Gollwitzer, A., Olcaysoy Okten, I., Pizarro, A. O., & Oettingen, G. (2022, May 19). Discordant Knowing: A Social Cognitive Structure Underlying Fanaticism. Journal of Experimental Psychology: General. Advance online publication. http://dx.doi.org/10.1037/xge00012

Discordant Knowing:

A Social Cognitive Structure Underlying Fanaticism

Anton Gollwitzer[1,2*] Irmak Olcaysoy Okten[3] Angel Osorio Pizarro[4] Gabriele Oettingen[3]

[1]Department of Leadership and Organizational Behavior, BI Norwegian Business School, Oslo, Norway

[2]Center for Adaptive Rationality, Max Planck Institute for Human Development, Berlin, Germany

[3]Department of Psychology, New York University, New York, NY, USA

[4]Department of Psychology, Yale University, New Haven, CT, USA

Author Note: All verbatim material files, data files, and data analysis files can be found here.

**Correspondence concerning this article should be addressed to: Anton Gollwitzer, Department of Leadership and Organizational Behaviour, BI Norwegian Business School, Nydalsveien 37, 0484 Oslo, Norway. Email: anton.gollwitzer@gmail.com*

Abstract

Examining the epistemic and social cognitive structures underlying fanaticism, radicalization, and extremism should shed light on how these harmful phenomena develop and can be prevented. In 9 studies (N = 3,277), we examined whether discordant knowing—felt knowledge about something that one perceives as opposed by most others—underlies fanaticism. Across multifaceted approaches, experimentally manipulating participants' views to fall under this framework (e.g., I am certain about X, but most other people think X is unknowable or wrong) heightened indicators of fanaticism, including aggression, determined ignorance, and wanting to join extreme groups in the service of these views. Additional analyses found that this effect occurs via threat-based mechanisms (Studies 1-7), can be intervened on to prevent fanaticism (Study 2), is conditional on the potency of opposition (Study 3), differs from effects on extremism (Study 4), and extends to mental representations of the self (Study 5). Generalizing these findings to real-world contexts, inducing participants with discordant knowledge about the 2020 U.S. Presidential Election and the morality of abortion heightened fanaticism regarding these topics (Studies 6 & 7). Additionally, anti-vaccine fanatics and followers of a real-world fanatical religious group exhibited greater discordant knowing than non-fanatical individuals (Studies 8 & 9). Collectively, the presented studies reveal that discordant knowing contributes to fanaticism, and further, highlight the potential of investigating constructs like fanaticism from an epistemic social cognitive perspective.

Keywords: social psychology, social cognition, fanaticism, discordant knowing, extremism, epistemology, reverse correlation, anti-vaccine, COVID-19

Discordant Knowing: A Social Cognitive Structure Underlying Fanaticism

People sometimes feel certain even in the face of majority opposition. Someone may be certain that God exists despite thinking that most people judge this claim as unknowable or wrong. Or someone may recognize others' claims that a romantic prospect is not interested in them, but nonetheless be convinced that they are meant to be. Or someone may "know" that a specific political ideology should be followed, despite others' claims to the contrary. We term this social cognitive structure—*discordant knowing*—felt knowledge or certainty about something that one perceives as opposed by the majority of other people, for instance, in terms of being judged as unknowable or inaccurate. Potentially, people adopt this isolating epistemic structure in an effort to satiate desires for certainty, control, and uniqueness (e.g., Gollwitzer & Oettingen, 2019; Hofstede, 1991; Imhoff & Lamberty, 2017; Kruglanski & Orehek, 2012; Webster & Kruglanski, 1994; Whitson & Galinsky, 2008).

Discordant knowing has two components. The first component entails felt knowledge or certainty—being sure about an opinion or viewpoint (Bagehot, 1871; Burton, 2008). Felt knowledge or certainty is captured by various psychological variables, including mental rigidity and dogmatism (e.g., Altemeyer, 2002; Bastian et al., 2015), overclaiming (Atir et al., 2015), belief superiority (Hall & Raimi, 2018), attitude certainty (Abelson, 1995; Gross et al. 1995; Krosnick & Petty, 1995), and moral convictions (Skitka, 2010), among other constructs. Notably, past research has differentiated felt knowledge or certainty from beliefs or views held with uncertainty and doubt. For instance, felt knowledge is held with greater rigidity and confidence than beliefs (e.g., DeRose, 2009; p. 186; Kirkpatrick et al., 1991). Additionally, while beliefs usually update in a Bayesian sense (based on new information), certainty is not impacted by, and may even increase, in response to counter-information (see Petty, 2021). Finally, while felt knowledge is

informed by people's desires and relates to greater intuitive thinking (Gollwitzer & Oettingen, 2019), beliefs are typically informed by experiences and facts (Bandura, 1977; Mischel, 1973; Oettingen, 2012; Schunk, 2012). Taken together, then, felt knowledge or certainty involves a higher degree of mental rigidity than belief and is less founded in one's experience and factual reality.

Notably, the past research on certainty has largely not considered whether people think their felt knowledge is opposed by others—the perceived level of agreement between one's certainty and the outside world. In this vein, the second component of discordant knowledge involves perceiving one's claim as being generally opposed by others, that is, judged as unknowable or wrong by the majority of other people. Some past work has examined the phenomenology of such opposition. Researchers have examined the experience of holding minority perspectives (e.g., Gardikiotis, 2011; Nemeth, 1986; Moscovici, 1980), stigmatized beliefs (e.g., Lantian et al., 2018), and views on the fringe of accepted belief systems (e.g., Barkun, 2015; Kruglanski et al., 2017). Yet, this research has predominantly failed to consider whether an opposed viewpoint or claim is held with certainty. Here, across 9 studies, we combine these two independently studied epistemic components, certainty and opposition, by examining discordant knowing: Being certain about a view or claim that one perceives as opposed by the majority of others, for instance, in terms of being judged as unknowable or wrong.

To better understand discordant knowing, consider several examples. In the political domain, discordant knowing can be captured by conspiracy theories or alternate facts that are held with certainty, as such theories entail secret information generally considered unknowable or inaccurate by most others (Lantian et al., 2017; Van Prooijen, 2018). Relatedly, religious views held with complete certainty (religious fundamentalism; e.g., Hill & Williamson, 2005;

Kirkpatrick et al., 1991) should also qualify as discordant knowing, if the individual perceives most others as opposing these views. And, in the romantic domain, stalking and unrequited romantic obsessions can entail discordant knowing because stalking often involves being certain about someone's affections despite the target and outside world opposing this certainty (see Cupach & Spitzberg, 2004). Additionally, and perhaps more generally, experiencing identity denial, the denial of one's personal or social identity, may qualify as discordant knowing (see Gollwitzer et al., 2013; Hogg et al., 2007; Sityaeva et al., 2020; McGregor et al., 2001) if the source of the denial is the majority of other people, and the targeted convictions about the self are held with certainty. Finally, we note that discordant knowing is not restricted to the described domains; it can span across various life-domains, including professional, interpersonal, health, political, and societal domains.

Discordant Knowing Underlies Fanaticism

Here, we introduce discordant knowing and hypothesize that this social cognitive structure underlies fanaticism. In particular, we propose that holding views with high certainty, juxtaposed with perceiving the outside world as doubting or denying these views, is one pathway to fanaticism. Though fanaticism has been defined in numerous ways, these definitions largely entail aggressive, all-encompassing, and uncritical zeal about a viewpoint, often associated with extreme beliefs, groups, or movements (see Marimaa, 2011). Specifically, in cultural anthropology, sociology, psychology, and moral philosophy, researchers have argued that behavioral indicators of fanaticism entail: **(1)** aggression—a willingness to "to destroy those who threaten the fanatically held belief" (Mead, 1977; p. 37) and harm those who have different opinions (Calhoun, 2004), **(2)** determined ignorance—a mindset that is "closed to argument and reason" (Milgram, 1977, p. 58), ignores facts and consequences (Perkinson, 2002), and entails closed eyes and ears to counter-

information (Hoffer, 1951), and **(3)** joining and adhering to extreme or fanatical groups and movements that support one's viewpoint ("true believers"; Hoffer, 1951; Stankov et al., 2010). In the present work, we directly examine whether discordant knowing incites these three behavioral indicators of fanaticism.

Theoretical Support

Past theorizing supports our hypothesis that the combination of certainty and majority opposition underlies fanaticism. For instance, theorists have separately discussed fanaticism as entailing the possession of absolute truths (Calhoun, 2004), holding beliefs that conflict with the outside world (Milgram, 1977), holding psychological tension between one's ideal and actual reality (e.g., Lehtsaar, 1997; Selengut, 2017), and holding a dualistic world view—us/me versus them (see Marimaa, 2011). Given that discordant knowing neatly captures these conflicting dualities, this past theorizing supports the present hypothesis.

Notably, past work has also framed fanaticism as residing in the mind (e.g., Marimaa, 2011; Milgram, 1977); extending this idea to the current work, we frame discordant knowing as subjective. That is, discordant knowing can induce fanaticism even if opposition is purely imagined and not rooted in real-life experiences. For instance, someone may be certain that climate change exists and perceive most others as denying this claim despite that, in reality, most people recognize climate change. As such, we posit that certainty combined with *perceived* opposition from the external world is sufficient to produce fanaticism.

Empirical Support

Empirical work also supports our predictions. For instance, judging one's attitudes as correct (attitude correctness) predicts greater competitiveness, anger, and confrontational tendencies (Niedbala et al., 2018; Rios et al., 2014)—constructs thematically related to

fanaticism—in response to opposition. Additionally, attitude certainty has been linked to resisting attitude change (Bassili, 1996; Tormala & Petty, 2002) as well as ignoring contradictory information (Jacks & Cameron, 2003; Knowles & Linn, 2004), constructs that qualify as determined ignorance, one of the noted indicators of fanaticism.

Our hypothesis also aligns with past research on moral convictions, social identity, and conspiracy theories. For instance, strong moral convictions, which fall under the first component of discordant knowing (certainty), are linked to fanatical outcomes, including aggression, violence, and terrorism (e.g., Mooijman et al., 2018). Additionally, identity denial and conspiracy theories, which include the second component of discordant knowing (external opposition; e.g., conspiracy theories are by definition discounted by most others), have been linked to greater fanatical attitudes (e.g., Gurr, 1970; Hogg et al., 2007; Sityaeva et al., 2020; McGregor et al., 2001; Rousis et al., 2020). Notably, these findings not only support a link between discordant knowing and fanaticism but, if our theorizing is correct, the present work should provide a general framework—discordant knowing—via which to understand exactly why and when these previously studied phenomena link to fanaticism.

Past work has also more directly linked discordant knowing to fanaticism. Specifically, an epistemic structure similar to discordant knowing—called paradoxical knowing—has been linked to fanaticism. Gollwitzer and Oettingen (2019) defined paradoxical knowing as certainty about something despite recognizing this thing as technically unknowable (e.g., "I am certain God exists even though I realize that this is technically unknowable"; akin to Moore's Paradox; see Moore, 1942 and Wittgenstein, 1953). To contrast discordant knowing with paradoxical knowing, while discordant knowing entails *external* social opposition—one's certainty being challenged by the external social world (opposition from *external others*), paradoxical knowing entails *internal*

opposition—internally recognizing one's own certainty as unfounded or lacking tangible evidence (opposition within the *self*; see Figure 1). As an example, consider certainty in God's existence. Discordant knowing would entail feeling certain about God's existence while perceiving external doubt—perceiving most others as denying or judging God's existence as unknowable. Paradoxical knowing, on the other hand, would entail feeling certainty about God's existence while holding internal doubt—acknowledging God's existence as technically unknowable (independently of others' judgments). Despite the noted differences, however, past work linking paradoxical knowing to fanaticism supports the current hypothesis given that discordant and paradoxical knowing conceptually overlap (in terms of the certainty component; see Figure 1).

But why should discordant knowing be a better predictor of fanaticism than paradoxical knowing? For one, all three indicators of fanaticism noted here (e.g., aggression) are social phenomena (by being inherently directed at "others"). Discordant knowing, unlike paradoxical knowing, entails external *social* opposition (i.e., "other people" opposing one's certainty) and thus may be a better fit in terms of predicting fanaticism. Indeed, paradoxical knowing may have solely been linked to indicators of fanaticism in past work because paradoxical knowing shares variance with discordant knowing. In a number of the studies presented here, we directly test whether discordant knowing is a better predictor of fanaticism than paradoxical knowing, and further, whether discordant knowing accounts for the previously observed link between paradoxical knowing and fanaticism.

Mechanism: Threat. Though past work supports an effect of discordant knowing on fanaticism, what processes drive this potential effect? Discordant knowing may contribute to fanaticism because it incites feelings of threat from the outside world. Because discordant knowing involves perceiving most others as incapable or refusing to substantiate one's felt knowledge or

certainty, it should incite feelings of epistemic threat and perhaps even persecution. And this experience of threat should be substantial—knowing can be thought of as a form of possession (a 'possession' of truth; Abelson, 1986), and people are particularly averse to losing possessions (Tversky & Kahneman, 1991). As such, people holding discordant knowledge should feel as if the outside world is attempting to steal a prized possession from them. In line with this reasoning, the three noted indicators of fanaticism directly parallel three major responses to threat: fight (aggression), flight (determined ignorance), and befriend (joining like-minded groups; e.g., Cannon, 1929; Taylor et al., 2000). Finally, we note that the proposed mediator, threat, should be specific to epistemic threat. For instance, encountering a tiger heightens threat but is unlikely to make you fanatical. Instead, specifically epistemic threat—threat induced by participants' felt knowledge being challenged—should evoke fanaticism.

Self-Other Conflict, Extremism, and Attitude Strength

Before turning to the conducted studies, we differentiate the current work from three related constructs—attitude strength, conflict between the self and others, and extremism.

Attitude Strength. Past work has linked attitude strength to select indicators of fanaticism, including attitude-relevant violent or defensive behaviors (e.g., the willingness to die for one's country, disregarding counter-attitudinal information; e.g., Gómez et al., 2020; Paredes et al., 2020; Shaw et al., 2011; Bassili, 1996; Tormala & Petty, 2002). Additionally, work on attitude certainty shows that external opposition to one's attitudes heightens constructs related to fanaticism (e.g., anger; Niedbala et al., 2018; Sawicki & Wegener, 2018). Unlike this past work, however, we examine fanaticism holistically rather than its single deleterious indicators (e.g., aggression), and further, examine threat as a mechanism leading to fanaticism and how intervening on threat can potentially prevent fanaticism. Additionally, unlike past work, we directly manipulate

the two components of discordant knowing—certainty and opposition. Finally, to directly differentiate discordant knowing from attitude strength, in the present work, we also tested whether discordant knowing predicts fanaticism above and beyond people's attitude strength (specifically, anti-vaccine attitudes).

Self-Other Conflict. Discrepancies similar to discordant knowing, in terms of conflict between the individual and the outside world, have been previously studied. For instance, research on the conflict between one's claims and others' opinions have been examined in terms of psychological reactance and defensiveness (e.g., Baumeister et al., 2000; Belanger et al., 2020), resistance to changing one's attitudes (Bassili, 1996; Knowles & Linn, 2004; Tormala & Petty, 2002) and ostracism (e.g., Hales & Williams, 2018). These past lines of work, however, did not examine fanaticism, and further, largely did not directly examine the two components of discordant knowing—certainty and opposition. For instance, the experience of conflict should be greatest when one is certain and perceives substantial (high potency) opposition from the outside world (see Lewin, 1946)—that is, when one holds discordant knowledge.

Extremism. Researchers have extensively examined a construct linked to fanaticism—extremism (holding non-normative views with intensity and firmness; e.g., Hogg, 2004; Klein & Kruglanski, 2013; Kruglanski et al., 2021; Wintrobe, 2006). Importantly, extremism differs from fanaticism, as defined here, in that one can exhibit extremism without exhibiting fanaticism. For instance, one can endorse a conspiracy theory with absolute firmness (extremism) but not engage in aggression to support this theory, not determinately ignore all counterinformation, and not actively seek out extreme groups supporting this theory (fanaticism; Calhoun, 2004; Hoffer, 1951; Mead, 1977; Milgram, 1977; Perkinson, 2002). As such, we conceptualize extremism and fanaticism as distinct (with the latter being more concerned with behavioral responses; Schuurman

& Taylor, 2018), and in turn, empirically test whether specific epistemic structures, including discordant knowing, differentially predict these two constructs.

The Present Research

In Studies 1 and 2, we tested whether experimentally manipulating people's views to fall under a discordant knowing framework (e.g., "I am certain about something that is judged by most others as unknowable") heightens the three noted indicators of fanaticism—aggression, determined ignorance, and wanting to join extreme groups in the service of these views. Additionally, we examined whether this potential effect is mediated by feeling threatened by the outside world, and whether intervening on this mediator can prevent the onset of fanaticism.

Studies 3 through 5 conceptually replicated and extended these results. In Study 3, we focused on the opposition component of discordant knowing. We tested whether discordant knowing incites fanaticism across varying forms of opposition (e.g., certainty judged by most others as "unknowable" versus "wrong"). And, we examined whether high opposition potency (certainty opposed by the majority) incites greater fanaticism than low opposition potency (certainty opposed by one person, as in past work; e.g., Niedbala et al., 2018). In Study 4, we tested whether our findings exhibit discriminant validity in terms of extremism. To do so, we examined whether discordant knowing and other epistemic frameworks differentially lead to political fanaticism (aggression, determined ignorance, and a willingness to join extreme groups in the service of a political viewpoint) versus political extremism (endorsing a non-normative political viewpoint with intensity and firmness). Finally, to examine whether our findings remain consistent across measurement methods, and when using more implicit measures, in Study 5, we used reverse correlation techniques (Dotsch & Todorov, 2012) to test whether inducing discordant knowing leads people to mentally represent themselves as more fanatical.

Finally, in Studies 6 through 9, we tested the generalizability of our findings. In Studies 6 and 7, we tested whether applying a discordant knowing framework to participants' political preference in the 2020 Presidential Election (Joe Biden or Donald Trump) and moral views on abortion (abortion is morally acceptable or wrong) incites fanatical responding regarding these views. Finally, Studies 8 and 9 extended our findings to real-world fanatical and fringe groups. In Study 8, we examined, within a population of anti-vaxxers, whether holding anti-vaccine views in a discordant knowing framework predicts greater anti-vaccine fanaticism and self-reported anti-vaccine behavior (including refusing a COVID-19 vaccine). And, in Study 9, we tested whether active members of a real-world fanatical religious group exhibit greater discordant knowing in terms of their religious views compared to non-fanatical controls.

Notably, in the present studies, we assessed the impact of discordant knowing on fanaticism as compared to three epistemic control structures: discordant believing (a lower degree of certainty opposed by the majority; e.g., "I believe but have doubts about something that most others oppose"), concordant knowing (certainty affirmed by the majority; e.g., "I am certain about something that most others affirm"), and concordant believing (a lower degree of certainty affirmed by the majority; e.g., "I believe but have doubts about something that most others affirm"). While discordant believing lacks the certainty component of discordant knowing (as it includes doubt), it contains the opposition component of discordant knowing.[1] And while concordant knowing lacks the opposition component of discordant knowing, it includes the certainty component of discordant knowing. Finally, concordant believing lacks both the certainty and the opposition components of discordant knowing. Given the 2x2 design implied by the two

[1] Notably, the lack of certainty in discordant believing solely implies having doubts rather than *not* believing something.

components of discordant knowing (certainty and opposition), these three epistemic structures are ideal controls.

Study 1: Establishing the Link Between Discordant Knowing and Fanaticism

In Study 1, we examined whether discordant knowing heightens fanaticism and whether this effect is mediated by epistemic threat. Participants were asked to report a claim or view important to them, and then we manipulated this view to fall under a discordant knowing, discordant believing, or concordant knowing framework (between-participants). To keep the study brief, concordant believing was not included (see Study 3 for this control structure).

Method

Participants. To observe a small-to-moderate effect (d = .35; 85% power) we needed 150 participants per condition (450 total). We aimed to recruit 525 participants (to account for exclusion) and ended up recruiting 522 participants (276 Female; M_{age} = 37.82, SD_{age} = 12.14; Mechanical Turk [MTurk]). The number of intended and recruited participants differed slightly because participants were collected in batches. Fifty-five cases were excluded for failing attention checks and one additional case for completing the study twice. There was no difference in attrition depending on condition, p = .435. The study pre-registration can be found here.[2] Verbatim materials, data files, and analysis files for all studies can be found here. All studies followed APA ethical standards and institutional approval was obtained.

Discordant Knowing. Participants were randomly assigned to one of three conditions: Discordant knowing, discordant believing, or concordant knowing (between-participants).

[2] Due to conceptual confusion when first running the studies, what is referred to as "paradoxical knowing" or "external paradoxical knowing" in the pre-registrations corresponds to "discordant knowing." Additionally, "internal paradoxical knowing" corresponds to "paradoxical knowing," and "believing" or "external believing" corresponds to "discordant believing."

Discordant Knowing. Participants first reported something important that they are certain about (the first component of discordant knowing). Participants read: "There are things in life that we feel like we know are true and are certain about. What is something you know and are certain about in life? What you feel like you know and are certain about can be general or can be more specific. What you feel like you know and are certain about should be important to you. Please write what you know and are certain about in life below."[3] After responding, participants imagined that most people in the world opposed their claim (the second component of discordant knowing). They read: "Imagine that most people in the world claim that what you know and are certain about is unknowable. That is, imagine that everyone is saying that what you know and are certain about, namely that (*what participants said they know piped in here*) is unknowable."[4] Similar vignette paradigms have been validated in past research (e.g., Finch, 1987; Lench et al., 2011; Salerno & Peter-Hagene, 2013).[5]

Discordant Believing. Participants in the discordant believing condition completed the identical prompts, except the first component of discordant knowing was altered to reduce certainty. To do so, the phrase "know and are certain about" was replaced with "believe but have doubts about." Specifically, they read: "There are things in life that we believe but have doubts about. What is something you believe but have doubts about in life…" Thereafter, participants read the opposition component of the discordant knowing prompt.

Concordant Knowing. Participants in the concordant knowing condition first completed the certainty prompt of the discordant knowing condition. They then completed the second

[3] The prompts in Studies 1 and 2 were purposely broad (i.e., "something in life"). Studies 3-5 conceptually replicated these findings when focusing on specific domains (e.g., politics).
[4] Our findings replicated across varying operationalizations of opposition (e.g., the majority judging one's claims as wrong rather than unknowable; see Studies 3 and 7).
[5] Such imagery or vignette paradigms are admittedly artificial. We address this concern in Studies 6 through 9.

discordant knowing prompt, but opposition was changed to affirmation. To do so, the word "unknowable" was changed to "knowable" (e.g., "Imagine that most people in the world claim that what you know and are certain about is knowable…"). See verbatim materials for full materials.

Certainty Manipulation Check. A 3-item manipulation check assessed the first component of discordant knowing—participants' certainty (e.g., "I feel certain that it is true"; Likert-scale: 1 = *Not at all agree* to 7 = *Strongly agree*). "It" in these items referred to what participants had reported being certain about/believing (depending on condition) in response to the manipulation prompts. The same is true for all other uses of "it" in measures across the presented studies.

Opposition Manipulation Check. A 3-item manipulation check assessed the second component of discordant knowing—the degree of majority opposition (e.g., "In the world described above, people are saying it is unknowable"; Likert-scale: 1 = *Not at all agree* to 7 = *Strongly agree*).

