RESEARCH ARTICLE

EASP WILEY

What they think of us: Meta-beliefs and solidarity-based collective action among the advantaged

Aya Adra[1,2] 🆔 | Mengyao Li[1] | Anna Baumert[1,2]

[1]Max Planck Institute for Research on Collective Goods, Bonn, Germany

[2]School of Education, Technical University Munich, München, Germany

Correspondence
Aya Adra, Max Planck Institute for Research on Collective Goods, Kurt Schumacher Strasse, 10, Bonn, Deutschland (DEU), DE 53113, Germany.
Email: adra@coll.mpg.de

Abstract

Collective action researchers have recently started investigating solidarity-based collective action by advantaged groups. This literature, however, has overlooked intergroup meta-beliefs (MBs, i.e., beliefs about the outgroup's beliefs), which we argue are crucial, since solidarity inherently involves protesting for the outgroup. In the context of racial inequality in the U.S., we focused on three MBs White Americans could hold: responsibility, inactivity, and allyship. In two studies (N_{total} = 648), we found that inactive and responsible MBs predicted higher collective action tendencies among low White identifiers via guilt and obligation to act. Conversely, we found that both predicted lower collective action tendencies among high White identifiers, via perceived unfairness. Finally, we found that ally MB was positively associated with collective action tendencies, regardless of identification. We highlight the importance of the meta-perspective in understanding solidarity-based collective action, and discuss conceptual and practical implications of these findings.

KEYWORDS
advantaged groups, collective action, meta-beliefs, social change, solidarity

1 | INTRODUCTION

Historically, the collective action literature has been centered on disadvantaged group members' willingness to take actions to better their conditions (see Becker & Tausch, 2015, for a review). That said, disadvantaged groups often garner support and protest participation from advantaged group members. Accordingly, social psychologists have started investigating advantaged group members' collective action on behalf of or in support of the disadvantaged (henceforth referred to as solidarity-based collective action; e.g., Iyer, Leach, & Crosby, 2003; Shepherd, Spears, & Manstead, 2013). So far, this work has largely focused on extending the findings of collective action studies among the disadvantaged, by showing that the central variables that have been shown to predict collective action

in that context (e.g., identification with the disadvantaged group, efficacy, injustice perceptions) are also useful in predicting solidarity-based collective action (e.g., van Zomeren, Postmes, Spears, & Bettache, 2011). In this existing literature, the focus has been on the role of perceptions, attitudes, and beliefs about the outgroup in shaping solidarity-based collective action.

In the current research, we extend the existing account of solidarity-based collective action by exploring the role of advantaged group members' beliefs about how the disadvantaged group think of them (i.e., intergroup meta-beliefs), specifically in relation to injustice and inequality. Given that solidarity-based collective action inherently involves working together with the disadvantaged, we argue that this meta-perspective is highly relevant for understanding advantaged group members' engagement in collective action. In

the context of racial inequality in the U.S., we focused on three meta-beliefs that members of the advantaged group, White Americans, likely hold about their role in perpetuating and redressing inequality: allyship, inactivity, and responsibility. In two studies, we sought to investigate the respective relationships between these meta-beliefs and solidarity-based collective action tendencies, the moderating role of ingroup identification, and the psychological mechanisms underlying some of these relationships.

2 | META-BELIEFS IN INTERGROUP CONTEXTS

While social psychologists have predominantly investigated the role of the beliefs and attitudes that people hold towards outgroups in shaping intergroup relations, recent research has shifted the focus to intergroup meta-beliefs (see Frey & Tropp, 2006, for a review). Some researchers have investigated meta-beliefs along the dimensions of the Stereotype Content Model (Fiske, 1998), by examining meta-warmth and meta-competence in different intergroup contexts (e.g., Figueiredo, Doosje, Valentim, & Zebel, 2010; Wout, Murphy, & Steele, 2010). Researchers have also examined meta-beliefs along a positive–negative continuum (e.g., Finchilescu, 2010), meta-beliefs derived from International Image Theory (O'Brien, Leidner, & Tropp, 2017), and meta-beliefs derived from acculturation models (António & Monteiro, 2015). These different streams of research have collectively highlighted the fact that individuals are influenced by their perception of others' beliefs about them, and that these perceptions contribute substantially to intergroup relations (Livingstone, Rothers, & Fernández, 2019). Indeed, some studies have even suggested that meta-beliefs play a more central role in shaping intergroup interactions than do beliefs about the outgroup (e.g., Vorauer & Kumhyr, 2001). Such findings have prompted researchers to claim that a truly social psychological approach to any number of intergroup phenomena must include an examination of the meta-level (e.g., Otten, 2002).

2.1 | Positive intergroup meta-beliefs

Research has shown that positive meta-beliefs generally improve intergroup relations (e.g., Vezzali, 2017). Indeed, the expectation of inclusion by outgroup members, which can be understood as a positive meta-belief, was shown to predict more positive attitudes towards the outgroup (Tropp & Bianchi, 2006). Directly relevant to our focus on advantaged group members, Vezzali (2017) demonstrated that the activation of positive meta-stereotypes among members of the dominant group (Italian high school students) led to the anticipation of greater enjoyment of an upcoming interaction with a member of the disadvantaged group (African immigrants). This effect was explained by an increase in positive feelings about contact and a decrease in concerns about being accepted (Vezzali, 2017). These findings fit nicely within a large body of work indicating that individuals tend

to reciprocate evaluations by others, whether positive or negative (Doosje & Haslam, 2005). This claim, however, is further complicated when considering negative intergroup meta-beliefs.

2.2 | Negative intergroup meta-beliefs

In general, negative meta-beliefs have been shown to predict negative outgroup attitudes (Putra & Wagner, 2017) and intergroup hostility (Issmer, Stellmacher, & Gollwitzer, 2013). The expectation of rejection by outgroup members, for example, was found to predict negative outgroup attitudes (Barlow, Louis, & Hewstone, 2009). These relationships have been replicated in the contexts of fictitious and real outgroups, and can even translate into support for aggression (O'Brien et al., 2017). While a plethora of such research has documented that when individuals feel their group is evaluated negatively, they oftentimes reciprocate the negativity (Kteily & Bruneau, 2015), this picture is likely more complex.

When members of the advantaged group believe that the disadvantaged view them in a negative light, they can act defensively (Kteily & Bruneau, 2015; Vorauer, 2003), particularly if the meta-beliefs are seen as inaccurate or offensive. However, they can act in a compensatory manner. For example, in a study of Norwegian majority members, Phelps (2013) found that the more majority members thought immigrants believed they (the majority) were cold, the more they showed willingness to accommodate immigrants. The author speculated that the underlying process involved collective guilt about the majority's moral shortcomings (Phelps, 2013). Lending support to this speculation, in the context of a past colonial conflict, Figueiredo et al. (2010) showed that Dutch participants who thought Indonesians held a negative belief towards them reported higher group-based guilt and compensatory behavioral intentions. These findings together demonstrate that negative meta-beliefs can have either negative or positive consequences for intergroup relations, depending on how advantaged group members perceive them. We will return to these diverging consequences when discussing negative meta-beliefs in the context of solidarity-based collective action.

3 | META-BELIEFS AND SOLIDARITY-BASED COLLECTIVE ACTION

In view of the research reviewed above, and the consistent finding that advantaged group members readily think at the meta-level (e.g., Vorauer, Hunter, Main, & Roy, 2000; Vorauer & Kumhyr, 2001), we argue that meta-beliefs are crucial for understanding solidarity-based collective action. Specifically, advantaged group members' meta-beliefs regarding how the disadvantaged view *their role in perpetuating and redressing intergroup inequality* should predict their tendencies to engage in collective action on behalf of the disadvantaged.

In the context of racial inequality in the U.S., we focused on three meta-beliefs that White Americans may hold, and that should

be particularly relevant to solidarity-based collective action. An *ally* meta-belief captures White Americans' belief that Black Americans think of them as allies in the fight against racial inequality. An *inactive* meta-belief captures White Americans' belief that Black Americans think of them as passive in that fight. A *responsible* meta-belief captures White Americans' belief that Black Americans think of them as responsible for the Black community's ongoing struggles.

Drawing on both real life examples and the literature on intergroup inequality, we reasoned that these three meta-beliefs should be highly relevant to solidarity-based collective action. Concepts of allyship, inactivity, and responsibility are pervasive in narratives attempting to position White people in the fight against racial inequality, among both White and Black Americans. For instance, White Americans have argued that "[White Americans] need to become trustworthy *allies*" (Morrison, 2013; emphasis added). Others have stressed that "[White Americans] cannot stay *silent*, because silence is acceptance" (Davidson, 2016; emphasis added). Still others have asserted that "[racism] was constructed and created by white people, and the ultimate *responsibility* lies with white people" (DiAngelo, in Iqubal, 2019; emphasis added). Such statements mirror those made by Black Americans in their positioning of White people in the fight against racial inequality.

