGBT bathroom rights," and political support for equal rights, marriage rights, bathroom rights, anti-discrimination policies, and LGBT individuals ($\alpha = 0.88$; in Phase 1, political support measured via follow-up survey, $N = 129$). These measures were highly correlated with our screening measure ($rs < 0.79$), and our sample highly endorsed relevant policies ($M = 5.88$, $SD = 1.14$), suggesting our screening measure was appropriate.

[7] In Study 2 Phase A, we also attempted to manipulate values. However, our manipulations had no effect on any of our measures (SOM). This is consistent with literature suggesting values are a stable trait, not a temporary mindset (Schwartz, 2006).

[8] This included statements from Phase A participants who were later removed as incomplete or duplicate responses.

[9] For each of the Phase B analyses, we also performed the same analysis adding audience identity (LGBT vs. cisgender-heterosexual) and its interaction with the independent variable. Then, we performed the same analyses adding audience identity (ally identification; linear) and its interaction with the independent variable. Unexpectedly, we found no interactions of audience identity in any of these models, $ps > 0.11$. We therefore report analyses not accounting for audience identity, using the same statistical models as in Study 1 (see Table 6).

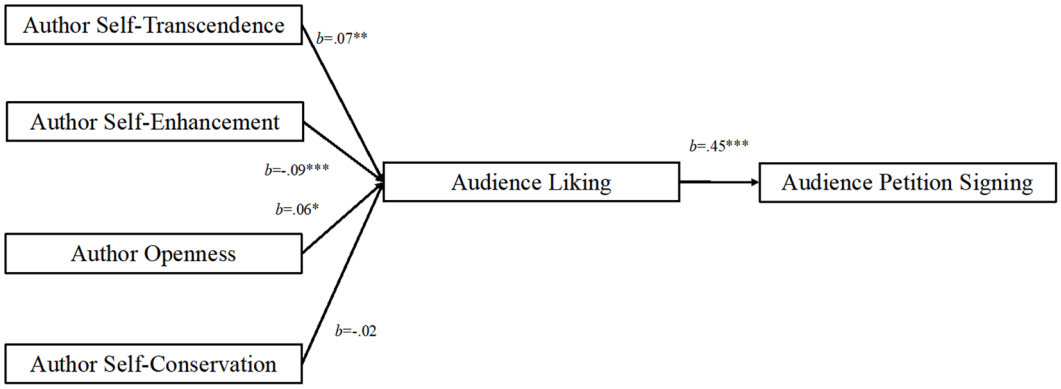


Fig. 3. Structural Equation Model for Study 2. Notes. *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$. All values standardized.

Unexpectedly, authors' self-enhancement again did not predict authors' petition signing interest or behavior, thus not supporting H2b (Study 2). One reason for this may be the relatively private nature of petition-signing: perhaps self-enhancing values only lead to forms of activism that are more publicly observable (e.g., Konrath et al., 2016; Kristofferson et al., 2013; Monin & Miller, 2001). Participants might have imagined that no one would see or admire their petition signatures, so this behavior would not satisfy their self-enhancement values. In contrast, the activism intentions scale also included public activities such as protesting and giving a speech, which may explain why self-enhancement values did relate to scores on this scale more consistently across studies.

In Phase B, allies' values again leaked through to affect audiences' impressions of their advocacy messages. Replicating the pattern in Study 1, author self-transcendence values were related to more positive audience impressions (increased liking), and author self-enhancement values were related to more negative audience impressions (here, both decreased liking and decreased activism intentions). These audience impressions affected audience members' own allyship: replicating Study 1, the more audience members liked the statements, the more they signed a relevant petition, supporting H3a and H3b.

Finally, focusing on a single issue in this study gave us the opportunity to see whether reactions to ally messages might differ between audience members who were versus were not part of the target group (here, LGBT people). Although we did not have specific predictions about any such differences, we were intrigued to find no differences in this sample. Future research could further investigate the intuitively plausible possibility that target group members may react differently to messages than dominant group members do (see also Jacoby-Senghor et al., 2021; Kutlaca, Radke, & Becker, 2022). In the current study, however, audience members across the board were more persuaded after reading statements from more self-transcending allies, and less persuaded after reading statements from more self-enhancing allies.

7. Study 3

Study 3 sought to replicate the results of Studies 1–2, with two changes to increase robustness. First, we used a time-lagged design for Phase A, first measuring values and three weeks later collecting the social media statement and measuring activism intentions. This method reduces common method bias (Podsakoff, MacKenzie, Lee, & Podsakoff, 2003) that might artificially inflate the correlation between measures. The time-lag between measures also addressed the possibility that, in earlier studies, the need to appear consistent across measures could have caused later measures to be influenced by earlier ones, such as values being influenced by writing a statement first, or activism intentions being influenced by reporting values first. Second, we offered only a single advocacy message to our Phase B audience participants, so that we could measure more precisely how a single statement might affect audience behavior. Study 3 (Phase B) was pre-registered: https://aspredicted.org/nz8dg.pdf.

7.1. Method

Participants. *Phase A (Authors):* we recruited 300 U.S. adults from Prolific to complete the Time 1 survey, then three weeks later invited these participants back again (via one message) to complete the Time 2 survey, which we left open for 2 days. We received 214 Time 2 responses, then removed 37 incompletes, for a final sample of $N = 177$.[10]

Phase B (Audience): In Phase B, we recruited 2000 U.S. adult volunteers from Prolific. We received 2094 responses, then removed 96 incompletes, yielding $N = 1998$.

