Comprehensive identity fusion theory (CIFT): New insights and a revised theory☆

William B. Swann Jr.✉, Jack W. Klein✉, and Ángel Gómez✉

✉Department of Psychology, University of Texas at Austin, Austin, TX, United States
✉Melbourne School of Psychological Sciences, University of Melbourne, Parkville, VIC, Australia
✉Department of Social and Organizational Psychology, Faculty of Psychology, Universidad Nacional de Educación a Distancia, UNED, Madrid, Spain
✉Corresponding author. e-mail address: swann@utexas.edu

Contents

1. Origin and nature of the identity fusion construct and its relationship to the social identity perspective 2
   1.1 The agentic-personal-self and identity-synergy principles 5
   1.2 Measuring identity fusion: The verbal fusion scale 8
   1.3 Fusion and extreme pro-group behavior 11
2. The developmental trajectory of identity fusion 13
   2.1 Attachment theory 14
   2.2 Self-verification 14
   2.3 Shared essence 16
   2.4 Emotionally compelling experiences 17
3. Targets of identity fusion 18
   3.1 Fusion with fellow humans (groups, leaders, and other individuals) 19
   3.2 Fusion with values, ideologies, and causes 20
   3.3 Fusion with brands and creatures 21
4. Properties of fusion targets that may entice potential fusers 22
5. Temporal stability of identity fusion 24
   5.1 Defusion 26
6. Related formulations: Devoted actors and an evolutionary approach to the causes of fusion 30
   6.1 Sacred values, fusion with values, and the devoted actor formulation 30
   6.2 Whitehouse’s evolutionary approach to identity fusion 32
7. The role of fusion in understanding key contemporary issues 35
   7.1 Motivated reasoning in the political sphere 35
   7.2 Fusion and terrorism 37
   7.3 Fusion and football hooliganism 40

☆ For helpful comments on this chapter, we thank Pelin Cunningham-Erdogdu, Philip Moniz and Alexandra Vázquez.
Abstract

We present a comprehensive identity fusion theory which revises and expands the original version. As did the original theory, the revision maintains that fusion—a synergistic union of the personal self and target of fusion—is marked by porous borders between the personal self and the target. These porous borders give rise to “identity synergy” wherein people synergistically channel their personal agency into behaviors that support the target. As a result, strongly fused people enact behaviors on behalf of the target even if it is costly to do so. The revised theory builds on its predecessor by (1) streamlining the theory by emphasizing the critical characteristics that distinguish it from the social identity perspective; (2) examining the developmental trajectory of identity fusion; (3) extending the theory to targets of fusion other than groups, especially values, ideologies, and individual persons; (4) discussing the properties of targets of fusion that entice potential fusers; (5) exploring the temporal stability of fusion, including strategies for defusion; (6) clarifying the relation of the theory to related formulations; and (7) analyzing applications of fusion to several important social issues.

1. Origin and nature of the identity fusion construct and its relationship to the social identity perspective

The seeds of identity fusion theory were planted in 2005 through a series of conversations between the first and last authors of this chapter. These conversations were triggered by tumultuous events that had recently occurred in their respective countries. For Swann, it was the attacks on the Twin Towers in 2001. For Gómez, it was the Madrid train bombings in 2004. Struggling to understand the motives of the terrorists who launched these attacks, the pair searched for leads in the literature. They discovered that attempts to account for extreme pro-group behavior had a long history in the behavioral sciences. Examples included Emile Durkheim’s (1964, 1995) concept of mechanical solidarity and collective effervescence; Victor Turner’s (1969) notion of spontaneous communitas (an intense feeling of togetherness and common humanity), and Alan Fiske’s (1991) notion of communal sharing (see also Sahlins, 1974). These formulations highlighted some of the emotional processes that might motivate extreme behavior. Nevertheless, they failed to illuminate the identity-related processes that Swann and Gómez came to believe were the root cause of the actions of terrorists.
The authors’ interest in the role of identity in extreme behavior took them to the social psychological literature and the *group identification* construct (Lewin, 1948). They discovered that most researchers agreed that identification involved feelings of allegiance to a collective (e.g., Brewer & Gardner, 1996; Postmes, Haslam, & Swaab, 2005; Prentice, Miller, & Lightdale, 1994). They also learned that the construct was typically associated with the social-identity perspective, which encompasses social identity theory (e.g., Tajfel & Turner, 1979) and its extension, self-categorization theory (e.g., Turner, Oakes, Haslam, & McGarty, 1994). Finally, they noted that the social identity perspective was not developed to explain extreme behavior. Instead, it was designed to explicate the role of intergroup comparisons in prejudice and discrimination. The insights it produced allowed it to dominate social psychological theorizing regarding group processes for over four decades (for recent examples, see Ellemers, Spears, & Doosje, 2002; Ellemers, De Gilder, & Haslam, 2004; Haslam, Reicher, & Platow, 2011; Hornsey & Jetten, 2004; Mackie, Devos, & Smith, 2000; Mackie, Smith, & Ray, 2008; Smith, Seger, & Mackie, 2007; Spears, Oakes, Ellemers, & Haslam, 1997).

Swann and Gómez embraced social identity theory’s distinction between the social self (referring to the groups to which one is aligned, such as “Spaniard” or “Democrat”) and the personal self (referring to unique qualities of the individual, such as assertive or careful). Nevertheless, their goals were distinct from those expressed in the classic versions of the social identity perspective (Tajfel & Turner, 1979; Turner et al., 1994). For example, whereas social identity theory is fundamentally a theory about intergroup processes with an emphasis on phenomena such as intergroup prejudice and discrimination, identity fusion theory is fundamentally a theory about how the personal self joins with various targets of fusion to motivate extreme behavior. Swann and Gómez also developed a distinctive understanding of the nature of personal identity and its relationship to social identity (we use “self” and “identity” interchangeably here). The classic social identity perspective embraced the idea of a sovereign social self that eclipses a feckless personal self when people align themselves with a group (the “depersonalization” hypothesis) and the related idea that the activation of social identities competes with the activation of the personal

---

1 While acknowledging the many informal revisions of the social identity perspective, we focus here on the classic versions because the revisions are inconsistent with one another and no single revision has been recognized as the successor to the original theory.
self (the “functional antagonism” hypothesis). Swann and Gómez rejected these ideas, contending instead that when strongly fused persons join groups, personal identities remain activated and work together with social identities to synergistically motivate behavior.

To be sure, Swann and Gómez’s critique of the depersonalization and functional antagonism constructs was reminiscent of earlier complaints voiced by revisionist social identity researchers (e.g., Abrams, 1994; Baray, Postmes, & Jetten, 2009; Greenaway et al., 2015; Pickett, Silver, & Brewer, 2002; Postmes & Jetten, 2006; Reid & Deaux, 1996; Simon, 2004; Spears, 2001; Stephenson, 1981). The pair went beyond these earlier complaints, however, by contending that important representations of the personal self are relatively stable and chronically activated. This allows the personal self to synergistically interact with social self-views to motivate extreme pro-group behaviors.

The synergistic union of the personal and social selves among strongly fused persons changes their mental representations of their relationship to fellow group members. Although continuing to acknowledge clear boundaries between themselves and others, strongly fused persons perceive these boundaries as porous and permeable. This leads to a powerful “sense of oneness” with the fusion target, defined by strong feelings of emotional closeness. These porous borders will, in turn, facilitate synergies between the personal and group identities wherein each strengthens the other. That is, rather than focusing on the group exclusively as a source of strength or self-enhancement (e.g., Abrams & Hogg, 1988; Martiny & Rubin, 2016; Tajfel & Turner, 1979), strongly fused persons will search for ways to strengthen the group. For such individuals, the relationship between the individual and group is reciprocal, with strongly fused persons actively supporting the group as well as drawing strength from the group. There is evidence supporting the construct validity of the reciprocal strength component. For example, the more people are fused with a group, the more group-related agency they experience (Gómez, Brooks et al., 2011) and the more they perceive both themselves and their group to be physically and spiritually formidable.

---

2 In the original publication on identity fusion, Swann, Gómez, Seyle, Morales, and Huici (2009) indicated that fusion causes the borders between the self and group to become blurred. Swann et al. (2012) later recognized that this is misleading because it implies that fusion causes people to lose sight of who they are. This latter possibility not only contradicts the theory and relevant evidence, it would theoretically compromise the capacity of the personal and social self to engage in mutual strengthening. They accordingly replaced the notion that fusion causes borders to become blurred with the idea that it causes the borders to become porous.
This empirical evidence supports the assumption that reciprocal strength is a defining component of identity fusion, and helps explain the strong relationship between identity fusion and a willingness to defend the group. In fact, mediational analyses suggest that perceptions of formidability are associated with the perception of ingroup trust, which fosters the will to fight for the group. Support for this causal chain has emerged in a wide array of participants, including Syrian refugees, imprisoned jihadists and gangs, U.S. military personnel, Ukrainians before and during the current war with Russia, European supporters of Ukraine, and individuals from the middle East and Africa (Gómez, Vázquez, & Atran, 2023).