The concepts of allyship, inactivity, and responsibility also largely map onto concepts that social psychologists have investigated in contexts of intergroup inequality and challenges to it. For instance, in a growing body of literature, researchers have investigated the ways in which the advantaged can act as allies (Louis et al., 2019, for a review), and be perceived as such (e.g., Kutlaca, Becker, & Radke, 2019). In parallel, research on the advantaged has highlighted that oftentimes, "[they] have little reason to respond to others' claims of relative deprivation" (Leach, Snider, Iyer, & Smith, 2002, p. 139), echoing the concept of inactivity. Finally, much of the literature on the role of group-based guilt in shaping attitudes towards racial compensation explores the focus of the advantaged on their groups' responsibility for wrongdoing (Iyer et al., 2003). Allyship, inactivity, and responsibility also emerge, at least in part, in theories of intergroup relations (e.g., International Image Theory; Alexander, Brewer, & Hermann, 1999), work on the phenomenology of advantage (e.g., Leach et al., 2002), and models conceptualizing solidarity by majority group members (e.g., Subašić, Reynolds, & Turner, 2008). The ubiquity of these concepts in real life and social psychological literatures therefore informed our decision to zoom in on them at the meta-level – that is, the extent to which the advantaged group believe that the disadvantaged view them as allies, inactive, or responsible in the context of intergroup inequality. While these three meta-beliefs might be far from exhaustive, we argue that they are highly relevant for our understanding of advantaged group members' collective action in support of the disadvantaged.

We also believe they are distinct from meta-beliefs previously explored in relation to attitudes towards intergroup inequality, such as meta-warmth and meta-competence (Phelps, 2013), in important ways. Unlike the global meta-warmth and meta-competence, the meta-beliefs of interest in the current studies are concretely rooted in the existing inequality between the two groups. We therefore reasoned that ally, inactive, and responsible meta-beliefs would be distinct from meta-warmth and meta-competence, and that their role in shaping solidarity would warrant special attention.

Specifically, we hypothesized that White Americans' endorsement of these three meta-beliefs would have differential relationships with their willingness to participate in collective action on behalf of Black Americans. The *ally* meta-belief not only portrays the advantaged group in a positive light, but also directly includes them in the collective action against racial inequality. Thus, we expected that if advantaged group members believe the disadvantaged think of them as allies, they might have greater willingness to participate in solidarity-based collective action in support of the disadvantaged.

In contrast, if advantaged group members believe that the disadvantaged think of them as *inactive* in the struggle against inequality, or *responsible* for their (disadvantaged) plight—both negative meta-beliefs—they might be more *or* less willing to participate in solidarity-based collective action. In line with the prior research on negative meta-beliefs, the relationships between the inactive and responsible meta-beliefs and collective action tendencies should depend on whether they trigger feelings of group-based guilt and a collective obligation to act on the one hand or, a sense of unfairness on the other. Indeed, group-based guilt has been shown to promote willingness to engage in a range of conciliatory intergroup behaviors (e.g., Brown & Cehajic, 2008; Figueiredo et al., 2010). Conversely, group members have also been shown to react defensively in response to the feeling that they are being unfairly cast in a negative light (e.g., Kteily & Bruneau, 2015). We argue that the extent to which advantaged group members identify with their group should moderate these divergent implications of responsible and inactive meta-beliefs.

3.1 | The moderating role of ingroup identification

A large literature examining social identification has consistently shown that individuals are motivated to defend their group to the extent that they identify with it (e.g., (Branscombe, Ellemers, Spears, & Doosje, 1999). Since high identifiers are motivated to uphold a positive ingroup image (Doosje & Branscombe, 2003; Doosje, Branscombe, Spears, & Manstead, 2006), they tend to justify or even deny the wrongfulness of their groups' actions to protect their identity (e.g., Bilali, Tropp, & Dasgupta, 2012; Li, Leidner, & Fernandez-Campos, 2019; Lowery, Knowles, & Unzueta, 2007). Low identifiers, by contrast, tend to be more ingroup critical, experience more group-based guilt, and are therefore more likely to react in a compensatory manner in response to negative portrayals of the ingroup (e.g., Doosje et al., 2006; Klein, Licata, & Pierucci, 2011; Roccas, Klar, & Liviatan, 2006).

In the current research, we focused on White identification as the relevant social identification that might moderate the associations between negative meta-beliefs and solidarity-based collective action tendencies. We expected the extent to which

FIGURE 1 Conceptual model depicting the moderating role of White identification in the link between inactive and responsible meta-beliefs and collective action tendencies
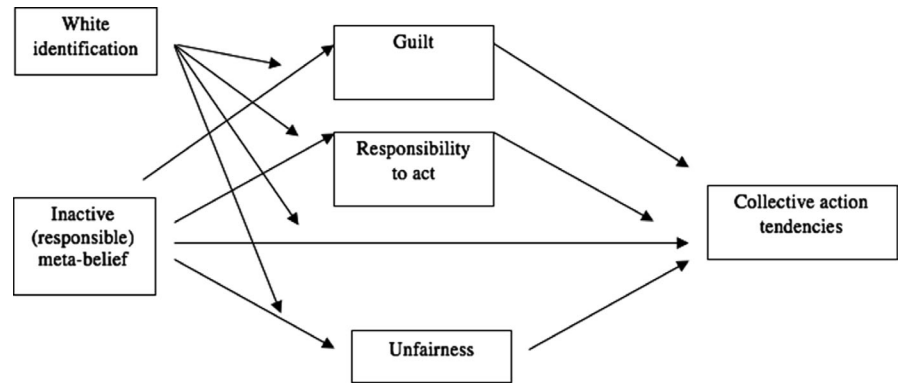
our White participants identified with White Americans to modulate the downstream processes from the inactive and responsible meta-beliefs. We hypothesized that among low identifiers, endorsing inactive and responsible meta-beliefs would be linked to higher collective action tendencies, explained by perceived collective guilt and obligation to act against racial inequality. Among high White identifiers, conversely, endorsing inactive and responsible meta-beliefs would be linked to lower collective action tendencies, explained by perceived unfairness. The conceptual model is displayed in Figure 1. Furthermore, we expected that endorsing an ally meta-belief would be positively associated with collective action tendencies, regardless of White identification.

4 | OVERVIEW OF CURRENT RESEARCH

In the two studies presented here, we tested whether the meta-beliefs regarding allyship, inactivity, and responsibility were distinct (a) from participants' own beliefs regarding the extent to which White Americans are allies, inactive, or responsible (Study 1), (b) from the more global and previously investigated concepts of meta-competence and meta-warmth (Study 2), and (3) from each other (Studies 1 and 2). Using these preparatory analyses, we then examined how the meta-beliefs of interest differentially predicted solidarity-based collective action among White Americans, and whether White identification moderated these relationships (Studies 1 and 2). Moreover, we investigated whether feelings of collective guilt and obligation to act explained the positive relationships between negative meta-beliefs and willingness to act among low White identifiers, and whether perceived unfairness explained their negative relationships among high White identifiers (Study 2, preregistered, all materials can be found on OSF).

This research therefore makes multiple contributions. We examined a novel set of meta-beliefs, which we argue are highly relevant to solidarity-based collective action. In doing so, we shed light on the importance of the meta-level as a crucially missing piece for understanding solidarity-based collective action. We also aimed to extend prior research on ingroup identification, by testing its role in moderating the implications of meta-beliefs.

5 | STUDY 1

The main goals of the study were twofold. First, we set out to examine whether the meta-beliefs regarding allyship, inactivity, and responsibility were distinct from each other, and from participants' own beliefs. Second, we investigated how these meta-beliefs differentially predicted collective action tendencies among White Americans, and whether White identification moderated these relationships. We expected that endorsing an ally meta-belief would be positively associated with collective action tendencies, regardless of White identification. We also expected that among low White identifiers, endorsing inactive and responsible meta-beliefs would be linked to higher collective action tendencies. Among high identifiers, conversely, endorsing inactive and responsible meta-beliefs would be linked to lower collective action tendencies.

5.1 | Method

5.1.1 | Participants

An a priori power analysis using G*Power (Faul, Erdfelder, Lang, & Buchner, 2007) showed that 266 participants in total should provide sufficient power (1–β = 0.80) for detecting an interaction effect of medium effect size[1] (i.e., f = 0.25; Cohen, 1969) at a critical alpha level of 0.05.