Procedure. First, in the Time 1 survey, Phase A participants ("authors") completed measures of values (as in Study 1) and demographic information. Three weeks later, in the Time 2 survey, authors identified an allyship cause (described as in Study 1), wrote a statement, and reported activism intentions (as in Study 1).

We used 167 of the 177 statements generated in Phase A as stimuli for Phase B. As in Study 1, most statements were about issues related to LGBTQ+ (75 statements), race (70), or gender (17). Unlike in Study 1, we included an additional 15 statements that were about various other issues, in order to better encompass the full range of allyship issues without excluding less commonly mentioned issues. We omitted 10 statements which were judged by two research assistants to be either off topic (e.g., "I probably would not write a post") or harmful.

Phase B ("audience") participants read a single statement selected at random. They then reported how persuasive the post was, how much they liked the post, and their activism intentions. Finally, they were offered the chance to sign a petition supporting the issue they had read about (by inputting their initials and email address, as in Studies 1–2, but here referring only to "the issue" rather than to LGBTQ+, racial, or women's rights specifically).

Measures. All measures were the same as in Study 1, with one exception. We included a separate 9-item scale of audience persuasion (1 not at all – 7 extremely; $\alpha = 0.91$; Thomas, Masthoff, & Oren, 2019) to complement our face-valid liking measure.

7.2. Results

Phase A: Authors. As expected, both author self-transcendence ($b = 0.64$, $SE = 0.12$, $t(172) = 5.14$, $p < .001$) and author self-enhancement ($b = 0.25$, $SE = 0.12$, $t(172) = 2.05$, $p = .042$) positively related to author activism intentions.

[10] We found that attrition from Time 1 to Time 2 was not significantly related to Schwartz values nor political ideology (p's > 0.279). This suggests we were successful in our goal of soliciting unbiased reports of values.

Phase B: Audience. As in previous studies, we used mixed model analyses, with a random intercept by message. The fixed effects were the author's four cardinal values (standardized) corresponding to each statement. We predicted persuasion, liking, and audience activism in separate models (Table 6).

As expected, audience persuasion was positively predicted by author self-transcendence ($b = 0.12$, $SE = 0.05$, $t(163) = 2.39$, $p = .018$). Likewise, audience liking was positively predicted by author self-transcendence ($b = 0.17$, $SE = 0.06$, $t(163) = 2.91$, $p = .004$). Unexpectedly, neither audience persuasion ($b = 0.06$, $SE = 0.05$, $t(164) = 1.20$, $p = .23$) nor liking ($b = 0.10$, $SE = 0.06$, $t(164) = 1.78$, $p = .08$) were related to author self-enhancement. Finally, audience activism intentions were not related to either author value ($ps > 0.101$).

Audience Petition Signing and Mediation by Liking. We next tested whether this relationship between author values and audience ratings would translate into actual differences in audience behavior. Author values had no total effects on audience petition signing ($ps > 0.126$). However, in separate models, audience persuasion ($b = 0.72$, $SE = 0.04$, $z = 16.71$, $p < .001$) and liking ($b = 0.80$, $SE = 0.04$, $z = 19.24$, $p < .001$) did predict petition signing.

Hence, similar to previous studies, we tested for an indirect effect of author values on audience petition signing via liking, and separately via persuasion. We modeled paths for the effect of each author cardinal value on each mediator (liking or persuasion); and a path for the effect of the mediator on petition signing (a similar path analysis to previous studies, with the change that each audience participant was associated with only one statement and only one author participant). This model showed high goodness-of-fit using persuasion as the mediator (Fisher's $C(8) = 12.41$, $p = .134$; BIC = 73.21) as well as using liking as the mediator (Fisher's $C(8) = 12.96$, $p = .113$; BIC = 73.76; see Fig. 4), consistent with significant indirect effects of author values—specifically, a positive indirect effect of self-transcendence—on audience petition signing, via audience persuasion and liking.

7.3. Discussion

In Study 3, we again found that both author self-transcendence and author self-enhancement predicted increased activism intentions, replicating our earlier results and supporting H1a-b. In turn, author self-transcendence also predicted more positive audience reactions to advocacy messages, including persuasion and liking, replicating our earlier results and supporting H3a. Finally, replicating Studies 1 and 2, audience liking and persuasion related to actual audience behavior (petition signing).

Across all three studies, author self-transcendence leaked through the messages resulting in increased audience liking. In Studies 1 and 2, author self-enhancement was negatively related to audience liking, as expected; however, this was not the case in Study 3, suggesting only mixed support for H3b. It is possible that the single-message design of Study 3, unlike the multiple-message design of Studies 1 and 2, is related to this different pattern of results. That is, it may be that author values are especially likely to leak through their messages when seen directly alongside other authors' messages. In many real-world contexts, people do encounter multiple social advocacy messages in succession (e.g. on social media), mirroring Studies 1 and 2, and thus may be sensitive to the negative effects of author self-enhancement as our audience members were in these studies. Overall, across three studies, we find that author values do indeed leak through to affect audiences – ultimately impacting the persuasiveness of authors' advocacy messages.