The larger point here is that the sense-of-oneness and reciprocal-strength components have come to assume center stage in identity fusion theory and research. The significance of each of these constructs clearly rests on the assumption that the personal self (i.e., a person’s unique characteristics) guides and motivates behavior even when a group identity is salient [see Swann et al.’s (2012) “agentic-personal-self” principle]. After all, if the personal self were not agentic, proposing that it combines synergistically with group identities would lose force. Fortunately for the identity synergy principle, the research literature has provided converging evidence that the personal self is in fact highly agentic (see also Greenaway, 2024, this volume).

1.1 The agentic-personal-self and identity-synergy principles

In recent years considerable evidence has emerged that personal identities are potent, more potent even than group identities. For example, people attribute more of “who they are” to their personal self and associate more future goals with their personal self (Heger, Voorhees, Porter, & Gaertner, 2023). People also ascribe greater worth to their personal identities than their social identities and have particularly strong reactions when these identities are threatened (for a review, see Sedikides, Gaertner, Luke, O’Mara, & Gebauer, 2013). This pattern emerges consistently across numerous moderators (various threats and identities, levels of group identification, individualistic vs. collectivistic cultures) and methodologies.

---

3 The formidability representation hypothesis (Fessler et al., 2012) states that formidability is the sum of an adversary’s tactical assets and liabilities compared to one’s own. Recently, some authors (Gómez et al., 2017; Gómez, Vázquez et al., 2023; Tossell et al., 2022) have distinguished two facets of formidability, physical (fighting ability or the capacity to inflict harm on others) and spiritual (the conviction and non-material resources including values, strength of beliefs, and character).
including experiments (Gaertner, Sedikides, & Graetz, 1999; Gaertner, Sedikides, & Cai, 2012; Nehrlich, Gebauer, Sedikides, & Abele, 2018), meta-analyses (Gaertner, Sedikides, Vevea, & Iuzzini, 2002), and cross-cultural studies (Chen, Boucher, Andersen, & Saribay, 2013; del Prado et al., 2007; Gaertner et al., 2012; Zhu, Wu, Yang, & Gu, 2016). This evidence of the relative potency of the personal self supports its significance in fusion theory and suggests that it may remain salient and consequential even when the social self is activated.

Direct tests of the identity synergy principle have been similarly supportive. Some of the strongest support for the principle was provided by investigations based on self-verification theory (Swann, 1983, 2012). The theory suggests that self-views become the center of their psychological universes, enabling them to make sense of events, guide behavior and predict what will happen next. The role of stable self-views in psychological life is so important that people seek and evoke self-verifying evaluations even if the self-view is negative (Swann, Pelham, & Krull, 1989), as support for both negative and positive identities foster the perception that the world is knowable and coherent and that others are likely to treat the individual in familiar and understandable ways (Ashokkumar & Swann, 2020; Swann, 2012).

Of particular relevance to the process of identity fusion, if people receive evaluations that challenge their self-views, they enact compensatory efforts to reaffirm these self-views (for reviews, see Ashokkumar & Swann, 2020; Kwang & Swann, 2010; Swann, 2012). Insofar as the boundaries between the personal and social self-views of strongly fused persons are highly permeable, self-verification theory suggests that challenges to either the personal or social self-views of such individuals should trigger compensatory efforts to reaffirm the other type of self-view. The first tests of this hypothesis focused on challenging participants by giving them unexpectedly positive feedback about personal qualities (Swann, et al., 2009). Following the challenge manipulation, the researchers assessed participants’

---

4 From this vantage point, efforts to verify subjectively accurate negative self-views is a feature rather than a bug of the self-verification process, for it is surely adaptive to reinforce negative self-views that are grounded in reality. If there is a “bug” associated with self-verification strivings, it is that they are sometimes guided by self-views that are inappropriately negative (e.g., “worthless”, “unlovable”). In such instances, individuals may seek negative evaluations that they do not deserve, causing depression, low self-esteem, and a tendency to foreclose possibilities that could improve their lives (Giesler, Josephs, & Swann, 1996; Swann, 1996).

5 Providing participants with unexpectedly negative feedback should also trigger compensatory activity, but such activities could reflect either self-verifying or self-enhancement motivations, an ambiguity that the investigators sought to avoid.
endorsement of pro-group behaviors. As predicted, challenging participants’ personal self-views increased subsequent endorsement of pro-group activity (fighting and dying for the group) among strongly fused persons but not among weakly fused persons (Swann et al., 2009; Experiments 1 and 2). Later studies replicated this effect using two different measures of identity fusion (Gómez, Brooks et al., 2011; Study 10; Paredes, Santos, Briñol, Gómez, & Petty, 2019) and measures of identity that focused on collective self-views (e.g., Chen, Chen, & Shaw, 2004; Chen, Shaw Taylor, & Jeung, 2006) and ingroup identities (Gómez, Seyle, Huici, & Swann, 2009; Gómez, Dovidio, Gaertner, Fernández, & Vázquez, 2013).

Consistent with the identity synergy principle, the desire of strongly fused persons to compensate for a challenge to their personal self-views amplified their efforts to protect their group. A further test of the identity synergy principle involved independently activating personal and social identities of strongly fused persons. Researchers first asked participants how they would respond to a physical attack either on themselves (a challenge to the personal self) or their group (a challenge to the social self). Later, the investigators assessed participants’ endorsement of pro-group actions. The results indicated that activating either the personal or social identities of strongly fused persons increased their subsequent propensity to endorse fighting or dying for the group. No such pattern emerged among weakly fused persons (Swann et al., 2009; Experiment 3).

Additional support for the identity synergy principle came from studies on the relationship between identity fusion and reactions to social ostracism. In several studies, ostracizing participants for either their personal preferences (i.e., personal self-views) or their group membership (i.e., social self-views) amplified their subsequent endorsement of three distinct types of compensatory activities: extreme actions for the group, resolve to remain in the group, and charitable donations to the group (Gómez, Morales, Hart, Vázquez, & Swann, 2011). As in the initial studies (Swann et al., 2009), the findings supported the identity synergy principle by showing that it was possible to amplify pro-group action by activating either the personal or social identities of strongly fused participants.

A related line of research was based on activating the repository of the personal self: the physical body. The researchers began by noting evidence that physical exercise bolsters autonomic arousal (Jacobs & Farel, 1971) and that identity fusion is linked to agency for the group (Gómez, Brooks et al., 2011; Swann, Gómez, Huici, Morales, & Hixon, 2010). With this evidence in hand, the researchers proposed that exercise would trigger arousal and compel strongly fused persons to enact pro-group behaviors. To test this idea, the researchers experimentally increased physiological arousal by
having high-school students either operate an exercycle, run wind sprints, or play dodgeball. As expected, increasing arousal amplified the tendency for strongly fused persons to augment pro-group behaviors such as donating personal funds to a needy group member or enacting pro-group motor activity (i.e., racing a group-related avatar). What’s more, feelings of agency statistically mediated the interactive effects of arousal and fusion on pro-group behavior (Swann, Gómez, Huici et al., 2010). A later investigation replicated this mediational finding using two distinct measures of identity fusion (Gómez, Brooks et al., 2011). A final follow-up study used an intergroup version of the trolley dilemma in which participants chose between either dying for the group or letting fellow group members die instead. The researchers found that inducing participants to express their immediate, gut-level reactions to the dilemma exaggerated the tendency of strongly fused persons to self-sacrifice for the ingroup (Spain) but not an outgroup (Northern Europe). Therefore, the impact of arousal on self-sacrifice was restricted to the target of fusion (Swann, Gómez et al., 2014).

The foregoing evidence for the sense-of-oneness and reciprocal-strength components supports two of fusion theory’s core ideas. That is, the research literature indicates that among strongly fused persons, the personal self and target of fusion become entwined, leading to strong feelings of emotional attachment to the fusion target. The personal self and target also work together synergistically, with each strengthening the other. In addition, this evidence has implications for how fusion should be measured.

### 1.2 Measuring identity fusion: The verbal fusion scale

The sense-of-oneness and reciprocal-strength components were key to the development of a verbal measure of identity fusion (Gómez, Brooks et al., 2011). First, like the pictorial measure, items were designed to capture a sense of oneness with the fusion target (“I am one with the group”). Most important, the reciprocal strength component led to the incorporation of items that emphasized the person’s contributions to the group (“I make my group strong”) as well as the group’s contributions to the self (“I am strong because of my group”). In fact, for strongly fused persons, the welfare of the group is as important as their own personal welfare. These feelings compel strongly fused individuals to endorse or enact behaviors that exemplify the group’s goals and values (Swann et al., 2009; Swann, Jetten, Gómez, Whitehouse, & Bastian, 2012), including even violence, self-sacrifice, and retribution against outgroup members (Fredman, Bastian, & Swann, 2017; Swann, Buhrmester et al., 2014; Vázquez et al., 2020).
As such, strongly fused identities pack more motivational “oomph” than could be provided by either the personal or group identity acting alone (Swann et al., 2012). Consistent with this reasoning, a growing body of research shows that identity fusion is an exceptionally strong predictor of violent pro-group behavior, consistently out-predicting dozens of rival variables including sacred values, moral convictions (Martel, Buhrmester, Gomez, Vazquez, & Swann, 2021), and group identification (for recent reviews, see Gómez, Chinchilla et al., 2020; Varmann et al., 2023; Wolfowicz, Litmanovitz, Weisburd, & Hasisi, 2021).