Mediator: Threat. Participants were asked how threatened they would feel if they were in the imagined world. Participants were asked: "How would you feel and act if you were in the above described world?" followed by a 3-item measure assessing their epistemic threat ("I would feel threatened," "I would feel like I am being boxed into a corner," "I would feel like people are out to get me"; Likert-scale: 1 = *Not at all agree* to 7 = *Strongly agree*).

Fanaticism. Directly after measuring threat, participants were asked how they would behave if they were in the imagined scenario. We assessed the three indicators of fanaticism outlined in the introduction: Aggression (3-items; e.g., "I would aggress [verbally or physically] towards others"), determined ignorance (3-items; e.g., "I would ignore other people"), and

willingness to join and adhere to groups/movements (3-items; e.g., "I would consider being part of an extreme group or movement in support of it"). All items: 1 = *Not at all agree* to 7 = *Strongly agree*.

Attention Checks. Participants completed three attention checks: (1) whether they were asked to report something they are certain about versus believe, (2) whether people opposed or affirmed their viewpoint, and (3) whether participants responded correctly after being told to ignore a set of response items and type "yes" into a text field. See Verbatim Materials for exact items.

Procedure. Participants completed the first portion of the manipulation, the certainty manipulation check, the second portion of the manipulation, the opposition manipulation check, the threat measure, the fanaticism measures (all items randomized), and the attention checks and demographics (in that order).[6]

Results

Manipulation Checks. General linear models (GLMs) observed an effect of Discordant Knowing (between-participants; 3-levels: discordant knowing, discordant believing, concordant knowing) on the manipulation checks. Planned pairwise comparisons revealed, as intended, that discordant knowing and concordant knowing incited greater certainty, ω_t = .97, than discordant believing, *p*s < .001. And, as intended, discordant knowing and discordant believing led to greater endorsement of the majority opposing one's view, ω_t = .93, than concordant knowing did, *p*s < .001 (Table 1).[7]

[6] One could argue that our certainty manipulation check was flawed because it assessed certainty before the end of the complete manipulation. Studies 5-7 found consistent results when measuring certainty after participants completed the entire manipulation (post-manipulation certainty).

[7] Unexpectedly, participants in the discordant believing condition reported greater perceived opposition than participants in the discordant knowing condition, $p = .003$, although, this effect was small ($d = 0.33$ as compared to $d \approx 2.00$ for the other condition comparisons) and was not observed in the other studies (e.g., Table 3). Importantly, all the reported findings remained when controlling for this difference.

Validation of the Manipulation. Notably, our opposition manipulation check may have been flawed in that it assessed opposition via the term "unknowable" rather than assessing direct opposition. To address this concern, we conducted a validation study (N = 74; MTurk; 23 Female; M_{age} = 37.34, SD_{age} = 10.89) in which we assessed opposition via the term "oppose" rather than "unknowable." Demonstrating that our manipulation induced perceived opposition, as intended, participants in the discordant knowing condition overwhelmingly judged their certainty as opposed by most others, M = 6.67, SD = 0.67 ("…people oppose what I know and am certain about."; 1 = *Not at all agree* to 7 = *Strongly agree*). Additionally, ensuring that our manipulation induced external opposition (discordant knowing) and not internal opposition (paradoxical knowing), participants in the discordant knowing condition did not judge their claimed knowledge as technically unknowable, M = 3.67, SD = 2.60 ("…I think that what I know and am certain about is technically unknowable."; 1 = *Not at all agree* to 7 = *Strongly agree*). Taken together, these results align with the manipulation inducing external opposition and discordant knowing (rather than internal opposition and paradoxical knowing).

Threat and Fanaticism. We averaged across the threat items, M = 2.81, SD = 1.77, ω_t = .91. We also averaged across the three fanaticism indicators, M = 2.83, SD = 1.38, after averaging across each indicators' respective items; aggression, ω_t = .82, determined ignorance, ω_t = .86, joining groups/movements, ω = .95. We did so as the three indicators strongly loaded onto a single factor (Eigenvalue of 1.96; Principal Axis Factor Analysis as extraction) and exhibited high inter-measure reliability, ω_t = .81.

GLMs indicated main effects of Discordant Knowing on threat and fanaticism, ps < .001 (Table 1).[8] As predicted, planned pairwise comparisons revealed that discordant knowing incited

[8] Effects on the individual components of fanaticism were also significant, ps < .040, see Supplements.

greater threat and fanaticism than discordant believing and concordant knowing, *p*s < .001. Additionally, and in line with discordant believing still inciting some degree of discordance (though less so than discordant knowing), discordant believing induced greater threat and fanaticism than concordant knowing, *p*s < .001 (Figure 2 and Table 1).

Our findings may be confounded by participants, depending on condition, reporting views of different valence or content. To account for this possibility, we content-analyzed participants' reported views in terms of valence (negative, neutral, positive; $k = .80$) and content-category (e.g., interpersonal, religion and spirituality; inter-rater reliability: $k = .85$; see Supplements). Though discordant beliefs were more positive and varied in content from discordant knowledge (e.g., more religious content; *p*s < .003; see similar findings in Gollwitzer & Oettingen, 2019), these differences did not account for the observed results (*p*s < .001, after accounting for valence and content).[9] Moreover, the observed effects on threat and fanaticism were not moderated by valence or content-category, *p*s > .091 (see Supplements).[10]

Mediation: Threat. We used the SPSS PROCESS macro (Hayes, 2012) to test the hypothesized mediation. Five thousand bootstrap samples were used to create 95% bias-corrected and accelerated (BCa) confidence intervals. As hypothesized, epistemic threat mediated the effect of discordant knowing on fanaticism. Said another way, participants who framed their viewpoints under a discordant knowing structure experienced heightened epistemic threat, which in turn, predicted greater fanatical responding. Epistemic threat accounted for 57% of the effect of discordant knowing versus discordant believing on fanaticism, and 75% of the effect of discordant knowing versus concordant knowing on fanaticism (Table 2). Notably, these results are limited in

[9] We did not examine whether the content differed between the discordant knowing and concordant knowing conditions because in these conditions the prompt asking for participants' views was identical.
[10] Additionally, in Studies 3 through 5, we pre-assigned the content-category and/or valence of participants' responses and again found consistent results.

that the threat to fanaticism pathway (path B->C) of the mediation was purely correlational (see Study 2 for a causal mediation).

Study 2: Intervening on Epistemic Threat to Prevent Fanaticism

Study 2 aimed to replicate Study 1 and test whether intervening on the identified mediator—epistemic threat—using cognitive reappraisal techniques (see Gross & John, 2003; Lazarus & Alfert, 1964) reduces or eliminates the effect of discordant knowing on fanaticism. A successful intervention on epistemic threat would support a causal mediation pathway (support a causal path between threat and fanaticism, e.g., Spencer et al., 2005), and further, may reveal one potential way to prevent people from adopting fanaticism.

Method

Participants. Applying the power analysis of Study 1, we aimed for 150 participants per condition (four conditions; 600 total). We attempted to recruit 700 participants (to account for exclusion) and ended with 704 participants (441 Female; M_{age} = 39.13, SD_{age} = 12.31; MTurk). One hundred thirty-four cases were excluded for failing attention checks, and two additional cases because participants completed the study twice. There was no difference in attrition depending on condition, p = .267. The study pre-registration can be found here (see footnote 2 regarding different terminology in the pre-registration). The verbatim materials can be found here.

The methods of Study 2 were as in Study 1 but for the addition of a discordant knowing intervention condition (*DK intervention*). This condition was identical to the discordant knowing condition except participants were prompted to cognitively reappraise their epistemic threat after the manipulation and before responding to the threat and fanaticism items: “You may feel threatened and uncomfortable when imagining that most people in the world claim that (*the participant’s reported viewpoint piped here*) is unknowable. Please try to reinterpret this situation

in a more positive light in an effort to reduce your feelings of threat and discomfort…" (adapted from Gross & John, 2003; see Supplements for full materials). An additional attention check assessed whether participants had read the threat reappraisal prompt or not.

Results

Manipulation Checks. As intended, participants in the discordant knowing condition reported greater certainty than those in the discordant believing condition, $p < .001$, but not than those in the concordant knowing and discordant knowing intervention conditions, ps $> .267$. Additionally, as intended, we observed greater perceived opposition in the discordant knowing condition than the concordant knowing condition, $p < .001$, but not than the discordant believing or the discordant knowing intervention conditions, ps $> .119$ (Table 3). See Supplements for all pairwise comparisons.

Threat and Fanaticism. GLMs revealed significant main effects of Discordant Knowing on epistemic threat ($\omega_t = .91$) and fanaticism ($\omega_t = .83$), ps $< .001$ (calculated as in Study 1; Eigenvalue: 2.08; Table 3). Replicating Study 1, planned pairwise comparisons revealed that discordant knowing incited greater threat and fanaticism than discordant believing and concordant knowing, ps $< .001$. Additionally, indicating a successful intervention on the mediator, we observed greater threat and fanaticism in the discordant knowing compared to the discordant knowing intervention condition, ps $< .001$ (Figure 3 and Table 3; see Supplements for all comparisons).

Mediation: Threat. Replicating Study 1, the effect of Discordant Knowing on fanaticism was mediated by epistemic threat. Specifically, 34% of the effect of discordant knowing versus discordant believing, and 64% of the effect of discordant knowing versus concordant knowing on fanaticism was accounted for by threat (Table 4). Importantly, demonstrating a successful

intervention at the level of the mediator, an indirect effect was also observed when comparing the discordant knowing to the discordant knowing intervention condition; 82% of the effect of discordant knowing versus discordant knowing intervention on fanaticism was accounted for by differences in epistemic threat. Collectively, these findings support a causal mediation pathway from discordant knowing to fanaticism via epistemic threat; by successfully intervening on epistemic threat, we provide causal support for the B pathway of the mediation model—epistemic threat causally heightens fanaticism (e.g., Spencer et al., 2005; see Supplements for additional analyses).

Study 3: The Role of Opposition Potency in Discordant Knowing

Building on Studies 1 and 2, Study 3 focused on the opposition component of discordant knowing. First, if majority opposition truly causes fanaticism, then having the majority explicitly versus indirectly oppose one's felt knowledge should still yield fanatical responding. To do so, we induced opposition more explicitly than in the previous studies—via the term "wrong" instead of "unknowable" (e.g., I am certain about something that the majority of others judge as "wrong"). Second, regarding opposition, we examined whether high potency opposition (certainty opposed by the majority of others, as in Studies 1 and 2) versus low potency opposition (certainty opposed by one person, as in previous work; Niedbala et al., 2018) leads to comparatively higher fanaticism. Doing so should help inform whether the potency of opposition (in terms of consensus) plays a meaningful role in producing fanaticism.

Study 3 also considered potential confounds. Specifically, we added concordant believing as a further control condition (to complete the full 2x2 design implied by the two components of discordant knowing), removed the manipulation check and threat measures (which may have primed participants' responding), and controlled for social desirability (to account for demand

bias). Finally, to conceptually replicate Studies 1 and 2, Study 3 focused on political and societal content. Participants reported a policy decision they are certain or believe would result in an improved society (e.g., greater gun control, greater restrictions on abortion). We chose this content because fanaticism often plays a role in political and societal convictions (e.g., Hoffer, 1951).

Method

The methods of Study 3 were identical to those of Study 1 except for several changes. First, to more directly induce opposition in the second component of the manipulation, we changed the term "unknowable" to "wrong" (and the term "knowable" to "right" in the case of affirmation). Second, we manipulated opposition potency by adding a low potency discordant knowing condition. In the second component of the manipulation, the phrase "most people" was changed to "one person."[11]

Third, we added an additional epistemic control condition: Concordant believing. This condition was identical to the discordant believing condition, except participants' beliefs were affirmed instead of opposed ("Imagine you are in a world where most people are saying that what you believe but are uncertain about is right…"; see Verbatim Materials). Fourth, to reduce potential demand effects, we removed the manipulation check and threat measures and, further, added a control measure of social desirability (Sârbescu et al., 2012).

Fifth, instead of something general "in life" (Studies 1 and 2), we had participants report a specific policy decision they are certain or believe would result in an improved society (e.g., greater gun control, greater restrictions on abortion). Sixth, we changed the believing prompt from "believe but have doubts about" to "believe but are uncertain about" because past work on extremism and radicalization has tended to focus on the term "uncertainty" rather than "doubt"

[11] Given that we were specifically interested in examining the impact of potency with regard to discordant knowing, potency was not manipulated in the other conditions.

(e.g., Hogg et al., 2013). Other small changes were made in line with the above changes (see Verbatim Materials).

Participants. Given the larger number of conditions in Study 3 (five conditions; see Table 5) we altered our power-analysis. Specifically, we considered the smallest observed effect-size in Studies 1 and 2 in terms of discordant knowing versus discordant believing predicting fanaticism (d = ~.43). A power-analysis indicated we needed ~99 participants per condition (five conditions; 495 total). We attempted to recruit 550 participants (to account for exclusion) and ended with 552 participants (318 Female; M_{age} = 34.25, SD_{age} = 12.59). Unlike the previous studies, participants were recruited on Prolific. Seventy-one cases were excluded for failing attention checks (see Verbatim Materials).[12] See here for pre-registration. See here for verbatim materials.

Results

Fanaticism. Because manipulation checks and threat were not assessed, we solely examined fanaticism (M = 2.98, SD = 1.26; Eigenvalue of 1.92; ω_t = .79). A GLM revealed a significant effect of Discordant Knowing on fanaticism, ps < .001 (Figure 4; Table 5). Replicating and extending Studies 1 and 2, planned pairwise comparisons revealed greater fanaticism in the high potency discordant knowing condition than in any of the other conditions, including the low potency discordant knowing condition, ps < .010. Additionally, indicating that the potency of opposition and certainty are similarly important in producing fanaticism, low potency discordant knowing predicted similar degrees of fanaticism as discordant believing, p = .258.[13] Finally,

[12] We also included two attention checks at the very start of the study to identify and remove bots. Any participants who were excluded from the study for failing this attention check were not included in the recruitment number.
[13] We replicated these potency effects in a study that manipulated discordant knowing via the term 'unknowable' instead of 'wrong' (as in Studies 1 and 2): High versus low potency discordant knowing: p = .001. In this study, however, and unlike the other studies we conducted (Studies 1 through 6), discordant knowing induced only marginally higher fanaticism than discordant believing, p = .087. This non-significant result may have been due to random noise or, potentially, a degradation in the quality of data collected on MTurk (indeed, as of 2020/2021 many researchers have turned to Prolific for higher quality data).

indicating that concordant believing (the added epistemic control condition) incites very low levels of fanaticism, concordant believing and concordant knowing resulted in similar levels of fanaticism, $p = .918$ (Figure 4; Table 5).

Study 4: Differentiating Fanaticism from Extremism

Study 4 extended the findings of Studies 1-3 by examining an extensively studied construct related to fanaticism—extremism (holding non-normative or controversial views with intensity and firmness; e.g., Hogg, 2004; Klein & Kruglanski, 2013; Kruglanski et al., 2021; Wintrobe, 2006). Extremism differs from fanaticism in that one can exhibit extremism without exhibiting fanaticism. For instance, one can endorse a conspiracy theory with intensity and firmness (extremism) but not engage in aggression supporting it, not determinately ignore counter-information, and not actively seek out groups supporting it (fanaticism; Calhoun, 2004; Hoffer, 1951; Mead, 1977; Milgram, 1977; Perkinson, 2002).

Based on this differentiation, we predicted that while solely discordant knowing heightens fanaticism (see Studies 1-3), both discordant knowing and concordant knowing heighten extremism. More specifically, while only discordant knowing induces threat and thus heightens fanaticism, discordant and concordant knowing both induce certainty, and thus should both heighten extremism and constructs related to extremism (e.g., intellectual arrogance—failing to recognize one's intellectual limitations; Stanley et al., 2019; Zmigrod et al., 2019). Notably, these predictions align with previous work showing that opposition to *as well as* affirmation of people's strong attitudes results in said attitudes becoming more extreme (e.g., more strongly endorsed;

Brehm, 1966; Bassili, 1996; Tormala & Petty, 2002; Petty, 2021; Petty et al., 2002), which is one ingredient of extremism (see Kruglanski et al., 2021).[14]

Method

Study 4 was identical to Study 1, except for three changes. First, participants reported a partisan political view that they are certain about versus believe (e.g., pro-gun control, anti-climate change). Second, on top of fanaticism, we assessed participants' extremism via their self-reported voting preferences and their momentary intellectual arrogance. Third, we assessed an additional mediator—participants' degree of certainty in their political viewpoint after the complete manipulation (post-manipulation certainty)—to examine whether such certainty mediates the proposed link between discordant and concordant knowing and extremism.

Participants. The power analysis of Study 1 was applied. We aimed for 150 participants per condition (450 total). We attempted to recruit 525 participants (to account for exclusion) and ended up recruiting 523 participants (242 Female; M_{age} = 40.16, SD_{age} = 13.10; MTurk). Forty-eight cases were excluded for failing attention checks and two for completing the study twice. See here for pre-registration. See here for Verbatim materials.

Manipulation Checks, Threat, and Fanaticism. The manipulation checks, threat, and fanaticism were assessed as in the previous studies.

Mediator: Post-Manipulation Certainty. Participants' certainty in their partisan political viewpoint after the complete manipulations was assessed via three items (e.g., "If I were in the scenario described above"…"I would feel like my perspective is the only right one"; 1 = *Not at all agree* to 7 = *Strongly agree*). We used different items to assess post-manipulation certainty than

[14] We note that extremism (e.g., Kruglanski et al., 2021) and attitude extremity (e.g., Petty, 2021) differ; while the former entails intense endorsement of deviant views the latter solely entails endorsing a view far from neutrality (i.e., favoring or disfavoring an attitude object).

the certainty manipulation check items (certainty assessed before the opposition/affirmation component of the manipulation) to ensure that participants did not simply stick with their earlier responses. The two mediators, threat and post-manipulation certainty, were assessed in randomized order.

Extremism. Political extremism was assessed via self-reported voting behavior and, more indirectly, via momentary intellectual arrogance. Regarding voting, participants were asked who they would vote for: "There is an election coming up. If you were in the scenario described above, who would you vote for?" (their political view was referred to as "it") 1 = *A Candidate Who is a Moderate Supporter of It*, 2 = *A Candidate Who is a Slightly Extreme Supporter of It*, 3 = *A Candidate Who is a Somewhat Extreme Supporter of It*, 4 = *A Candidate Who is a Very Extreme Supporter of It*. Regarding intellectual arrogance, we included a 6-item adapted version of a validated scale by Leary and colleagues (2017; reverse-coded; e.g., "I would recognize the value in opinions that are different from my own"; 1 = *Not at all agree* to 7 = *Strongly agree*). Fanaticism and extremism were measured after the proposed mediators in randomized order. woof

Results

Manipulation Checks. As in the previous studies, the manipulation was successful (see Table S4).

Threat and Fanaticism. GLMs revealed main effects of Discordant Knowing on threat and fanaticism, *p*s < .001. Conceptually replicating Studies 1-3, discordant knowing incited greater epistemic threat (ω_t = .93) and fanaticism (Eigenvalue of 2.08; ω_t = .81) than discordant believing and concordant knowing, *p*s < .003 (Figure 5; Table 6). As in Studies 1 and 2, the effect of discordant knowing on fanaticism was mediated by epistemic threat (Table 7; see Table S5 for all mediation effects).

Post-Manipulation Certainty and Extremism. GLMs revealed main effects of Discordant Knowing on post-manipulation certainty and extremism, *p*s < .001. Regarding post-manipulation certainty, and as expected, both discordant knowing and concordant knowing incited greater post-manipulation certainty in participants' political views (M = 5.11, SD = 1.54, ω_t = .88) than discordant believing, *p*s < .001 (Figure 5 and Table 6). Concordant knowing also incited higher post-manipulation certainty than discordant knowing, though, this effect was marginal, p = .058.[15] These results indicate, as predicted, that while only discordant knowing produces epistemic threat, both discordant knowing and concordant knowing produce or maintain certainty.

Unlike fanaticism, and as predicted, both concordant and discordant knowing (as compared to discordant believing) heightened extremism in the form of greater self-reported voting for a more extreme candidate, *p*s < .001, and greater momentary intellectual arrogance, ω_t = .95, *p*s < .001 (Figure 5; Table 6). Additionally, as predicted, participants' post-manipulation certainty in their political view predicted greater extreme voting and intellectual arrogance, both within the discordant knowing condition, $r(163) = .24$, $p = .002$, and $r(164) = .30$, $p < .001$, and within the concordant knowing condition, $r(470) = .41$, $p < .001$, and $r(471) = .47$, $p < .001$. Aligning with these results, post-manipulation certainty mediated the links between both discordant and concordant knowing (versus discordant believing) and extremism, as assessed by voting and intellectual arrogance (see Tables 7 and S5). Notably, unlike post-manipulation certainty, epistemic threat neither predicted more extreme voting nor intellectual arrogance in the discordant knowing condition, $r(163) = .08$, $p = .328$, and $r(164) = -.01$, $p = .920$ (the same was also true

[15] Controlling for participants' post-manipulation certainty did not change the effects of discordant knowing on fanaticism, *p*s < .003. Additionally, the effect of discordant knowing on fanaticism was not mediated by post-manipulation certainty, indirect effect: β = -.050, SE = .027, 95% CI: [-0.108, 0.0002].

across conditions, $r(470) = .01$, $p = .911$ and $r(471) = -.08$, $p = .070$).[16] Collectively, these findings indicate that while discordant knowing and concordant knowing both heighten extremism by inducing or maintaining certainty (post-manipulation certainty), only discordant knowing heightens fanaticism because only discordant knowing heightens epistemic threat (see Figure 6 for a visual model).

Study 5: Discordant Knowing and Fanatical Representations of the Self

Study 5 aimed to extend our results beyond self-report measures. Using reverse correlation methods, we tested whether discordant knowing (as compared to discordant believing and concordant knowing) leads participants to mentally represent themselves, in terms of their physical appearance, as more fanatical (as evaluated by independent raters).

Method

The methods of Study 5 were identical to Study 1 except for two changes. First, to extend Studies 3 and 4 to explicitly negative political and societal views, we had participants report something they are certain versus believe is currently a problem in society (e.g., climate change, a lack of gun control, the prevalence of abortion). Second, in addition to assessing fanaticism via the measures of Studies 1-4, we assessed fanaticism via reverse correlation methods. Reverse correlation is a psychophysical paradigm that allows researchers to build visualizations of people's mental representations by having them choose between randomly varying, noisy images (Ahumada & Lovell, 1971; Eckstein & Ahumada, 2002). The technique has been used to visualize people's mental representations of gender, race, ethnicity, stereotypes, and personality traits (see

[16] These specific findings were counter to our pre-registered predictions (see pre-registration here). Unlike predicted, only post-manipulation certainty (and not threat) mediated the links between discordant knowing and heightened extremism and intellectual arrogance.

Brinkman et al., 2017). Here, we used this technique to visualize participants' mental representations of themselves (their own appearance) when induced with discordant knowing.

Participants. Applying the power analysis of Study 1, we aimed for 150 participants per condition (450 total). We attempted to recruit 500 participants (to account for exclusion) and ended up recruiting 502 participants (241 Female; M_{age} = 38.99, SD_{age} = 12.17; MTurk). Seventy-two cases were excluded for failing attention checks, 6 additional cases for taking the study twice, and 2 cases because participants did not respond to all items. Of the remaining 422 participants, 414 completed the online reverse correlation task. See here for pre-registration. See here for verbatim materials.