8. General discussion

Across three two-phase studies, we tested how basic human values (Schwartz, 1992) relate to ally engagement and effectiveness. As we expected, self-professed allies reported higher engagement when they valued self-transcendence (across all three studies), and when they valued self-enhancement (in two of three studies). Furthermore, self-transcendence predicted not only activism intentions, but also real activism behavior (petition signing to support target group; Study 2).

Allies' values also had an impact on how persuasive they were in promoting their cause, appearing to "leak through" into the way they formulated their advocacy messages. Across all three studies, audiences expressed greater liking for messages from ally authors who valued self-transcendence, and (in two of three studies) *less* liking for messages from authors who valued self-enhancement. In turn, the more audiences liked these messages, the more likely they were to sign petitions supporting the cause. Together, our findings suggest that although self-enhancement and self-transcendence values may both undergird ally engagement, each has different effects on allyship expression and effectiveness.

8.1. Theoretical contributions

This work contributes to the field's understanding of allyship in two key ways. First, whereas an emerging literature examines how allies' motives relate to their alignment with the cause or likelihood of taking action (e.g., Brown & Ostrove, 2013; Cortland et al., 2017; Drury & Kaiser, 2014), ours is the first investigation to demonstrate how allies' motives relate to their persuasiveness. Our work draws on theoretical insights from work examining the motivation and development of allies (e.g., Edwards, 2006; Schultz, 2001; Radke et al., 2020), to shed light on the downstream effects of those allies' engagement. Our work thus reinforces the relevance of the basic human values framework (Schwartz, 1992) to understanding allyship (Edwards, 2006; Schultz, 2001). Our work suggests that knowing an ally's underlying values can help predict how effective they will be at rallying others to join the cause.

Second, we use validated, universally relevant constructs from the values framework (Sagiv & Schwartz, 2022; Sagiv et al., 2017; Schwartz,

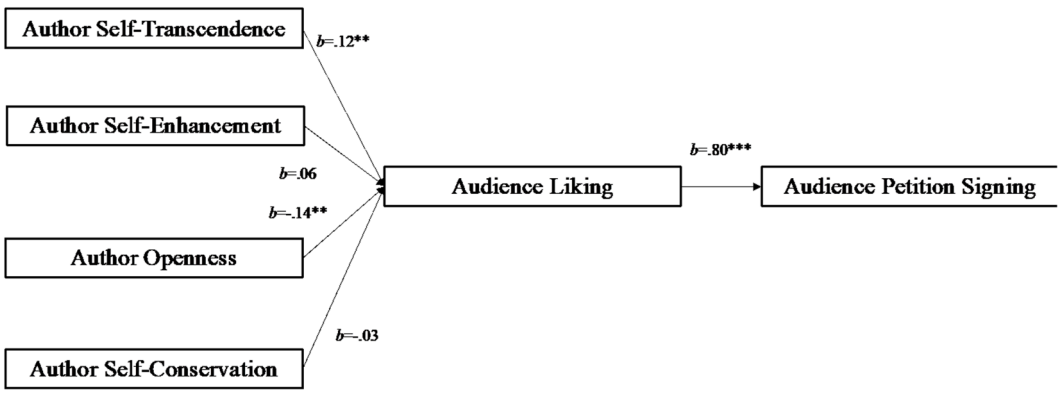


Fig. 4. Structural Equation Model for Study 3. Notes. *$p < 0.05$; **$p < 0.01$; *** $p < 0.001$. All values standardized.

1992, 1994) to shed light on a debate about allyship both inside and outside academia: Do self-enhancing allies truly help the cause, or does their desire for personal status prevent them from helping? On the one hand, some past findings suggest that good and helpful behavior can spring from self-centered concerns (Jordan & Rand, 2019) or pursuing personal goals (Cortland et al., 2017). On the other hand, other findings suggest that allies primarily concerned with performative self-presentation will not stay meaningfully engaged with the cause (Kristofferson et al., 2013). Our findings are consistent with both perspectives. We find that self-enhancement values make allies more engaged, but also less persuasive in their communication. We find that self-transcendence values, in contrast, motivate ally engagement, while also enhancing (not reducing) persuasiveness. Thus, our work highlights the utility of considering how different values relate to different actions and effectiveness, not just whether they motivate action at all. It also raises the intriguing possibility (not tested directly in our work) that more performative allies may lose motivation over time because they are not as effective or persuasive, causing them to lose self-efficacy or become frustrated—an alternative way of explaining the prior finding that such allies are less likely to stick with the cause.

8.2. Limitations and future directions

Our work has some limitations, and leaves open intriguing questions, that future work could address. First, the effects of self-transcendence were consistent across all three studies, but the effects of self-enhancement were less consistent, with the author engagement effect and the audience liking effect each disappearing in one of our three studies. This pattern may suggest that the effects of self-enhancement are weaker or more context-dependent, and future work could investigate contextual factors that make a difference. To take one possible example, recent work suggests that activism behaviors may cluster along dimensions of relative visibility and effortfulness (Brown, Badaan, Craig, & Saunders, 2023); perhaps self-enhancing allies are more likely to be engaged with high visibility behaviors, but self-transcending allies are more engaged across the board. In general, future work could explore how values relate to a wider range of different allyship behaviors, as well as allies' effectiveness at enacting different behaviors (beyond their persuasiveness when writing an advocacy statement).