One contribution of the verbal measure of identity fusion was to exemplify the difference between identity fusion and identification in ways that previous measures had failed to do. Consider the initial article on identity fusion in which the researchers employed a pictorial measure of fusion (Swann et al., 2009). Although the measure did predict endorsement of extreme behavior, it provided little insight into how respondents perceived the target of fusion—for example, whether the respondent believed that it was incumbent on group members to strengthen the group or merely rely on the group for self-enhancement. As such, the pictorial measure was poorly suited for discriminating between fusion and identification. For these reasons, researchers subsequently developed a verbal measure of fusion that was specifically designed to capture the unique elements of fusion. For example, to capture reciprocal strength, the researchers included items that emphasized reciprocal strength, namely the respondent’s contributions to the group (“I make my group strong”) as well as the group’s contributions to the respondent (“I am strong because of my group”).

A ten-study validation article of the verbal measure of fusion (Gómez, Brooks et al., 2011) provided converging evidence for the independence of measures of fusion and identification. A factor analysis (N = 1000) revealed that all the items on the fusion measure loaded on the fusion factor but not on the identification factor. Moreover, all the items on the identification measure loaded on the identification factor but not on the fusion factor. In addition, Gomez et al. (2011) found that the correlations between the verbal fusion measure and three prominent identification measures were modest to moderate: for Jetten, Branscombe, Schmitt, and Spears (2001): $r (112) = 0.26$, for Mael and Ashforth (1992): $r (198) = 0.56$, $p < 0.01$, and for Tropp and Wright (2001): $r (248) = 0.23$. Further, the verbal fusion scale predicted endorsement of extreme pro-group behaviors with greater fidelity than did Swann et al. (2009) pictorial measure of identity fusion, which was, in turn, superior to a measure of group identification (but see Varmann et al., 2023).
The researchers also replicated previous evidence that the personal and social selves of strongly fused persons are functionally equivalent, and that feelings of agency and invulnerability mediated the effects of fusion on extreme behavior. Finally, although both the verbal and pictorial measures of fusion were significant predictors of pro-group behavior when entered separately into regressions, only the verbal scale remained significant when both scales were included in the regression. Apparently, both scales measure alignment with the fusion target but the verbal scale assessed something more—presumably reciprocal strength.

In the decade following the creation of the verbal measure of fusion, researchers reported that it was an exceptionally strong predictor of pro-group behavior (for recent reviews, see Bortolini, Newson, Natividade, Vázquez, & Gómez, 2018; Gómez, Chinchilla et al., 2020; Varmann et al., 2023; White, Newson, Verrelli, & Whitehouse, 2021; Wolfowicz et al., 2021). This led some to speculate that identity fusion is “identification on steroids.” Although this characterization of fusion does acknowledge that it out-predicts measures of identification, it overlooks qualitative differences in the mechanisms presumed to underlie the two constructs. For example, strongly fused persons are assumed to be motivated by a relatively stable social and personal self, and they devote themselves to the group as an end in itself rather than as a means of self-enhancement. In contrast, strongly identified (but weakly fused) persons are theoretically motivated by the social self only and devote themselves to the group as a means of enhancing themselves. The existence of these distinct mechanisms makes it inappropriate to assert that fusion is merely strong identification. By analogy, conceptualizing galloping as “fast walking” would fail to acknowledge the distinct mechanisms involved in the two forms of movement; when horses begin to gallop, they shift from using slow-twitch to fast-twitch muscle fibers, they become intermittently airborne, and their gait is qualitatively different.

Even if one were circumspect regarding the unique properties of identity fusion, measures of fusion have two pragmatic advantages. First, there is consensus that the verbal measure of fusion is the optimal way to assess fusion unless practical factors render it impossible [in which case one can employ a single pictorial item (Swann et al., 2009), or a computer-based variant of the pictorial item (the Dynamic Identity Fusion Index; DIFI, Jiménez et al., 2015)]. In contrast, there are dozens of measures of identification but no consensus regarding the optimal measure. Second, whereas identity fusion scales can be readily adapted to different targets,
investigators who have sought to measure identification with targets other than social groups have felt compelled to create unique scales for each target (e.g., identification with animals, Amiot, Sukhanova, & Bastian, 2020; brands, Tildesley & Coote, 2009; or leaders, Kark, Shamir, & Chen, 2003).

The foregoing considerations aside, the ultimate test of the utility of measures of fusion is whether they can predict the extreme pro-group behaviors that the construct was developed to explain. We consider this issue next.

1.3 Fusion and extreme pro-group behavior
The early fusion papers were the first to establish a link between fusion and willingness to defend the ingroup through violence (e.g., Gómez, Brooks et al., 2011; Swann et al., 2009). Fusion has subsequently been invoked to explain a large number of violent phenomena, including fighting by frontline Libyan revolutionaries (Whitehouse, McQuinn, Buhrmester, & Swann, 2014), the formation of Viking war bands (Raffield, Greenlow, Price, & Collard, 2015), Israeli retaliatory activities during the 2015 Palestinian Stabbing intifada (Fredman et al., 2017), Brazilian and British football hooliganism (Newson et al., 2018; Newson, 2021), inter-tribal warfare in Cameroon (Buhrmester, Zeitlyn, & Whitehouse, 2020), endorsement of honor violence including slapping or disowning one’s daughter (Ashokkumar & Swann, 2023), and hostilities initiated by several religious and secular terrorist organizations (Gómez et al., 2017; Gómez et al., 2020; Gómez, Bélanger et al., 2021; Gómez, Atran et al., 2022; Gómez, Chiclana et al., 2022; Gómez et al., 2023). These field studies have been supplemented by experimental evidence of the causal relationship between fusion and key outcome variables. For example, experimentally increasing fusion to an ingroup increased self-reported willingness to fight and die for that ingroup and other forms of self-sacrifice (Gómez et al., 2019). Simply put, the linkage between identity fusion and willingness to violently defend the ingroup may be the most robust finding in the fusion literature (for a recent meta-analysis, see Varmann et al., 2023).

Precisely why fusion is such a reliable predictor of pro-group behaviors such as violence and self-sacrifice has recently come into focus, with particular emphasis placed on relevant mediating mechanisms (for a review, see Gómez, Vázquez, Blanco, & Chinchilla, 2024). Another fundamental issue has arisen in a debate over whether altruism plays a role in compelling highly fused people to sacrifice themselves for the group. For instance, Whitehouse asserted (2018, p. 3; but later equivocated, p. 47) that the link between fusion and violence is altruistically motivated. Others countered
with an egoistic account of self-sacrifice (Gaertner, Heger, & Sedikides, 2018), citing evidence that participants primed to consider a group they are fused with still prefer personal over collective pronouns (Heger et al., 2023). This egoistic account dovetails with the simple idea that strongly fused people are motivated to defend the group because their representations of themselves are deeply intertwined with their representations of the group. That is, it makes just as much sense to ask why a strongly fused person is motivated to defend their group as it would be to ask why a strongly fused person is motivated to protect their arm or leg. This helps explain evidence that fusion predicts a willingness to fight for the group and to have the group fight for the self. Nevertheless, the fact that fusion does not predict a willingness to sacrifice the group for the self (Heger & Gaertner, 2018; Swann, Gómez et al., 2014) suggests that the personal self is not always prioritized over the group.

Despite the plethora of evidence linking fusion to efforts to protect the ingroup through violence, it is important to recognize that strongly fused persons are not inherently aggressive, nor does fusion invariably lead to violence (Kiper & Sosis, 2018; Xygalatas, 2018). To the contrary, strong fusion tends to predict endorsement of violence only when it involves defending the group against an imminent threat. Note, for example, that items in the commonly used fight and die scale [the outcome measure in Varmann et al. (2023) meta-analysis], were couched in terms of violent action to defend the group (e.g., “Hurting other people is acceptable if it means protecting my group”). Similarly, longitudinal research found that Palestinian attacks against Jews moderated endorsement of hostility against Palestinians by strongly fused Jews (Fredman et al., 2017). Moreover, linguistic indicators of fusion covaried with existential threat narratives in the manifestos of violent jihadists and gang members (Ebner, Kavanagh, & Whitehouse, 2022). And when fusion does predict violence, it is not indiscriminate violence. Fusion with Spain, for example, predicts fighting and dying to defend Spain but not Europe as a whole (Swann, Gómez, Dovidio, Hart, & Jetten, 2010), and fusion to a football team predicts violence toward rival fans but not other targets (e.g., family members; Newson et al., 2018).

The critical role of threat has been explicitly incorporated in several fusion-based theories, including the devoted actor model (Atran, 2016), the threat-plus-fusion model (Whitehouse, 2018), and the fusion-secure base hypothesis (Klein & Bastian, 2023). In each of these formulations, outgroup threat is understood to “flip the switch” that regulates the relationship of
fusion to intergroup violence. In fact, when the outgroup is familiar and non-threatening, strongly fused members are positively disposed toward its members and express positive sentiments toward them (Klein, Greenaway, & Bastian, 2024; Vázquez, Gómez, López-Rodríguez, & Swann, 2023).