The sample consisted of 304 White Americans. Participants were sampled via Amazon's Mechanical Turk using the panel option to ensure two sets of inclusion criteria: (1) White/Caucasian and (2) U.S. citizens or permanent residents of the U.S. Our screening of the data resulted in the exclusion from the data of 15 participants who self-reported to be either not White/Caucasian or not U.S. citizens or permanent residents of the U.S., and of 16 who did not pay sufficient attention (indicated by their incorrect answers to the

---

[1]To our knowledge, there are no previous studies exploring an association between intergroup meta-beliefs and collective action tendencies. The closest relevant result is arguably that of Figueiredo et al. (2010), who reported a correlation of |0.22| (corresponding to f = 0.23) between meta-perceptions and intentions to compensate for past colonial conflict.

attention check questions[2]). Our final sample consisted of 273 participants (60% women; age M = 40.00, SD = 13.01, range = 19–73).

5.1.2 | Procedure and measures

Participants were invited to take part in a survey about social issues in the U.S. After consenting to participate in the study, they completed the following measures in the order presented below. All items were measured on 9-point visual analogue scales.

Meta-beliefs and beliefs

Meta-beliefs. Participants were randomly presented with 18 meta-belief items. They were asked to indicate how likely or unlikely (1 = *not likely at all*; 9 = *very likely*) they thought it was that *most Black people in the U.S.* would hold the following beliefs about White people. Six items captured each of the three meta-beliefs of interest respectively: the *ally* meta-belief (e.g., "In general, Black people believe that White people have been an important force in the movement against discrimination"; "In general, Black people believe that White people have been active participants in the movement against discrimination"; α = 0.91), the *inactive* meta-belief (e.g., "In general, Black people believe that White people are largely silent regarding Anti-Black discrimination in the U.S."; "In general, Black people believe that White people are typically inactive when it comes to the struggle for racial equality"; α = 0.93), and the *responsible* meta-belief (e.g., "In general, Black people believe that White people contribute to Anti-Black discrimination"; "In general, Black people believe that White people are responsible for the struggles faced by the Black community"; α = 0.89).

Beliefs. In random order, participants were presented with 18 belief items corresponding to the 18 meta-beliefs. They were asked to indicate how much they *personally* agreed or disagreed (1 = *strongly disagree*; 9 = *strongly agree*) with the following statements. Each statement was identical to a meta-belief item, but "In general, Black people believe that…" was substituted with "In general I believe that…". The resulting scale included six items capturing the ally belief (α = 0.78), the inactive belief (α = 0.97), and the responsible belief (α = 0.95).

Collective action. Adapted from Selvanathan, Techakesari, Tropp, and Barlow (2018), five items measured participants' willingness to engage in different forms of collective action against racial injustice (e.g., "For each of the following actions, please indicate how willing you are to engage in it in the future: attend demonstrations,

protests, or rallies against racial injustice alongside Black people"). A composite score for collective action tendencies was formed by averaging across all five items (α = 0.88).

White identification. Adapted from Leach et al.'s (2008) multidimensional model, seven items measured two dimensions, centrality and importance, of White identification. When factor analyzed, all items loaded onto one factor. We therefore created a composite score for White identification combining both dimensions (e.g., "Being a White person is an important part of how I see myself" and "I am glad to be White"; α = 0.91).

5.2 | Results

5.2.1 | Preparatory analyses

As a preparatory step, we first aimed to demonstrate that the meta-beliefs of interest (a) can be meaningfully distinguished from White participants' own beliefs about their ingroup's role as allies or inactive in the fight against racial inequality, and as responsible for racial inequality, and (b) can be meaningfully distinguished from each other. To do that, we conducted a series of confirmatory factor analyses (CFAs),[3] first on all the meta-belief and belief items, then on all the meta-belief items. Maximum Likelihood Estimation was used to estimate model parameters. Full results including item loadings and factor correlations are provided in Appendix S1.

To scrutinize whether meta-beliefs were distinct from beliefs, as we predicted, we tested two alternative measurement models for the meta-belief and belief items. First, we tested a model specifying three correlated latent factors, with the meta-belief and belief items of one type (i.e., ally, inactive, and responsible, respectively) loading together onto one factor. The model fit was not acceptable, χ^2(591) = 6,247.03, p < .001; RMSEA = 0.19, SRMR = 0.27, CFI = 0.93. Next, we tested a model with six latent factors in which the meta-belief items of each type (i.e., ally, inactive, and responsible) and the belief items of each type loaded onto separate factors. This model yielded an acceptable fit, χ^2(579) = 1,084.56, p < .001; RMSEA = 0.06, SRMR = 0.05, CFI = 0.95.

To scrutinize whether ally, inactive, and responsible meta-beliefs were distinct from one another, we tested two alternative measurement models for the meta-belief items. First, we tested a model specifying two correlated latent factors in which the positive (i.e., ally meta-belief) and negative (i.e., inactive and responsible) meta-beliefs loaded onto separate factors. The model fit was not acceptable, χ^2(134) = 423.32, p < .001; RMSEA = 0.09, CFI = 0.93, although it did meet the SRMR criterion, SRMR = 0.05. Next, we tested a model with three latent factors in which the meta-belief items of each type (i.e., ally, inactive, and responsible)

---

[2]At two separate points in each study; once in the middle and once towards the end, we inserted attention check items that looked identical to the other items of the survey, but asked participants "Please move the slider to the extreme left (Not likely at all) for this question" and "Please move the slider to the extreme right (Very likely) for this question". We also asked participants to indicate, at the end of the study, how attentive they were while taking the survey, how seriously they took the survey, and whether they were distracted while taking the survey. Full exclusion criteria based on attention checks can be found on OSF.

[3]We also conducted exploratory factor analyses (EFAs), first on all the meta-belief and belief items, then on all the meta-belief items. Full results of the EFAs are provided in Appendix S1.

TABLE 1 Means, standard deviations, and bivariate correlations between variables of interest

	Mean (SD)	Inactive MB	Ally MB	White ID	CA tendencies
Responsible MB	7.29 (1.33)	0.79**	−0.61**	−0.02	0.06
Inactive MB	6.99 (1.51)	1	−0.70**	0.02	0.06
Ally MB	3.91 (1.69)		1	−0.05	0.11
White identification	5.31 (1.92)			1	−0.27**
Collective action tendencies	5.06 (2.20)				1

**p < .001.
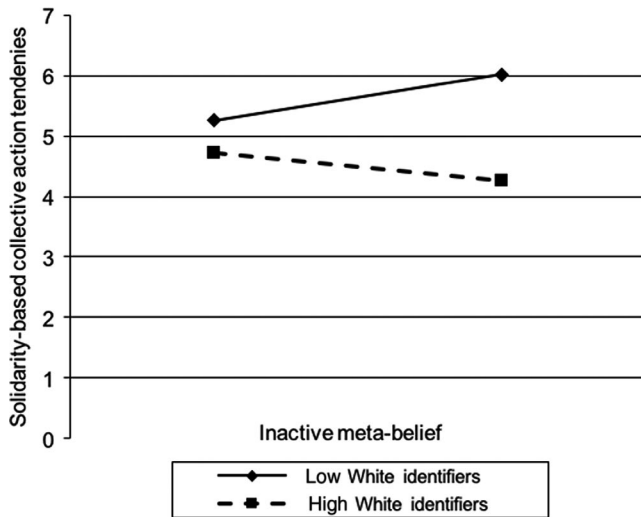
FIGURE 2 Graph depicting the moderating role of White identification in the link between inactive meta-belief and collective action tendencies
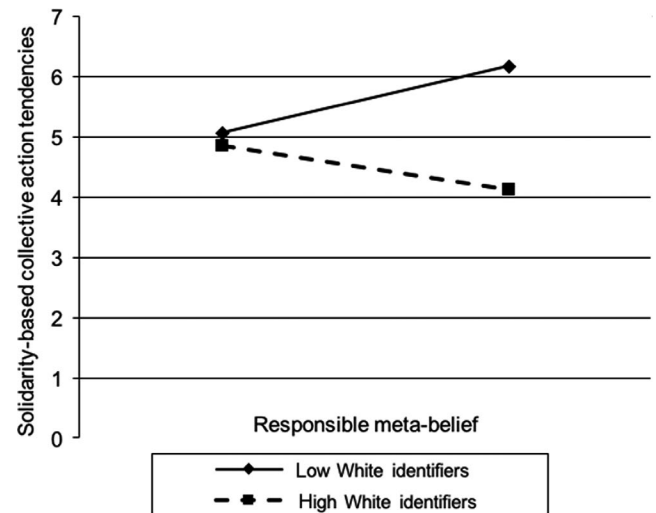
FIGURE 3 Graph depicting the moderating role of White identification in the link between responsible meta-belief and collective action tendencies

loaded onto separate factors. This model yielded an acceptable fit, $\chi^2(132) = 283.16$, $p < .001$; RMSEA = 0.06, SRMR = 0.04, CFI = 0.96.