A second question is how, specifically, authors' values are "leaking through" in a way that is detectable by audiences. As we show in exploratory analyses presented in the SOM, we did not find that audiences responded differently based on authors' pronoun use or emotional valence in particular (although, replicating previous work, we did find some evidence that authors' values were related to these elements of language use; Chen et al., 2014). However, many other aspects of authors' language use not examined in the current analyses could account for the differing responses from audiences. Future work could thus help understand what kind of residue authors' personal values left in their messages (e.g., certainty, clarity; Cheatham & Tormala, 2017; e.g., trustworthiness; Knowlton et al., 2022; Park, Vani, Saint-Hilaire, & Kraus, 2022). Moreover, it is possible that allies' values may leak through via strategic decisions: for instance, self-transcending values (compared to self-enhancing values) may be associated with different mental models of what might be persuasive to an audience, in turn shaping the messaging decisions allies make. Future work might also explore allies' persuasiveness in longer statements, or differences in how they enact other ally behaviors (e.g., nonverbal behavior; real-time conversation).

A third issue is how different stakeholder audiences may respond to different allies. For instance, recent laboratory work finds that explicitly stated motivations for allyship behavior can affect the self-esteem of marginalized group targets of said allyship (e.g., "that goes against my personal values;" Chu & Ashburn-Nardo, 2022; see also Kutlaca et al., 2022). In our studies, when we tested group differences in audience reactions to allies' naturalistic posts, we were surprised to find that fellow allies, non-allies, and marginalized group members themselves tended to respond to posts similarly (see SOM). This suggests that the "leakage" audiences are detecting in the messages may be fairly broad – such as overall positive or negative tone – as opposed to the nuances of the particular argument or the underlying values themselves, to which target group members may be more sensitive (Jacoby-Senghor et al., 2021). However, it will be important for future work to probe this pattern in more detail, and examine marginalized group members' experiences of a wider range of ally behaviors. Future research might also examine whether members of target groups can distinguish the motives of self-enhancing versus self-transcendent allies, and if language plays the same role.

A fourth question for further research is how values other than self-transcendence and self-enhancement may relate to allyship. For theoretical reasons, we focused on self-transcendence and self-enhancement, but we also included the other two cardinal values of conservation and openness following recommended analytic practices for the values survey. We had no theoretical predictions about these values, and we found mixed effects (SOM). Another possibility we did not test is that different values may have interacting effects on allies' behavior and persuasiveness (e.g., perhaps self-enhancement has a more negative impact on persuasiveness when allies also have low levels of self-transcendence values).

A final limitation is that we measured, and did not manipulate, allies' values, precluding strong causal conclusions. Although we tried to manipulate allies' orientations, e.g. by asking them to focus on creating a just world vs. being a good person, we found that participants did not report any changes in their personal values after completing these activities (Footnote 5; SOM). It is unsurprising that we could not manipulate values in a brief study, because values are theorized to be stable individual traits from early childhood or even birth (Schwartz, 2006; Vecchione et al., 2016). However, in real activism contexts, leaders may be able to work with allies over an extended period and hence shift their primary motivation for engaging in the work, in ways we could never accomplish in a brief study. Indeed, work in educational studies suggests ally motivations can develop and mature with training (Edwards, 2006). Thus, if possible, leaders may find it useful to develop allies from a self-enhancing towards a more self-transcending orientation. For example, emphasizing justice narratives (e.g., anti-racism) rather than personal morality narratives may help shift allies' focus and therefore their behaviors (Phillips & Lowery, 2018). Future work could further examine how to train or change allies' values over time to increase their engagement and effectiveness.

8.3. Practical implications

Leaders working toward diversity, equity, and inclusion goals often debate what role allies should play, whether allies do more harm than good, and which allies can be trusted to be more than performative. Underlying this debate might be deeper questions about the underlying motives of dominant group members who think of themselves as allies. Our results suggest that different personal values may spur allyship—but allyship of different styles. In the recruiting and managing of allies in organizational diversity and inclusion efforts, simple value surveys can match individuals to the roles best suited to their motivational frames, and in which they may be most likely to succeed and remain motivated over time. Specifically, our results suggest that people who value self-transcendence and those who value self-enhancement are both likely to *intend* to engage in activism, but self-transcending allies may follow through more on those intentions, and they may also be better able to persuade others to support the cause. Lack of follow through has the potential to erode trust amongst those the ally means to support and thus, even lead to inadvertent sabotage of the cause. Thus, it may be optimal for leaders to assign more self-transcending allies into outreach/messaging roles, while garnering commitments from more self-enhancing allies to turn their intentions into real action. More

generally, leaders should recognize that the allies who express the most dedication may not be the ones who will actually behave the most effectively.

9. Conclusion

We investigated what motivates dominant group members to act as allies, i.e. support a marginalized group's cause. Using the Schwartz values framework, we found that both a self-transcending orientation toward universalism and benevolence, and a self-enhancing orientation toward personal status, achievement, and power, relate to increased ally engagement among self-professed allies. Self-transcendence was also associated with increased effectiveness—here, successfully persuading others to support the cause—while self-enhancement was associated with decreased effectiveness. These results suggest that it may be useful for leaders and organizations to focus on cultivating allies' self-transcendence values while ensuring that allies' self-enhancing values lead to engagement without being merely performative. In this way, scholars and activists alike might look beyond merely motivating allies, and toward enhancing allies' effectiveness.