Fanaticism. We assessed fanaticism as in Studies 1-4 and via a reverse correlation task (the code to run this online-task—which was created for this article—can be found open source here). In the reverse correlation task, participants were presented with 100 trials each depicting two randomly generated pixelated faces. Within each trial, participants were asked to select the face they thought would look more like them if they were in the world described in the manipulation (e.g., discordant knowing condition: "Continue to imagine that you are in the world where most people are saying that what you know and are certain about is unknowable…select the face in each pair that you think looks more like how you would feel/look if you were in this world"). Notably, the pairs of faces were noise-generated and heavily pixelated, meaning that participants' choices are assumed to be largely subconscious. Finally, to create reverse correlation output faces, participants' face selections were (1) averaged across participants in each condition, respectively (using the R reverse correlation package; Dotsch, 2016), resulting in 3 composite faces (1 per condition; see Figure 7), and (2) averaged within each participant to create individual

faces for each participant.[17] Finally, these output faces were evaluated by independent participants in terms of how fanatical they appear. The full task was hosted on our server and can be viewed here (the linked version is the discordant knowing condition).

Results

Manipulation Checks. As in the previous studies, the manipulation was successful (Table S6).

Fanaticism. We replicated the findings of Studies 1-4. Discordant knowing incited greater fanaticism than discordant believing and concordant knowing on the fanaticism items, *ps* < .001 (Figure S1 and Table S6). Again, this effect was mediated by heightened threat (Table S7).

We next examined the reverse correlation results. We first created composite faces for each condition by averaging across all the face-pair choices of participants in each condition, respectively. Given that each participant made 100 face selections and there were ~100 participants per condition, the three composites faces were each comprised of ~10,000 selections (see Figure 7 for composite faces).

Independent raters (MTurk; N = 501) evaluated the fanatical appearance of the three composite faces. Each of the three composite faces were pitted against one another in a side-by-side format (the side of the screen was randomized) and raters were asked to compare the faces based on aggression ("Which person do you think is more aggressive"), determined ignorance ("Which person do you think is more determined to ignore people that stand against his beliefs?)[18], and willingness to join an extreme group or movement ("Which person do you think is more likely to join an extreme group or movement?"). Paired sample *t*-tests found that raters judged the

[17] Analyses based off of the individual participant faces were suggested by a Reviewer in response to concerns about propagation error and were thus not included in our pre-registration.
[18] The term "his" was used because the generated faces looked more male than female (Figure 7).

discordant knowing face as more fanatical (averaged across the three fanatical indicators) than the discordant believing, 1 = *Participant Chose the Discordant Knowing Face*, 0 = *Participant Chose the Discordant Believing Face*, $M = 0.63$, $SD = 0.30$, $t(500) = 9.39$, $p < .001$, $d = 0.42$, and the concordant knowing faces, 1 = *Participant Chose the Discordant Knowing Face*, 0 = *Participant Chose the Concordant Knowing Face*, $M = 0.82$, $SD = 0.24$, $t(500) = 30.27$, $p < .001$, $d = 1.35$.

Recent work, however, indicates that creating composite faces based on condition and having raters evaluate them, as in the analysis above, creates propagation error and Type I error (Cone et al., 2020). To address this, a further group of independent raters (Prolific; $N = 500$) evaluated the single faces generated from individuals' responses (rather than the composite faces). Each rater was presented with 50 trials each including a pair of faces, one face selected from the discordant knowing condition and one face from the discordant believing condition (screen order randomized; faces were randomly selected, without replacement, from the total individual faces in the two conditions).[19] Raters then selected which face was more fanatical in each presented pair (in terms of the three fanatical indicators). In line with the composite-face results, a one-sample *t*-test found that raters selected a greater proportion of discordant knowing than discordant believing faces as fanatical, $t(499) = 5.57$, $p < .001$, $d = .25$.

Study 6: Discordant Knowing and the 2020 Presidential Election

In Studies 6 through 9, we extended the generalizability of our findings to real-world settings. In Study 6, we focused on a political event with broad societal implications—the 2020 U.S. presidential election. Before the 2020 election results, we examined whether being certain that one's preferred candidate (Joe Biden or Donald Trump) would be a better president—juxtaposed with having this view opposed by the majority of other people—heightens fanatical

[19] Concordant knowing faces were not included because propagation error is unlikely to account for the extreme differences between the concordant and discordant knowing composite faces.

responding. Notably, instead of manipulating certainty as in the previous studies, we assessed certainty continuously by asking participants how sure they were that their preferred candidate (Joe Biden or Donald Trump) would be a better president. To manipulate opposition, we then presented participants with a scientific article that had purportedly found that most people judged the participant's view as either unknowable or knowable (e.g., most people claimed that it is unknowable that Biden would be a better president than Trump).

Method

Design. Unlike the previous studies, the design of Study 6 entailed participants' certainty that their preferred candidate would be a better president (assessed pre-manipulation) as a continuous predictor, and a 2x2 between-subjects design: Candidate Preference (pro-Biden vs. pro-Trump) and Opposition (opposition vs. affirmation). Given this design, participants who exhibited high pre-manipulation certainty and were in the opposition condition were classified as high in discordant knowledge. We predicted that our findings would be consistent across Candidate Preference (pro-Biden vs. pro-Trump).

Participants. A power analysis based on identifying a small-to-medium effect size indicated that we needed 350 participants. We attempted to recruit 425 participants (to account for exclusion) and ended with 427 participants (182 Female; M_{age} = 37.69, SD_{age} = 11.80). Sixty-two cases were excluded for failing attention checks and two additional cases because participants completed the study twice. See pre-registration here. See verbatim materials here.

Presidential Preference. Participants first reported who they thought would be the better president, Joe Biden or Donald Trump (order randomized).

Pre-Manipulation Certainty. After assessing presidential preference, three items assessed participants' certainty that their preferred candidate would be a better president (e.g., "I

am certain that Joe Biden [Donald Trump] would be a better president than Donald Trump [Joe Biden]"; 1 = *Not at all agree* to 7 = *Strongly agree*). This measure aligns with previous work measuring certainty (see Petrocelli et al., 2007).

Opposition: Opposition vs. Affirmation. Participants were thereafter told that the purpose of the study was to gather feedback on a media post covering a forthcoming scientific publication. The media post, which was hosted on an external website and integrated into the survey, was identical to an actual media post announcing a scientific publication of a U.S. northeastern university, but the content was altered. Depending on Opposition (opposition vs. affirmation), participants read that researchers had found that most people judged it as unknowable (opposition) versus knowable (affirmation) that the participant's preferred candidate would be a better president.

Threat. Three items assessed how threatened participants felt by the media post (e.g., "The post and the reported findings make me feel threatened"; 1 = *Not at all agree* to 7 = *Strongly agree*).

Fanaticism. Fanaticism was assessed as in the previous studies, except that the items were adapted to relate to the media post. Aggression (e.g., "I feel like telling the author of the post that he is wrong and that he should stop peddling pseudoscience"), determined ignorance (e.g., "Articles like the above post should be ignored as their content is misleading"), and willingness to join fanatical groups/movements (e.g., "The post makes me want to join people who stand against such misleading content, even if these people use heavy-handed tactics"). All scales: 1 = *Not at all agree* to 7 = *Strongly agree*.

Additional Measures. Directly after the manipulation (before assessing threat), we also assessed (randomized-order): Participants' post-manipulation certainty, an opposition

manipulation check, and distractor items about the quality of the media post (e.g., "The post is well written"). Finally, we assessed two attention checks (see Supplements).

Results

Manipulation Check and Post-Manipulation Certainty. GLMs indicated that the manipulation was successful. Participants in the opposition versus affirmation condition reported greater opposition against their view that Biden (Trump) would be a better president, $p < .001$ (Table S8). Additionally, and similarly to Study 4, the manipulation did not appear to alter participants' certainty in their view—the opposition versus affirmation conditions did not differ in terms of their post-manipulation certainty that Biden (Trump) would be a better president, $p = .872$ (Table S8).

Threat and Fanaticism. Discordant knowing was quantified as high certainty that one's preferred candidate would be a better president (pre-manipulation) while also being in the opposition condition. As predicted, GLMs revealed that pre-manipulation certainty (continuous predictor; $M = 5.75$, $SD = 1.63$, $\omega_t = .92$) interacted with Opposition (2-level between-participants factor; opposition versus affirmation) to predict epistemic threat ($\omega_t = .90$), and fanaticism (Eigenvalue of 2.71; $\omega_t = .95$; Figure 8; Table 8), ps $< .001$. Unwrapping this interaction revealed that the effect of Opposition on epistemic threat and fanaticism appeared for participants high ($M = 7.38$; +1 SD)[20] in pre-manipulation certainty, ps $< .001$, but not for those low ($M = 4.12$; -1 SD) in pre-manipulation certainty, ps $> .109$. That is, as hypothesized, participants holding discordant knowledge—those who were highly certain that their preferred candidate would be a better president and also received majority opposition against this view—exhibited the highest degree of epistemic threat and fanaticism.

[20] +1 SD extended above the 7-point scale.

Several robustness checks supported our findings. First, our results did not differ as a function of participants' candidate preference (pro-Biden [59%] vs. pro-Trump [41%]), *p*s > .499 (see Supplements). Second, excluding participants who doubted that the media post was real (in an open-ended comments section; 5% percent) did not change the observed effect-sizes. Third, accounting for the possibility that our results were driven by measuring certainty continuously (which can arguably inflate the possibility of interpreting this measure as a confidence rather than a certainty measure), the effects of Opposition on threat and fanaticism were also observed when recoding the certainty measure in a binary manner (7 out of 7 on the certainty scale coded as 1; all other responses coded as 0), *p*s < .001 (see Supplements for details). Finally, suggesting that discordant knowing requires complete certainty to induce fanaticism, an effect of Opposition on threat and fanaticism was not observed for participants who were not completely certain (any responses < 7; *n* = 180), and the observed two-way interactions between certainty and Opposition predicting threat and fanaticism were no longer significant when only including such participants, *p*s > .457.

Moderated Mediation. Because of the continuous measure of certainty, we conducted a moderated mediation to test the proposed threat mediation pathway. Our model included Opposition as the categorical predictor, pre-manipulation certainty as the moderator (between the A->B path), threat as the mediator, and fanaticism as the dependent variable (Figure 9). The predicted moderated mediation was observed—the Opposition manipulation heightened fanaticism via threat, but only for participants who reported being certain that their preferred candidate would be a better president (Table 9).

Study 7: Discordant Knowing and Abortion

Study 7 aimed to conceptually replicate Study 6 by focusing on an additional source of fanaticism in the United States—abortion (Joffe, 2010). We examined whether placing people's abortion views (pro- versus anti-abortion) in a discordant knowing framework heightens fanatical responding. Given the results of Study 6, we predicted that participants high in discordant knowing—those who are certain of their abortion views and whose views are opposed by a majority of others—would exhibit the highest degree of fanaticism.

The design of Study 7 was identical to Study 6 except for that opposition towards participants' abortion views was manipulated in three different ways. Participants were told that the majority of other people either: (1) explicitly judged the participant's view on abortion as unknowable or knowable (knowability; as in Study 6), (2) associated abortion with being morally acceptable or wrong on implicit measures (implicit), or (3) judged abortion as morally acceptable or wrong when asked to respond honestly and anonymously (anonymity). In this way, we examined whether solely perceiving the outside world as subconsciously or secretly opposing one's certainty also heightens fanaticism.

We also accounted for two potential confounds. First, the results of Study 6 may have been driven by generally heightened trait certainty and mental rigidity rather than certainty specific to the reported political viewpoints. To address this concern, we assessed participants' general trait intellectual arrogance and controlled for this variable. Second, the self-report measures of threat in Studies 1-6 may have actively directed participants' attention towards threat and thus produced our results. To address this concern, we assessed threat using a free-response measure and analyzed participants' responses via natural language processing (NLP) methods in Study 7.

Method

Participants. A power analysis based on pilot data indicated that we needed ~323 participants for 95% power. We attempted to recruit 400 participants on Prolific (to account for exclusion) and ended with 399 participants (158 Female; M_{age} = 32.56, SD_{age} = 10.83). Fifty cases were excluded for failing attention checks and four additional cases because participants completed the survey twice. See pre-registration here. See verbatim materials here.

Abortion View. Participants first reported whether they view abortion as morally acceptable versus wrong (binary choice).

Pre-Manipulation Certainty (Binary). After reporting their view on abortion, participants completed a binary measure assessing their certainty of this view. 1 = "*I am certain that abortion is morally acceptable (wrong)*", 0 = "*I am NOT certain that abortion is morally acceptable (wrong).*"

Pre-Manipulation Certainty (Continuous). Thereafter, similar to Study 6, a three-item continuous measure assessed participants' certainty in their abortion view (e.g., "I am certain that abortion is morally acceptable [wrong]"; 1 = *Not at all agree* to 7 = *Strongly agree*).

Intellectual Arrogance. We assessed trait intellectual arrogance via two reverse-coded trait intellectual humility measures (Leary et al., 2017; Porter & Schumman, 2018).

Opposition: Opposition vs. Affirmation. As in Study 6, participants then read what they were told was a forthcoming media post covering a scientific article. Depending on condition, participants were told that the researchers had found that most people either opposed or affirmed their abortion views.

Opposition Type. Within the opposition and affirmation conditions, the type of opposition versus affirmation (Opposition Type) varied in three ways (between-participants): knowability, implicit, or anonymity.

***Opposition Type: Knowability*.** Depending on condition (opposition or affirmation) and Abortion View (pro- or anti-abortion), the media post reported that most people had judged the participant's view on abortion as unknowable (opposition) or knowable (affirmation). For instance, participants holding pro-abortion views in the opposition condition were told that most people judged the moral acceptability of abortion as technically unknowable.

Opposition Type: Implicit. Depending on condition (opposition or affirmation) and Abortion View (pro- or anti-abortion), the media post reported that most people had judged abortion as morally acceptable or wrong on implicit measures (e.g., "tasks where people have to respond very quickly and can't control their responses"). For instance, participants holding pro-abortion views in the opposition condition were told that most people judged abortion as being morally wrong when these judgments were measured implicitly.

***Opposition Type: Anonymous*.** Depending on condition (opposition or affirmation) and Abortion View (pro- or anti-abortion), the media post reported that most people had judged abortion as morally acceptable or wrong when these judgments were assessed honestly and anonymously. For instance, participants holding pro-abortion views in the opposition condition were told that most people judged abortion as morally wrong when they were asked to honestly and anonymously report their abortion views.

Threat. Epistemic threat was assessed via a free-response item (min 25 words): "How do you feel in response to the article? How does the content of the article make you feel? Please write your thoughts and feelings below." Participants' threat was quantified in two ways. First, we utilized the LIWC software (Pennebaker et al., 2015); the emotional tone summary variable was used as a proxy for threat (0-100), as pre-registered. Lower numbers in emotional tone reveal greater anxiety, sadness, and hostility (reverse-recoded here). Second, two independent raters

(blind to condition) manually coded responses in terms of threat (1 = *No Threat* to 7 = *High Threat*; inter-rater reliability: ICC = .88). See OSF project page for details.

Fanaticism. Fanaticism was assessed as in Study 6.

Other Measures. Participants' post-manipulation certainty, the opposition manipulation check, and attention checks were adapted from Study 6.

Results

Manipulation Check and Post-Manipulation Certainty. The manipulation was successful. Participants in the opposition (versus affirmation) condition perceived greater opposition towards their views on abortion, $p < .001$ (Table S9). Unlike Studies 4 and 6, however, opposition (versus affirmation) induced lower post-manipulation certainty, $p < .001$ (Table S9). Notably, though, this was driven by participants who were not certain in their abortion view (pre-manipulation) becoming even less certain after experiencing opposition; when including only participants who were certain of their abortion view pre-manipulation ($n = 152$), a difference in post-manipulation certainty was no longer observed, $p = .166$.

Threat and Fanaticism. Similar to Study 6, high pre-manipulation certainty combined with experiencing opposition (collapsed across Opposition Type: knowability, implicit, anonymity) was classified as discordant knowledge. Conceptually replicating Study 6, two GLMs found that pre-manipulation certainty (continuous predictor; $M = 5.78$, $SD = 1.56$, $\omega_t = .93$) interacted with Opposition (opposition versus affirmation) to predict epistemic threat (Tone variable from LIWC) and fanaticism (Eigenvalue of 2.63; $\omega_t = .93$), ps $< .003$ (Figure 10; Table 10). As expected, the effect of Opposition on epistemic threat and fanaticism appeared for participants high ($M = 7.34$; +1 SD)[21] in pre-manipulation certainty, ps $< .001$, but not for those

[21] +1 *SD* extended above the 7-point scale.

low (*M* = 4.22; -1 *SD*) in pre-manipulation certainty, *p*s > .101. That is, as hypothesized, participants holding discordant knowledge—those who were certain in their abortion views and experienced opposition against these views—exhibited the highest degree of threat and fanaticism.

Robustness Checks

Several analyses supported the robustness of these findings. First, the observed findings neither differed across pro- and anti-abortion views nor across Opposition Type (knowability, implicit, and anonymity), *p*s > .187. Second, indicating that certainty regarding a specific content (abortion) and not trait certainty heightens fanaticism, adding trait intellectual arrogance to the reported models did not change our results, *p*s < .003, and interactions between intellectual arrogance and Opposition predicting threat and fanaticism were not observed, *p*s > .514. Third, controlling for the observed differences in certainty *after* the manipulation (post-manipulation certainty) did not impact the results, *p*s < .003. Fourth, the observed findings also remained consistent when replacing threat (from LIWC) with manually coded threat, $p < .001$. Finally, and in line with Study 6, our findings remained consistent when replacing the continuous certainty with the binary certainty measure, *p*s < .005 (see Supplements for analysis details).

As in Study 6, we examined whether discordant knowing may require complete certainty to induce fanaticism. Replicating Study 6, Opposition heightened threat and fanaticism for completely certain participants (7-out-of-7 on the continuous certainty measure), *p*s < .001, but not for participants who were not completely certain (all other responses), *p*s > .067.[22] Additionally, the two-way interaction between pre-manipulation certainty and Opposition on fanaticism was no longer observed when excluding completely certain participants, $p = .238$, $\eta p^2 = .008$, though the interaction remained significant for threat, $p = .001$, $\eta p^2 = .055$. At the same

[22] The $p = .068$ was found for predicting threat—the effect was in the *opposite* direction than for completely certain participants (i.e., opposition slightly reduced threat in not completely certain participants).

time, though, when excluding certain participants based off of the binary certainty measure, the two-way interactions between pre-manipulation certainty and Opposition neither predicted threat, $p = .529$, $\eta p^2 = .007$, nor fanaticism, $p = .742$, $\eta p^2 = .002$. Overall, these results further suggest that discordant knowing requires complete certainty to lead to fanaticism.

Moderated-Mediation Model. We conducted the same moderated mediation as in Study 6 to test the proposed threat mediation pathway (see Figure 9 for a visualization). The predicted moderated mediation was observed—opposition heightened fanaticism via threat, but only for participants who were certain in their view on abortion (Table 11).

Study 8: Discordant Knowing and Anti-Vaccine Fanaticism

Study 8 continued to examine real-world generalizability. We tested whether individuals who hold anti-vaccine claims in a discordant knowing framework (e.g., I am certain that vaccines are dangerous, but most people oppose this view) exhibit higher levels of anti-vaccine fanaticism. Notably, such results could have substantial implications. Anti-vaccine attitudes and movements, which have proliferated in the past decade, have caused viral outbreaks of largely eradicated diseases resulting in unnecessary suffering and death (e.g., Measles outbreak of 2019; Givetash, 2019; Offit, 2011). And, during the COVID-19 pandemic, researchers warn that opposition to vaccinations will amplify future outbreaks (Johnson et al., 2020).

In Study 8, we also tested whether altering the opposition component of discordant knowing to nonsocial internal opposition (from external social opposition) impacts our results. That is, someone could feel certain about an anti-vaccine view while also recognizing a lack of evidence supporting this view (nonsocial internal opposition), regardless of whether they perceive others as opposing this view (social opposition). Here we investigated to what extent certainty combined with such internal nonsocial opposition—termed paradoxical knowing in past research

(e.g., "I feel certain that God exists even though this is technically unknowable"; see Gollwitzer & Oettingen, 2019)—also predicts anti-vaccine fanaticism, and further, whether discordant knowing still predicts fanaticism when accounting for such paradoxical knowing (see Introduction for a more detailed description of paradoxical knowing; Figure 1).

Study 8 also extended the past studies in additional ways. Past theorizing suggests that fanaticism is conceptually linked to missionary activity (Rosińska, 2020). Study 8 thus tested whether anti-vaccine discordant knowledge predicts pushing anti-vaccine views onto others. Supporting this possibility, attitude certainty has been shown to predict greater intentions to persuade and force one's attitude onto others (Cheatham & Tormala, 2015; Rios et al., 2014). Finally, to assess actual vaccine behavior, we tested whether discordant knowing predicts refusing a potential COVID-19 vaccine if it were available (at the time of study completion, COVID-19 vaccines were not yet available).

Method

Design. Unlike the previous studies, certainty and opposition were measured continuously. As such, discordant knowing was quantified as holding high certainty and perceiving a high degree of opposition against this certainty. We also assessed believing (holding some doubt) continuously as a control variable.

Participants. A power analysis based on a small-to-medium effect size indicated that we needed 138 participants (95% power). We attempted to recruit 175 participants (to account for exclusion) and ended with 206 participants (82 Female; M_{age} = 38.57, SD_{age} = 11.87).[23] Twenty-three cases were excluded for failing attention checks and 41 additional cases because participants

[23] The final participant number was higher than intended because MTurk hits were reposted for participants who pretended to be anti-vaxxers (participants who restarted the study and changed their answer to a pre-selection anti-vaccine question).

repeated the study or restarted the study and changed their answer on an anti-vaccine pre-selection question. See pre-registration here. See verbatim materials here.

Anti-Vaccine Pre-Selection Question. We only included participants holding anti-vaccine views. To do so, at the start of the study, participants read: "Vaccines for diseases such as measles, mumps, and rubella can be unsafe for healthy children" (1 = *Strongly disagree* to 7 = *Strongly agree*). Participants who responded 4 or above were categorized as endorsing anti-vaccination to some extent and qualified for the study (~31%).

Anti-Vaccine Attitudes. Participants' anti-vaccine attitudes were assessed via a 6-item scale (Horne et al., 2015; e.g., "Children do not need vaccines for diseases that are not common anymore"; 1 = *Strongly disagree* to 7 = *Strongly agree*).

Certainty and Opposition. Participants reported their certainty, perceived social opposition (used to quantify discordant knowing), and nonsocial opposition (used to quantify paradoxical knowing) in response to three different vaccine-contents: (1) vaccines being generally dangerous, (2) a future COVID-19 vaccine being dangerous, and (3) vaccines linking to autism.

Certainty. Participants' anti-vaccine certainty was assessed via three items ("I am certain that some vaccines are dangerous", "I am certain that the coronavirus vaccine will be dangerous", "I am certain that some vaccines are linked to autism"; 1 = *Not at all agree* to 7 = *Strongly agree*).

***Believing*.** Participants' beliefs were assessed via a matched 3-item measure that, unlike the certainty measure, included doubt (e.g., "I think it is probably true that some vaccines are dangerous, but I'm not sure"; 1 = *Not at all agree* to 7 = *Strongly agree*).

Social Opposition. For each of the three vaccine contents, three items (9-items total) assessed social opposition—the degree to which participants perceived others as opposing their

anti-vaccine views (e.g., “Most people refuse to acknowledge that some vaccines are dangerous”; 1 = *Not at all agree* to 7 = *Strongly agree*).