5.2.2 | Main analyses

Means, standard deviations, and bivariate correlations between our variables are reported in Table 1.

To test our main hypotheses, we conducted three moderated regression analyses using PROCESS (Hayes, 2018, Model 1) with 5,000 bootstrap samples and 95% confidence intervals. In each analysis, we entered ally, inactive, or responsible meta-belief as the respective predictor, collective action tendencies as the outcome, and White identification as a continuous moderator. The predictor and the moderator were both mean centered. When the interaction between the meta-belief and White identification was significant, we conducted simple slope analyses.[4]

---

[4]We also tested all three models while entering each corresponding belief as a covariate. The patterns of results were largely consistent with those reported here. The full results are reported in Appendix S1.

Ally meta-belief

The overall model was significant, $F(1, 273) = 9.18$, $p < .001$, $R^2 = .09$. Ally meta-belief did not significantly predict collective action tendencies ($b = 0.15$, $SE = 0.08$, CI95 [−0.004, 0.296]). There was no significant interaction between ally meta-belief and White identification ($b = 0.04$, $SE = 0.04$, CI95 [−0.026, 0.113]). White identification negatively predicted collective action tendencies ($b = −0.30$, $SE = 0.07$, CI95 [−0.433, −0.168]).

Inactive meta-belief

The overall model was significant, $F(1, 273) = 10.46$, $p < .001$, $R^2 = .10$. Inactive meta-belief did not significantly predict collective action tendencies ($b = 0.05$, $SE = 0.08$, CI95 [−0.116, 0.218]). As hypothesized, there was a significant interaction between inactive meta-belief and White identification in predicting collective action tendencies ($b = −0.10$, $SE = 0.04$, CI95 [−0.180, −0.030]; see Figure 2). As expected, inactive meta-belief positively predicted collective action tendencies among those *low* (i.e., 1 *SD below* the mean) on White identification ($b = 0.25$, $SE = 0.11$, CI95 [0.043, 0.459]). In contrast, inactive meta-belief did not predict collective action tendencies among those *high* (i.e., 1 *SD above* the mean) on

White identification (b = −0.18, SE = 0.18 CI95 [−0.425, 0.069]). White identification negatively predicted collective action tendencies (b = −0.30, SE = 0.07, CI95 [−0.432, −0.170]).

Responsible meta-belief

The overall model was significant, $F(1, 273) = 13.27$, $p < .001$, $R^2 = .13$. Responsible meta-belief did not significantly predict collective action tendencies (b = 0.07, SE = 0.09, CI95 [−0.113, 0.245]). As hypothesized, there was a significant interaction between responsible meta-belief and White identification in predicting collective action tendencies (b = −0.17, SE = 0.04, CI95 [−0.259, −0.088]; see Figure 3). As expected, responsible meta-belief positively predicted collective action tendencies among those *low* on White identification (b = 0.40, SE = 0.12, CI95 [0.150, 0.642]). Conversely, responsible meta-belief negatively predicted collective action tendencies among those *high* on White identification (b = −0.31, SE = 0.13 CI95 [−0.568, −0.058]). White identification again negatively predicted collective action tendencies (b = −0.29, SE = 0.07, CI95 [−0.422 −0.165]).

5.3 | Discussion

In this first study, set in the context of racial inequality in the U.S., we investigated the relationships between three meta-beliefs that White Americans could hold: allyship, inactivity, and responsibility, and their solidarity-based collective action tendencies, while testing the moderating role of White identification. Our preparatory analyses indicated that ally, responsible, and inactive meta-beliefs were partially overlapping, yet psychologically distinguishable from participants' own corresponding beliefs, and that they were meaningfully distinguishable from one another. Moreover, results of our main analyses were partially consistent with our predictions. As hypothesized, inactive and responsible meta-beliefs predicted collective action tendencies, and the direction critically depended on White identification as moderator. Among individuals who did not strongly identify as White, inactive and responsible meta-beliefs were positively associated with collective action tendencies. This was in line with our expectation of non-defensive and compensatory mechanisms among this subgroup. Conversely, among individuals who strongly identified as White, responsible meta-belief was negatively associated with collective action tendencies. This was in line with our expectation of defensive mechanisms among this subgroup. Surprisingly, inactive meta-belief did not significantly predict collective action tendencies among high identifiers. Also, inconsistent with our prediction, ally meta-belief was not significantly associated with collective action tendencies. Both of these associations, however, were in the predicted directions (negative and positive, respectively). We thus tested our hypotheses regarding these relationships again in Study 2.

6 | STUDY 2

In this second study, we set out to replicate the distinctions between ally, inactive, and responsible meta-beliefs, and to examine whether they are distinct from the more global and previously investigated meta-warmth and meta-competence. As in Study 1, we again tested our hypotheses regarding the links between the three meta-beliefs of interest and solidarity-based collective action among White Americans, including the moderating role of White identification. In addition to replicating Study 1, we also examined the psychological mechanisms underlying the relationships between inactive and responsible meta-beliefs on the one hand, and collective action tendencies on the other. We expected that among low White identifiers, endorsing inactive and responsible meta-beliefs would be linked to higher collective action tendencies via perceived collective guilt and obligation to act against racial inequality. Conversely, among high White identifiers, endorsing inactive and responsible meta-beliefs would be linked to lower collective action tendencies via perceived unfairness. Finally, we again tested the hypothesis that endorsing ally meta-belief would be positively associated with collective action tendencies, regardless of White identification.

6.1 | Method

All materials for Study 2 are available on OSF.

6.1.1 | Participants

We conducted a power simulation using the software R, using parameters from Study 1. The simulation suggested that a sample size of 400 participants would be adequate (critical alpha 0.05; $1-\beta > 0.80$).

Four hundred and eight participants were recruited via Amazon's Mechanical Turk using the same inclusion criteria as in Study 1. Our screening of the data resulted in the exclusion from the data of 19 participants who self-reported to be either not White/Caucasian or not U.S. citizens or permanent residents of the U.S., and of 14 who did not pay sufficient attention (indicated by their incorrect answers to the attention check questions). Our final sample consisted of 375 participants (60% women; age M = 41, SD = 12.56, range = 19–74).

6.1.2 | Procedure and measures

Participants followed the same procedure as in Study 1 and completed the following measures in the order presented below, unless specified otherwise. All items were measured on 9-point visual analogue scales.

Meta-warmth and meta-competence

Participants were asked to indicate how likely they thought that most Black people in the U.S. would hold a number of beliefs about White people. Derived from the Stereotype Content Model (Fiske, 1998), four items tapped into meta-warmth (e.g., "In general, Black people believe that White people are kind"; $\alpha = 0.97$) and four items tapped into meta-competence (e.g., "In general, Black people believe that White people are intelligent"; $\alpha = 0.95$).

TABLE 2 Means, standard deviations, and bivariate correlations between variables of interest

	Mean (SD)	Inactive MB	Ally MB	Collective guilt	Obligation to act	Unfairness	White ID	CA tendencies
Responsible MB	7.24 (1.48)	0.75**	−0.50**	0.09	0.11*	0.12*	−0.06	0.04
Inactive MB	6.94 (1.52)	1	−0.54**	0.07	0.04	0.11*	−0.04	0.02
Ally MB	4.25 (1.75)		1	0.10	0.08	−0.06	0.09	0.13*
Collective guilt	5.06 (2.54)			1	0.55**	−0.59**	−0.21**	0.58**
Obligation to act	7.09 (1.94)				1	−0.32**	−0.29**	0.63**
Unfairness	6.67 (1.82)					1	0.28**	−0.39**
White identification	5.67 (1.71)						1	−0.29**
Collective action tendencies	4.90 (2.37)							1

*p < .01.

**p < .001.

Meta-beliefs

Participants were presented with 9 meta-belief items, in random order, taken from Study 1. Three items captured each of the three meta-beliefs of interest: the ally meta-belief (α = 0.83), the inactive meta-belief (α = 0.86), and the responsible meta-belief (α = 0.88).

Mediators

The presentation order of the mediators was counterbalanced, such that participants were randomly assigned to either respond first to the items measuring the non-defensive mediators (i.e., collective guilt, obligation to act) and then the defensive mediator (i.e., unfairness), or vice versa.

Collective guilt. Adapted from Branscombe, Slugoski, and Kappen (2004), three items measured participants' collective guilt (e.g., "I feel guilty about White Americans' harmful actions towards Black Americans"; α = 0.90).