CRediT authorship contribution statement

L. Taylor Phillips: Writing – original draft, Methodology, Formal analysis, Conceptualization. **Tamar A. Kreps:** Conceptualization, Methodology, Writing – review & editing. **Dolly Chugh:** Writing – review & editing, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

Data will be made available on request.

Acknowledgments

We thank Sean Malahy for his thoughtful comments, as well as the participants of the Dismantling Bias Conference at Purdue University (2022). We appreciate the grant support of the Center for Global Economy and Business at Stern School of Business, New York University. Please send correspondence to: L. Taylor Phillips, ltaylorphillips@nyu.edu.

Appendix A. Supplementary material

Supplementary data to this article can be found online at https://doi.org/10.1016/j.obhdp.2024.104333.

References

Ashburn-Nardo, L. (2018). What can allies do? In A. J. Colella, & E. B. King (Eds.), *The Oxford handbook of workplace discrimination* (pp. 373–386). Oxford University Press.
Bai, Y., Maruskin, L. A., Chen, S., Gordon, A. M., Stellar, J. E., McNeil, G. D., & Keltner, D. (2017). Awe, the diminished self, and collective engagement: Universals and cultural variations in the small self. *Journal of Personality and Social Psychology, 113*(2), 185.
Bain, K., Kreps, T. A., Meikle, N. L., & Tenney, E. R. (2021). Amplifying voice in organizations. *Academy of Management Journal, 64*, 1288–1312.
Barasch, A., Levine, E. E., Berman, J. Z., & Small, D. A. (2014). Selfish or selfless? On the signal value of emotion in altruistic behavior. *Journal of Personality and Social Psychology, 107*, 393.
Bardi, A., & Schwartz, S. H. (2003). Values and behavior: Strength and structure of relations. *Personality and Social Psychology Bulletin, 29*, 1207–1220.
Bergsieker, H. B., Leslie, L. M., Constantine, V. S., & Fiske, S. T. (2012). Stereotyping by omission: Eliminate the negative, accentuate the positive. *Journal of Personality and Social Psychology, 102*(6), 1214.
Bergsieker, H. B., Shelton, J. N., & Richeson, J. A. (2010). To be liked versus respected: Divergent goals in interracial interactions. *Journal of Personality and Social Psychology, 99*, 248.
Besta, T., & Zawadzka, A. M. (2019). Expansion of the self of activists and nonactivists involved in mass gatherings for collective action. *Group Processes & Intergroup Relations, 22*, 182–199.
Boyd, R., Wilson, S., Pennebaker, J., Kosinski, M., Stillwell, D., & Mihalcea, R. (2015). Values in words: Using language to evaluate and understand personal values. In *Proceedings of the international AAAI conference on web and social media*, *9*(1), Article 1.
Boyd, R. L., & Pennebaker, J. W. (2017). Language-based personality: A new approach to personality in a digital world. *Current Opinion in Behavioral Sciences, 18*, 63–68.
Broockman, D., & Kalla, J. (2016). Durably reducing transphobia: A field experiment on door-to-door canvassing. *Science, 352*, 220–224.
Brown, R. M., Badaan, V., Craig, M. A., & Saunders, B. A. (2023). *An effort-visibility typology of actions to reduce racial inequality.* Unpublished Manuscript.
Brown, K. T., & Ostrove, J. M. (2013). What does it mean to be an ally? The perception of allies from the perspective of people of color. *Journal of Applied Social Psychology, 43*, 2211–2222.
Brunswik, E. (1956). *Perception and the representative design of psychological experiments*. Berkeley: University of California Press.
Cheatham, L. B., & Tormala, Z. L. (2017). The curvilinear relationship between attitude certainty and attitudinal advocacy. *Personality and Social Psychology Bulletin, 43*, 3–16.
Chen, J., Hsieh, G., Mahmud, J. U., & Nichols, J. (2014, February). Understanding individuals' personal values from social media word use. In *Proceedings of the 17th ACM conference on computer supported cooperative work & social computing* (pp. 405-414). ACM.