Nonsocial Opposition. Matched items assessed nonsocial opposition—the degree to which participants questioned the evidence underlying anti-vaccine views. These items made no mention of other people denying the participants’ anti-vaccine claims (e.g., “One cannot really know whether vaccines are dangerous or not—it is technically unknowable”; 1 = *Not at all agree* to 7 = *Strongly agree*).

Threat. Epistemic threat was assessed as in the previous studies.

Fanaticism. Fanaticism was assessed as in Studies 1-5, except that the following prompt was displayed before the items: “Please think about how you feel and think about vaccines and how the rest of the world feels and thinks about vaccines. How strongly do you agree with the following statements?” (e.g., “I would aggress [verbally or physically] towards others”, “I want to ignore other people”). Additionally, the joining groups items were adapted to the anti-vaccine content (e.g., “I have considered being part of or am part of an anti-vaccine group or movement”).

Missionary Activism. A 9-item measure assessed participants’ desire to spread their anti-vaccine views (3 items per vaccine content; e.g., “I try to help other people understand that some vaccines are linked to autism,” “I [would] have a responsibility to share the potential dangers of this coronavirus vaccine with others”; 1 = *Not at all agree* to 7 = *Strongly agree*).

Anti-Vax Behavior. Two items assessed participants’ self-reported anti-vaccine behavior (“If there was a Coronavirus vaccine, I would refuse or elect to forgo it,” “If I was given the choice now, I would refuse or elect to forgo vaccines”). Two further items assessed participants’ self-reported anti-vaccine behavior in terms of their children (e.g., “I refuse or elect to forgo vaccines

for my children [if you do not have children, imagine that you do and respond accordingly]"). 1 = *Not at all agree* to 7 = *Strongly agree*.

Results

Data Preparation.

Discordant Knowing. Discordant knowing was quantified as high certainty in one's anti-vaccine views while also perceiving these views as opposed by the majority of others (social opposition). We collapsed across the three vaccine contents (see Supplements). Because both certainty and opposition were measured continuously, discordant knowing was calculated according to an ambivalence formula adapted from Thompson et al. (1995): ([Certainty + Social Opposition]/2 – ABS[Certainty – Social Opposition]; M = 3.87, SD = 1.61; see Gollwitzer & Oettingen, 2019, for similar methods). Higher ambivalence scores represent greater discordant knowing.

Paradoxical Knowing. Paradoxical knowing was calculated in the same manner but with nonsocial rather than social opposition: ([Certainty + Nonsocial Opposition]/2 – ABS[Certainty – Nonsocial Opposition]; M = 3.02, SD = 1.92). In line with discordant and paradoxical knowing conceptually overlapping, the two variables correlated moderately, r = .45.[24]

Discordant and Paradoxical Believing. We also calculated discordant believing (ambivalence score between belief and social opposition) and paradoxical believing (ambivalence score between belief and nonsocial opposition) as control variables (see Supplements for descriptive statistics).

[24] Why would individuals even exhibit paradoxical knowing? That is, why would someone feel certain about something but simultaneously recognize that what they feel certain about is unknowable? For potential antecedents of paradoxical knowing, see Gollwitzer and Oettingen (2019). Also, see Friesen, Campbell, and Kay (2015) who posit that people may claim that their views are technically unknowable in an effort to defend these claims from potential opposition.

Data Analysis. We conducted three sets of multivariate regressions. Each of these three sets included four models; these four models predicted epistemic threat, fanaticism, missionary activism, and self-reported anti-vaccine behavior, respectively. Model Set 1 included discordant knowing as the predictor of interest (Table 12). The control variables included discordant believing, overall vaccine attitudes, age, gender, education level, and political orientation (see Verbatim Materials). Model Set 2 was identical except discordant knowing and discordant believing were replaced with paradoxical knowing and paradoxical believing (Table 13). Finally, Model Set 3 included all four epistemic structures as predictors (Table 14).

In Model Set 1, discordant knowing positively predicted threat, fanaticism, missionary activity, and self-reported anti-vaccine behavior, $.24 < \beta < .55$, ps $< .016$ (Table 12).[25] Notably, discordant knowing was as good a predictor of the assessed anti-vaccine outcomes as a validated anti-vaccine attitude measure, and further, predicted additional variance beyond this measure (Horne et al., 2015). In Model Set 2, paradoxical knowing also positively predicted the four anti-vaccine outcomes, $.30 < \beta < .38$, ps $< .013$ (Table 13). However, importantly, in Model Set 3, which included both discordant and paradoxical knowing as predictors, discordant knowing still predicted the anti-vaccine outcomes, $.23 < \beta < .54$, ps $< .016$ (Figure 11), while paradoxical knowing no longer did so, $-.04 < \beta < .27$, ps $> .056$ (Table 14). These results align with specifically discordant knowing producing fanaticism and suggest that previously observed links between paradoxical knowing and fanatical indicators were likely driven by paradoxical knowing overlapping with discordant knowing (and not the reverse).

Study 9: Discordant Knowing Among Jehovah's Witnesses

[25] We found inconsistent results on a binary item measuring self-reported anti-vaccine behavior (see Supplements).

Study 9 continued to examine generalizability. One major area in which fanaticism plays a role is religion (e.g., Marimaa, 2011). Unlike religious faith (which allows for doubt), religious fanaticism or fundamentalism entails cognitive certainty and a closed belief system (Hill & Williamson, 2005; Kirkpatrick et al., 1991). In Study 9, we thus tested whether active members of a fanatical religious group, Jehovah's Witnesses (Friedson, 2015; Testoni et al., 2019), hold their religious claims in a discordant knowing framework (as compared to non-fanatical religious individuals). Additionally, in line with Study 8, Jehovah's Witnesses should hold their religious views in a discordant knowing rather than paradoxical knowing framework. That is, they should feel certain in their religious claims while perceiving others as opposing these views (social opposition), rather than perceiving a lack of evidence for these views (evidence-based, nonsocial opposition). Indeed, paradoxical knowing, in terms of religious views, is likely to be eliminated by a rigid echo-chamber promoting a fundamentalist religious view on God's proven existence (see Iyengar & Hahn, 2009).

But are Jehovah's Witnesses fanatical? Jehovah's Witnesses are best known for their refusal of blood transfusions (even when the alternative is death; Ott & Cooley, 1977) and their belief that Armageddon is imminent (Penton, 2015). Scholars largely consider the group to be fanatical—a fundamentalist religion that holds a marked determined ignorance towards the outside world; for instance, the group discourages socializing with non-believers (aside for missionary purposes; e.g., Friedson, 2015; Testoni et al., 2019). Additionally, Jehovah's Witnesses are not allowed to read other religious teachings (see their 141 rules here), shun and cast out ex-members, including family members (Gross, 2016), and construe others as being "short-sighted" (Watchtower Online Library, 2014). Finally, Study 9 may be particularly impactful given that Jehovah's Witnesses entail a sizable congregation around the world (over ~8 million; Lawson &

Xydias, 2020); yet, little to no research has been able to survey its active members (e.g., Kristin et al., 2017).

Method

Participants. Active members of a Jehovah's Witnesses congregation were recruited. The experimenter clearly expressed their purpose (to collect scientific data) before recording any data and promoted a transparent dialogue (as was true for all studies presented here, informed consent was collected). We recruited as many Jehovah's Witnesses as possible during the Summer of 2019. Because of the difficulty of collecting this data, we ended with responses from only 14 Jehovah's Witnesses. This sample size is similar to the only other psychological study (to our knowledge) that has examined active Jehovah's Witnesses (Kristin et al., 2017).[26] Despite this limitation, we conducted our planned analyses (see Ginges et al. 2011, for the importance of small group analyses in understanding violent extremism).

We also recruited a control sample of religious, assumably non-fanatic participants (MTurk; n = 20). Because the religious claims of Jehovah's Witnesses are loosely based on Christian principles, only participants who identified as Christian were included (one Buddhist participant was excluded). The control group was composed of eight Catholics, seven Christians (sub-denomination left unspecified), two orthodox Christians, one Baptist, and one born again Christian. A sensitivity power-analyses indicated that with the final sample (N = 33) we could detect an effect-size of: f = .58 (90% power). Though not ideal, we still had enough power to detect a large effect. And indeed, one would imagine epistemic differences between fanatical and non-fanatical religious followers to be quite large. The study was not pre-registered. See verbatim materials here.

[26] Eight additional participants were recruited on a previous version of the study (a pilot version) that did not include all the measures we ultimately assessed. These participants were not included in the analyses.

Certainty. Two items assessed religious certainty. One assessed certainty in God's existence and the other in the Bible being the true word of God (e.g., "I am certain that the Bible is the true word of God"; 1 = *Not at all agree* to 7 = *Strongly agree*).

Believing. Matched control items assessed religious beliefs—endorsing the two religious claims but also holding some doubt (e.g., "It is likely that the Bible is the true word of God, but I'm not sure"). See Boyd (2013), Fowler (1981), and Miller-Perrin and Mancuso (2015) for similar distinctions between religious certainty and religious faith.

Social Opposition. Two items assessed social opposition—participants' perception of the outside world opposing the two religious claims (e.g., "People outside of my community claim that whether the Bible is the true word of God is technically unknowable"; 1 = *Not at all agree* to 7 = *Strongly agree*.).

Nonsocial Opposition. Two matched items assessed nonsocial opposition—participants' perception of the two religious claims as lacking evidence (e.g., "Whether the Bible is the true word of God is technically unknowable"; 1 = *Not at all agree* to 7 = *Strongly agree*.). [27]

Results

Data Preparation

As in Study 8, discordant knowing and paradoxical knowing were quantified via ambivalence scores. Participants who reported both high religious certainty and high religious social opposition were categorized as exhibiting high discordant knowing. Discordant knowing was calculated: ([Certainty + Social Opposition]/2 – ABS[Certainty – Social Opposition]). Paradoxical knowing was calculated in the same manner but with nonsocial instead of social opposition: ([Certainty + Nonsocial Opposition]/2 – ABS[Certainty – Nonsocial Opposition]). As

[27] Additional measures were assessed. As these measures are not central to the arguments made here these measures are discussed in the Supplements.

in Study 8, we also included discordant believing (ambivalence score between religious belief and social opposition) and paradoxical believing (ambivalence score between religious belief and nonsocial opposition) as control variables.

Data Analysis

The Jehovah's Witnesses and MTurk control samples did not differ in terms of age, gender, education, or the number of years they had been part of the religious group/denomination, $p > .139$. In line with Studies 1-8, Jehovah's Witnesses exhibited significantly higher discordant knowing in their religious claims, $M = 4.64$, $SD = 1.66$, than the non-fanatical control group, $M = 2.83$, $SD = 1.95$, $F(1, 31) = 7.91$, $p = .008$, $\eta p^2 = .203$ (Figure 12). This effect remained when controlling for participants' discordant beliefs, $p = .029$. Notably, the observed effect was driven by a combination of Jehovah's Witnesses being absolutely certain in their religious claims (every participant responded 7 on a 7-point scale, $M = 7.00$; $M = 4.87$ in the MTurk controls) while also perceiving people outside of their community as denying their religious "knowledge" ($M = 5.43$ on a 7-point scale; $M = 4.42$ in the MTurk controls).

Jehovah's Witnesses did not, however, exhibit higher paradoxical knowing. Instead, they exhibited significantly *lower* paradoxical knowing, $M = -1.20$, $SD = 1.36$,[28] than the MTurk controls, $M = 2.20$, $SD = 1.97$, $F(1, 31) = 30.71$, $p < .001$, $\eta p^2 = .498$ (Figure 12). This effect remained when adjusting for participants' paradoxical beliefs, $p < .001$. Notably, the observed effect was driven by Jehovah's Witnesses perceiving little nonsocial opposition—followers overwhelmingly rejected the idea that their religious claims were technically unknown ($M = 1.54$ on a 7-point scale; $M = 4.74$ in the MTurk controls). That is, Jehovah's Witnesses overwhelmingly claimed that tangible real-world evidence substantiated their religious claims.

[28] Given the way ambivalence scores are calculated (see Results section), these values can be negative.

General Discussion

Across 9 studies, a simple social-cognitive framework we refer to as discordant knowing—certainty about something one perceives as opposed by the majority of others—predicted greater fanaticism. For instance, experimentally manipulating participants' views to fall under a discordant knowing framework (e.g., "I am certain about X, but most other people think X is unknowable or wrong") heightened behavioral indicators of fanaticism, including aggression, determined ignorance, and wanting to join extreme groups in the service of one's view. Helping to explain this effect, process analyses revealed that adopting discordant knowledge leads people to feel epistemically threatened by the outside world (e.g., "I feel like people are out to get me"), in turn activating fanaticism.

Studies 1 and 2 established the hypothesized phenomenon. Experimentally manipulating participants' viewpoints to fall under discordant knowing (versus discordant believing or concordant knowing) heightened fanaticism in service of these viewpoints. Additionally, this effect was partially mediated by epistemic threat, and further, intervening on said threat, via cognitive reappraisal techniques, prevented the onset of fanaticism.

Studies 3-5 conceptually replicated and extended these findings in terms of the opposition component of discordant knowing (Study 3), differentiating fanaticism from extremism (Study 4), and utilizing more implicit measures (Study 5). In Study 3, both direct and indirect opposition (others judging one's felt knowledge as "wrong" versus "unknowable") induced fanaticism. Additionally, high potency opposition (majority opposition) appeared necessary for discordant knowing to incite fanaticism—low potency opposition (opposition from one person) did not convincingly heighten fanaticism. In Study 4, both discordant knowing and a different epistemic structure, concordant knowing (the majority *affirming* one's certainty), heightened political

extremism (by inducing certainty), but only discordant knowing heightened political fanaticism (by inducing threat). Finally, in Study 5, reverse correlation techniques indicated that holding discordant knowledge leads people to mentally represent themselves (their own physical appearance) as more fanatical, as judged by independent observers.

Studies 6-9 generalized our findings to real-world settings. In Study 6, inducing discordant knowing about an impactful real-world event (the 2020 U.S. Presidential election) incited greater fanatical responding regarding that event. Participants who were certain that their preferred candidate in the 2020 election would be a better president (Biden or Trump) responded more fanatically after reading a scientific article purportedly showing that the majority of others judged this view as unknowable. Study 7 replicated these results in another heavily polarized political topic—abortion. Participants who were certain about their moral views on abortion exhibited greater fanaticism in response to an article indicating that most others opposed their moral view.

Finally, Studies 8 and 9 examined fringe and fanatical groups. In Study 8, anti-vaxxers who hold their anti-vaccine views in a discordant knowing framework exhibited greater epistemic threat, anti-vaccine fanaticism, desire to share anti-vaccine claims with others, and self-reported vaccine hesitancy. Finally, in Study 9, followers of a fanatical religious group (Jehovah's Witnesses) exhibited higher discordant knowing in their religious claims than matched control participants of non-fanatical denominations.

Our findings provide several key contributions. First, by approaching belief-systems from a two-component perspective (certainty juxtaposed with majority opposition) we provide a novel social-cognitive approach to understanding fanaticism and related constructs. Second, we demonstrated that discordant knowing and fanaticism are linked via epistemic threat, and that fanaticism can potentially be attenuated by reducing said threat. Third, we documented that high

potency opposition (majority vs. single-person opposition) is an important ingredient for discordant knowing to heighten fanaticism. Fourth, we differentiated fanaticism from extremism, showing that the former is driven by perceived opposition and threat, while the latter can be induced by mere affirmation. Fifth, we showed that the social-cognitive ingredients comprised in discordant knowing—certainty and perceived opposition—predict individuals' fanaticism in support of specific political and moral views (e.g., views on abortion). Finally, our results indicate that discordant knowing can predict the degree of fanaticism in fringe groups and their members (e.g., anti-vaxxers; Jehovah's Witnesses).

The Two Components of Discordant Knowing

Felt Knowledge and Certainty. The present findings suggest that certainty or felt knowledge is a critical component of discordant knowing and fanaticism. Holding views with doubt or uncertainty did not generate fanaticism, and this was true even when these views were strongly opposed by most others (average effect-size difference between discordant knowing and discordant believing on fanaticism: d = .55). Additionally, in Studies 6 and 7, the effect of discordant knowing on fanaticism was observed only for completely certain participants (7 out of 7 on the certainty scale). And, in Study 9, every member of a fanatical religious group—Jehovah's Witnesses—reported being completely certain in their religious views. Finally, general trait certainty did not account for our findings—in Study 7, opposition towards participants' certainty about a specific view (abortion) still predicted fanaticism when controlling for participants' general trait certainty.

The present research is not the first to suggest, however, that felt knowledge and certainty play a role in constructs related to fanaticism, such as extremism, dogmatism, and self-righteousness (e.g., Berger & Zijderveld, 2009; Dunning, 2011; Hogg et al., 2013; Mitzen &

Schweller, 2011; Schub, 2016; Van Prooijen & Krouwel, 2019). Indeed, in Study 4, concordant knowing (affirmation of one's certainty) as well as discordant knowing, which both entail certainty, heightened political extremism, intellectual arrogance, and post-manipulation certainty.[29] These findings align with past work indicating that affirmation *as well as* opposition to people's strong attitudes can result in said attitudes becoming more extreme (i.e., less neutral; e.g., Brehm, 1966; Bassili, 1996; Tormala & Petty, 2002; Petty, 2021; Petty et al., 2002), which is one ingredient of extremism (see Kruglanski et al., 2021). Importantly, however, the current work extends this literature by showing that while affirmation of one's certainty (e.g., echo-chambers; Iyengar & Hahn, 2009) may be sufficient to heighten extremism and related constructs, such affirmation (all else being equal) does not appear to incite fanaticism. Instead, only opposition to one's certainty appears to heighten fanaticism.

The present work also supports past findings that opposition towards strongly (vs. weakly) held attitudes generates greater anger and negative responding (e.g., Niedbala et al., 2018; Sawicki & Wegener, 2018). Our studies additionally complement this research by examining fanaticism holistically and by demonstrating that discordant knowing incites fanaticism even when opposition is purely imagined (Studies 1-5). Additionally, we document that high potency opposition (opposition from the majority) is a necessary ingredient for discordant knowing to convincingly incite fanaticism; in Study 3, low potency (opposition from one person, as in Niedbala et al., 2018 and Sawicki & Wegener, 2018) produced similar levels of fanaticism as holding a majority opposed viewpoint *without* certainty (discordant believing). Finally, on a broader level, our work contributes to attitude opposition research by identifying one context in which attitude opposition

[29] Relating these findings to a Bayesian framework, it appears that viewpoints held in a discordant knowing framework fail to follow 'normal' Bayesian reasoning—individuals' priors fail to update in response to new information (alternatively, the new information is immediately considered implausible, that is, P(B) is perceived as 0 in P(A/B)).

should predict fanaticism, that is, when attitudes are held with high certainty and are perceived as opposed by the majority.

Opposition. We also found that majority opposition—perceiving most others as opposing one's view—is necessary for discordant knowing to convincingly produce threat and fanaticism. Indeed, in our studies, discordant knowing induced greater levels of threat and fanaticism than affirming views held with certainty (concordant knowing) and affirming views held with uncertainty (concordant believing). And, these results were found across participants merely imagining majority opposition (Studies 1-5), facing real-world majority opposition (Studies 6 and 7), and real-world fanatical individuals' perceived degree of majority opposition (Studies 8 and 9). As such, perceiving majority opposition against one's view appears to be an important ingredient to produce fanaticism.

Potency. In Study 3, high potency opposition (opposition from the majority) most powerfully incited fanaticism (as compared to opposition from a single other). Still, it remains possible that opposition from a single other induces high levels of fanaticism if that person is of great importance (e.g., of high status or importance, such as a mentor or close family member). Future research should examine this possibility.

Types of Opposition. Our findings remained consistent across varying opposition types. For instance, discordant knowing heightened fanaticism when opposition was framed more directly and more indirectly (others judging one's viewpoint as "wrong" versus "unknowable"). Additionally, in Study 7, discordant knowing induced fanaticism when opposition was framed as solely subconscious (most people opposing one's certainty on implicit measures) or solely anonymous (most people opposing one's certainty when asked to respond honestly and anonymously). Collectively, these findings support the conceptual replicability of our results, and

further, indicate that discordant knowing heightens fanaticism even when opposition is far from explicit.

Social Versus Nonsocial. Though discordant knowing heightened fanaticism across varying opposition types, we did observe a major boundary condition—opposition must be social to induce fanaticism. In Study 8, paradoxical knowing—certainty (about anti-vaccine claims) combined with perceived nonsocial internal opposition (i.e., perceiving a lack of evidence supporting one's claim; Gollwitzer & Oettingen, 2019)—failed to predict anti-vaccine fanaticism after controlling for discordant knowing. Additionally, in Study 9, a fanatical religious group, Jehovah's Witnesses, perceived most outsiders as opposing their religious claims (social opposition) and did not perceive their religious claims as lacking evidence or as unknowable (nonsocial opposition). These results suggest that past findings linking paradoxical knowing to indicators of fanaticism (Gollwitzer & Oettingen, 2019) were likely driven by discordant knowing and paradoxical knowing overlapping, rather than by paradoxical knowing per se. Fanaticism, it appears, is a social phenomenon driven by social opposition.

Notably, this conclusion aligns with past work linking social exclusion to constructs related to fanaticism (e.g., increased likelihood of joining radical groups; Hales & Williams, 2018; Renström et al., 2020; Williams et al., 2019). And further, the present studies may inform this work in terms of the moderating effects of certainty. That is, those who feel certain about belonging to a group and are rejected by the majority of the group should be more likely to exhibit fanatical responses compared to those who hedge some doubts about their belongingness.

The Content of Discordant Knowledge. In Studies 1 and 2, discordant knowing heightened fanaticism when participants were free to report viewpoints in any domain (e.g., personal, societal, political). And, in these studies, the domain and valence of the reported

viewpoints neither accounted for nor interacted with discordant knowing to predict fanaticism. Finally, we demonstrated that discordant knowing also incites fanaticism in specific domains relevant to fanaticism (e.g., Hoffer, 1951); it leads to and predicts fanaticism in terms of societal concerns (Studies 3 and 5), political views (Studies 4 and 6), moral convictions (Study 7), health claims (Study 8), and religious views (Study 9).

Still, certain aspects or contents may amplify the degree to which discordant knowing heightens fanaticism. For instance, discordant knowing may more strongly induce fanaticism when individuals' views have moral flavors. Indeed, moral convictions predict antisocial tendencies even after controlling for indices of attitude strength (e.g., importance, extremity, certainty, centrality; Skitka et al., 2005), and the moralization of social and ideological issues has been linked to fanaticism in terms of violent protest (Mooijen et al., 2018). Additionally, past work indicates that moral views are more likely to be held with certainty, that is, as absolute truths or facts (e.g., Heiphetz & Young, 2017; Skitka, 2010). In line with this reasoning, in Study 7, we found that moral convictions held in a discordant knowing framework (e.g., abortion is morally wrong) heightened fanaticism, and indeed, a large number of participants, approximately 80%, reported being certain in their views on abortion. Given these findings, future research should more closely examine the interplay between moral convictions, discordant knowing, and fanaticism.

Theoretical Contributions

Our findings directly contribute to literature on fanaticism by empirically testing whether certainty combined with perceived majority opposition, as indirectly suggested by numerous theorists (e.g., Calhoun, 2004; Milgram, 1977), contributes to fanaticism. We hope doing so renews interest in fanaticism as a psychological construct. Fanaticism, though originally discussed by major scientists in psychology (e.g., Milgram, 1977; Festinger et al., 1956), has received

surprisingly little attention over the last few decades, especially in terms of its social cognitive components (conversely, there has been much research on extremism, e.g., Hogg & Blaylock, 2011; Kruglanski et al., 2021).