Having said this, we acknowledge that, although threat may be the strongest moderator of the link between fusion and pro-group violence, it is not the only such moderator. Efforts to identify other moderators of the link between fusion and pro-group violence are underway. For example, the moral justifiability of violence appears to influence whether strongly fused people will act. The positive relationship between fusion and violent pro-group behavior interacts with moral concerns, with fusion predicting violence only among those who are convinced of its justifiability (Chinchilla, Vázquez, & Gómez, 2021) and who value parochial over universal values (Sheikh et al., 2014). Although under-researched, it is very likely that the culture of the group itself flavors the manifestation of pro-group behavior among strongly fused people. If we accept that certain groups (e.g., Quakers, pacifists) are more likely to avoid violent solutions compared to other groups (e.g., terrorist groups, militaries), then fusion to these groups should be weakly—or perhaps even negatively—associated with willingness to use physical violence in defense of the ingroup. This may help explain why the relationship between fusion to country and a willingness to violently defend it ranges from $r = 0.61$ in Poland to 0.32 in China (Swann, Buhrmester et al., 2014). Conceivably, some cultures are more comfortable with forthright endorsement of state-based violence (for a discussion of other potential moderators of the relationship between fusion and self-sacrifice for the group, see Gómez et al., 2024).

To be sure, violence is not the only way for strongly fused persons to advance their group’s agenda. In fact, fusion is also related to non-violent pro-group behavior (Chinchilla et al., 2021). The larger point here, however, is that measures of fusion do indeed predict the extreme pro-group behaviors that the creators of identity fusion theory hoped to explain. Given this, it is important to understand how people become fused to begin with.

2. The developmental trajectory of identity fusion

For newborns, one of their first orders of business is forming attachment relationships with their primary caregivers. These attachment
relationships likely provide the psychological architecture on which identity fusion is built. Once this architecture is in place, various factors can foster fusion, after which it will remain fairly stable. We cover each of these developments in turn.

2.1 Attachment theory
Healthy attachment relationships with the primary caregiver (e.g., Bowlby, 1980) are fusion-like in that they are marked by a powerful affinity for the caregiver. Nevertheless, early attachment relationships lack a critical ingredient of identity fusion, for infants do not possess representations of self as distinct individuals. This changes at about 18 months of age. Once rudimentary personal self-views are in place, children who are securely attached to their primary caregiver may become fused with that caregiver.

To be sure, some have proposed that attachment and fusion are orthogonal constructs with very different developmental trajectories and outcomes. Reese & Whitehouse (2021), for example, contend that fusion in early childhood arises from the recognition of shared biology (via phenotypic matching), while attachment develops via social interaction. These assumptions led them to speculate that an adopted child who rarely sees their biological mother would become attached to their adopted mother but fused to their biological mother. We are skeptical of this claim. On the contrary, we suspect that the quality of caregiver-child interactions contributes much more to fusion than the recognition of biological relations. More generally, we believe that the overlap between fusion and attachment is likely far greater than Reese & Whitehouse (2021) implied. For example, both forms of alignment involve an emotionally intense feeling of oneness with the other. In addition, both are highly consequential. In fact, not only do healthy attachment relationships ensure the baby’s comfort, nourishment, and support in the short term (e.g., Sorce & Emde, 1981), such relationships also increase the likelihood of psychological well-being in the long term (Bowlby, 1980; Marrero-Quevedo, Blanco-Hernández, & Hernández-Cabrera, 2018; Mikulincer & Shaver, 2012).

2.2 Self-verification
One reason why attachment relationships have lasting effects is that they encourage children to re-create the original attachment relationship in the future (Sroufe & Fleeson, 1986). Emerging representations of the self (“self-views”) may mediate this phenomenon (Bowlby, 1980). Self-verification theory (Swann, 1983; 2012) is relevant here. The theory assumes that by 18 months of age, children internalize the reactions of others into
corresponding self-views (Mead, 1934). Once formed, self-views begin to play a crucial role in children’s lives both epistemically (enabling them to make sense of their worlds) and pragmatically (enabling them to get along with others). The role of stable self-views in psychological life is so important that by the third grade, children take steps to maintain them by seeking evaluations that verify their negative and positive self-views (Cassidy, Aikins, & Chernoff, 2003; Cassidy, Ziv, Mehta, & Feeney, 2003).

People who receive a steady supply of self-verification from fellow group members will enjoy feelings of being known and understood. These feelings will strengthen their interpersonal (“relational”) ties to their fellow group members. In exceptional circumstances, feelings of being known and understood may even culminate in the close “familial” ties that people have toward family members. In fact, Gómez et al. (2024) have reported that strongly fused gang members often use family metaphors in describing their ties to fellow gang members (e.g., “Members of the gang are my family”, “We are a family, a brotherhood, we will never fail each other”, “With my gang I feel like I am with my family”).

Researchers have recently identified links direct between self-verification and identity fusion. Preliminary support for this linkage was provided by evidence that members of small groups reported stronger attachment to fellow group members when they received self-verifying evaluations from them (Swann, Jr., Milton, & Polzer, 2000). A recent series of eight studies provides more direct support for the link between self-verification and identity fusion (Gómez et al., 2024). Three correlational studies showed that increased self-verification from three distinct targets of fusion (a group, value, and leader) was associated with higher identity fusion and willingness to engage in extreme acts for each of these targets. Further, two experiments demonstrated that manipulating perceived verification increased participants’ fusion with the group, which in turn predicted willingness to fight and die on the group’s behalf. Finally, in interviews with incarcerated street and organized crime gang members, the researchers found that feelings of being known and understood by other gang members predicted their fusion with the gang; fusion, in turn, predicted willingness to engage in costly sacrifices for the gang. These studies therefore offered evidence of a causal connection between self-verification, identity fusion, and extreme behaviors for the target of fusion.

Another recent article provided further evidence of links between self-verification and identity fusion. Rousis, Martel, Bosson, & Swann (2023) focused on self-verification among incels (involuntary celibates who feel
that they have been wrongfully deprived of sex due to the shallow priorities of women). Two studies showed that incels who perceived that other incels verified them (i.e., saw them as they saw themselves) were more fused to the incel group. A third study revealed that self-verification was positively associated with identity fusion which, in turn, was positively associated with online harassment of women. Finally, in these studies and earlier research on self-verification, these effects emerged regardless of whether the self-view under scrutiny was positive or negative.  

Evidence that verification of negative self-views contributes to identity fusion is particularly significant here because it challenges the first of three general assumptions of social identity theory: “Individuals strive to maintain or enhance their self-esteem: They strive for a positive self-concept” (Tajfel & Turner, 1979, p. 40; see also Abrams & Hogg, 1988; Martiny & Rubin, 2016). This raises the possibility that self-verification motivates identity fusion and self-enhancement strivings motivate identification. Alternatively, self-verification may motivate both fusion and identification. In either case, self-verification is clearly implicated in identity fusion.

2.3 Shared essence

Of course, self-verification is merely one of several processes that can cultivate identity fusion. Researchers have also tested the hypothesis that

---

6 Although many studies indicate that people prefer positive evaluations (Alicke & Sedikides, 2009), these findings could reflect self-verification strivings to an unknown extent. That is, given that most participants (about 70%) have positive self-views (Diener & Diener, 1995), a tendency for most participants to prefer positive evaluations on average could reflect the disproportionate influence of the upper 70% who seek positive evaluations that are both self-verifying and self-enhancing. When researchers measure self-views, they routinely discover that those who score in the lower tercile prefer and seek negative evaluations—a clear refutation of self-enhancement theory (e.g., Bosson & Swann, 1999; Chen et al., 2004; Giesler et al., 1996; Swann & Read, 1981; Swann et al., 1989; Swann, Hixon, Stein-Seroussi, & Gilbert, 1990; Swann, Wenzlaff, & Tafarodi, 1992). In addition, when people’s personal identities are challenged by feedback that is more positive than they expect, they compensate by subsequently intensifying their efforts to obtain evaluations that verify their negative self-views (e.g., Swann & Hill, 1982; Swann & Read, 1981; Swann et al., 1992; Swann, et al., 2009). Evidence of self-verification strivings is especially striking in field studies. For example, when married people with negative identities find themselves with partners who see them more positively (or negatively) than they see themselves, they withdraw by divorcing (e.g., Cast & Burke, 2002) or becoming less intimate with their partner (e.g., Burke & Stets, 1999; De La Ronde & Swann, 1998; Murray, Holmes, Dolderman, & Griffin, 2000; Ritts & Stein, 1995; Schäfer, Wickrama, & Keith, 1996; Swann, De La Ronde, & Hixon, 1994). Similarly, when members of work groups receive verification of their negative or positive self-views, they are more likely to be creative within those groups (Ayduk, Gyurak, Akinola, & Mendes, 2013; Swann et al., 2000), and more inclined to remain in those groups (Wieserfeld, Swann, Brockner, & Bartel, 2007). Not surprisingly, such self-verification activities tend to stabilize people’s self-views (e.g., Swann & Hill, 1982).
perceiving that one shares characteristics with other group members is associated with identity fusion. One series of studies explored the links between shared genes, fusion, and willingness to sacrifice for one’s twin (Vázquez, Gómez, Ordoñana, Swann, & Whitehouse, 2017). Relative to dizygotic twins, monozygotic twins reported stronger fusion and elevated desire to have contact and share experiences with their twin (Study 1), to forgive and grant favors to their twin after being disappointed by them (Study 2) and make sacrifices for their twin (Study 3). Moreover, fusion with the twin mediated the impact of zygosity on these outcomes. Apparently, genetic relatedness fosters fusion which, in turn, predicts sharing, tolerance, and self-sacrificial behavior.