Chow, R. M., Lowery, B. S., & Hogan, C. M. (2013). Appeasement: Whites' strategic support for affirmative action. *Personality and Social Psychology Bulletin, 39*(3), 332–345.
Chu, C., & Ashburn-Nardo, L. (2022). Black Americans' perspectives on ally confrontations of racial prejudice. *Journal of Experimental Social Psychology, 101*, Article 104337.
Chugh, D. (2018). *The person you mean to be: How good people fight bias*. HarperCollins.
Cialdini, R. B., & Goldstein, N. J. (2004). Social influence: Compliance and conformity. *Annual Review of Psychology, 55*, 591–621.
Corning, A. F., & Myers, D. J. (2002). Individual orientation toward engagement in social action. *Political Psychology, 23*, 703–729.
Correll, S. J., Weisshaar, K. R., Wynn, A. T., & Wehner, J. D. (2020). Inside the black box of organizational life: The gendered language of performance assessment. *American Sociological Review, 85*(6), 1022–1050.
Cortland, C. I., Craig, M. A., Shapiro, J. R., Richeson, J. A., Neel, R., & Goldstein, N. J. (2017). Solidarity through shared disadvantage: Highlighting shared experiences of discrimination improves relations between stigmatized groups. *Journal of Personality and Social Psychology, 113*, 547.
Crano, W. D., & Prislin, R. (2006). Attitudes and persuasion. *Annual Review of Psychology, 57*, 345–374.
Crittle, C. S. (2017). Performative "Wokeness": Exploring blacks' perceptions of white allies who confront anti-black racism (Doctoral dissertation, Tufts University).
Czopp, A., & Monteith, M. (2003). Confronting prejudice (literally): Reactions to confrontations of racial and gender bias. *Personality and Social Psychology Bulletin, 29*, 532–544.
DeSteno, D., Dasgupta, N., Bartlett, M. Y., & Cajdric, A. (2004). Prejudice from thin air: The effect of emotion on automatic intergroup attitudes. *Psychological Science, 15*, 319–324.
Drury, B. J., & Kaiser, C. R. (2014). Allies against sexism: The role of men in confronting sexism. *Journal of Social Issues, 70*, 637–652.
Dupree, C. H., & Fiske, S. T. (2019). Self-presentation in interracial settings: The competence downshift by White liberals. *Journal of Personality and Social Psychology, 117*(3), 579.
Edwards, K. E. (2006). Aspiring social justice ally identity development: A conceptual model. *NASPA Journal, 43*, 39–60.
Effron, D. A., & Miller, D. T. (2012). How the moralization of issues grants social legitimacy to act on one's attitudes. *Personality and Social Psychology Bulletin, 38*(5), 690–701.
Feinberg, M., & Willer, R. (2013). The moral roots of environmental attitudes. *Psychological Science, 24*(1), 56–62. https://doi.org/10.1177/0956797612449177
Fischer, R., & Schwartz, S. (2011). Whence differences in value priorities? Individual, cultural, or artifactual sources. *Journal of Cross-Cultural Psychology, 42*(7), 1127–1144.
Foster-Gimbel, O. A., Pillemer, J., & Phillips, L. T. Authentic-ally? Feeling authentic increases allyship behavior via psychological standing. *Manuscript under review.*
Fredrickson, B. L. (2004). The broaden–and–build theory of positive emotions. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences, 359*, 1367–1377.
Gosling, S. D., Ko, S. J., Mannarelli, T., & Morris, M. E. (2002). A room with a cue: Personality judgments based on offices and bedrooms. *Journal of Personality and Social Psychology, 82*(3), 379.
Gouveia, V. V., Vione, K. C., Milfont, T. L., & Fischer, R. (2015). Patterns of value change during the life span: Some evidence from a functional approach to values. *Personality and Social Psychology Bulletin, 41*, 1276–1290.
Graham, J., Nosek, B. A., Haidt, J., Iyer, R., Koleva, S., & Ditto, P. H. (2011). Mapping the moral domain. *Journal of Personality and Social Psychology, 101*, 366.