Obligation to act. Collective obligation to act was measured using three items asking participants how much they felt that White Americans have a collective obligation to take action against racial inequality (e.g., "I believe that White Americans should act together against Anti-Black discrimination"; α = 0.94).

Unfairness. Adapted from Sullivan, Landau, Branscombe, and Rothschild (2012), unfairness was measured using three items (e.g., "It is unfair if people think that just because I'm a White American, I should feel guilty for the suffering of Black Americans"; α = 85).

Collective action

The same five items from Study 1 measured participants' willingness to engage in collective action against racial injustice (α = 0.90).

White identification

The same seven items from Study 1 measured participants' White identification (α = 0.88).

6.2 | Results

6.2.1 | Preparatory analyses

As a preparatory step, we first aimed to replicate the distinctions between ally, inactive, and responsible meta-beliefs, and to additionally examine whether these meta-beliefs were distinct from the more global and previously investigated meta-warmth and meta-competence. Full results of CFAs[5] are provided in Appendix S1.

With the items of ally, inactive, and responsible meta-beliefs, we tested the same two models as in Study 1. Again, the model with positive items (i.e., ally meta-beliefs) and negative items (i.e., inactive and responsible meta-beliefs) loading onto two latent factors, did not yield acceptable fit, $\chi^2(26) = 180.92$, $p < .001$; RMSEA = 0.13, CFI = 0.93, although it did meet the SRMR criterion, SRMR = 0.06. The model with the meta-belief items of each type (i.e., ally, inactive, and responsible) loading onto separate factors, however, yielded an acceptable fit, $\chi^2(24) = 78.01$, $p < .001$; RMSEA = 0.08, SRMR = 0.05, CFI = 0.97.

To scrutinize whether ally, inactive, and responsible meta-beliefs were distinct from the previously studied meta-warmth and meta-competence, we tested a model with five latent factors, and items of meta-ally, meta-inactive, meta-responsible, meta-warm, and meta-competence loading onto a separate factor. This model yielded an acceptable fit, $\chi^2(109) = 224.37$, $p < .001$; RMSEA = 0.05, SRMR = 0.04, CFI = 0.98.

6.2.2 | Main analyses

Means, standard deviations, and bivariate correlations between our variables are reported in Table 2.

---

[5] We also conducted exploratory factor analyses (EFAs), first on all the meta-belief, then on all the meta-belief items and meta-warmth and meta-competence. Full results of the EFAs are provided in Appendix S1.

TABLE 3 Direct and indirect effects of responsible and inactive meta-beliefs on collective action tendencies through obligation to act and guilt among low White identifiers

	Obligation to act	Guilt
	Coefficient, CI95	Coefficient, CI95
Responsible MB → Mediator (a)	0.27* [0.089 0.443]	0.30* [0.063 0.538]
Mediator → CA tendencies (b)	0.52* [0.410, 0.627]	0.30* [0.222, 0.385]
Responsible MB → CA tendencies (c)	0.01 [−0.160 0.171]	
Responsible MB → Mediator → CA tendencies (ab)	0.14* [0.051 0.290]	0.09* [0.005 0.174]
Index of moderated mediation	$b = -0.04$ [−0.095 0.001]	$b = -0.03$ [−0.059 0.004]
Inactive MB → Mediator (a)	0.16* [0.005 0.313]	0.26* [0.042 0.487]
Mediator → CA tendencies (b)	0.52* [0.410 0.628]	0.31* [0.224 0.387]
Inactive MB → CA tendencies (c)	0.08 [0.188 0.120]	
Inactive MB → Mediator → CA tendencies (ab)	0.08* [0.004 0.188]	0.08* [0.0008 0.160]
Index of moderated mediation	$b = -0.04$ [−0.085 0.003]	$b = -0.03$ [−0.060 0.002]

*Significant effects at 95% CI.

To test the relationship between ally meta-belief and collective action tendencies, moderated by White identification, we conducted the same moderated regression as in Study 1 (Hayes, 2018, Model 1). To test our hypotheses that inactive and responsible meta-beliefs would predict (a) more collective action tendencies via perceived collective guilt and obligation to act among low identifiers, and (b) fewer collective action tendencies via perceived unfairness via high identifiers, we then conducted two sets of moderated mediation analyses as described below (Hayes, 2018, Model 8). All analyses were conducted using 5,000 bootstrap samples and 95% confidence intervals. The predictor and the moderator were again mean centered.[6]

Ally meta-belief

The overall model was significant, $F(1, 375) = 15.74$, $p < .001$, $R^2 = .11$. As we expected, ally meta-belief positively predicted collective action tendencies ($b = 0.21$, $SE = 0.07$, CI95 [0.075, 0.338]). There was no significant interaction between ally meta-belief and White identification ($b = 0.04$, $SE = 0.03$, CI95 [−0.030, 0.104]). In addition, White identification negatively predicted collective action tendencies ($b = -0.42$, $SE = 0.07$, CI95 [−0.554, −0.286]).

Next, we tested the indirect effects of inactive and responsible meta-beliefs on collective action tendencies via collective guilt and obligation to act, moderated by White identification. In each model, inactive or responsible meta-belief was introduced as the predictor, collective guilt and obligation to act as parallel mediators, identification as the moderator, and collective action tendencies as the outcome (Hayes, 2018, model 8). The results of these moderated mediations are summarized in Table 3.

Inactive meta-belief

As hypothesized, the indirect effect of inactive meta-belief on collective action tendencies through collective guilt and obligation to act were both significant among *low* White identifiers. The more strongly low White identifiers held an inactive meta-belief, the more they experienced collective guilt and perceived an obligation to act. In turn, collective guilt and obligation to act both positively predicted collective action tendencies. Among *high* White identifiers, neither of the indirect effects (collective guilt, $b = -0.05$, $SE = 0.07$, CI95 [−0.188, 0.071]; obligation to act, $b = -0.02$, $SE = 0.04$, CI95 [−0.108, 0.043]) was significant.

Responsible meta-belief

As hypothesized, the indirect effects of responsible meta-belief on collective action tendencies through collective guilt and obligation to act were both significant among *low* White identifiers. The more strongly low White identifiers held a responsible meta-belief, the more they experienced collective guilt and perceived a collective obligation to act. In turn, collective guilt and obligation to act both positively predicted collective action tendencies. Among *high* White identifiers, in contrast, neither of the indirect effects (collective guilt, $b = -0.01$, $SE = 0.04$, CI95 [−0.096, 0.061]; obligation to act, $b = -0.004$, $SE = 0.06$, CI95 [−0.130, 0.114]) was significant.

We also tested indirect effects of inactive and responsible meta-beliefs on collective action tendencies via perceptions of unfairness, moderated by White identification.[7] We again conducted two

---

[6]We also tested all five models while entering meta-warmth and meta-competence as a covariate. The patterns of results are largely consistent with those reported here, with the exception of the indirect effects of inactive and responsible meta-beliefs through unfairness among high identifiers. The full additional results are reported in Appendix S1.

[7]We also tested the full model displayed in Figure 1, where we tested the indirect effects of inactive and responsible meta-beliefs, separately, on collective action tendencies via collective guilt, obligation to act, and unfairness as parallel mediators, moderated by White identification. The indirect effects of inactive and responsible meta-beliefs on collective action tendencies via collective guilt and obligation to act among low White identifiers were significantly positive. The indirect effects of inactive and responsible meta-beliefs on collective action tendencies via unfairness among high White identifiers were non-significant, but in the expected negative direction.

TABLE 4 Direct and indirect effects of responsible and inactive meta-beliefs on collective action tendencies through unfairness among high White identifiers

	Unfairness
	Coefficient, CI95
Responsible MB → Unfairness (a)	0.16* [0.040, 0.281]
Unfairness → CA tendencies (b)	−0.44* [−0.554, −0.324]
Responsible MB → CA tendencies (c)	0.64* [0.213, 1.072]
Responsible MB → Unfairness → CA tendencies (ab)	−0.10* [−0.212, −0.016]
Index of moderated mediation	b = −0.01 [−0.048 0.024]
Inactive MB → Unfairness (a)	0.14* [0.022, 0.255]
Mediator → CA tendencies (b)	−0.44* [−0.554 −0.324]
Inactive MB → CA tendencies (c)	0.35 [−0.086 0.794]
Inactive MB → Unfairness → CA tendencies (ab)	−10* [−0.203 −0.014]
Index of moderated mediation	b = −0.02 [−0.049 0.013]

*Significant effects at 95% CI.

moderated mediation analyses where inactive and responsible meta-beliefs were respectively introduced as the predictor, unfairness as a mediator, identification as a moderator, and collective action tendencies as the outcome (Hayes, 2018, model 8). Results of these moderated mediations are summarized in Table 4.