Other researchers demonstrated that the perception of psychological similarities (e.g., values) as well as biological similarities (e.g., genetic) foster fusion and associated behaviors (Swann, Buhrmester et al., 2014). They introduced a priming manipulation that encouraged participants to focus on either shared psychological or biological characteristics. For strongly fused participants, priming either type of shared characteristic increased their familial ties to other group members. These familial ties, in turn, mediated the influence of fusion on the endorsement of extreme sacrifices for the group.

### 2.4 Emotionally compelling experiences

Shared experiences may also trigger fusion with fellow group members. Several researchers have argued that traumatic shared experiences with ingroup members (e.g., painful rituals or the horrors of frontline combat) can trigger negative affect that promotes fusion to the associated group (Whitehouse & Lanman, 2014; Whitehouse et al., 2017). Whitehouse and colleagues buttressed these hypotheses with evidence of links between shared experiences and fusion in a host of special populations, from Libyan insurgents, Muslim fundamentalists in Indonesia to Brazilian football hooligans (Whitehouse, 2018).

Similarly, survivors of a natural disaster may become fused when they develop mutual understandings of the pain caused by the event. Experimental evidence indicates that recollections of the 2011 Christchurch increased fusion to the local community (Segal, Jong, & Halberstadt, 2018; see also Henríquez & Urrúa, 2023; Henríquez, Urzúa, & López-López, 2022). Nevertheless, if the citizens of Christchurch did not work together during the earthquake recovery, and instead engaged in looting and violence, the ordeal could well have undermined fusion.
However, shared experiences need not be traumatic to foster fusion. In fact, fusion may emerge during interactions that are cooperative. Behavioral synchrony is one example of this phenomenon: asking participants to perform a simple rhythmic movement in time with others led to more fusion than asking them to perform the movement out of time (Reddish, Tong, Jong, Lanman, & Whitehouse, 2016). Perceived emotional synchrony—which emerges in collective gatherings when shared attention and behavioral synchrony produce a shared emotional state—is also associated with fusion (Zumeta et al., 2020). Similarly, feelings of collective efficacy can also foster fusion when group members perceive that they are emotionally synchronized in a cooperative task, such as athletic competition (Zumeta, Oriol, Telletxea, Amutio, & Basabe, 2016). Complementary behaviors can promote fusion even if they are merely conceptually related. For example, Berscheid’s (1991) theory of emotional attachment contends that, in close relationships, developing, and pursuing complementary goals can promote fusion-like emotional attachments to partners. This is important, because such emotional attachments to lovers may last longer than those produced by relatively ephemeral phenomena such as sexual attraction.

Discussions of strong attachments to lovers is also significant because it suggests that it may be appropriate to broaden the scope of fusion research by exploring targets of fusion other than groups. As we discuss below, in recent years researchers have explored the hypothesis that people can fuse to targets of fusion that do not directly involve groups and familial ties. Such evidence of fusion with targets other than groups suggests that familial ties are a contributory but not a necessary component of fusion. Instead, it appears that the essential component of fusion is the synergistic relationship between the personal self and the target of fusion.

### 3. Targets of identity fusion

As noted above, identity fusion theory was developed to address aspects of group functioning that were downplayed by social identity theory (Tajfel & Turner, 1979). Not surprisingly, then, the early research on identity fusion emphasized group functioning in general and fusion with groups in particular. Nevertheless, Swann et al. (2012) explicitly acknowledged the possibility of fusion with other targets:

“Individuals may also fuse with abstractions. Feelings of oneness with God, for example, may compel people to devote their lives to the priesthood.
Similarly, people may develop a “calling”—a powerful urge to pursue some professional or recreational activity—that is propelled by feelings of fusion. Further, people may feel fused with brands or products. In each of these instances, strongly fused persons may experience feelings of oneness and connectedness with the target of their devotion even though there is no group associated with the target of their attachment”. (p. 448).

The notion that people can fuse with abstractions other than groups further distinguishes fusion theory from classic Social Identity Theory and its insistence on the sovereignty of the social self. Within our more expansive framework, fusion is no longer confined to the union of the personal self with the social self; instead, any abstraction can be the target of fusion. For example, people who are strongly fused with the value “liberty” perceive it to be core to their self-definition. As a result, they will be highly motivated to sacrifice in the service of liberty. Moreover, when thoughts of liberty are salient, strongly fused individuals do not lose sight of other aspects of the self. Finally, because such alignments only indirectly involve other individuals, it is inappropriate to implicate the social self or claim that fusion derives some of its power from relational ties to fellow group members. Instead, the synergistic relationship between the personal self and the target of fusion defines the fusion construct and is exclusively responsible for the extreme sacrifices it inspires.

As theory and research on identity fusion have escaped the shadow of social identity theory and its emphasis on group functioning, the scope of the phenomena under fusion theory’s purview has broadened accordingly. In expanding the definition of identity fusion to include targets other than groups, we note that the dictionary defines “identity” quite broadly as “the fact of being who or what a person or thing is.” This broad definition suggests that, in addition to groups, any entity that has distinctive and readily identifiable qualities can be a target of fusion. Below we briefly consider several such targets of fusion.

3.1 Fusion with fellow humans (groups, leaders, and other individuals)

Of all the targets of fusion that have been studied, groups have been the most popular by far (for a meta-analytic review, see Varmann et al., 2023). In part, this reflects the history of the fusion construct; fusion theory was developed to explain the actions of terrorists who were acting on behalf of their group. As such, the first several articles on identity fusion (Gómez, Brooks et al., 2011; Gómez, Morales et al., 2011; Swann et al., 2009;
Swann et al., 2012; Swann, Buhrmester et al., 2014; Swann, Gómez et al., 2014; Swann, Gómez, Dovidio et al. 2010; Swann, Gómez, Huici et al., 2010) focused on fusion with a specific group: one’s country.

Historical considerations aside, the enduring popularity of fusion with groups is likely due to the exceptional power of this target of fusion. In part, this reflects the fact that fusion with groups likely satisfies a potent need to belong (Baumeister & Leary, 1995) and to have personal significance (Gómez, Martínez et al., 2021; Kruglanski, Fernandez, Factor, & Szumowska, 2019). In addition, fusion with groups uniquely invokes alignments between the personal self and (a) a collection of people who form the group, (b) a leader of the group; and (c) values and ideas associated with the group. Furthermore, fusion with groups may be exceptionally powerful because group membership can be self-verifying (Rousis, Martel, Bosson, & Swann, 2023; Gómez et al., 2024). That is, insofar as group members are surrounded by fellow group members who share their values and perspectives, they are likely to receive self-verification from these group members. As noted earlier in this chapter, feeling self-verified buttresses feelings of fusion. Finally, being surrounded by like-minded others may polarize group members and motivate extreme behavior on behalf of the group (Melton & Motyl, 2018). Enacting extreme behaviors may, in turn, reinforce feelings of fusion through a self-perception process (Bem, 1972).

People can also fuse with specific individuals. For example, individuals have been shown to fuse to siblings (Vázquez, Gómez, Ordoñana, & Paredes, 2015; Vázquez, Gómez, Ordoñana et al., 2017; Vázquez, Ordoñana, Whitehouse, & Gómez, 2019) and to romantic partners (Joo & Park, 2017; Walsh & Neff, 2018). Leaders may be a particularly appealing target of fusion. Often, fusion with leaders is linked to fusion with the group that the leader represents. Nevertheless, if forced to choose between the leader or group, either the leader or group may prevail. For example, whereas most Trump supporters were more loyal to Trump than the Republican party (Kunst, Dovidio, & Thomsen, 2019), Clinton supporters reported higher fusion to the Democratic party than Clinton (Misch, Fergusson, & Dunham, 2018).

3.2 Fusion with values, ideologies, and causes

As mentioned above, when people become fused to abstractions such as values, ideologies, and causes, these alignments can complement their attachments to groups and individuals. For example, when compared to individuals imprisoned for crimes unrelated to terrorism, imprisoned terrorists displayed elevated levels of admiration toward Islamic activists, and admiration
was positively associated with fusion with religion (Gómez, Bélanger et al., 2021). Fusion, in turn, was positively associated with costly sacrifices for religion. Other research indicated that for imprisoned jihadists, both fusion with their most cherished group (i.e., Muslims) and fusion with their most cherished value (i.e., religion) were associated with high levels of willingness to self-sacrifice. A parallel pattern emerged among imprisoned gang members (Gómez, Atran, et al., 2022). Furthermore, regression analyses indicated that in a sample of Spaniards, fusion with a group (Ukraine), a value (freedom), and a leader (President Zelenskyy), independently predicted the will to fight for Ukraine or for freedom (Gómez et al., 2023).

Fusion with values may sometimes eclipse fusion with a group. For example, imprisoned jihadists expressed higher levels of fusion with religion than with Muslims, and they endorsed more willingness to sacrifice for their religion than for Muslims (Gómez, Atran, et al., 2022). Similarly, when members of ISIS and members of ISIS resistance groups were asked to choose between their most cherished group and their most cherished value, those who preferred the value expressed the strongest will to sacrifice (Gómez et al., 2017).