Hekman, D. R., Johnson, S. K., Foo, M.-D., & Yang, W. (2017). Does diversity-valuing behavior result in diminished performance ratings for non-White and female leaders? *Academy of Management Journal, 60*(2), 771–797.
Hofstede, G. (2001). *Culture's consequences: Comparing values, behaviors, institutions and organizations across nations*. Sage.
Hofstede, G. (2011). Dimensionalizing cultures: The Hofstede model in context. *Online Readings in Psychology and Culture, 2*(1), 8.
House, R. J., Hanges, P. J., Javidan, M., Dorfman, P. W., & Gupta, V. (Eds.). (2004). *Culture, leadership, and organizations: The GLOBE study of 62 societies*. Sage Publications.
Jacoby-Senghor, D. S., Rosenblum, M., & Brown, N. D. (2021). Not all egalitarianism is created equal: Claims of nonprejudice inadvertently communicate prejudice between ingroup members. *Journal of Experimental Social Psychology, 94*, Article 104104.
Jordan, J. J., & Rand, D. G. (2019). Signaling when nobody is watching: A reputation heuristics account of outrage and punishment in one-shot anonymous interactions. *Journal of Personality and Social Psychology*.
Judd, C. M., Westfall, J., & Kenny, D. A. (2012). Treating stimuli as a random factor in social psychology: A new and comprehensive solution to a pervasive but largely ignored problem. *Journal of Personality and Social Psychology, 103*(1), 54.
Jun, S., Phillips, L. T., & Foster-Gimbel, O. A. (2023). The missing middle: Asian employees' experience of workplace discrimination and pro-Black workplace allyship. *Journal of Applied Psychology*.
Kendi, I. X. (2019). *How to be an antiracist*. One World.
Klar, M., & Kasser, T. (2009). Some benefits of being an activist: Measuring activism and its role in psychological well-being. *Political Psychology, 30*, 755–777.
Knowles, E. D., Lowery, B. S., Chow, R. M., & Unzueta, M. M. (2014). Deny, distance, or dismantle? How white Americans manage a privileged identity. *Perspectives on Psychological Science, 9*(6), 594–609.
Knowlton, K., Carton, A. M., & Grant, A. M. (2022). Help (Un) wanted: Why the most powerful allies are the most likely to stumble—and when they fulfill their potential. *Research in Organizational Behavior*, 100–180.
Konrath, S., Ho, M. H., & Zarins, S. (2016). The strategic helper: Narcissism and prosocial motives and behaviors. *Current Psychology, 35*, 182–194.
Kristofferson, K., White, K., & Peloza, J. (2013). The nature of slacktivism: How the social observability of an initial act of token support affects subsequent prosocial action. *Journal of Consumer Research, 40*, 1149–1166.
Kruglanski, A. W., & Thompson, E. P. (1999). Persuasion by a single route: A view from the unimodel. *Psychological Inquiry, 10*(2), 83–109.
Kutlaca, M., Radke, H. R., & Becker, J. C. (2022). The impact of including advantaged groups in collective action against social inequality on politicized identification of observers from disadvantaged and advantaged groups. *Political Psychology, 43*(2), 297–315.
Leach, C. W., Snider, N., Iyer, A., & Smith, H. (2002). Poisoning the consciences of the Fortunate: The experience of relative advantage and support for social equality. In I. Walker, & H. Smith (Eds.), *Relative deprivation: Specification, development and integration* (pp. 136–163).
Lee, Y. H., & Hsieh, G. (2013, April). Does slacktivism hurt activism? The effects of moral balancing and consistency in online activism. In *Proceedings of the SIGCHI conference on human factors in computing systems* (pp. 811–820). ACM.
Lefcheck, J. S. (2016). piecewiseSEM: Piecewise structural equation modelling in r for ecology, evolution, and systematics. *Methods in Ecology and Evolution, 7*(5), 573–579.
Lin, S. C., Schaumberg, R. L., & Reich, T. (2016). Sidestepping the rock and the hard place: The private avoidance of prosocial requests. *Journal of Experimental Social Psychology, 64*, 35–40.
Lindeman, M., & Verkasalo, M. (2005). Measuring values with the short Schwartz's value survey. *Journal of Personality Assessment, 85*, 170–178.
Madera, J. M., Hebl, M. R., & Martin, R. C. (2009). Gender and letters of recommendation for academia: Agentic and communal differences. *Journal of Applied Psychology, 94*(6), 1591.
Magee, J. C., & Smith, P. K. (2013). The social distance theory of power. Personality and social. *Psychology Review, 17*, 158–186.
McGinnies, E., & Ward, C. D. (1980). Better liked than right: Trustworthiness and expertise as factors in credibility. *Personality and Social Psychology Bulletin, 6*(3), 467–472.
Monin, B., & Miller, D. T. (2001). Moral credentials and the expression of prejudice. *Journal of Personality and Social Psychology, 81*, 33.
Mullen, E., & Monin, B. (2016). Consistency versus licensing effects of past moral behavior. *Annual Review of Psychology, 67*, 363–385.
Nabi, R. L. (2002). Discrete emotions and persuasion. In *The persuasion handbook: Developments in theory and practice* (p. 308).
Nolan, J. M., Schultz, P. W., Cialdini, R. B., Goldstein, N. J., & Griskevicius, V. (2008). Normative social influence is underdetected. *Personality and Social Psychology Bulletin, 34*, 913–923.
Park, J. W., Vani, P., Saint-Hilaire, S., & Kraus, M. W. (2022). Disadvantaged group activists' attitudes toward advantaged group allies in social movements. *Journal of Experimental Social Psychology, 98*, Article 104226.
Pennebaker, J. W., Booth, R. J., & Francis, M. E. (2007). *Linguistic inquiry and word count: LIWC [Computer software]*, 135.
Petty, R. E., & Cacioppo, J. T. (1986). Advances in experimental social psychology. In *The elaboration likelihood model of persuasion* (pp. 123–205).
Phillips, L. T., & Lowery, B. S. (2018). Herd invisibility: The psychology of racial privilege. *Current Directions in Psychological Science, 27*, 156–162.
Piff, P. K., Dietze, P., Feinberg, M., Stancato, D. M., & Keltner, D. (2015). Awe, the small self, and prosocial behavior. *Journal of Personality and Social Psychology, 108*(6), 883.