Inactive meta-belief

As hypothesized, the indirect effect of inactive meta-belief on collective action tendencies through unfairness was significantly negative among *high* White identifiers. This indirect effect indicates that high White identifiers' endorsement of an inactive meta-belief predicted more perceived unfairness, which in turn predicted less willingness to engage in collective action. The indirect effect of inactive meta-belief on collective action tendencies through unfairness, by contrast, was not significant among *low* White identifiers, b = −0.03, SE = 0.04, CI95 [−0.107, 0.037]).

Responsible meta-belief

As hypothesized, the indirect effect of responsible meta-belief on collective action tendencies through unfairness was significantly negative among *high* White identifiers. This indirect effect indicates that high White identifiers' endorsement of a responsible meta-belief predicted more perceived unfairness, which in turn predicted less willingness to engage in collective action. The indirect effect of responsible meta-belief on collective action tendencies through unfairness, by contrast, was not significant among *low* White identifiers (b = −0.05, SE = 0.04, CI95 [−0.144, 0.020]).

6.3 | Discussion

In this study, results of the factor analyses replicated the factor structure of the three meta-beliefs that we found in Study 1,

suggesting again that they are meaningfully distinguishable from each other. Additionally, they also lent support to our argument that ally, inactive, and responsible meta-beliefs are distinct from meta-warmth and meta-competence, two intergroup meta-perceptions frequently examined in previous research.

Furthermore, our main analyses provided consistent support for our hypotheses. In line with our prediction, we found a significant positive association between endorsement of the ally meta-belief and collective action tendencies, regardless of identification (see a mini meta-analysis of effect across studies in Appendix S1; Goh, Hall, & Rosenthal, 2016). We also found evidence for our hypothesis that perceived collective guilt and obligation to act would explain the positive association between inactive and responsible meta-beliefs on the one hand, and collective action tendencies on the other, among *low* White identifiers. In other words, among individuals who weakly identified as White, endorsing an inactive or a responsible meta-belief positively predicted their experiences of guilt and their feeling that White Americans had an obligation to act to redress racial inequality. Guilt and obligation to act in turn predicted their higher willingness to engage in solidarity-based collective action. Conversely, also as expected, a sense of unfairness explained the negative associations between inactive and responsible meta-beliefs on the one hand, and collective action tendencies on the other, among *high* White identifiers. In other words, among individuals who strongly identified as White, endorsing an inactive or responsible meta-belief predicted a stronger sense that they were judged unfairly, which in turn predicted their lower willingness to engage in solidarity-based collective action.

7 | GENERAL DISCUSSION

In the two studies presented here, we extended the research on solidarity-based collective action, by investigating the role of advantaged group members' beliefs about how the disadvantaged group think of them (i.e., intergroup meta-beliefs) in shaping their willingness to engage in action in support of the disadvantaged. We argued that a truly social psychological approach to solidarity-based collective action, which inherently involves working for and with the outgroup, should include the meta-level. Specifically, we reasoned that advantaged group members' meta-beliefs regarding how the disadvantaged view their role in perpetuating and redressing intergroup inequality predict their collective action tendencies.

In the context of racial inequality in the U.S., we examined three meta-beliefs that members of the advantaged group, White Americans, likely hold: allyship, inactivity, and responsibility. The results largely supported our predictions. First, the more White Americans believed that Black Americans think of them as allies in the fight against racial inequality, the more they were willing to engage in collective action in support of Black Americans, regardless of their level of identification (see a mini meta-analysis of effect across studies in Appendix S1). It might be worth noting that the items we used to measure the ally meta-belief included a mix of past perfect

and present tenses. This could have potentially posed an issue for our prediction that the ally meta-belief would positively predict solidarity. The use of the past perfect tense might have a moral licensing effect, such that the belief that Black Americans think of them as having contributed to anti-discrimination efforts might reduce White Americans' intentions for future actions. Our results across the two studies, however, showed that it was not the case.

Second, the relationships between inactive and responsible meta-beliefs and collective action were moderated by White identification. Among low White identifiers, the more they believed that Black Americans think of them as inactive in the fight against racial inequality, or responsible for Black Americans' ongoing struggles, the more they felt collective guilt and an obligation to act. This in turn predicted their higher willingness to engage in collective action in support of Black Americans. Among high White identifiers, in contrast, the more they believed that Black Americans think of them as inactive or responsible, the more they felt a sense of unfairness. This in turn predicted their lower willingness to engage in collective action in support of Black Americans. These findings make both theoretical and practical contributions.

Indeed, our findings highlight the importance of considering the meta-perspective in our accounts of solidarity-based collective action, by suggesting that advantaged group members' beliefs regarding what the disadvantaged think of them substantially contribute to their willingness to engage in solidarity. This opens up an avenue for research on solidarity-based collective action to further investigate a previously ignored set of cognitions: meta-beliefs. We argue that such research can enrich our understanding of how advantaged group members come to decide to engage in collective action in support of a disadvantaged outgroup. While the existing literature has focused on instrumental, emotional, and ideological predictors of such engagement, it has overlooked its relational nature. Specifically, when the advantaged choose to engage in solidarity, they are participating in an undertaking in support of a group, the disadvantaged, to which they do not belong. Furthermore, this disadvantaged group is potentially in direct contestation with their own (advantaged) group. To respond to the disadvantaged and their challenges to inequality, it is reasonable to assume that the advantaged gauge the disadvantaged's attitudes and beliefs towards them, and what role they are perceived to play in the status quo and its potential change. Our findings indicate that it might be crucial to factor in these perceptions, if we are to fully understand the motivation or hesitation of the advantaged to join solidarity-based collective action. Interestingly, our results suggest that the relationships between the negative intergroup meta-beliefs and solidarity were explained by collective guilt and a sense of obligation to act. These findings suggest that intergroup emotions and cognitions can be triggered by perceptions of what other people think, not just by individuals' own perceptions of a situation. This influence of others' supposed beliefs is arguably a novel demonstration of the "social life of emotions" (Tiedens, Leach, & Oatley, 2004), highlighting the ways in which people's experience of the world is partially shaped by their understanding of others' experiences.

Our findings also contribute to the literature on meta-beliefs, which has investigated more general intergroup meta-perceptions, such as meta-warmth and meta-competence, rooted in the stereotype content model (Fiske, 1998). We extended this research by examining a new set of meta-beliefs, which are more concrete and relational in nature. Our results suggest that people hold such concrete, relational meta-beliefs, encompassing concepts such as allyship, inactivity in the fight for equality, and intergroup responsibility, and that these meta-beliefs are linked to their collective action tendencies.

Importantly, we also showed that the links between the meta-beliefs and solidarity-based collective action were moderated by ingroup identification. This highlights a third contribution of the current work. It extended to the realm of meta-beliefs the previous research on ingroup identification as moderating intergroup attitudes and beliefs. Indeed, ingroup identification has been shown to modulate a large array of intergroup phenomena, such that high identifiers generally tend to be defensive in the face of their groups' wrongdoings, while low identifiers tend to act in compensatory ways (e.g., Bilali et al., 2012). Our findings indicate that this moderating role of ingroup identification also applies to the meta-level. Whether people dealt defensively or in a compensatory manner with their belief that outgroup members think of them negatively was contingent on how much they identified with their group.

7.1 | Practical implications

These conceptual links could potentially offer practical recommendations to disadvantaged group members, or activists more generally, who are looking to increase support against intergroup inequality among the ranks of the advantaged. It seems that it is crucial to factor in identification as an important psychological moderator when targeting advantaged group members, perhaps by using different messaging on weakly versus strongly identified individuals. Our results point to the possibility that individuals who do not strongly identify with the advantaged group might be responsive to narratives centering on the inactivity or responsibility of their ingroup, while this same messaging might backfire among high identifiers. Interestingly, our results also point to the possibility that strongly identified individuals, along with weakly identified, might be motivated by the portrayal of their ingroup as an integral part of the fight against social inequality. Thus, the findings illuminate important ways in which one type of messaging might lead to both more and less engagement in collective action, depending on the audience.

7.2 | Limitations and future directions

Several limitations of the current research are also worth mentioning. First, the correlational nature of our data prevents us from making causal claims. It would therefore be useful for future research

to attempt to directly manipulate intergroup meta-beliefs to experimentally test their effects on willingness to engage in solidarity-based collective action. This effort is especially relevant to the previously suggested practical recommendations. While our studies explored the association between personally endorsed meta-beliefs and collective action tendencies, experimental studies can investigate the effects of activating or inducing ally, responsible, and inactive meta-beliefs using tailored messaging. It would be important to explore whether the psychological correlates of advantaged group members' endorsed meta-beliefs, which were the focus of the current work, map onto the consequences of exposing them to information about what the disadvantaged ostensibly think of them.