One advantage of recognizing that people fuse to causes, values, and ideologies is that such alignments may explain why people sometimes fuse with outgroups and outgroup leaders (e.g., Gómez et al., 2023; Kunst et al., 2019). For example, in the wake of Russia’s invasion of Ukraine, many Spaniards became strongly fused with Ukraine based on a common ideology—a commitment to democracy and the right to defend one’s country against foreign invaders (Gómez et al., 2023). In these and similar instances, acknowledging that people can fuse with shared ideologies avoids the untenable claim that an outgroup is an ingroup (Crimston & Hornsey, 2018).

3.3 Fusion with brands and creatures

Brands are often symbolic representations of people’s values, ideologies, and causes. People who have a strong affinity for brands may become fused with that brand and display corresponding behaviors such as brand loyalty (Krishna & Kim, 2020, 2021; Lin & Sung, 2013). Similarly, people may become fused with animals that symbolize certain values. For example, several years ago an American trophy hunter brutally killed “Cecil” the Lion. The incident turned into a viral media event in which thousands of people donated to a conservation organization called WildCRU. Strong fusion with Cecil predicted whether people donated to WildCRU nine months later (Buhrmester, Burnham et al., 2018; see also Amiot & Bastian, 2017). Events
like this demonstrate how people, animals, brands, or events can come to represent certain values or beliefs. From this vantage point, fusion with brands and creatures could be viewed as a subset of the larger class of fusion with values, ideologies, and causes.

Of course, the properties of the target of fusion itself may influence people’s experiences with the target and the levels of fusion that result. Although this topic is still ripe for investigation, we can point to several characteristics of fusion targets that will likely influence the levels of fusion people develop toward the target.

4. Properties of fusion targets that may entice potential fusers

Some targets of fusion have properties that make them particularly attractive to potential fusers. Although there are many target properties that could encourage fusion, we focus on three particularly important ones here: clarity, consistency, and congruence.

Targets of fusion that are high in **clarity**—readily identifiable and well-defined—will be more likely to elicit feelings of fusion than targets that are lower in clarity—fuzzier or poorly demarcated. For example, a value that is amorphous (e.g., happiness) or a leader or brand that has no defining features will be associated with weaker fusion than values, leaders, or brands that are well defined (e.g., patriotism). Similarly, when the target of fusion is a group, one determinant of clarity is the group’s entitativity—the degree to which a group has unity, coherence, and internal organization (Blanchard, Caudill, & Walker, 2020; Campbell, 1958; Lickel et al., 2000). Whether members have direct contact with other group members may also influence the clarity of a group. That is, direct contact tends to provide group members with relatively rich, concrete, and elaborate information that gives rise to robust “local fusion.” In contrast, indirect contact tends to provide relatively impoverished, abstract, and vague information about fellow group members resulting in “extended fusion”. Direct contact may therefore foster higher clarity (for a discussion of local and extended fusion, see Swann et al., 2012).

A second property of targets that should influence fusion is its **consistency**—the degree to which the target faithfully provides whatever it was that attracted the individual in the first place. Consider a group or brand that is associated with a specific set of values. The group should inspire
fusion among loyalists only insofar as they continue to profess the values that are emblematic of the group. For example, imprisoned jihadists revealed that they joined the jihad due to its purported values but later became disillusioned when they observed the actual values of group members. They left the group as a result (Gómez, Chiclana, et al., 2022). Such instances in which people become fused to certain values or ideologies contrast sharply with membership in groups that attract members by offering them a safe haven or “secure base” (Bowlby, 1982; Klein & Bastian, 2023). Friendship groups, for example, attract members through the promise of mutual support and an implicit promise to “have one another’s back”. Here again, a key to success in commanding the allegiance of group members is unwavering consistency. Over time, such consistency will promote strong relational ties and trust in the target of fusion.

A third property that influences fusion is the congruence of the target with other aspects of the self, including various social identities or values. For example, it is more tenable to fuse with a religious group if one is already fused to the relevant god, or to fuse to a social movement if one is already fused to its ideological principles. In these cases where endorsement of a value, group, or other target is aligned with other pre-existing aspects of the personal or social self (i.e., is self-verifying), fusion becomes more likely (see, for example, Atran, 2016). In fact, people often exhibit “fusion clusters” involving multiple congruent targets, such as a Jihadist who reports feeling fused to Islam, fellow terrorists, the prophet, and God, with each fusion target reinforcing the strength of the others. For example, research examining Spanish support for Ukraine in its conflict with Russia found that fusion to the Ukrainian president Zelensky and fusion to freedom were significantly correlated (Gómez et al., 2023). Apparently, fusion with two related fusion targets can be mutually reinforcing.

It is important to distinguish our fusion-cluster proposal from the claim that identity fusion is a personality trait. That is, research has shown that identity fusion is unrelated to any of the Big Five traits, $rs < 0.07, ps > 0.10$ (Gómez & Vázquez, 2015). Moreover, when people fuse to clusters of congruent targets (e.g., a value, a group, and leader), they do so out of a commitment to one overarching idea. In contrast, if fusion is a personality trait, one would expect “fusers” to align themselves with a variety of potentially unrelated targets at the same level of abstraction. Empirical research suggests otherwise (Gómez, Brooks et al., 2011). That is, the tendency to fuse with one group (e.g., one’s religion) was uncorrelated with the tendency to fuse with other groups (e.g., one’s country; all $rs < 0.11$, ns). Of course, if
there was considerable overlap between religion and country, as in a theocracy, we may expect fusion to both to be correlated, but this is another example of a fusion-cluster centered around a central concept. The target-specificity of fusion means that it resembles an attitude rather than a trait.

Having considered the independent contributions of properties of targets to identity fusion, we note that these properties may interact with one another. For instance, the consistency property could compel a strongly fused person to abandon the group in the wake of a change in group policy or beliefs (e.g., the Protestant reformation, the split of the IRA over Northern Ireland). This desire for consistency works together with the congruence principle, in which defusion from the changed group is spurred by its newfound lack of alignment with a value deemed important to the self. Naturally, the degree to which a desire for consistency or alignment is relevant is dictated by the property’s clarity. If a group or value is “fuzzy”, then an individual will not care whether it is congruent with other aspects of the self.

Whereas clarity and consistency will influence how strongly people fuse to most targets, some properties of targets will apply only to a limited class of targets. For example, the verisimilitude of specific ideas (e.g., Jesus was the Son of God) underlying certain ideologies (e.g., Christianity) may uniquely influence the strength of fusion with such ideologies. Similarly, the physical attractiveness or values of a love interest may determine whether we become fused with that person. In these and related instances, the properties that make a target alluring will determine the strength of fusion with that target.

One upshot of this discussion is that the fuzziness or inconsistency of targets will diminish fusion. This raises the more general issue regarding the stability of identity fusion over time.

5. Temporal stability of identity fusion

Identity fusion theorists have generally assumed that fusion tends to remain stable. In fact, the original statement of the theory (Swann et al., 2012) asserted that fusion was “irrevocable”. In part, this hypothesis was based on theoretical considerations. Given that identity fusion involves synergies between social identities and relatively stable personal identities, fusion should be similarly stable. Moreover, insofar as fusion is associated with porous borders between personal and social identities, renouncing
one’s feelings of fusion would be perilously close to renouncing the personal self. Finally, repudiating fusion would require substantial re-structuring of the relevant self-views, close relationships and even behaviors associated with the target of fusion—a task that could trigger emotional vulnerability if one enacted extreme behavior for the target of fusion. Independent of these conceptual considerations, the assumption that identity fusion was irrevocable was based on evidence. For example, among strongly fused persons, the rank orderings of fusion scores remained stable over months ($r = 0.61, 0.59, 0.62, 0.54$, Swann et al., 2014, p. 447).

The foregoing conceptual and empirical reasons for believing in the irrevocability hypothesis notwithstanding, recent evidence has suggested that it is too strong. When researchers examined the reactions of Spanish participants to three negative historic events (a corruption scandal involving the Royal Family and two separatist efforts by a prosperous region of Spain), they discovered that average fusion scores declined following these events (Vázquez, Gómez, & Swann, 2017). Closer inspection, however, revealed that these declines were limited to sentiments toward the group category (“collective ties”); the negative events did not tarnish sentiments toward individual group members (“relational ties”), nor did these events diminish willingness to fight and die for Spain.

The results of experimental studies added additional nuance to the stability question. Gómez et al. (2019) reported that experimentally compromising either collective ties (i.e., sentiments toward the group as a whole) or relational ties (i.e., sentiments toward individual group members) lowered state identity fusion (fusion at the moment) but not trait fusion (the relatively stable feelings of fusion assessed with no temporal instructions). These findings suggest that state fusion may vary while trait fusion remains stable.