Additionally, we provide a framework under which previous theories and findings can be understood. For instance, the present work may help explain why conspiracy theories are sometimes linked to fanatical indicators (e.g., aggression; e.g., Abalakina-Paap et al., 1999; Golec de Zavala & Federico, 2018; van Prooijen & van Vugt, 2018). Simply put, conspiracy theories neatly fall under the framework of discordant knowing—they involve certainty (at least in some cases) about a claim that is generally opposed by the outside world (judged as wrong or unknowable; e.g., Imhoff, Lamberty, & Klein, 2018; Lantian et al., 2017; van Prooijen, 2018). Similarly, the current work may clarify why experiencing identity denial can lead people to join extreme groups and exhibit fanatical attitudes (e.g., Hogg et al., 2007; Sityaeva et al., 2020; McGregor et al., 2001). Akin to conspiracy theories, experiencing identity denial involves having something one feels like one *knows* (i.e., one's own identity) challenged by the outside world. Finally, the present work may explain why attitude opposition predicts constructs related to fanaticism (e.g., anger, aggression; Niedbala et al., 2018); opposition to one's attitudes—when those attitudes are held with certainty and opposition is potent—falls under the framework of discordant knowing, and thus may heighten fanatical indicators.

Importantly, the present work not only provides a framework to understand past work, but also directly informs these areas of research (e.g., conspiracy theories) by elucidating the specific context in which fanaticism is likely to arise. We find fanaticism most likely to emerge when an individual is certain about a specific viewpoint *and* perceives majority opposition against this viewpoint. For instance, holding a conspiracy theory with high certainty and perceiving this theory

as challenged by the majority of others should most heighten fanaticism. Merely believing (not being certain about) about a conspiracy theory or not perceiving opposition to this conspiracy theory, on the other hand, should not induce fanatical responding (e.g., aggression).

Finally, the present work aligns with recent claims in philosophy and international relations that misplaced certainty leads to deleterious outcomes related to fanaticism (e.g., violence, terrorism; e.g., Berger & Zijderveld, 2009; Mitzen & Schweller, 2011; Schub, 2016). For instance, Mitzen and Schweller (2011) argue that misplaced certainty, and not uncertainty, is what motivates international conflicts and wars. In line with this reasoning, we found opposed certainty rather than uncertainty—discordant knowing rather than discordant believing—to induce the highest levels of fanaticism. This finding, at first, appears to oppose previous work in psychology arguing that uncertainty contributes to constructs related to fanaticism (e.g., extremism, terrorism; e.g., Hogg & Blaylock, 2011). Yet, this would be a hasty assumption; indeed, uncertainty may drive individuals to adopt minority perspectives with certainty "to be free of feelings of uncertainty" (Reginster, 2003), and as such, may be an antecedent of discordant knowing, and in turn, fanaticism. The present results, then, may help inform one pathway via which uncertainty can contribute to constructs like fanaticism.

Applied Implications

The current studies provide several applied and methodological contributions.

Fanaticism Interventions. Our results reveal that people can attenuate the effect of discordant knowing on fanaticism by cognitively reappraising the epistemic threat that discordant knowing induces. Specifically, prompting people to reinterpret this threat in a positive light (using cognitive reappraisal) significantly reduced reported threat, and in turn, fanatical tendencies. These results suggest that interventions targeting fanatical responding (e.g., aggression) should focus on

intervening on the epistemic threat that discordant knowing incites rather than trying to change people's viewpoints per se (for a similar argument, see Schuurman & Taylor, 2018). Indeed, as observed in Studies 5-7 and supported by past work (e.g., Brehm, 1966; Bassili, 1996), attempts to change certainty-held views and attitudes often counter productively makes these views more extreme. Intervening directly on epistemic threat, then, should allow individuals to hold extreme views but *not act on them* in terms of fanatical responses (e.g., aggression).

Politics and Society. Studies 3-6 revealed that discordant knowing leads to fanatical responding in domains where fanaticism is especially relevant—political and societal domains (e.g., Hoffer, 1951). For instance, in Studies 6 and 7, participants induced with discordant knowing exhibited greater fanaticism in response to a news article opposing their political views (on the 2020 U.S. Election and on abortion). Given the current polarized state of partisanship in the U.S (e.g., Gollwitzer et al., 2020; Mason, 2018), these findings—which elucidate the cases in which political convictions can result in deleterious fanatical behaviors—may be of particular significance.

Anti-Vaccination. In Study 8, holding anti-vaccine views in a discordant knowing framework predicted participants' desire to share these views with others, as well as their refusal to get vaccinated (including against COVID-19). These links were quite large and were similar in size to the observed link between general anti-vaccine attitudes and these anti-vaccine outcomes ($r \sim .35$). As such, the current results may inform anti-vaccine movements, which have proliferated in the past decade and have caused outbreaks of past and current diseases (e.g., COVID-19; e.g., Givetash, 2019; Offit, 2011).

Limitations and Caveats

Methodological Limitations. First, Studies 1-5 treated certainty in a solely categorical manner. Studies 6-9, however, found consistent results when certainty was assessed in a continuous manner (e.g., Abelson, 1995; Gross et al. 1995; Petty & Krosnick, 1995). That being said, participants may have interpreted the continuous items (e.g., "I am certain that my candidate would be a better president") as assessing confidence rather than certainty. Discounting this possibility, however, past work has assessed certainty using similar items (e.g., Petrocelli et al., 2007), and, in Study 7, our results replicated when assessing certainty via a binary rather than continuous certainty item.

Second, in Studies 8 and 9, discordant knowing was calculated via ambivalence scores between certainty and opposition. This method fails to account for certainty or opposition (on their own) being the main driver in predicting fanaticism. However, in our experimental studies, we demonstrated that the specific combination of certainty and opposition heightens fanaticism to a much greater degree than the two components do independently. Third, our assessment of extremism in Study 4 was limited to the included measures—endorsing an extreme political candidate and intellectual arrogance. Given that extremism has been defined in numerous ways (e.g., Klein & Kruglanski, 2013), future work should examine whether these varying conceptualizations differentially relate to discordant knowing and fanaticism. Fourth, the sample size of Study 9 (Jehovah's Witnesses) was quite small due to the special population of interest. Given the small sample size, these findings should be approached cautiously, and further, null-effects in Study 9 should be approached cautiously given the high risk of Type II error.

Fifth, our results may be driven by demand or response bias. Several findings argue against this possibility, however. Our findings: (1) replicated when applying implicit and spontaneous measures (e.g., reverse correlation and free-response NLP methods), (2) were neither accounted

for nor moderated by participants' general social desirability concerns, (3) remained when removing potential demand-primes (e.g., manipulation checks), and (4) were found for high potency but not low potency opposition (the between-participants study design should have made participants equally subject to demand across the two conditions). For elaborations on these points and additional reasons, see Supplements.

Conceptual Limitations. We consider a number of conceptual limitations. First, we failed to examine whether discordant knowing heightens fanaticism if someone does not perceive opposition, despite such opposition existing. For instance, someone might be certain that climate change is a hoax and perceive others as affirming this view (via social projection), despite that in reality, most people do not endorse this view. In such cases, the catalyzer for fanaticism (the perception of threat) is assumably not generated, and, as such, fanaticism should not develop.

Second, we did not explicitly consider cases in which discordant knowing can have "positive" fanatical consequences. For instance, being certain that a marginalized group is being persecuted (e.g., Rohingya in Burma), while perceiving the majority of others as denying this claim, may lead individuals to fight for the rights of the persecuted group. These responses, in turn, can ultimately result in a more fair and equitable society.

Third, discordant knowing may heighten variables seemingly opposed to fanaticism, such as open-mindedness and collaboration. This is likely only the case, however, when these variables align with the goal of threat-reduction (see Kossowska et al., 2018 and Roets et al., 2015). To explicate, discordant knowing should induce openness towards information or people supporting one's felt-knowledge as this best reduces threat (by re-securing the felt-knowledge). Indeed, in our studies, discordant knowing produced an openness to joining extreme groups that support one's certainty. In sum, though discordant knowing may induce greater openness or approach behavior

(which at first glance may appear counter to fanatical responding), this openness should be in the service of the felt-knowledge, rather than entail general openness.

Fourth, though we considered some alternate predictors of fanaticism (e.g., attitude strength), numerous other variables were not considered. Factors such as peer pressure, need for closure, and significance quest, for instance, should drive individuals towards fanaticism without necessarily activating discordant knowing (e.g., Webber et al., 2018; Kruglanski et al., 2014; Kruglanski et al., 2017). Examining whether and how these pathways to fanaticism relate to discordant knowing should be examined in future work.

Fifth, and in the same vein, discordant knowing does not always produce fanaticism. Our models did not exhibit perfect fit, for instance, some people may be immune to the effects of discordant knowing. Indeed, Study 2 suggested that individuals skilled at regulating threat-responses may be shielded from developing fanaticism. Future research should seek to uncover the individual differences that underlie the unexplained variance in our models.

Sixth, the relationship between discordant knowing and fanaticism may be bi-directional. That is, fanatical tendencies may lead people to start holding their views in a discordant knowing framework. For instance, joining extreme groups may increase individuals' certainty in their viewpoints (e.g., via echo-chambers; Conover et al., 2012; Livingstone et al., 2019). And, perpetrators of aggression are often avoided by others, in turn perhaps leading them to perceive others as more opposed to their viewpoints. Nonetheless, the potential bi-directionality of the observed effects does not negate the causal effects of discordant knowing on fanaticism observed here. Still, future work should prioritize examining these bi-directional links to gain a better understanding of how discordant knowing and fanaticism are linked.

Seventh, we did not identify antecedents or moderators of discordant knowing. Regarding antecedents, chronic paranoia, resentment, and defensive narcissism (variables linked to fanaticism; Cassorla, 2019) may increase people's susceptibility to adopting discordant knowing. Additionally, a need for uniqueness (Imhoff & Erb, 2009) and quest for personal significance (e.g., Kruglanski et al., 2014) may encourage individuals to adopt discordant knowing as such knowledge entails certainty that most people are not privy to. Certain motivational and social factors may also precede discordant knowing. For one, high incentives or the desire for alternate realities may lead individuals to cling onto views (with certainty) in the face of opposition (see Gollwitzer & Oettingen, 2019). For another, experiencing threat or trauma may lead individuals to adopt discordant knowing to deny a changing or altered reality. For instance, recent racial demographic shifts in the United States threatening White people's dominance may motivate White individuals to feel certain that they are being discriminated against despite most others' rejection of this claim (see Parker, 2021).

We also did not consider potential moderators of our results. For instance, the observed effects may be limited to important claims or views. Indeed, majority-opposition against a view one cares little about (even if one is certain about this view) may not induce fanatical responding (e.g., "most people oppose my certainty that chairs have four legs"). In line with motivational models of extremism; e.g., Klein & Kruglanski, 2013), then, the observed effects may require a motivational component (in terms of importance and commitment). Additionally, personality attributes linked to fanaticism (e.g., narcissism; Cassorla, 2019), may moderate our results. Though we did not examine such attributes in detail, in Study 7, participants' trait intellectual arrogance neither accounted for nor moderated our findings. These results suggest that discordant knowing

heightens fanaticism across trait certainty—even the intellectually humble seem to fall prey to the effects of discordant knowing on fanaticism.

Furthermore, it remains unclear whether the observed effects are temporally stable; does inducing discordant knowing heighten fanaticism only temporarily or over a longer time-period? Additionally, future work should identify mediators other than threat that underlie the observed effect. For instance, discordant knowing may lead individual to experience greater uniqueness (due to being privy to 'secret' knowledge) which in turn could promote fanaticism (see Kruglanski et al., 2014). And finally, future work should examine whether group-support and tight group networks ("Network" in Kruglanski et al., 2019) play a role in discordant knowing producing fanaticism.

Conclusion

From interpersonal frictions to acts of terrorism, fanaticism has deleterious consequences for individual lives and for society more generally. By focusing on the potential epistemic and social cognitive structures underlying fanaticism, we were able to shed light on the phenomenology and origins of fanaticism. Across 9 studies, we found that a specific structure, discordant knowing— felt knowledge or certainty about something that one perceives as opposed by most others—heightens fanaticism by making individuals feel threatened by the outside world. Collectively, our results contribute to our understanding of fanaticism, raise possible ways to reduce fanaticism in society, and offer a novel perspective on how to study fanaticism and related constructs.

References

Abelson, R. P. (1986). Beliefs are like possessions. *Journal for the Theory of Social Behaviour, 16*, 223–250.

Abelson, R. P. (1988). Conviction. *American Psychologist*, 43, 267–275

Abelson, R. P. (1995). Attitude extremity. In R. E. Petty & J. A. Krosnick (Eds.*), Attitude strength: Antecedents and consequences* (pp. 25–41). Mahwah, NJ: Erlbaum.

Ahumada Jr, A., & Lovell, J. (1971). Stimulus features in signal detection. *The Journal of the Acoustical Society of America*, *49*, 1751-1756.

Altemeyer, B. (2002). Dogmatic behavior among students: Testing a new measure of dogmatism. *The Journal of Social Psychology, 142*, 713–721.

Abalakina-Paap, M., Stephan, W. G., Craig, T., & Gregory, W. L. (1999). Beliefs in conspiracies. *Political Psychology*, *20*(3), 637-647.

Atir, S., Rosenzweig, E., & Dunning, D. (2015). When knowledge knows no bounds: Self-perceived expertise predicts claims of impossible knowledge. *Psychological Science*, *26*, 1295–1303.

Bagehot, W. (1871). On the emotion of conviction. *The Contemporary Review*, *17*, 32–40.

Bandura, A. (1977). Self-efficacy: toward a unifying theory of behavioral change. *Psychological Review*, *84*(2), 191-215.

Barkun, M. (2015). Conspiracy theories as stigmatized knowledge. *Diogenes*, *62*, 114-120.

Bassili, J. N. (1996). Meta-judgmental versus operative indexes of psychological attributes: The case of measures of attitude strength. *Journal of Personality and Social Psychology*, *71*(4), 637-653.

Bastian, B., Bain, P., Buhrmester, M. D., Gómez, Á., Vázquez, A., Knight, C. G., & Swann, W. B. (2015). Moral Vitalism: Seeing good and evil as real, agentic forces. *Personality and Social Psychology Bulletin*, *41*, 1069–1081.

Berger, P., & Zijderveld, A. (2009). *In praise of doubt*. New York: Harper Collins.

Boyd, G. A. (2013). *Benefit of the doubt: Breaking the idol of certainty*. Grand Rapids, MI: Baker Books.

Brehm, J. W. (1966). *A theory of psychological reactance.* Academic Press.

Brinkman, L., Todorov, A., & Dotsch, R. (2017). Visualising mental representations: A primer on noise-based reverse correlation in social psychology. *European Review of Social Psychology*, *28*, 333-361.

Burton, R.A. (2008). *On being certain: Believing you are right even when you're not.* (1st ed.). St. Martin's Griffin.

Calhoun, L. (2004). An anatomy of fanaticism. *Peace Review*, *16*, 349-356.

Cannon, W. B. (1929). Bodily changes in pain, hunger, fear, and rage (2nd ed.). Appleton-Century.

Cassorla, R. (2019). Fanaticism: Reflections based on phenomena in the analytic field. *The International Journal of Psychoanalysis*, *100*, 1338-1357.

Cone, J., Brown-Iannuzzi, J. L., Lei, R., & Dotsch, R. (2021). Type I error is inflated in the two-phase reverse correlation procedure. *Social Psychological and Personality Science*, *12*, 760-768.

Conover, M. D., Gonçalves, B., Flammini, A., & Menczer, F. (2012). Partisan asymmetries in online political activity. *EPJ Data Science,* 1, 1–19. https://doi.org/10.1140/epjds6

Cupach, W. R., & Spitzberg, B. H. (2004). *The dark side of relationship pursuit: From attraction to obsession and stalking*. Mahwah, NJ: Lawrence Erlbaum Associates.

DeRose, K. 2009. *The Case for Contextualism: Knowledge, Skepticism and Context*. Oxford: Oxford University Press.

Dotsch, R., Wigboldus, D. H., Langner, O., & van Knippenberg, A. (2008). Ethnic out-group faces are biased in the prejudiced mind. *Psychological Science*, *19*(10), 978-980.

Dotsch, R. (2016). rcicr: Reverse correlation image classification toolbox. *R package version 0.3.4.1*.

Dotsch, R., & Todorov, A. (2012). Reverse correlating social face perception. *Social Psychological and Personality Science, 3,* 562-571.

Dunning, D. (2011). The Dunning–Kruger effect: On being ignorant of one's own ignorance. In *Advances in Experimental Social Psychology* (Vol. 44, pp. 247-296). Academic Press.

Eckstein, M. P., & Ahumada, A. J. (2002). Classification images: A tool to analyze visual strategies. *Journal of Vision, 2,* i.

Festinger, L., Riecken, H. W., & Schachter, S. (1956). *When prophecy fails.* Minneapolis: University of Minnesota Press.

Finch, J. (1987). The vignette technique in survey research. *Sociology, 21,* 105-114.

Fowler, J. W. (1981). *Stages of faith: The psychology of human development and the quest for meaning.* New York, NY: HarperCollins.

Friedson, M. L. (2015). Psychotherapy and the Fundamentalist client: The aims and challenges of treating Jehovah's Witnesses. *Journal of Religion and Health, 54,* 693-712.

Gardikiotis, A. (2011). Minority influence. *Social and Personality Psychology Compass, 5,* 679-693.

Ginges, J., Atran, S., Sachdeva, S., & Medin, D. (2011). Psychology out of the laboratory: The challenge of violent extremism. *American Psychologist, 66,* 507-519.

Givetash, L. (2019). Global measles cases surge amid stagnating vaccinations. *NBC News http://www. nbcnews. com/news/amp/ncna1096921 (6 December 2019)*.

Golec de Zavala, A., & Federico, C. M. (2018). Collective narcissism and the growth of conspiracy thinking over the course of the 2016 United States presidential election: A longitudinal analysis. *European Journal of Social Psychology*, *48*(7), 1011-1018.

Gollwitzer, A. (2021, July 8). *Discordant Knowing: A Social Cognitive Structure Underlying Fanaticism.* Retrieved from osf.io/x4km9

Gollwitzer, A., Martel, C., Brady, W. J., Pärnamets, P., Freedman, I. G., Knowles, E. D., & Van Bavel, J. J. (2020). Partisan differences in physical distancing are linked to health outcomes during the COVID-19 pandemic. *Nature Human Behaviour*, *4*(11), 1186-1197.

Gollwitzer, A., & Oettingen, G. (2019). Paradoxical knowing: A shortcut to knowledge and its antisocial correlates. *Social Psychology, 50,* 145-161.

Gollwitzer, P. M., Marquardt, M. K., Scherer, M., & Fujita, K. (2013). Identity-goal threats: Engaging in distinct compensatory efforts. *Social Psychological and Personality Science*, *4*, 555-562.

Gómez, A., Chinchilla, J., Vázquez, A., López-Rodríguez, L., Paredes, B., & Martínez, M. (2020). Recent advances, misconceptions, untested assumptions and future research agenda for identity fusion theory. *Social and Personality Psychology Compass*. 1-15.

Gross, T. (Host). (2019, June 24). *'I couldn't continue on': A former Jehovah's Witness on leaving the faith* [Audio podcast]. National Public Radio.

https://www.npr.org/2019/06/24/735352084/i-couldn-t-continue-on-a-former-jehovah-s-witness-on-leaving-the-faith.

Gross, S. R., Holtz, R., & Miller, N. (1995). Attitude certainty. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 215–246). Mahwah, NJ: Erlbaum.

Gross, J. J., & John, O. P. (2003). Individual differences in two emotion regulation processes: Implications for affect, relationships and well-being. *Journal of Personality and Social Psychology, 85,* 348–362.

Gurr, T. (1970). Why men rebel. Princeton, NJ: Princeton University Press.

Hales, A. H., & Williams, K. D. (2018). Marginalized individuals and extremism: The role of ostracism in openness to extreme groups. *Journal of Social Issues*, *74*(1), 75-92.

Hall, M. P., & Raimi, K. T. (2018). Is belief superiority justified by superior knowledge?. *Journal of Experimental Social Psychology*, *76*, 290-306.

Hayes, A. F. (2012). Process SPSS macro. *[Computer Software and Manual]. Available from http://www. afhayes. com/public/process. pdf*.

Heiphetz, L., & Young, L. L. (2017). Can only one person be right? The development of objectivism and social preferences regarding widely shared and controversial moral beliefs. *Cognition*, *167*, 78-90.

Hill, P. C., & Williamson, W. P. (2005). *The psychology of religious fundamentalism*. New York, NY: Guilford Press.

Hoffer, E. (1951). *The true believer*. Perennial.

Hofstede, G. (1991). *Cultures and organizations: Software of the mind*. McGraw-Hill.

Hogg, M. A. (2004). Uncertainty and extremism: Identification with high entitativity groups under conditions of uncertainty. In V. Yzerbyt, C. M. Judd, & O. Corneille (Eds.), *The psychology of group perception: Perceived variability, entitativity, and essentialism* (pp. 401–418). New York: Psychology Press.

Hogg, M. A., & Blaylock, D. L. (Eds.). (2011). *Extremism and the psychology of uncertainty* (Vol. 3). John Wiley & Sons.

Hogg, M. A., Kruglanski, A., & van den Bos, K. (2013). Uncertainty and the roots of extremism. *Journal of Social Issues*, *69*, 407-418.

Hogg, M. A., Sherman, D. K., Dierselhuis, J., Maitner, A. T., & Moffitt, G. (2007). Uncertainty, entitativity, and group identification. *Journal of Experimental Social Psychology*, *43*(1), 135-142.

Horne, Z., Powell, D., Hummel, J. E., & Holyoak, K. J. (2015). Countering antivaccination attitudes. *Proceedings of the National Academy of Sciences*, *112*, 10321-10324.

Imhoff, R., & Erb, H. P. (2009). What motivates nonconformity? Uniqueness seeking blocks majority influence. *Personality and Social Psychology Bulletin*, *35*(3), 309-320.

Imhoff, R., & Lamberty, P. K. (2017). Too special to be duped: Need for uniqueness motivates conspiracy beliefs. *European journal of social psychology*, *47*, 724-734.

Iyengar, S., & Hahn, K. S. (2009). Red media, blue media: Evidence of ideological selectivity in media use. *Journal of Communication*, *59*(1), 19-39.

Joffe, C. E. (2010). *Dispatches from the abortion wars: the costs of fanaticism to doctors, patients, and the rest of us*. Beacon Press.

Johnson, N. F., Velásquez, N., Restrepo, N. J., Leahy, R., Gabriel, N., El Oud, S., ... & Lupu, Y. (2020). The online competition between pro-and anti-vaccination views. *Nature*, 1-4.

Kirkpatrick, L. A., Hood, R. W., Jr., & Hartz, G. (1991). Fundamentalist religion conceptualized in terms of Rokeach's theory of the open and closed mind: New perspectives on some old ideas. In M. Lynn & D. Moberg (Eds.), *Research in the social scientific study of religion* (Vol. 3, pp. 157–179). Greenwich, CT: JAI Press.

Klein, K. M., Kruglanski, A. W. (2013). Commitment and extremism: A goal systemic analysis. *Journal of Social Issues, 69,* 419-435.

Knowles, E. S., & Linn, J. A. (2004). *Resistance and Persuasion*. Psychology Press.

Kossowska, M., Szumowska, E., Dragon, P., Jaśko, K., & Kruglanski, A. W. (2018). Disparate roads to certainty processing strategy choices under need for closure. *European Review of Social Psychology, 29,* 161-211.