Second, the three meta-beliefs are clearly not exhaustive, and there might exist others that could be relevant to solidarity-based collective action. For instance, it is plausible that advantaged group members might believe that the disadvantaged think of them as benefiting from the status quo of inequality, as morally unbothered by it, or even as actively supportive of it. Future research, by extending our current findings that concrete and relational meta-beliefs are important for solidarity-based collective action, could explore whether and how other potentially existing meta-beliefs differentially predict solidarity. Importantly, while we conceptually and statistically distinguished between the two kinds of negative meta-belief, inactive and responsible, we predicted and showed that they had similar downstream associations with solidarity. Still, we argue that this should not be interpreted as an indication that the relationships between all possible negative meta-beliefs and collective action tendencies should be expected to replicate our findings. The more general negative meta-beliefs that are not rooted in inequality or relevant for collective action, for example, might not substantially contribute to solidarity-related action tendencies. It would be useful for future research to explore other meta-beliefs that, similar to inactive and responsible, could be particularly relevant for collective action against different forms of injustice.

Third, we can speculate that the intergroup meta-beliefs that are relevant and the extent to which they contribute to shaping solidarity-based collective action might be dependent on the particularities of the intergroup context under investigation. For example, it seems possible that in situations of protracted conflict, where there is very little positive interaction between groups, the ally meta-belief might be less relevant. Still, responsible meta-belief could potentially promote support for compensatory actions, including solidarity-based collective action, among weakly identified advantaged group members. Future research could explore the role of the meta-level in predicting solidarity-based collective action in a wider array of intergroup contexts, to provide boundary conditions for when we should expect intergroup meta-beliefs to matter for solidarity, and when not. Similarly, we expect the moderating role of identification with the advantaged group to depend on the context and/or the content of identification (e.g., Roccas et al., 2006). For example, the interaction between advantaged identification and different meta-beliefs might be contingent on the norms associated with the identity. If a central aspect of an advantaged identity is a commitment to equality, the expectations derived for high identifiers will likely diverge from our hypotheses. Future studies can explore how various advantaged identities, depending on their intricacies, contribute to shaping the relationship between intergroup meta-beliefs and solidarity. Furthermore, it is worth noting that other inter-individual difference variables, beyond identification with the advantaged, might moderate the links between different meta-beliefs and collective action. Other candidates for future research could be, for example, political orientation or system justification, the latter of which has been shown to explain defensive versus non-defensive attitudes in advantaged group members (Hässler, Shnabel, Ullrich, Arditti-Vogel, & SimanTov-Nachlieli, 2019).

Finally, in the current work, we chose to explore the association between the three meta-beliefs and collective action against racial inequality, without specifying to our participants what the exact manifestation of this racial inequality is. Our results are therefore limited to White Americans' collective action tendencies against racial inequality as defined broadly, and cannot, for instance, distinguish between actions organized in response to structural, overt (e.g., blatant racism), or covert (e.g., macroaggressions) instances of inequality. Future research can refine our results by exploring the potentially differential role of intergroup meta-beliefs in shaping solidarity aimed at responding to different forms of racial inequality.

8 | CONCLUSION

In sum, across two studies among advantaged group members, we found that the endorsement of an ally meta-belief was positively associated with their collective action tendencies in support of the disadvantaged, although the effect was not large and not significant in Study 1. We also found that among individuals who weakly identify with their ingroup, the more they endorsed an inactive or a responsible meta-belief, the higher their solidarity-based collective action tendencies, and this relationship was explained by their experiences of collective guilt and obligation to act. In contrast, among individuals who strongly identify with their ingroup, the more they endorsed an inactive or a responsible meta-belief, the lower their solidarity-based collective action tendencies, and this relationship was explained by a sense of unfairness. This work demonstrated the importance of the meta-level for our understanding of solidarity-based collective action among the advantaged.

CONFLICT OF INTEREST
The authors declare that there is no conflict of interest regarding the publication of this article.

ETHICS STATEMENT
The authors confirm that the research was conducted ethically, results are reported honestly, the submitted work is original and not (self-)plagiarized, authorship reflects individuals' contributions and conflicts of interest are disclosed.

TRANSPARENCY STATEMENT

The materials and data are publicly available at https://osf.io/4b57e/.

ORCID

Aya Adra https://orcid.org/0000-0001-9648-2021

REFERENCES

Alexander, M. G., Brewer, M. B., & Hermann, R. K. (1999). Images and affect: A functional analysis of out-group stereotypes. *Journal of Personality and Social Psychology*, Advance online publication. 77(1), 78. https://doi.org/10.1037/0022-3514.77.1.78

António, J. H. C., & Monteiro, M. B. (2015). Minorities' acculturation and social adjustment: The moderator role of meta-perceptions of majority's acculturation attitudes. *International Journal of Psychology*, 50(6), 422–430. https://doi.org/10.1002/ijop.12214

Barlow, F. K., Louis, W. R., & Hewstone, M. (2009). Rejected! Cognitions of rejection and intergroup anxiety as mediators of the impact of cross-group friendships on prejudice. *British Journal of Social Psychology*, 48(3), 389–405. https://doi.org/10.1348/014466608X387089

Becker, J. C., & Tausch, N. (2015). A dynamic model of engagement in normative and non-normative collective action: Psychological antecedents, consequences, and barriers. *European Review of Social Psychology*, 26(1), 43–92. https://doi.org/10.1080/10463283.2015.1094265

Bilali, R., Tropp, L. R., & Dasgupta, N. (2012). Attributions of responsibility and perceived harm in the aftermath of mass violence. *Peace and Conflict: Journal of Peace Psychology*, 18(1), 21–39. https://doi.org/10.1037/a0026671

Branscombe, N. R., Ellemers, N., Spears, R., & Doosje, B. (1999). The context and content of social identity threat. In N. Ellemers, R. Spears & B. Doosje (Eds.), *Social identity* (pp. 35–58). Oxford, UK: Blackwell.

Branscombe, N. R., Slugoski, B., & Kappen, D. M. (2004). Collective guilt: What it is and what it is not. In N. R. Branscombe & B. Doosje (Eds.), *Collective guilt: International perspectives* (pp. 16–34). Cambridge, UK: Cambridge University Press. Retrieved from https://researchonline.jcu.edu.au/7329/

Brown, R., & Cehajic, S. (2008). Dealing with the past and facing the future: Mediators of the effects of collective guilt and shame in Bosnia and Herzegovina. *European Journal of Social Psychology*, 38(4), 669–684. https://doi.org/10.1002/ejsp.466

Cohen, J. (1969). *Statistical power analysis for the behavioral sciences*. New York, NY: Academic Press.

Davidson, J. (2016). *White silence is not an option. Huffpost*. Retrieved from https://www.huffpost.com/entry/white-silence-is-not-an-o_b_10927210

Doosje, B., & Branscombe, N. R. (2003). Attributions for the negative historical actions of a group. *European Journal of Social Psychology*, 33(2), 235–248. https://doi.org/10.1002/ejsp.142

Doosje, B., Branscombe, N. R., Spears, R., & Manstead, A. S. R. (2006). Antecedents and consequences of group-based guilt: The effects of ingroup identification. *Group Processes & Intergroup Relations*, 9(3), 325–338. https://doi.org/10.1177/1368430206064637

Doosje, B., & Haslam, S. A. (2005). What have they done for us lately? The dynamics of reciprocity in intergroup Contexts1. *Journal of Applied Social Psychology*, 35(3), 508–535. https://doi.org/10.1111/j.1559-1816.2005.tb02133.x

Faul, F., Erdfelder, E., Lang, A.-G., & Buchner, A. (2007). G*Power 3: A flexible statistical power analysis program for the social, behavioral, and biomedical sciences. *Behavior Research Methods*, 39(2), 175–191. https://doi.org/10.3758/BF03193146

Figueiredo, A. M., Doosje, B., Valentim, J. P., & Zebel, S. (2010). Dealing with past colonial conflicts: How perceived characteristics of the victimized outgroup can influence the experience of group-based guilt. *International Journal of Conflict and Violence*, 4(1), 89–105. https://doi.org/10.4119/UNIBI/ijcv.64

Finchilescu, G. (2010). Intergroup anxiety in interracial interaction: The role of prejudice and metastereotypes. *Journal of Social Issues*, 66(2), 334–351. https://doi.org/10.1111/j.1540-4560.2010.01648.x

Fiske, S. T. (1998). Stereotyping, prejudice, and discrimination. In D. T. Gilbert, S. T. Fiske & G. Lindzey (Eds.), *The handbook of social psychology* (Vols. 1–2, 4th ed., pp. 357–411). New York, NY: McGraw-Hill.