A recent study added an additional wrinkle to the stability-of-fusion issue. The investigators (Gómez et al., 2024) reported that they successfully changed trait fusion by increasing perceived self-verification. This appears to be the first evidence that trait fusion can be experimentally increased in a longitudinal design (cf. Gómez, Bélanger et al., 2021). Although this is an undeniably important finding, we should add three caveats. First, although increasing perceived verification increased absolute fusion (i.e., assessed by comparing average fusion scores of the experimental and control groups), it had minimal impact on relative fusion (assessed by correlating scores before and after the treatment), $r = 0.39$ and $0.43$, baseline vs. experimental conditions, respectively. These results indicate that rank orderings of fusion remain stable over time, even as absolute fusion values vary in response to...
situational influences (cf. Mathieu & Gosling, 2012). Second, although the researchers succeeded in increasing trait fusion, they did not attempt to decrease fusion; experimentally decreasing trait fusion may be more difficult due to processes such as loss aversion (Kahneman, Knetsch, & Thaler, 1991). Third, given the earlier evidence that fusion tends to remain stable over time (Swann et al., 2014), we suspect that there are many factors that stabilize fusion in naturally occurring settings. For example, the highly agentic nature of strongly fused people may motivate them to actively resolve any dissatisfaction with the group, thereby precluding defusion. To the extent that this is true, any changes produced in laboratory experiments will likely be reversed in the days and weeks following the experimental intervention.

Considering the foregoing caveats, we suspect that Gómez et al. (2024) success in changing fusion may say more about the potency of their perceived self-verification manipulation than the fragility of identity fusion. Conceivably, their manipulation was especially effective because it targeted the heart of identity fusion: the personal self. In any event, the results of the Gómez et al. (under review) study, together with earlier evidence of success in changing fusion or its components (Gómez et al., 2019; Vázquez, Gómez, & Swann, 2017), indicate that Swann et al. (2012) claim that identity fusion is irrevocable was too strong. Although fusion may be resilient, it does change. This raises a further question: if fusion can be altered, how might one defuse individuals who are strongly fused with belligerent groups such as terrorists, gangs, or cults?

5.1 Defusion

There are many contexts in which defusion—that is, the reversal of the fusion process—is desirable. A case in point is violent extremism, with defusion among terrorists certainly commanding the most attention. Nevertheless, defusion may also represent an effective method of combatting less serious forms of extremism (e.g., football hooliganism, gangs), and may also prove useful in more benign contexts. For example, just as politicians can be motivated to defuse supporters of other candidates or parties (consider the role of ‘opposition research’ and negative advertising), marketers may seek to degrade loyalty to competing brands.

The following discussion should be read with the knowledge that defusion—both in empirical research and conceptualization—is still in infancy, at least partly due to the inherent difficulties in working with extremist populations (see Gómez, Vázquez, Chinchilla et al., 2023). While
future research will undoubtedly shine further light on this burgeoning area of fusion theory, for now we are limited to offering a few speculations.

We begin by noting that, because the state of fusion is so far reaching, defusion may be thwarted by an inability to sufficiently compartmentalize experiences (Hugenberg & Bodenhausen, 2004). The resulting single-mindedness of strongly fused individuals could impair their ability to display the role flexibility needed to maintain components of their life that are unrelated to the target of fusion. This problem will be magnified when the target of fusion is deviant in nature. For example, former gang members who commit crimes as part of initiation rituals may become social pariahs who cannot win acceptance within settings that welcomed them prior to their fusion with the gang. Knowledge of such possibilities may encourage strongly fused individuals to contemplate defusion with trepidation. Should they move forward with defusion, they will face substantial restructuring of the personal self, including the very meaning of their actions and their relation to others.

With these caveats in mind, we propose that there are three broad strategies that could produce defusion: reduction, removal, and replacement. Perhaps the most obvious defusion strategy is through the reduction of a person’s ties to the fusion target. If fusion is to a group, reduction could be accomplished by degrading ties to fellow group members. For example, an early discussion of defusion (Fredman et al., 2015) cites the case of Mosab Hassan Yousef, a Hamas fighter who became disillusioned after witnessing that leadership was responsible for hurting fellow ingroup members. This observation led to his eventual defusion and exit from the group. This mirrors the results of qualitative research which has similarly highlighted experiences of betrayal or mistreatment as motivating terrorists to abandon the group (Gómez, Chiclana et al., 2022; Reinares, 2011). In fact, the successful defusion manipulation used by Gómez et al. (2019) involved asking participants to recall an experience that made them question their relationship with the ingroup. This finding points to the potential viability of reduction strategies as a deradicalization tool.

Nonetheless, a reduction approach should be considered with some caution. At least some research points to a possible ‘blowback effect’, in which exposure to counter-narratives (e.g., attacking the ingroup’s social, political, or religious legitimacy) actually increased support for the extremist groups among those with greatest radicalization risk (Bélanger et al., 2020). Of course, one difference between the Gómez et al. (2019) and Bélanger et al. (2020) research is the source of the negative information; while the
former asked participants to self-generate instances of betrayal, the latter delivered the instances via external means (e.g., an Imam, ISIS defector). Given this, it seems that degrading perceptions of the fusion target can indeed cause defusion, but this strategy is risky when external agents are the source of such degradation.

*Removal* from the group may also cause defusion. This commonly occurs in the context of prison sentences for terrorists or other violent extremists but could include any instance of prolonged separation from the group. Previous terrorist literature has identified prison as offering inmates ‘space to think’ and reassess their commitment to their group and ideology (Ferguson & McAuley, 2020; Ferguson, Burgess, & Hollywood, 2015), particularly in relation to their familial commitments (Reinares, 2011). That said, the effect of incarceration on defusion has been somewhat mixed, with research among former Jihadist women suggesting prison time decreases fusion to Muslims and Islam (Gómez, Chiclana et al., 2022), while research on a male Jihadist sample seemingly suggested that prison had no effect on fusion (Gómez, Atran et al., 2022). Nevertheless, it is worth highlighting the extremely small samples in these studies (reflecting the inherent difficulties associated with obtaining extremist samples). Moreover, prison experiences can vary greatly. For example, although solitary confinement may be conducive to defusion from a group, the time it affords for reflection could conceivably intensify fusion with values. In fact, the perpetrator of the 2011 Norway terrorist attacks (Anders Breivik) recently claimed that the time he spent in solitary confinement has left him more radicalized than he had been (BBC News, 2017).

Removal may also be effective in instances in which the ingroup ceases to exist, with the success of defusion likely related to depend on the destruction of relational ties between former members of the ingroup. Researchers studied this phenomenon in the context of the Tamil Tigers, a Sri Lankan terrorist group that was eliminated in May 2009. When the smoke had cleared, the government required some former fighters to pass through rehabilitation centers prior to being released into general society. The findings indicated that those who completed the rehabilitation program were less extreme than their counterparts, but that those who maintained connections with their fellow former terrorists harbored more extreme views than those who did not (Webber et al., 2018). These findings suggest that relational ties should be severed if the removal strategy is to be viable.

Finally, *replacement* is perhaps the most attractive method of achieving defusion. This strategy involves the substitution of a problematic fusion target with a less problematic one. Yousef, the former Hamas fighter described
previously, replaced Hamas with a Christian bible group (Fredman et al., 2015). Similarly, many former terrorists report cutting ties with their comrades and violent ideology as they started a family of their own (Ferguson et al., 2015; Reinares, 2011). Through immersing themselves in a new support structure, highly fused people ‘switch worlds’ and alter their identities accordingly. There are obvious advantages to this approach, chief among them the retention of a fusion target which can support them through their deradicalization process. A striking example of this strategy is the ‘Twinning Project’, in which incarnated individuals were rehabilitated through replacing fusion to gangs or prisoners with fusion to a local football club (Newson & Whitehouse, 2020). While clearly a viable strategy in principle, enticing people to abandon worlds that have come to define them is a hugely significant undertaking.

Given the challenges associated with developing and implementing successful defusion strategies, it is worth asking whether defusion is the most important, or even a necessary, component of a deradicalization strategy. Because strongly fused people rely on their ingroup for support (Klein & Bastian, 2023), as well as their very sense of self, instigating defusion will certainly take a psychological toll on the individual. There are also complexities involved in determining the target of defusion. While defusion from an extremist group like ISIS is surely a worthwhile goal, it is less clear that defusion from ‘Islam’ or ‘all Muslims’ is desirable, useful, or ethical (see Gómez, Atran et al., 2022; Gómez, Chiclana et al., 2022). These caveats suggest reduction and removal strategies should be employed cautiously, and ideally in the context of a broader replacement strategy.

There is also reason to believe that these strategies may be more effective when employed against groups than ideologies or values. For one, it is not clear how one can be separated from an idea, or that an idea can be ‘destroyed’, suggesting that the removal strategy is a non-starter. Likewise, it is difficult to conceive of how one can be actively ‘betrayed’ by an idea—a core component of the reduction strategy—although the disillusionment associated with an attractive idea’s lack of real-world efficacy (e.g., communism) may have a similar effect. There is some limited evidence (Gómez, Atran et al., 2022) that once incarcerated, Jihadists remain more fused to their religion (i.e., Islam) than to fellow group members (i.e., Muslims). If this proves to be a robust effect, it will suggest that defusing from ideas is more difficult than defusing from groups.