Kristin, R. H., Stålsett, G., Hegstad, H., & Danboltd, L. J. (2017). Emotional forecasting of happiness: Emotion regulation strategies among members of end-time focused Jehovah's Witnesses. *Archive for the Psychology of Religion, 39,* 312-343.

Krosnick, J. A., & Petty, R. E. (1995). Attitude strength: An overview. In R. E. Petty & J. A. Krosnick (Eds.), *Attitude strength. Antecedents and consequences* (pp. 1-24). Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.

Kruglanski, A. W., Bélanger, J. J., & Gunaratna, R. (2019). *The three pillars of radicalization: Needs, narratives, and networks*. Oxford University Press, USA.

Kruglanski, A. W., Gelfand, M. J., Bélanger, J. J., Sheveland, A., Hetiarachchi, M., & Gunaratna, R. (2014). The psychology of radicalization and deradicalization: How significance quest impacts violent extremism. *Political Psychology, 35,* 69-93.

Kruglanski, A. W., Jasko, K., Chernikova, M., Dugas, M., & Webber, D. (2017). To the fringe and back: Violent extremism and the psychology of deviance. *American Psychologist*, *72*, 217-230.

Kruglanski, A. W., & Orehek, E. (2012). The need for certainty as a psychological nexus for individuals and society. M.A. Hogg, D.L. Blaylock, M.A. Hogg, D.L. Blaylock (Eds.), *Extremism and the psychology of uncertainty*, pp. 1-18. Wiley-Blackwell.

Kruglanski, A. W., Szumowska, E., Kopetz, C. H., Vallerand, R. J., & Pierro, A. (2021). On the psychology of extremism: How motivational imbalance breeds intemperance. *Psychological Review*, *128*, 264-189.

Lantian, A., Muller, D., Nurra, C., & Douglas, K. M. (2017). "I know things they don't know!": The role of need for uniqueness in belief in conspiracy theories. *Social Psychology*, *48*, 160–173.

Lantian, A., Muller, D., Nurra, C., Klein, O., Berjot, S., & Pantazi, M. (2018). Stigmatized beliefs: Conspiracy theories, anticipated negative evaluation of the self, and fear of social exclusion. *European Journal of Social Psychology*, *48*, 939-954.

Lawson, R., & Xydias, G. K. (2020). Reassessing the size of mormons, adventists and witnesses: Using census data to test the reliability of membership data and accounting for the disparate patterns found. *Review of Religious Research*, *62*(3), 413-437.

Lazarus, R. S., & Alfert, E. (1964). Short-circuiting of threat by experimentally altering cognitive appraisal. *The Journal of Abnormal and Social Psychology*, *69*, 195-205.

Leary, M. R., Diebels, K. J., Davisson, E. K., Jongman-Sereno, K. P., Isherwood, J. C., Raimi, K. T., ... & Hoyle, R. H. (2017). Cognitive and interpersonal features of intellectual humility. *Personality and Social Psychology Bulletin*, *43*, 793-813.

Lehtsaar, T. (1997). Äärmusliku usugrupi psühholoogia. *Tartu: Tõnu Lehtsaar kirjastus*.

Lench, H., Flores, S., & Bench, S. (2011). Discrete emotions predict changes in cognition, judgment, experience, behavior, and physiology: A meta-analysis of experimental emotion elicitations. *Psychological Bulletin*, 137, 834–855. doi:10.1037/a0024244

Lewin, K. (1946). Behavior and development as a function of the total situation. In L. Carmichael (Ed.), *Manual of child psychology* (pp. 791–844). John Wiley & Sons Inc.

Livingstone, A. G., Fernández Rodríguez, L., & Rothers, A. (2020). "They just don't understand us": The role of felt understanding in intergroup relations. *Journal of Personality and Social Psychology*, *119*(3), 633-656.

Marimaa, K. (2011). The many faces of fanaticism. *KVÜÕA Toimetised, 14,* 29-55.

Mason, L. (2018). *Uncivil agreement: How politics became our identity.* University of Chicago Press.

McGregor, I., Zanna, M. P., Holmes, J. G., & Spencer, S. J. (2001). Compensatory conviction in the face of personal uncertainty: Going to extremes and being oneself. *Journal of Personality and Social Psychology, 80*, 472–488.

Mead, M. (1977). Fanatical thinking: The panhuman disorder. *ETC: A Review of General Semantics*, *34*, 35–38.

Milgram, S. (1977). The social meaning of fanaticism. *ETC: A Review of General Semantics*, 58-61.

Miller-Perrin, C., & Mancuso, E. K. (2015). *Faith from a positive psychology perspective.* Dordrecht, The Netherlands: Springer Science+Business Media.

Mischel, W., Ebbesen, E. B., & Zeiss, A. R. (1973). Selective attention to the self: Situational and dispositional determinants. *Journal of Personality and Social Psychology, 27*(1), 129-142.

Mitzen, J., & Schweller, R. L. (2011). Knowing the unknown unknowns: Misplaced certainty and the onset of war. *Security Studies*, *20*(1), 2-35.

Mooijman, M., Hoover, J., Lin, Y., Ji, H., & Dehghani, M. (2018). Moralization in social networks and the emergence of violence during protests. *Nature Human Behaviour, 2*, 389-396.

Moore, G. E. (1942). "Reply to my Critics" in P. Schilpp ed. *The Philosophy of G. E. Moore*. La Salle, Illinois: Open Court.

Moscovici, S. (1980). Toward a theory of conversion behavior. In L. Berkowitz (Ed.), *Advances in experimental social psychology* (Vol. 13, pp. 209-239). New York: Academic Press.

Nemeth, C. J. (1986). Differential contributions of majority and minority influence. *Psychological Review, 93*, 23-32.

Niedbala, E. M., Hohman, Z. P., & Elleby, J. S. (2018). When I'm right you're wrong: Attitude correctness facilitates anger and approach motivation toward opposing individuals. *Social Influence, 13*(3), 150-162.

Oettingen, G. (2012). Future thought and behaviour change. *European Review of Social Psychology*, *23*(1), 1-63.

Offit, P. A. (2011). *Deadly Choices: How the Anti-Vaccine Movement Threatens Us All.* New York: Basic Books.

Ott, D. A., & Cooley, D. A. (1977). Cardiovascular surgery in Jehovah's Witnesses: report of 542 operations without blood transfusion. *Jama, 238*, 1256-1258.

Paredes, B., Briñol, P., Petty, R. E., & Gómez, Á. (2020). Increasing the predictive validity of identity fusion in leading to sacrifice by considering the extremity of the situation. *European Journal of Social Psychology*, 1-12.

Parker, C. S. (2021). Status threat: Moving the right further to the right?. *Daedalus, 150*, 56-75.

Pennebaker, J.W., Booth, R.J., Boyd, R.L., & Francis, M.E. (2015). Linguistic inquiry and word count: LIWC2015.

Penton, M. J. (2015). *Apocalypse delayed: the story of Jehovah's Witnesses*. University of Toronto Press.

Perkinson, H. J. (2002). Fanaticism: flight from fallibility. *ETC: A Review of General Semantics*, *59*, 170-174.

Petrocelli, J. V., Tormala, Z. L., & Rucker, D. D. (2007). Unpacking attitude certainty: Attitude clarity and attitude correctness. *Journal of Personality and Social Psychology*, *92*, 30-41.

Petty, R. E. (2021, February). *The role of threat in predicting extreme attitudes and behavior* [Conference Presentation]. SPSP 2021 The Psychology of Extremism Virtual Preconference.

Petty, R. E., Briñol, P., & Tormala, Z. L. (2002). Thought confidence as a determinant of persuasion: The self-validation hypothesis. *Journal of Personality and Social Psychology, 82*(5), 722–741.

Petty, R. E., & Krosnick, J. A. (Eds.). (1995). Attitude strength: Antecedents and consequences. Hillsdale, NJ: Erlbaum.

Porter, T., & Schumann, K. (2018). Intellectual humility and openness to the opposing view. *Self and Identity, 17*, 139-162.

Renström, E. A., Bäck, H., & Knapton, H. M. (2020). Exploring a pathway to radicalization: The effects of social exclusion and rejection sensitivity. *Group Processes & Intergroup Relations*, *23*(8), 1204-1229.

Reginster, B (2003). What is a free spirit? Nietzsche's critique of fanaticism, *Archiv fur Geschichte der Philosophie*, *85*, 51–85.

Rios, K., DeMarree, K. G., & Statzer, J. (2014). Attitude certainty and conflict style: Divergent effects of correctness and clarity. *Personality and Social Psychology Bulletin, 40*, 819–830.

Roets, A., Kruglanski, A. W., Kossowska, M., Pierro, A., & Hong, Y. Y. (2015). The motivated gatekeeper of our minds: New directions in need for closure theory and research. In *Advances in experimental social psychology* (Vol. 52, pp. 221-283). Academic Press.

Rosińska, Z. (2020). The Phenomenon of Fanaticism. *Dialogue and Universalism*, 159-164.

Rousis, G. J., Richard, F. D., & Wang, D. Y. D. (2020). The truth is out there: The prevalence of conspiracy theory use by radical violent extremist organizations. *Terrorism and Political Violence*, 1-19.

Salerno, J., & Peter-Hagene, L. (2013). The interactive effect of anger and disgust on moral outrage and judgments. *Psychological Science*, 24, 2069–2078.

Sârbescu, P., Costea, I., & Rusu, S. (2012). Psychometric properties of the Marlowe-Crowne social desirability scale in a Romanian sample. *Procedia-Social and Behavioral Sciences*, *33*, 707-711.

Sawicki, V., & Wegener, D. T. (2018). Metacognitive reflection as a moderator of attitude strength versus attitude bolstering: Implications for attitude similarity and attraction. *Personality and Social Psychology Bulletin*, *44*(5), 638-652.

Schub, R. J. (2016). *Certainty and war.* Doctoral Dissertation, Harvard University, Graduate School of Arts & Sciences.

Schunk, D. H. (2012). Social cognitive theory. In K. R. Harris, S. Graham, T. Urdan, C. B. McCormick, G. M. Sinatra, & J. Sweller (Eds.), *APA educational psychology handbook, Vol. 1. Theories, constructs, and critical issues* (pp. 101–123).

Selengut, C. (2017). *Sacred fury: Understanding religious violence*. Rowman & Littlefield.

Shaw, M., Quezada, S. A., & Zárate, M. A. (2011). Violence with a conscience: Religiosity and moral certainty as predictors of support for violent warfare. *Psychology of Violence*, *1*, 275-186.

Schuurman, B., & Taylor, M. (2018). Reconsidering radicalization: Fanaticism and the link between ideas and violence. *Perspectives on Terrorism*, *12*, 3-22.

Sityaeva S.M., Yaremtchuk S.V., Tarasenko V.V. (2020) Relationship between identity fusion and extremist attitudes of young. In: Solovev D., Savaley V., Bekker A., Petukhov V. (eds) *Proceeding of the International Science and Technology Conference "FarEastCon 2019"*. Smart Innovation, Systems and Technologies, vol 172. Springer, Singapore. https://doi.org/10.1007/978-981-15-2244-4_64

Skitka, L. J. (2010). The psychology of moral conviction. *Social and Personality Psychology Compass*, *4*, 267-281.

Skitka, L. J., Bauman, C. W., & Sargis, E. G. (2005). Moral conviction: Another contributor to attitude strength or something more?. *Journal of Personality and Social Psychology, 88*(6), 895-917.

Spencer, S. J., Zanna, M. P., & Fong, G. T. (2005). Establishing a causal chain: why experiments are often more effective than mediational analyses in examining psychological processes. *Journal of Personality and Social Psychology, 89*, 845-851.

Stankov, L., Higgins, D., Saucier, G., & Knežević, G. (2010). Contemporary militant extremism: A linguistic approach to scale development. *Psychological Assessment, 22*, 246-258.

Stanley, M. L., Sinclair, A. H., & Seli, P. (2020). Intellectual humility and perceptions of political opponents. *Journal of Personality, 88*(6), 1196-1216.

Taylor, S. E., Klein, L. C., Lewis, B. P., Gruenewald, T. L., Gurung, R. A., & Updegraff, J. A. (2000). Biobehavioral responses to stress in females: Tend-and-befriend, not fight-or-flight. *Psychological Review*, *107*, 411–428.

Testoni, I., Bingaman, K., Gengarelli, G., Capriati, M., De Vincenzo, C., Toniolo, A., ... & Zamperini, A. (2019). Self-Appropriation between social mourning and individuation: A qualitative study on psychosocial transition among Jehovah's witnesses. *Pastoral Psychology*, *68*, 687-703.

Thompson, M. M., Zanna, M. P., & Griffin, D. W. (1995). Let's not be indifferent about (attitudinal) ambivalence. In R. E. Petty& J. A. Krosnick (Eds.), *Attitude strength: Antecedents and consequences* (pp. 361-386). Hillsdale, NJ: Erlbaum.

Tormala, Z. L., & Petty, R. E. (2002). What doesn't kill me makes me stronger: The effects of resisting persuasion on attitude certainty. *Journal of Personality and Social Psychology*, *83*(6), 1298-1313.

Tversky, A., & Kahneman, D. (1991). Loss aversion in riskless choice: A reference-dependent model. *The Quarterly Journal of Economics*, *106*, 1039–1061.

Van Prooijen, J. W. (2018). *The psychology of conspiracy theories*. Routledge.

Van Prooijen, J. W., & Krouwel, A. P. (2019). Psychological features of extreme political ideologies. *Current Directions in Psychological Science*, *28*, 159-163.

van Prooijen, J. W., & Van Vugt, M. (2018). Conspiracy theories: Evolved functions and psychological mechanisms. *Perspectives on Psychological Science*, *13*(6), 770-788.

Watchtower Online Library. (2014). Do you appreciate what you have received?: The watchtower announcing Jehovah's kingdom—2014. Retrieved from https://wol.jw.org/en/wol/d/r1/lp-e/2014929#h=28.

Webber, D., Babush, M., Schori-Eyal, N., Vazeou-Nieuwenhuis, A., Hettiarachchi, M., Bélanger, J. J., ... & Gelfand, M. J. (2018). The road to extremism: Field and experimental evidence that significance loss-induced need for closure fosters radicalization. *Journal of personality and social psychology*, *114*, 270.

Webster, D. M., & Kruglanski, A. W. (1994). Individual differences in need for cognitive closure. *Journal of Personality and Social Psychology*, *67*(6), 1049-1062.

Whitson, J. A., & Galinsky, A. D. (2008). Lacking control increases illusory pattern perception. *Science*, *322*, 115-117.

Williams, K. D., Hales, A. H., & Michels, C. (2019). Social ostracism as a factor motivating interest in extreme groups. In *Current Directions in Ostracism, Social Exclusion, and Rejection Research* (pp. 18-31). Routledge.

Wintrobe, R. (2006). *Rational extremism: The political economy of radicalism*. Cambridge University Press.

Wittgenstein, L. (1953). *Philosophical investigations* (G. E. M. Anscombe, Trans.). New York: Macmillan.

Zmigrod, L., Rentfrow, P. J., & Robbins, T. W. (2019, August 5). The partisan mind: Is extreme political partisanship related to cognitive inflexibility? *Journal of Experimental Psychology: General.* Advance online publication. http://dx.doi.org/10.1037/xge0000661

Figures and Tables

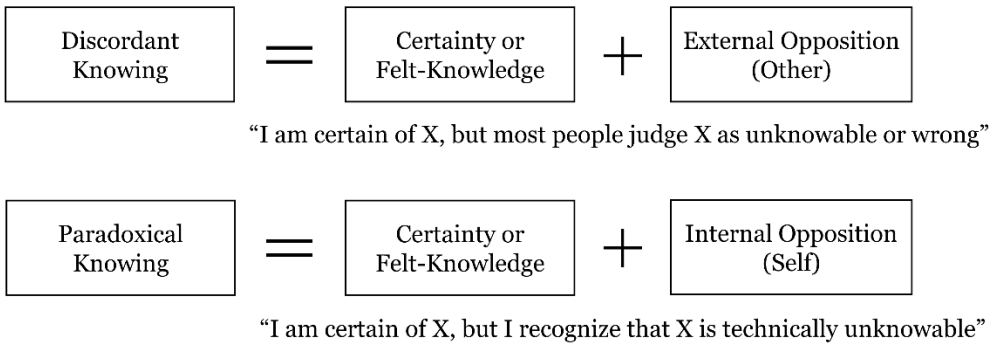

Figure 1. Conceptual diagram differentiating discordant knowing from paradoxical knowing.

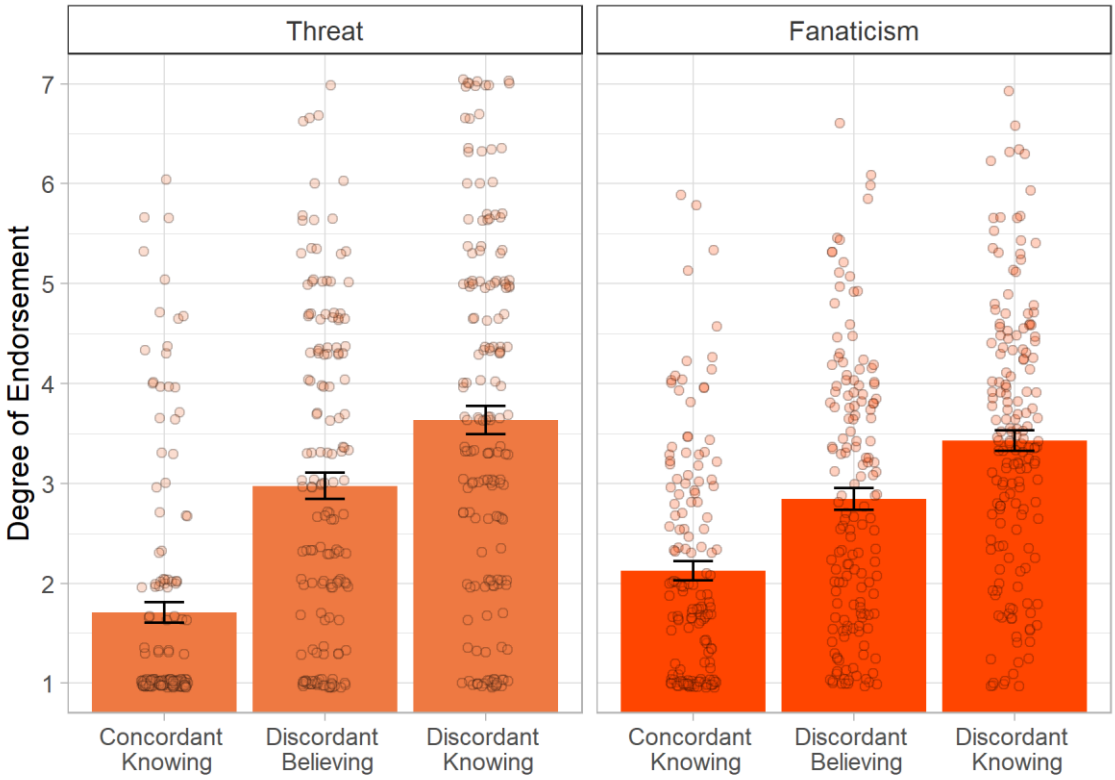

Figure 2. Effects of Discordant Knowing on threat and fanaticism in Study 1. Error bars: +-1 *SE*.

Table 1

Study 1. Effects of Discordant Knowing.

	Discordant Knowing (*M* and *SD)*	Discordant Believing (*M* and *SD)*	Concordant Knowing (*M* and *SD)*	Significance Test		
Study 1	*n* = 167	*n* = 152	*n* = 146	Test-Statistic	*p* value	Effect-Size
Manipulation Checks						
Certainty	6.80, 0.56	4.69, 1.31	6.73, 0.60	$F(2, 462) = 280.93$	$p < .001$	$\eta^2 = .549$
	◊	◊		$t(317) = 21.14$	$p < .001$	$d = 2.09$
	◊		◊	$t(311) = 0.69$	$p = .490$	$d = 0.12$
		◊	◊	$t(296) = 19.84$	$p < .001$	$d = 2.00$
Opposition	5.27, 2.03	5.83, 1.32	2.25, 1.62	$F(2, 462) = 193.20$	$p < .001$	$\eta^2 = .455$
	◊	◊		$t(317) = 2.96$	$p = .003$	$d = 0.33$
	◊		◊	$t(311) = 15.76$	$p < .001$	$d = 1.64$
		◊	◊	$t(296) = 18.30$	$p < .001$	$d = 2.42$
Mechanism						
Epistemic Threat	3.63, 1.81	2.98, 1.62	1.71, 1.23	$F(2, 462) = 58.54$	$p < .001$	$\eta^2 = .202$
	◊	◊		$t(317) = 3.69$	$p < .001$	$d = 0.38$
	◊		◊	$t(311) = 10.69$	$p < .001$	$d = 1.24$
		◊	◊	$t(296) = 6.89$	$p < .001$	$d = 0.88$
Dependent Variable						
Fanaticism	3.43, 1.33	2.85, 1.35	2.13, 1.13	$F(2, 462) = 40.78$	$p < .001$	$\eta^2 = .150$
	◊	◊		$t(317) = 4.08$	$p < .001$	$d = 0.43$
	◊		◊	$t(311) = 9.06$	$p < .001$	$d = 1.05$
		◊	◊	$t(296) = 4.88$	$p < .001$	$d = 0.58$

Note. ◊ signifies inclusion in a pairwise comparison. Cohen's *d* calculated in terms of the individual pairwise comparisons (i.e., each condition comparison treated individually; not using mean-square-error across conditions).

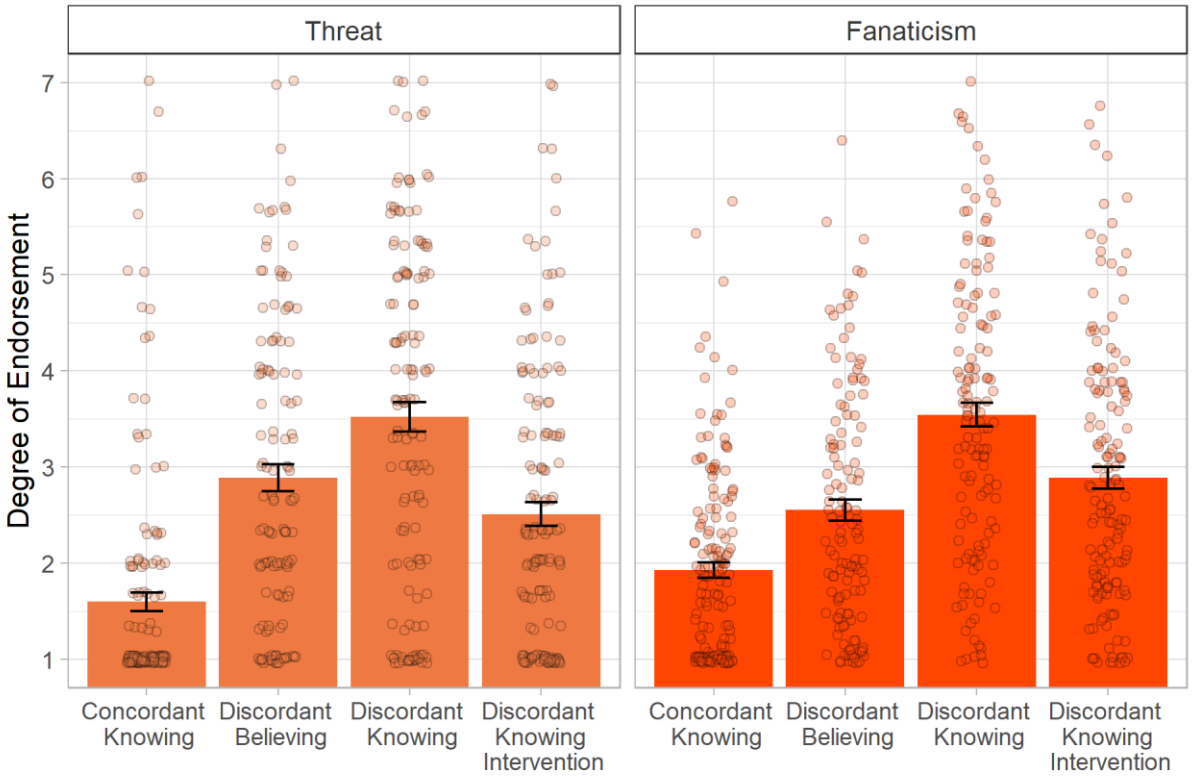

Figure 3. Effects of Discordant Knowing on threat and fanaticism in Study 2. Error bars: +-1 *SE*.