Frey, F. E., & Tropp, L. R. (2006). Being seen as individuals versus as group members: Extending research on metaperception to intergroup contexts. *Personality and Social Psychology Review*, 10(3), 265–280. https://doi.org/10.1207/s15327957pspr1003_5

Goh, J. X., Hall, J. A., & Rosenthal, R. (2016). Mini meta-analysis of your own studies: Some arguments on why and a primer on how. *Social and Personality Psychology Compass*, 10(10), 535–549. https://doi.org/10.1111/spc3.12267

Hässler, T., Shnabel, N., Ullrich, J., Arditti-Vogel, A., & SimanTov-Nachlieli, I. (2019). Individual differences in system justification predict power and morality-related needs in advantaged and disadvantaged groups in response to group disparity. *Group Processes & Intergroup Relations*, 22(5), 746–766. https://doi.org/10.1177/1368430218773403

Hayes, A. F. (2018). Partial, conditional, and moderated moderated mediation: Quantification, inference, and interpretation. *Communication Monographs*, 85(1), 4–40. https://doi.org/10.1080/03637751.2017.1352100

Iqubal, N. (2019). Academic Robin DiAngelo: "We have to stop thinking about racism as someone who says the N-word". *The Guardian*. Retrieved from https://www.theguardian.com/world/2019/feb/16/white-fragility-racism-interview-robin-diangelo

Issmer, C., Stellmacher, J., & Gollwitzer, M. (2013). When disadvantaged adolescents strike out: The impact of negative metastereotypes on delinquency. *Journal of Criminal Psychology*, 3(1), 4–18. https://doi.org/10.1108/20093821311307721

Iyer, A., Leach, C. W., & Crosby, F. J. (2003). White guilt and racial compensation: The benefits and limits of self-focus. *Personality and Social Psychology Bulletin*, 29(1), 117–129. https://doi.org/10.1177/0146167202238377

Klein, O., Licata, L., & Pierucci, S. (2011). Does group identification facilitate or prevent collective guilt about past misdeeds? Resolving the paradox. *British Journal of Social Psychology*, 50(3), 563–572. https://doi.org/10.1111/j.2044-8309.2011.02028.x

Kteily, N. S., & Bruneau, E. G. (2015). They see us as less than human: Meta-dehumanization predicts intergroup conflict via reciprocal dehumanization. *Journal of Personality and Social Psychology*, 110(3), 343.

Kutlaca, M., Becker, J., & Radke, H. (2019). A hero for the outgroup, a black sheep for the ingroup: Societal perceptions of those who confront discrimination. *Journal of Experimental Social Psychology*, 88, 103832. https://doi.org/10.1016/j.jesp.2019.103832

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). Cambridge, UK: Cambridge University Press.

Leach, C. W., van Zomeren, M., Zebel, S., Vliek, M. L. W., Pennekamp, S. F., Doosje, B., … Spears, R. (2008). Group-level self-definition and self-investment: A hierarchical (multicomponent) model of in-group identification. *Journal of Personality and Social Psychology*, 95(1), 144–165. https://doi.org/10.1037/0022-3514.95.1.144

Li, M., Leidner, B., & Fernandez-Campos, S. (2019). Stepping into perpetrators' shoes: How ingroup transgressions and victimization shape support for retributive justice through perspective-taking with perpetrators. *Personality and Social Psychology Bulletin*, 46(3), 424–438. https://doi.org/10.1177/0146167219858652

Livingstone, A., Rothers, A., & Fernández, L. (2019). "They just don't understand us": The role of felt understanding in intergroup relations. *Journal of Personality and Social Psychology*, Advance online publication. https://doi.org/10.1037/pspi0000221

Louis, W. R., Thomas, E., Chapman, C. M., Achia, T., Wibisono, S., Mirnajafi, Z., & Droogendyk, L. (2019). Emerging research on intergroup prosociality: Group members' charitable giving, positive contact, allyship, and solidarity with others. *Social and Personality Psychology Compass*, *13*(3), e12436. https://doi.org/10.1111/spc3.12436

Lowery, B. S., Knowles, E. D., & Unzueta, M. M. (2007). Framing inequity safely: Whites' motivated perceptions of racial privilege. *Personality and Social Psychology Bulletin*, *33*(9), 1237–1250. https://doi.org/10.1177/0146167207303016

Morrison, M. S. (2013, n.a.). *Becoming trustworthy White allies. Reflections: A magazine of theological and ethical inquiry from Yale Divinity School.* Retrieved from https://reflections.yale.edu/article/future-race/becoming-trustworthy-white-allies

O'Brien, T., Leidner, B., & Tropp, L. R. (2017). Are they for us or against us? How intergroup metaperceptions shape foreign policy attitudes. *Group Processes & Intergroup Relations*, *21*(6), 941–961.

Otten, S. (2002). Towards a more social social psychology: The case of meta-stereotypes. *Revista de Psicología Social (International Journal of Social Psychology)*, *17*(3), 287–292. https://doi.org/10.1174/02134740260373008

Phelps, J. M. (2013). *Societal psychology in Norway* (pp. 1–5). Retrieved from https://phs.brage.unit.no/phs-xmlui/handle/11250/174702

Putra, I. E., & Wagner, W. (2017). Prejudice in interreligious context: The role of metaprejudice and majority–minority status. *Journal of Community & Applied Social Psychology*, *27*(3), 226–239. https://doi.org/10.1002/casp.2305

Roccas, S., Klar, Y., & Liviatan, I. (2006). The paradox of group-based guilt: Modes of national identification, conflict vehemence, and reactions to the in-group's moral violations. *Journal of Personality and Social Psychology*, *91*(4), 698–711. https://doi.org/10.1037/0022-3514.91.4.698

Selvanathan, H. P., Techakesari, P., Tropp, L. R., & Barlow, F. K. (2018). Whites for racial justice: How contact with Black Americans predicts support for collective action among White Americans. *Group Processes & Intergroup Relations*, *21*(6), 893–912. https://doi.org/10.1177/1368430217690908

Shepherd, L., Spears, R., & Manstead, A. S. R. (2013). "This will bring shame on our nation": The role of anticipated group-based emotions on collective action. *Journal of Experimental Social Psychology*, *49*(1), 42–57. https://doi.org/10.1016/j.jesp.2012.07.011

Subašić, E., Reynolds, K. J., & Turner, J. C. (2008). The political solidarity model of social change: Dynamics of self-categorization in intergroup power relations. *Personality and Social Psychology Review*, *12*(4), 330–352. https://doi.org/10.1177/1088868308323223

Sullivan, D., Landau, M. J., Branscombe, N. R., & Rothschild, Z. K. (2012). Competitive victimhood as a response to accusations of ingroup harm doing. *Journal of Personality and Social Psychology*, *102*(4), 778–795. https://doi.org/10.1037/a0026573

Tiedens, L. Z., Leach, C. W., & Oatley, K. (2004). *The social life of emotions*. Cambridge, UK: Cambridge University Press.

Tropp, L. R., & Bianchi, R. A. (2006). Valuing diversity and interest in intergroup contact. *Journal of Social Issues*, *62*(3), 533–551. https://doi.org/10.1111/j.1540-4560.2006.00472.x

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. https://doi.org/10.1177/1368430210395637

Vezzali, L. (2017). Valence matters: Positive meta-stereotypes and interethnic interactions. *The Journal of Social Psychology*, *157*(2), 247–261. https://doi.org/10.1080/00224545.2016.1208140

Vorauer, J. D. (2003). Dominant group members in intergroup interaction: Safety or vulnerability in numbers? *Personality and Social Psychology Bulletin*, *29*(4), 498–511. https://doi.org/10.1177/0146167202250917

Vorauer, J. D., Hunter, A. J., Main, K. J., & Roy, S. A. (2000). Meta-stereotype activation: Evidence from indirect measures for specific evaluative concerns experienced by members of dominant groups in intergroup interaction. *Journal of Personality and Social Psychology*, *78*(4), 690–707.

Vorauer, J. D., & Kumhyr, S. M. (2001). Is this about you or me? Self-versus other-directed judgments and feelings in response to intergroup interaction. *Personality and Social Psychology Bulletin*, *27*(6), 706–719. https://doi.org/10.1177/0146167201276006

Wout, D. A., Murphy, M. C., & Steele, C. M. (2010). When your friends matter: The effect of White students' racial friendship networks on meta-perceptions and perceived identity contingencies. *Journal of Experimental Social Psychology*, *46*(6), 1035–1041. https://doi.org/10.1016/j.jesp.2010.06.003

SUPPORTING INFORMATION

Additional supporting information may be found online in the Supporting Information section.