Plant, E. A., & Devine, P. G. (1998). Internal and external motivation to respond without prejudice. *Journal of Personality and Social Psychology, 75*, 811.
Podsakoff, P. M., MacKenzie, S. B., Lee, J. Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology, 88*(5), 879.
Radke, H. R., Kutlaca, M., Siem, B., Wright, S. C., & Becker, J. C. (2020). Beyond allyship: Motivations for advantaged group members to engage in action for disadvantaged groups. *Personality and Social Psychology Review, 24*(4), 291–315.
Rai, T. S., & Fiske, A. P. (2011). Moral psychology is relationship regulation: Moral motives for unity, hierarchy, equality, and proportionality. *Psychological Review, 118*, 57.
Rasinski, H. M., & Czopp, A. M. (2010). The effect of target status on witnesses' reactions to confrontations of bias. *Basic and Applied Social Psychology, 32*, 8–16.
Rokeach, M. (1977). *The nature of human values*. Free Press.
Rucker, D. D., Preacher, K. J., Tormala, Z. L., & Petty, R. E. (2011). Mediation analysis in social psychology: Current practices and new recommendations. *Social and Personality Psychology Compass, 5*(6), 359–371.
Russell, G. M. (2011). Motives of heterosexual allies in collective action for equality. *Journal of Social Issues, 67*(2), 376–393.
Sagiv, L., & Roccas, S. (2021). How do values affect behavior? Let me count the ways. *Personality and Social Psychology Review, 25*(4), 295–316.
Sagiv, L., Roccas, S., Cieciuch, J., & Schwartz, S. H. (2017). Personal values in human life. *Nature Human Behaviour, 1*(9), 630–639.
Sagiv, L., & Schwartz, S. H. (2022). Personal values across cultures. *Annual Review of Psychology, 73*, 517–546.
Salter, P., & Adams, G. (2013). Toward a critical race psychology. *Social and Personality Psychology Compass, 7*, 781–793.
Sandy, C. J., Gosling, S. D., Schwartz, S. H., & Koelkebeck, T. (2017). The development and validation of brief and ultrabrief measures of values. *Journal of Personality Assessment, 99*(5), 545–555.
Schultz, P. W. (2001). The structure of environmental concern: Concern for self, other people, and the biosphere. *Journal of Environmental Psychology, 21*(4), 327–339.
Schwartz, S. H. (1992). Universals in the content and structure of values: Theoretical advances and empirical tests in 20 countries. In *Advances in experimental social psychology* (Vol. 25, pp. 1–65). Academic Press.
Schwartz, S. H. (1994). Are there universal aspects in the structure and contents of human values? *Journal of Social Issues, 50*, 19–45.
Schwartz, S. H. (1999). A theory of cultural values and some implications for work. *Applied Psychology, 48*, 23–47.
Schwartz, S. (2006). A theory of cultural value orientations: Explication and applications. *Comparative Sociology, 5*, 137–182.
Schwartz, S. H. (2007). Basic human values: Theory, measurement, and applications. *Revue française de sociologie, 47*(4), 929.
Schwartz, S. H. (2011). Studying values: Personal adventure, future directions. *Journal of Cross-Cultural Psychology, 42*, 307–319.
Schwartz, S. H. (2012). An overview of the Schwartz theory of basic values. *Online readings in Psychology and Culture, 2*(1), 2307–10919.
Schwartz, S. H., & Boehnke, K. (2004). Evaluating the structure of human values with confirmatory factor analysis. *Journal of Research in Personality, 38*(3), 230–255.
Sherf, E. N., Tangirala, S., & Weber, K. C. (2017). It is not my place! Psychological standing and men's voice and participation in gender-parity initiatives. *Organization Science, 28*, 193–210.
Sidanius, J., & Pratto, F. (1999). *Social dominance: An intergroup theory of social hierarchy and oppression*. Cambridge: Cambridge University Press.
Simmons, J. P., Nelson, L. D., & Simonsohn, U. (2018). False-positive citations. *Perspectives on Psychological Science, 13*, 255–259.
Smith, P. B., Peterson, M. F., & Schwartz, S. H. (2002). Cultural values, sources of guidance, and their relevance to managerial behavior: A 47-nation study. *Journal of Cross-cultural Psychology, 33*(2), 188–208.
Sverdlik, N., Roccas, S., & Sagiv, L. (2012). Morality across cultures: A values perspective. In M. Mikulincer, & P. R. Shaver (Eds.), *Herzliya series on personality and social psychology. The social psychology of morality: Exploring the causes of good and evil* (pp. 219–235).
Tamir, M., Schwartz, S. H., Cieciuch, J., Riediger, M., Torres, C., Scollon, C., Dzokoto, V., Zhou, X., & Vishkin, A. (2016). Desired emotions across cultures: A value-based account. *Journal of Personality and Social Psychology, 111*, 67–82.
Thomas, R. J., Masthoff, J., & Oren, N. (2019). Can i influence you? Development of a scale to measure perceived persuasiveness and two studies showing the use of the scale. *Frontiers in Artificial Intelligence, 2*, 24.
Thomas, E. F., McGarty, C., Reese, G., Berndsen, M., & Bliuc, A. M. (2016). Where there is a (collective) will, there are (effective) ways: Integrating individual-and group-level factors in explaining humanitarian collective action. *Personality and Social Psychology Bulletin, 42*(12), 1678–1692.
Tiedens, L. Z. (2001). Anger and advancement versus sadness and subjugation: The effect of negative emotion expressions on social status conferral. *Journal of Personality and Social Psychology, 80*, 86.
Uzefovsky, F., Döring, A. K., & Knafo-Noam, A. (2016). Values in middle childhood: Social and genetic contributions. *Social Development, 25*, 482–502.
Van Zomeren, M., Postmes, T., Spears, R., & Bettache, K. (2011). Can moral convictions motivate the advantaged to challenge social inequality? Extending the social identity model of collective action. *Group Processes & Intergroup Relations, 14*(5), 735–753.
Vecchione, M., Schwartz, S., Alessandri, G., Döring, A. K., Castellani, V., & Caprara, M. G. (2016). Stability and change of basic personal values in early adulthood: An 8-year longitudinal study. *Journal of Research in Personality, 63*, 111–122.
Wilkins, C. L., & Kaiser, C. R. (2014). Racial progress as threat to the status hierarchy: Implications for perceptions of anti-White bias. *Psychological Science, 25*(2), 439–446.