Table 2

Study 1: The Effect of Discordant Knowing on Fanaticism was Mediated by Epistemic Threat.

Variables:	Independent Variable	Mediator	Dependent Variable
Study 1 (N = 465)	Discordant Knowing	Epistemic Threat	Fanaticism
	Relative Effect: Discordant Knowing (coded: 1) Versus Discordant Believing (coded: 0)		
Total Effect	β = .422, SE = .104, t = 4.08, p < .001, 95% CI: [0.219, 0.626]		
Direct Effect	β = .181, SE = .082, t = 2.22, p = .027, 95% CI: [0.021, 0.341]		
Indirect Effect	β = .241, SE = .072, 95% CI: [0.102, 0.388]		
	Relative Effect: Discordant Knowing (coded: 1) Versus Concordant Knowing (coded: 0)		
Total Effect	β = .945, SE = .105, t = 9.03, p < .001, 95% CI: [0.739, 1.151]		
Direct Effect	β = .238, SE = .091, t = 2.62, p = .009, 95% CI: [0.060, 0.417]		
Indirect Effect	β = .707, SE = .074, 95% CI: [0.569, 0.854]		

Note. Standardized Betas (β) are calculated in terms of z-score normalization. They are not the unstandardized slope.

Table 3

Study 2. Effects of Discordant Knowing.

	Discordant Knowing (*M* and *SD)*	Discordant Believing (*M* and *SD)*	Concordant Knowing (*M* and *SD)*	Discordant Knowing Intervention (*M* and *SD)*	Significance Test		
Study 2	*n* = 138	*n* = 126	*n* = 158	*n* = 146	Test-Statistic	*p* value	Effect-Size
Manipulation Checks							
Certainty	6.86, 0.60	4.59, 1.42	6.75, 0.59	6.79, 0.53	$F(3, 564) = 227.01$	$p < .001$	$\eta^2 = .547$
	◊	◊			$t(262) = 22.02$	$p < .001$	$d = 2.08$
	◊		◊		$t(294) = 1.10$	$p = .268$	$d = 0.18$
	◊			◊	$t(282) = 0.71$	$p = .484$	$d = 0.12$
Opposition	5.61, 1.75	5.67, 1.51	1.86, 1.35	5.30, 2.02	$F(3, 564) = 182.75$	$p < .001$	$\eta^2 = .493$
	◊	◊			$t(262) = 0.31$	$p = .759$	$d = 0.04$
	◊		◊		$t(294) = 19.22$	$p < .001$	$d = 2.40$
	◊			◊	$t(282) = 1.56$	$p = .120$	$d = 0.16$
Mechanism							
Epistemic Threat	3.52, 1.81	2.89, 1.58	1.60, 1.22	2.51, 1.49	$F(3, 564) = 41.14$	$p < .001$	$\eta^2 = .180$
	◊	◊			$t(262) = 3.38$	$p = .001$	$d = 0.37$
	◊		◊		$t(294) = 10.81$	$p < .001$	$d = 1.24$
	◊			◊	$t(282) = 5.61$	$p < .001$	$d = 0.61$
Dependent Variable							
Fanaticism	3.54, 1.48	2.55, 1.24	1.93, 1.00	2.89, 1.36	$F(3, 564) = 41.09$	$p < .001$	$\eta^2 = .179$
	◊	◊			$t(262) = 6.32$	$p < .001$	$d = 0.73$
	◊		◊		$t(294) = 10.85$	$p < .001$	$d = 1.27$
	◊			◊	$t(282) = 4.35$	$p < .001$	$d = 0.46$

Note. ◊ signifies inclusion in a pairwise comparison. Cohen's *d* calculated in terms of the individual pairwise comparisons (i.e., each condition comparison treated individually; not using mean-square-error across conditions).

Table 4

Mediation Effects in Study 2: The Effect of Discordant Knowing on Fanaticism was Mediated by Epistemic Threat

Variables:	**Independent Variable**	**Mediator**	**Dependent Variable**
Study 2 (*N* = 568)	Discordant Knowing	Epistemic Threat	Fanaticism

	Relative Effect: Discordant Knowing (coded: 1) Versus Discordant Believing (coded: 0)
Total Effect	β = .708, *SE* = .112, *t* = 6.32, *p* < .001, 95% CI: [0.488, 0.927]
Direct Effect	β = .466, *SE* = .087, *t* = 5.36, *p* < .001, 95% CI: [0.295, 0.637]
Indirect Effect	β = .242, *SE* = .082, 95% CI: [0.086, 0.405]
	Relative Effect: Discordant Knowing (coded: 1) Versus Concordant Knowing (coded: 0)
Total Effect	β = 1.152, *SE* = .106, *t* = 10.88, *p* < .001, 95% CI: [0.944, 1.360]
Direct Effect	β = .420, *SE* = .089, *t* = 4.70, *p* < .001, 95% CI: [0.244, 0.596]
Indirect Effect	β = .732, *SE* = .086, 95% CI: [0.571, 0.912]
	Relative Effect: Discordant Knowing (coded: 1) Versus Discordant Knowing Intervention (coded: 0)
Total Effect	β = .468, *SE* = .108, *t* = 4.34, *p* < .001, 95% CI: [0.256, 0.680]
Direct Effect	β = .082, *SE* = .085, *t* = 0.96, *p* = .337, 95% CI: [-0.086, 0.250]
Indirect Effect	β = .386, *SE* = .081, 95% CI: [0.228, 0.550]

Note. Standardized Betas (β) are calculated in terms of *z*-score normalization. They are not the slope.

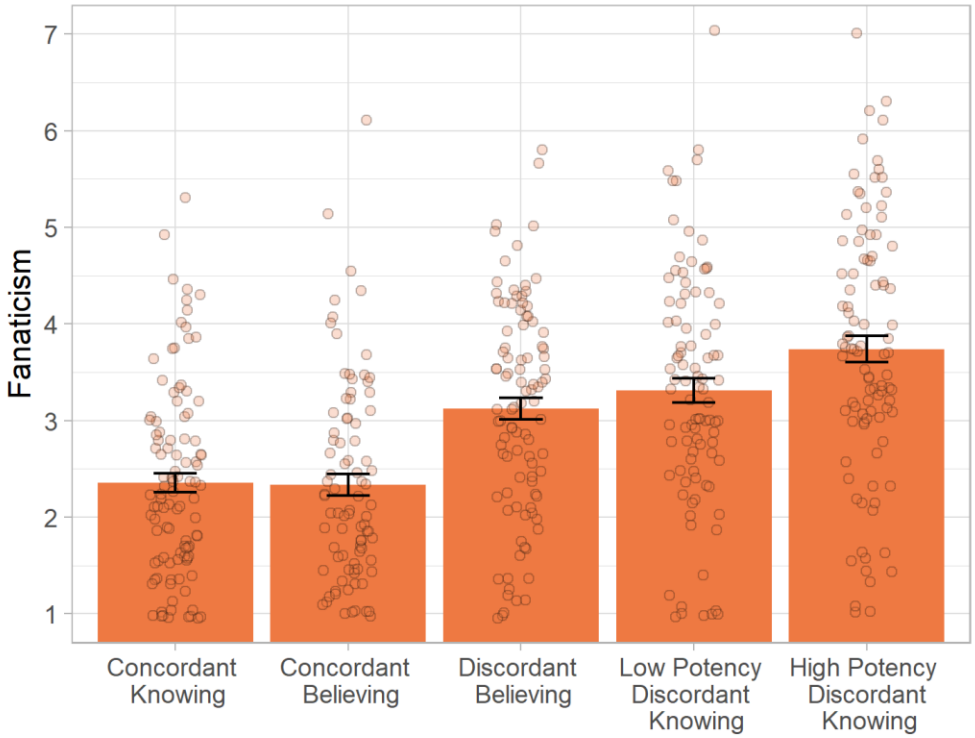

Figure 4. Effects of Discordant Knowing on fanaticism in Study 3. Error bars: +-1 *SE*.

Table 5

Study 3. Effects of Discordant Knowing on Fanaticism

	High Potency Discordant Knowing (*M* and *SD*)	Low Potency Discordant Knowing (*M* and *SD*)	Discordant Believing (*M* and *SD*)	Concordant Believing (*M* and *SD*)	Concordant Knowing (*M* and *SD*)	Significance Test		
Study 3	*n* = 97	*n* = 94	*n* = 98	*n* = 87	*n* = 105	Test-Statistic	*p* value	Effect-Size
Dependent Variable								
Fanaticism	3.74, 1.31	3.31, 1.22	3.12, 1.11	2.34, 1.04	2.36, 0.99	$F(4, 476) = 27.98$	$p < .001$	$\eta^2 = .190$
	◊	◊				$t(189) = 2.61$	$p = .009$	$d = 0.34$
	◊		◊			$t(193) = 3.78$	$p < .001$	$d = 0.51$
	◊			◊		$t(182) = 8.35$	$p < .001$	$d = 1.18$
	◊				◊	$t(200) = 8.66$	$p < .001$	$d = 1.19$
		◊	◊			$t(190) = 1.13$	$p = .258$	$d = 0.16$
				◊	◊	$t(190) = 0.10$	$p = .918$	$d = 0.02$

Note. ◊ signifies inclusion in a pairwise comparison. Cohen's *d* calculated in terms of the individual pairwise comparisons (i.e., each condition comparison treated individually; not using mean-square-error across conditions).

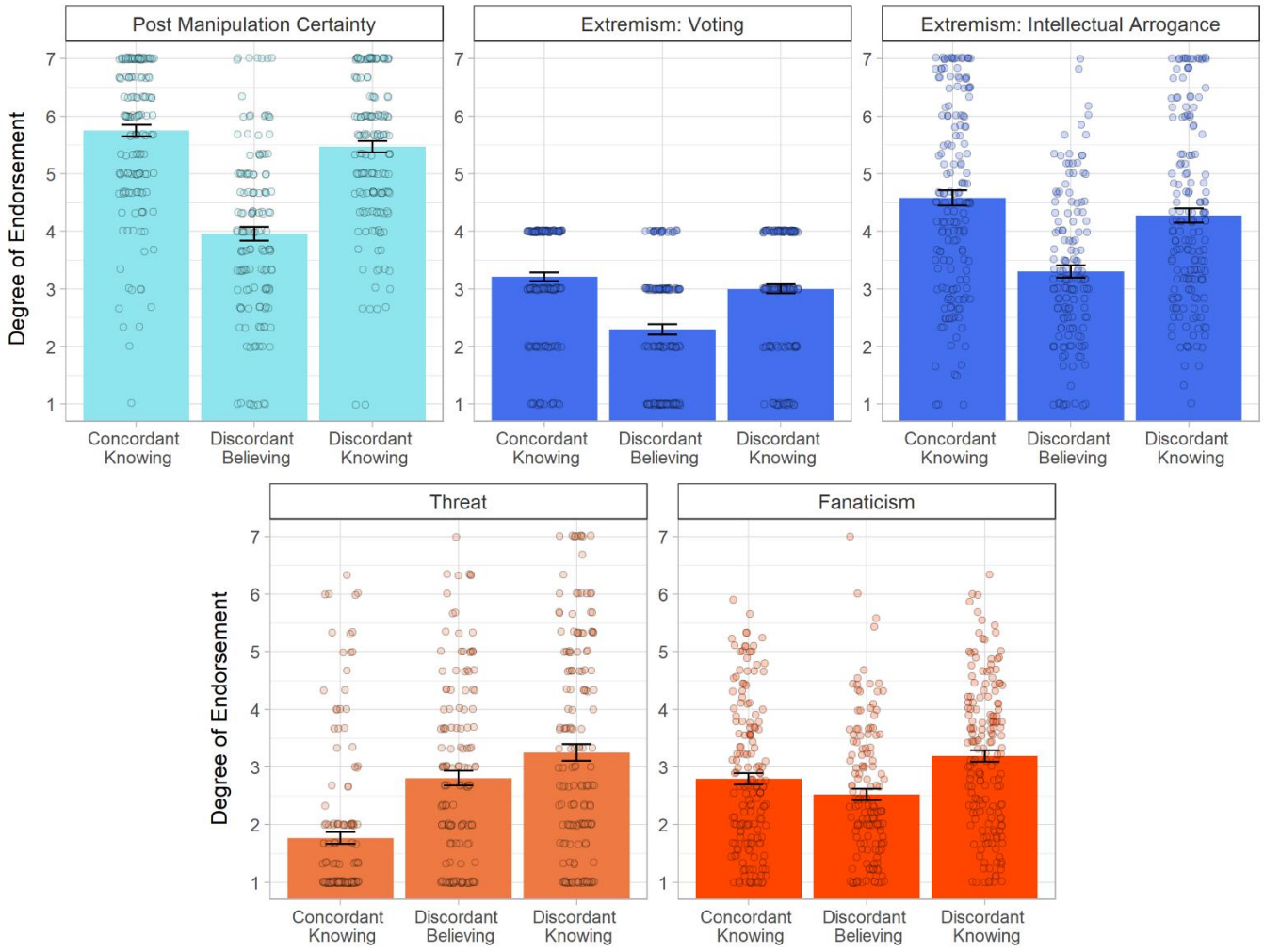

Figure 5. Effects of Discordant Knowing on post-manipulation certainty, extremism, threat, and fanaticism in Study 4. Error bars: +-1 *SE*.

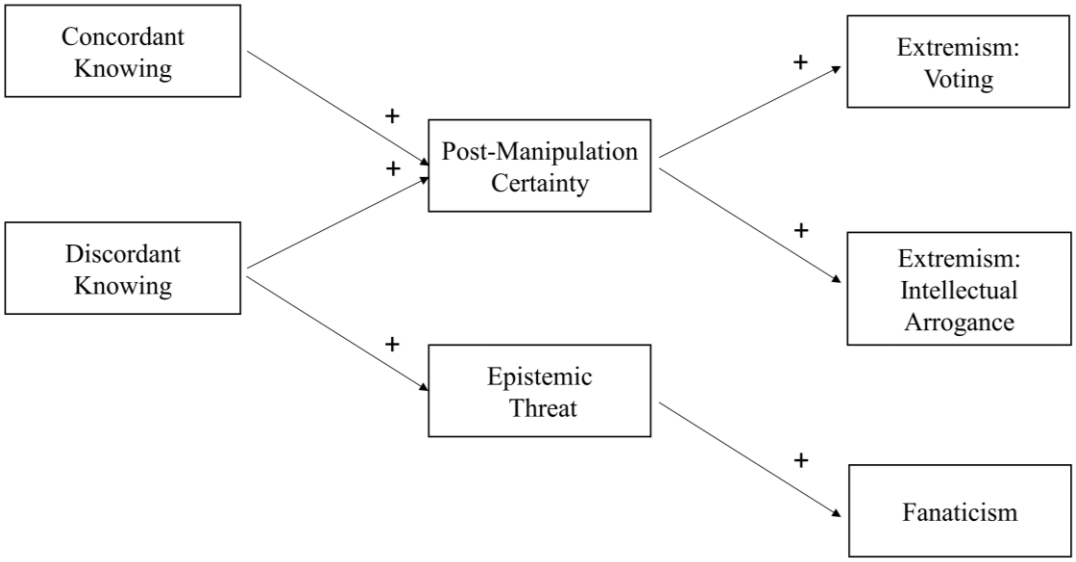

Figure 6. A visual model of the observed links in Study 4: Both discordant knowing and concordant knowing heightened post-manipulation certainty, which was linked to increased extremism. Only discordant knowing heightened epistemic threat, however, which was linked to heightened fanaticism.

Table 6

Study 4. Effects of Discordant Knowing

	Discordant Knowing (*M* and *SD*)	Discordant Believing (*M* and *SD*)	Concordant Knowing (*M* and *SD*)	Significance Test		
Study 4	*n* = 165	*n* = 144	*n* = 162	Test-Statistic	*p* value	Effect-Size
Mechanisms						
Epistemic Threat	3.26, 1.87	2.81, 1.57	1.77, 1.33	$F(2, 468) = 36.75$	$p < .001$	$\eta^2 = .136$
	◊	◊		$t(307) = 2.46$	$p = .014$	$d = 0.26$
	◊		◊	$t(325) = 8.38$	$p < .001$	$d = 0.92$
		◊	◊	$t(304) = 5.65$	$p < .001$	$d = 0.71$
Post-Manipulation Certainty	5.48, 1.29	3.96, 1.44	5.76, 1.29	$F(2, 468) = 78.91$	$p < .001$	$\eta^2 = .252$
	◊	◊		$t(307) = 9.99$	$p < .001$	$d = 1.11$
	◊		◊	$t(325) = 1.89$	$p = .058$	$d = 0.22$
		◊	◊	$t(304) = 11.76$	$p < .001$	$d = 1.32$
Dependent Variables						
Fanaticism	3.20, 1.28	2.52, 1.17	2.79, 1.29	$F(2, 468) = 11.60$	$p < .001$	$\eta^2 = .047$
	◊	◊		$t(307) = 4.75$	$p < .001$	$d = 0.55$
	◊		◊	$t(325) = 2.96$	$p = .003$	$d = 0.32$
		◊	◊	$t(304) = 1.90$	$p = .059$	$d = 0.22$
Extremism: Voting	3.00, 1.02	2.30, 1.08	3.23, 0.94	$F(2, 468) = 34.22$	$p < .001$	$\eta^2 = .128$
	◊	◊		$t(307) = 6.04$	$p < .001$	$d = 0.67$
	◊		◊	$t(325) = 2.04$	$p = .043$	$d = 0.23$
		◊	◊	$t(304) = 8.02$	$p < .001$	$d = 0.92$
Extremism: Intellectual Arrogance	4.28, 1.56	3.30, 1.29	4.58, 1.64	$F(2, 468) = 29.27$	$p < .001$	$\eta^2 = .111$
	◊	◊		$t(307) = 5.67$	$p < .001$	$d = 0.68$
	◊		◊	$t(325) = 1.77$	$p = .077$	$d = 0.19$
		◊	◊	$t(304) = 7.34$	$p < .001$	$d = 0.87$

Note. ◊ signifies inclusion in a pairwise comparison. Two participants were excluded from these analyses for missing data.

Table 7

Mediation Effects in Study 4. The Effect of Discordant Knowing on Fanaticism was Mediated by Epistemic Threat. The Effects of Discordant Knowing and Concordant Knowing on Extremism were Mediated by Post-Manipulation Certainty.

Variables:	Independent Variable	Mediator	Dependent Variable
Study 4 (N = 473)	Discordant Knowing	Epistemic Threat	Fanaticism
	Relative Effect: Discordant Knowing (coded: 1) Versus Discordant Believing (coded: 0)		
Total Effect	β = .522, SE = .112, t = 4.68, p < .001, 95% CI: [0.303, 0.741]*		
Direct Effect	β = .395, SE = .099, t = 3.99, p = .001, 95% CI: [0.200, 0.589]*		
Indirect Effect	β = .128, SE = .057, 95% CI: [0.016, 0.239]*		
	Relative Effect: Discordant Knowing (coded: 1) Versus Concordant Knowing (coded: 0)		
Total Effect	β = .311, SE = .108, t = 2.88, p = .004, 95% CI: [0.099, 0.523]*		
Direct Effect	β = -.116, SE = .102, t = -1.13, p = .259, 95% CI: [-0.316, 0.085]		
Indirect Effect	β = .427, SE = .065, 95% CI: [0.309, 0.563]*		
	Discordant Knowing	Post-Manipulation Certainty	Extremism: Voting
	Relative Effect: Discordant Knowing (coded: 1) Versus Discordant Believing (coded: 0)		
Total Effect	β = .645, SE = .107, t = 6.03, p < .001, 95% CI: [0.435, 0.855]*		
Direct Effect	β = .338, SE = .113, t = 3.00, p = .003, 95% CI: [0.117, 0.560]*		
Indirect Effect	β = .307, SE = .061, 95% CI: [0.194, 0.436]*		
	Concordant Knowing	Post-Manipulation Certainty	Extremism: Voting
	Relative Effect: Concordant Knowing (coded: 1) Versus Discordant Believing (coded: 0)		
Total Effect	β = .843, SE = .107, t = 7.86, p < .001, 95% CI: [0.632, 1.053]*		
Direct Effect	β = .481, SE = .117, t = 4.11, p < .001, 95% CI: [0.251, 0.711]*		
Indirect Effect	β = .362, SE = .067, 95% CI: [0.235, 0.500]*		

Note. Standardized Betas (β) are calculated in terms of z-score normalization. They are not the slope. For mediation results for Intellectual Arrogance see Supplements, Table S5.

Discordant Knowing | Discordant Believing | Concordant Knowing

Figure 7. Study 5: Participants' mental representations of how they would look after their viewpoint regarding societal concerns was manipulated to fall under a discordant knowing, discordant believing, or concordant knowing framework (assessed via reverse correlation methods). Independent raters judged the discordant knowing face as more fanatical than the discordant believing and concordant knowing composite faces.

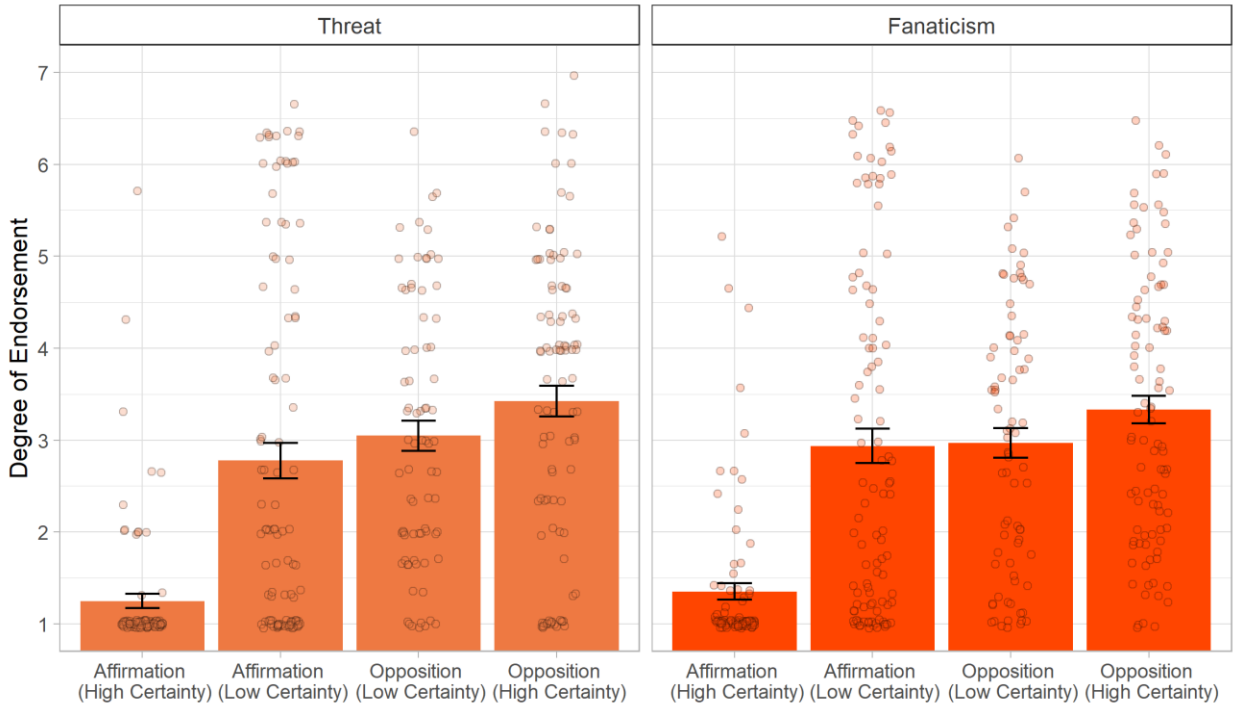

Figure 8. Effects of pre-manipulation certainty and Opposition (affirmation vs. opposition) on threat and fanaticism in Study 6. High versus low certainty was calculated via a median split (median: 7) for the purposes of this figure. High certainty combined with opposition captures discordant knowing. Error bars: +-1 *SE*.

Table 8

Study 6. Effects of Discordant Knowing on Epistemic Threat and Fanaticism. Discordant Knowing was Quantified as High Certainty + Majority Opposition.

Two-Way Interaction Terms Between Certainty and Opposition (Opposition vs. Affirmation) Predicting Epistemic Threat and Fanaticism:		
	Certainty x Opposition → Threat: $F(1, 359) = 21.37$, $p < .001$, $\eta p^2 = .056$	
	Certainty x Opposition → Fanaticism: $F(1, 359) = 21.43$, $p < .001$, $\eta p^2 = .056$	
Simple Effects of Opposition (Opposition coded 1; Affirmation coded 0) at High and Low Certainty		
High Certainty (+1.00 *SD*)	Opposition → Threat: $F(1, 359) = 68.62$, $p < .001$, $\eta^2 = .160$	
	Opposition → Fanaticism: $F(1, 359) = 57.19$, $p < .001$, $\eta^2 = .137$	
Low Certainty (-1.00 *SD*)	Opposition → Threat: $F(1, 359) = 2.57$, $p = .110$, $\eta^2 = .007$	
	Opposition → Fanaticism: $F(1, 359) = 0.79$, $p = .376$, $\eta^2 = .002$	
Predicted Means and Standard Deviations for Threat and Fanaticism in Each Cell (Certainty x Opposition)		
	Affirmation Condition	Opposition Condition
High Certainty (+1.00 *SD*)	Threat: $M = 1.60$, $SE = 0.17$	Threat: $M = 3.58$, $SE = 0.17$
	Fanaticism: $M = 1.70$, $SE = 0.16$	Fanaticism: $M = 3.44$, $SE = 0.16$
Low Certainty (-1.00 *SD*)	Threat: $M = 2.47$, $SE = 0.15$	Threat: $M = 2.86$, $SE = 0.19$
	Fanaticism: $M = 2.63$, $SE = 0.15$	Fanaticism: $M = 2.84$, $SE = 0.18$

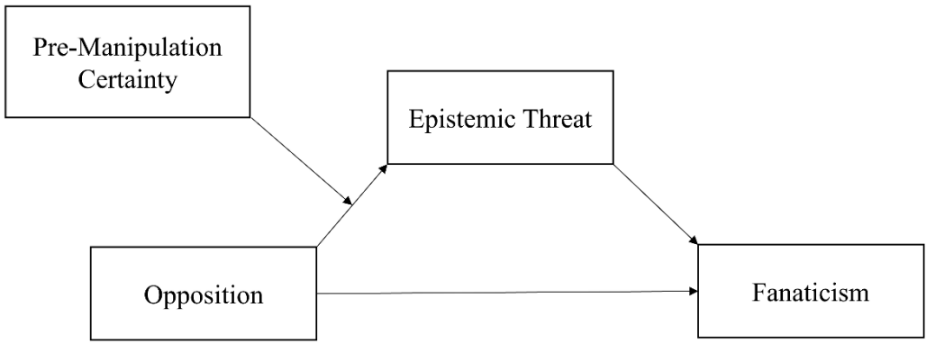

Figure 9. A conceptual diagram of the moderated mediation in Study 6. The effect of Opposition on political fanaticism via threat was only observed for participants who were certain that their preferred candidate would be a better president.

Table 9

Mediation Effects in Study 6: A Moderated Mediation was Observed. The Effect of Opposition on Fanaticism via Epistemic Threat was Only Observed for Participants High in Pre-Manipulation Certainty.

Variables:	Independent Variable	Moderator	Mediator	Dependent Variable
Study 6 (N = 363)	Opposition	Pre-Manipulation Certainty	Threat	Fanaticism
	Conditional Indirect Effects of Opposition on Fanaticism:			
	Opposition (coded: 1) Versus Affirmation (coded: 0) at Low Certainty (z-scored) = -1.28			
Indirect Effect	β = .082, SE = .131, 95% CI: [-0.203, 0.319]			
	Opposition (coded: 1) Versus Affirmation (coded: 0) at High Certainty (z-scored) = 0.77			
Indirect Effect	β = .869, SE = .093, 95% CI: [0.690, 1.056]*			
	Index of Moderated Mediation:			
Moderated Mediation:	β = .384, SE = .077, 95% CI: [0.246, 0.549]*			

Note. Standardized Betas (β) are calculated in terms of z-score normalization. They are not the slope.

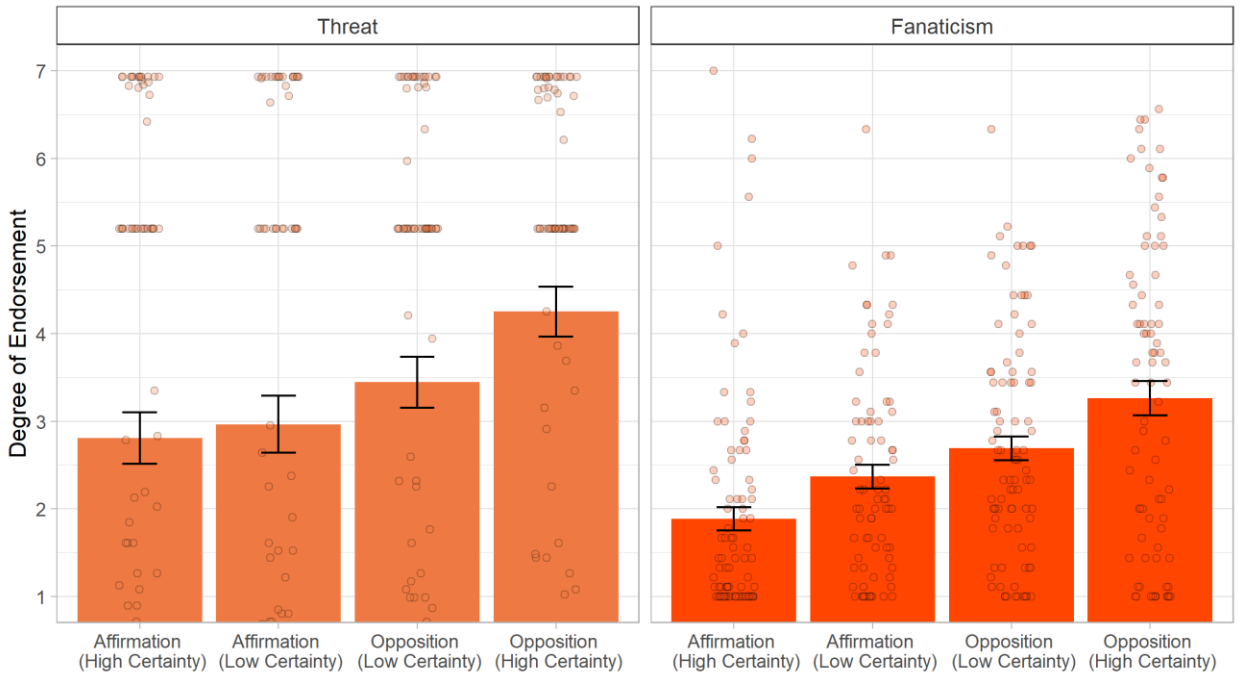

Figure 10. Effects of pre-manipulation certainty and Opposition (affirmation vs. opposition) on threat and fanaticism in Study 7. High certainty combined with opposition represents discordant knowing. High versus low certainty was calculated via a median split (median: 7) for this figure. Threat was recoded onto a 1-7 scale for this figure. Error bars: +-1 *SE*.

Table 10

Study 7. Effects of Discordant Knowing on Epistemic Threat and Fanaticism. Discordant Knowing was Quantified as High in Pre-Manipulation Certainty + Majority Opposition.

Two-Way Interaction Terms Between Certainty and Opposition (Opposition vs. Affirmation) Predicting Epistemic Threat and Fanaticism: Certainty x Opposition → Threat: $F(1, 336) = 10.08$, $p = .002$, $\eta p^2 = .029$ Certainty x Opposition → Fanaticism: $F(1, 336) = 10.66$, $p = .001$, $\eta p^2 = .031$		
Simple Effects of Opposition (Opposition coded 1; Affirmation coded 0) at High and Low Certainty		
High Certainty (+1.00 *SD*)	Opposition → Threat: $F(1, 336) = 21.57$, $p < .001$, $\eta^2 = .060$ Opposition → Fanaticism: $F(1, 336) = 40.92$, $p < .001$, $\eta^2 = .109$	
Low Certainty (-1.00 *SD*)	Opposition → Threat: $F(1, 336) < 0.01$, $p = .954$, $\eta^2 < .001$ Opposition → Fanaticism: $F(1, 336) = 2.69$, $p = .102$, $\eta^2 = .008$	
Predicted Means and Standard Deviations for Threat and Fanaticism in Each Cell (Certainty x Opposition)		
	Affirmation Condition	Opposition Condition
High Certainty (+1.00 *SD*)	Threat: $M = 37.13$, $SE = 4.19$ Fanaticism: $M = 1.93$, $SE = 0.15$	Threat: $M = 64.96$, $SE = 4.29$ Fanaticism: $M = 3.30$, $SE = 0.15$
Low Certainty (-1.00 *SD*)	Threat: $M = 46.26$, $SE = 4.78$ Fanaticism: $M = 2.34$, $SE = 0.17$	Threat: $M = 46.62$, $SE = 3.84$ Fanaticism: $M = 2.70$, $SE = 0.14$

Table 11

Mediation Effects in Study 7: A Moderated Mediation was Observed. The Effect of Opposition on Fanaticism via Epistemic Threat was Only Observed for Participants High in Pre-Manipulation Certainty.

Variables:	**Independent Variable**	**Moderator**	**Mediator**	**Dependent Variable**
Study 7 (N = 338)	Opposition	Pre-Manipulation Certainty	Threat	Fanaticism
	Conditional Indirect Effects of Opposition on Fanaticism:			
	Opposition (coded: 1) Versus Affirmation (coded: 0) at Low Certainty (- 1.14 SD)			
Indirect Effect	β = -.007, SE = .030, 95% CI: [-0.073, 0.051]			
	Opposition (coded: 1) Versus Affirmation (coded: 0) at High Certainty (+0.78 SD)			
Indirect Effect	β = .103, SE = .043, 95% CI: [0.031, 0.198]*			
	Index of Moderated Mediation:			
Moderated Mediation:	β = .057, SE = .028, 95% CI: [0.012, 0.121]*			

Note. Standardized Betas (β) are calculated in terms of z-score normalization. They are not the slope.

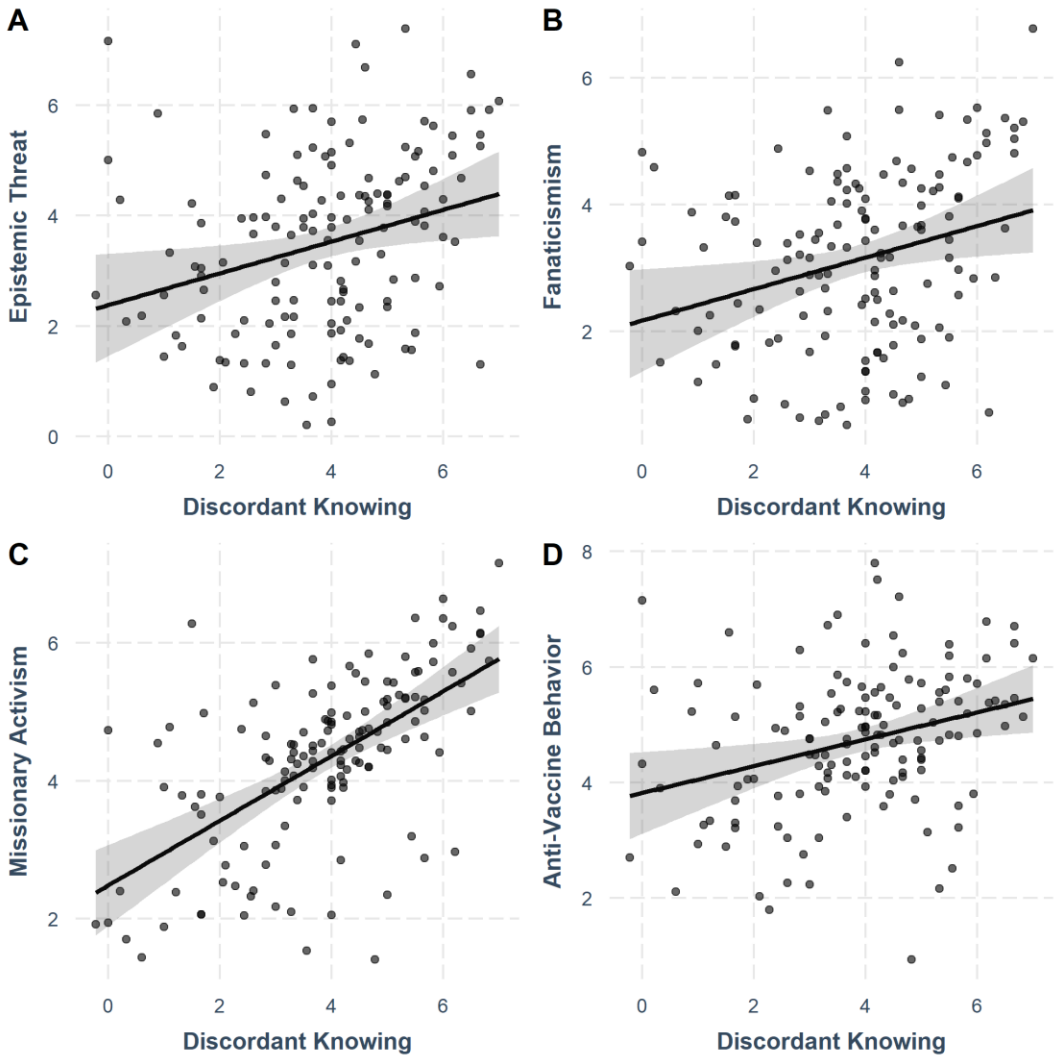

Figure 11. Partial residual effect plots depicting discordant knowing predicting epistemic threat, fanaticism, missionary activism, and anti-vaccine behavior in Study 8. Model Set 3 was used, see Table 13. Error bands: *CI*s (using geom_ribbon in R).

Table 12

Study 8. Discordant Knowing Predicting Epistemic Threat, Fanaticism, Missionary Activism, and Self-Reported Anti-Vaccine Behavior (Multivariate Regressions Including Covariates).

Dependent Variables:	Epistemic Threat	Fanaticism	Missionary Activism	Anti-Vaccine Behavior
Total Variance Explained:	$R^2 = .27$	$R^2 = .26$	$R^2 = .50$	$R^2 = .46$
Predictors in the Model:				
Discordant Knowing	$\beta = .36$, $p < .001$	$\beta = .34$, $p < .001$	$\beta = .55$, $p < .001$	$\beta = .28$, $p < .001$
Discordant Believing	$\beta = .11$, $p = .175$	$\beta = .14.$, $p = .091$	$\beta = .09$, $p = .183$	$\beta = .15$, $p = .028$
Vaccine Attitudes	$\beta = .16$, $p = .059$	$\beta = .02$, $p = .830$	$\beta = .23$, $p = .001$	$\beta = .49$, $p < .001$
Age	$\beta = -.13$, $p = .093$	$\beta = -.19$, $p = .018$	$\beta = -.08$, $p = .209$	$\beta = .09$, $p = .171$
Gender	$\beta = -.10$, $p = .201$	$\beta = -.07$, $p = .341$	$\beta = -.02$, $p = .771$	$\beta = .03$, $p = .609$
Education	$\beta = .10.$ $p = .172$	$\beta = .23$, $p = .002$	$\beta = .12$, $p = .050$	$\beta = .04$, $p = .517$
Politics	$\beta = .09$, $p = .251$	$\beta = .06$, $p = .485$	$\beta = .01$, $p = .843$	$\beta = .02$, $p = .821$

Table 13

Study 8. Paradoxical Knowing Predicting Threat, Fanaticism, Missionary Activism, and Self-Reported Anti-Vaccine Behavior (Multivariate Regressions Including Covariates).

Dependent Variables:	Epistemic Threat	Fanaticism	Missionary Activism	Anti-Vaccine Behavior
Total Variance Explained:	$R^2 = .29$	$R^2 = .29$	$R^2 = .36$	$R^2 = .44$
Predictors in the Model:				
Paradoxical Knowing	$\beta = .38$, $p = .001$	$\beta = .30$, $p = .012$	$\beta = .31$, $p = .007$	$\beta = .30$, $p = .004$
Paradoxical Believing	$\beta = .06$, $p = .580$	$\beta = .17$, $p = .135$	$\beta = .12$, $p = .273$	$\beta = .03$, $p = .746$
Vaccine Attitudes	$\beta = .35$, $p < .001$	$\beta = .20$, $p = .007$	$\beta = .50$, $p < .001$	$\beta = .63$, $p < .001$
Age	$\beta = -.06$, $p = .466$	$\beta = -.11$, $p = .167$	$\beta = -.01$, $p = .866$	$\beta = .15$, $p = .033$
Gender	$\beta = -.11$, $p = .123$	$\beta = -.09$, $p = .209$	$\beta = -.03$, $p = .722$	$\beta = .02$, $p = .709$
Education	$\beta = .03$, $p = .711$	$\beta = .16$, $p = .038$	$\beta = .07$, $p = .301$	$\beta = -.01$, $p = .870$
Politics	$\beta = .12$, $p = .121$	$\beta = .08$, $p = .319$	$\beta = .09$, $p = .247$	$\beta = .04$, $p = .524$

Table 14

Study 8. Discordant Knowing and Paradoxical Knowing Predicting Epistemic Threat, Fanaticism, Missionary Activism, and Self-Reported Anti-Vaccine Behavior (Multivariate Regressions Including Covariates).

Dependent Variables:	Epistemic Threat	Fanaticism	Missionary Activism	Anti-Vaccine Behavior
Total Variance Explained:	$R^2 = .31$	$R^2 = .31$	$R^2 = .51$	$R^2 = .47$
Predictors in the Model:				
Discordant Knowing	$\beta = .24$, $p = .015$	$\beta = .24$, $p = .015$	$\beta = .53$, $p < .001$	$\beta = .23$, $p = .010$
Paradoxical Knowing	$\beta = .27$, $p = .056$	$\beta = .17$, $p = .215$	$\beta = -.04$, $p = .734$	$\beta = .12$, $p = .334$
Discordant Believing	$\beta = -.07$, $p = .455$	$\beta = -.05$, $p = .611$	$\beta = .02$, $p = .833$	$\beta = .07$, $p = .399$
Paradoxical Believing	$\beta = .11$, $p = .346$	$\beta = .22$, $p = .065$	$\beta = .19$, $p = .054$	$\beta = .05$, $p = .607$
Vaccine Attitudes	$\beta = .23$, $p = .007$	$\beta = .09$, $p = .320$	$\beta = .25$, $p = .001$	$\beta = .53$, $p < .001$
Age	$\beta = -.07$, $p = .395$	$\beta = -.12$, $p = .124$	$\beta = -.06$, $p = .396$	$\beta = .12$, $p = .079$
Gender	$\beta = -.12$, $p = .108$	$\beta = -.10$, $p = .189$	$\beta = -.03$, $p = .631$	$\beta = .02$, $p = .719$
Education	$\beta = .03$, $p = .709$	$\beta = .16$, $p = .035$	$\beta = .10$, $p = .126$	$\beta = .01$, $p = .904$
Politics	$\beta = .08$, $p = .316$	$\beta = .04$, $p = .645$	$\beta = -.001$, $p = .986$	$\beta = .01$, $p = .895$

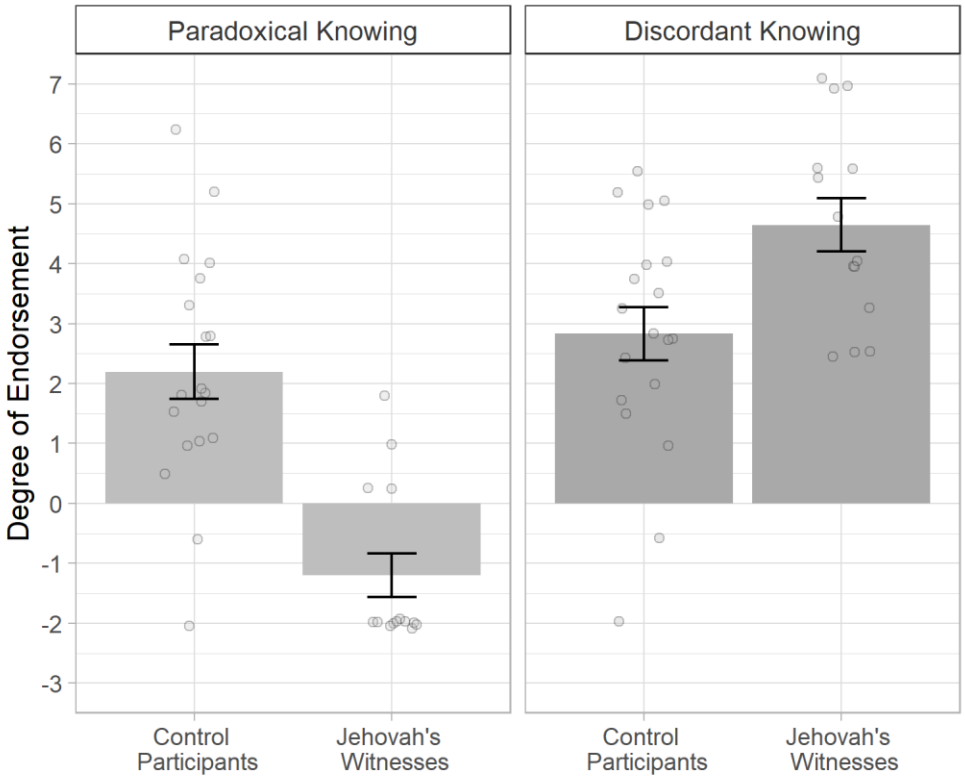

Figure 12. Levels of discordant and paradoxical knowing observed in Jehovah's Witnesses versus non-fanatical control participants. Error bars: +-1 *SE*.