In any event, alternative strategies not focused on defusion per se, but still informed by fusion theory, could prove to be effective methods of deradicalization. There is now a wealth of evidence that strongly fused
people are motivated to violence by perceptions of outgroup threat (see the *Fusion and extreme pro-group behavior* section above). This suggests that reducing these threat perceptions would likewise reduce the perceived need for violence (Vázquez, Gómez et al., 2023). Similarly, encouraging the adoption of prosocial norms, increasing the salience of peaceful alternatives to violence, or increasing the cost of violence could all weaken the link between fusion and violence. These outcomes may in fact follow from the practical strategies outlined above, with some former terrorists noting that their openness to non-violent strategies increased while in prison (Ferguson & McAuley, 2020; Ferguson et al., 2015). Detailing the pros, cons, and practical challenges of these methods lies outside the scope of the present chapter, but we simply raise them here to highlight that fusion theory may inform deradicalization strategies even in the absence of defusion. In fact, an effective strategy may be to find ways to maintain fusion while changing violent behaviors enacted for the target of fusion. Here, the goal would be to produce disengagement rather than deradicalization, which has proven to be optimal in some instances (e.g., Horgan, 2008).

In summary, research conducted since the original statement of fusion theory has led to a more nuanced understanding of the stability of the construct. It is now clear that although fusion is not irrevocable, it is quite resilient. This feature of identity fusion, in combination with its predictive power, has led other researchers to contemplate its implications for understanding the extreme behaviors of contemporary terrorists, as well as the war-making activities of humans throughout our evolutionary history.

### 6. Related formulations: Devoted actors and an evolutionary approach to the causes of fusion

Identity fusion theory has inspired two closely related formulations. First, the devoted actor approach pairs identity fusion with sacred values in an effort to predict extreme behavior with greater fidelity. Second, Whitehouse’s variation on identity fusion theory adopts an evolutionary perspective on the causes of fusion. We consider each of these formulations in turn.

#### 6.1 Sacred values, fusion with values, and the devoted actor formulation

Tetlock (2003) proposed that values are sacred when they are associated with an absolutist prohibition on transgressing the value, a prohibition that
motivates outrage at transgressions even if they are merely hypothetical. Following Tetlock and others (e.g., Rappaport, 1971), Atran (2016) combined the sacred values construct with identity fusion in attempting to predict terrorism. He coined the term “devoted actors” to refer to people who are both aligned with values that they deem “sacred” and fused with a group. The putative virtue of his approach is its capacity to leverage the predictive power of alignment to both values and groups (i.e., sacred values and identity fusion) to explain violent extremism. Consistent with this approach, considerable evidence indicates that devoted actors display extraordinarily high levels of costly sacrifices for the value and group (Atran, Sheikh, & Gomez, 2014; Gómez et al., 2017; Gómez et al., 2023; Sheikh, Gómez, & Atran, 2016).

Empirical support for the devoted actor approach notwithstanding, it suffers from two weaknesses, both stemming from the sacred values construct. One problem is that sacred values have typically been defined in terms of their consequences—such as an unwillingness to negotiate; insensitivity to material costs, risks, rewards, consequences; blindness to exit strategies in conflict; linkage to distinct neural signatures in the brain, and so on. This circularity means that we do not know what makes a value sacred (or, by implication, what makes devoted actors devoted). We only know that sacred values are associated with certain consequences.

The other weakness of the devoted actor approach is empirical. Specifically, when researchers developed an alternative, fusion-based measure of alignment with a value, they discovered that it consistently out-predicted the traditional measure of sacred values. That is, in six studies, researchers discovered that when a traditional measure of sacred values and a measure of fusion with a value competed in a foot-race in which the outcome measure was willingness to fight and die for a target cause (i.e., pro-gun or anti-gun, pro-choice, or pro-life), the fusion-based measure consistently out-predicted the measure of sacred values (Martel et al., 2021). This pattern occurred regardless of the target cause (gun or abortion rights), the participant’s position on the cause, or nationality of participants (American vs Spanish). A subsequent study lent further support for this conclusion (Gómez et al., 2023).

At least two possibilities may explain the relatively anemic performance of the measure of sacred values. One possibility is suggested by evidence that measures of fusion with a value are associated with the personal self, but measures of sacred values are not. That is, researchers attempting to predict willingness to fight and die for a cause found that a threat to the
personal self amplified the effects of fusion with a value but not a measure of sacred values (Martel et al., 2021). Alternatively, or in addition, measures of sacred values may be hamstrung by the fact that such measures assess moral prohibitions against “selling out” (i.e., abdicating one’s values for material gain). Given that people are poor estimators of their ability to resist social pressures (e.g., Milgram, 1963), answers to questions about selling out may be inherently unreliable.

In short, the sentiment that motivated the development of the Devoted Actor approach rings true. Measuring alignment with values as well as groups does seem to increase predictive validity over measuring alignment with groups only. Nevertheless, given the importance of this insight, it is crucial to utilize measures of alignment with values that are conceptually crisp and empirically sound. Measures of fusion with values fill both priorities. That is, they avoid the circularity inherent in sacred value measures, as they are defined independently of the outcome they predict (i.e., as the union of the personal self and the value) rather than exclusively in terms of the outcome. In addition, measures developed to tap fusion with values predict relevant outcomes with greater fidelity than traditional measures of sacred values (Martel et al., 2021; see also Gómez et al., 2023). As a measure of alignment with values, then, it appears that the fusion–with–values measure is superior on both conceptual and empirical grounds.

6.2 Whitehouse’s evolutionary approach to identity fusion

Anthropologist Harvey Whitehouse (2018) has developed an intriguing account of the causes of identity fusion that emphasizes its adaptive utility. He contends that in humanity’s evolutionary past, fusion functioned as a social glue that gave warbands an edge. He also points to a pair of proximal causes of fusion. Specifically, he suggests that fusion grows out of intense collective experiences (e.g., dysphoric rituals, combat experiences) and perceptions of shared biology. Whitehouse marshals impressive support for his ideas. For example, multiple studies have highlighted the link between fusion and shared dysphoria (e.g., Jong, Whitehouse, Kavanagh, & Lane, 2015; Whitehouse et al., 2017), and perceptions of family (Swann, Buhrmester et al., 2014). Nevertheless, we contend that Whitehouse’s preoccupation with the function of identity fusion has led him to devise a relatively narrow account of its nature, causes, and consequences. Our revised theory is designed to address these shortcomings in several ways.

Regarding the nature of fusion, Whitehouse emphasizes a single target of fusion—groups. As noted above, our revised theory examines additional
targets of fusion such as values, ideologies, other individuals, and brands. Moreover, our revised theory introduces new phenomena. For example, our proposed *congruence* property of fusion suggests that multiple fusion targets can be mutually-reinforcing.

Regarding the causes of fusion, Whitehouse’s evolutionary framework leads him to focus on distal causes that are untestable (given current technologies) and proximal causes that are narrow and limited. For example, the proximal causes he emphasizes are shared negative experiences and autobiographical memories. In emphasizing these causes, he underestimates the complexity of the pathways to fusion. For example, when shared experiences are non-verifying, or when they trigger verification of a quality that diminishes group functioning (such as selfishness), they may decrease fusion. The claim that shared dysphoric experiences can trigger fusion is also overly simplistic. Although this may be true under some circumstances, considerable evidence indicates that dysphoric experiences can also lead to isolation and block the formation of a shared identity. For example, shared traumatic experiences—such as combat exposure—are associated with posttraumatic stress disorder symptomology, moral injury, and suicide ideation (Morissette, DeBeer, Kimbrel, Meyer, & Gulliver, 2018; Russell, Ferguson, Lucas, & Mussap, 2023; Schwartz, Halperin, & Levi-Belz, 2022). Moreover, a recent longitudinal study on hazing practices in university settings indicated that positive rather than negative experiences predicted ingroup bonding (Cimino & Thomas, 2022).

To be clear, we agree with Whitehouse that dysphoric experiences can act as a social glue under certain circumstances, but stress that the moderators needed to produce such an outcome are currently undertheorized.

Whitehouse’s exclusive focus on emotionally intense experiences, particularly dysphoric ones, and biological recognition as the root causes of fusion is also undercut by evidence of alternative paths. For example, researchers have found that the perception of shared values can contribute to fusion (Swann, Buhrmester et al., 2014). Furthermore, there is emerging evidence that fusion can also grow out of euphoric experiences, including admiration (Gómez, Bélanger et al., 2021) and awe (Song et al., under review), raising doubts that dysphoric experiences have cornered the market on the causes of fusion.

Whitehouse’s focus on dysphoric experiences also overlooks the fact that fusion effects continue to surface in a modern world in which many of the painful rituals he features have been banished. For example, contemporary adolescents routinely display high levels of fusion with friends.
seemingly in the absence of shared dysphoric experiences. Whitehouse and colleagues have attempted to address these concerns, contending that modern events such as raves may effectively function as modern-day fusion-building rituals (Newson, Khurana, Cazorla, & van Mulukom, 2021). The fact that raves are euphoric rather than dysphoric aside, we suspect that a single, powerful event may not be a necessary condition for identity fusion to emerge. Surely many of the strongest bonds people form with one another arise gradually as they co-experience life over months, years, or decades.

We also suspect that at least some of the phenomena Whitehouse observes could be explained by a mechanism that he overlooks: self-verification. That is, sharing experiences with other group members contributes to self-verification (Swann et al., 2000) and perceived verification causes fusion (Gómez et al., under review). This may help explain the phenomena that Whitehouse features, as periods of adversity (think combat zones) will likely encourage peers to provide one another with self-verification. In short, although we agree that dysphoric events may foster fusion, we doubt that they should be given primacy over alternative, less dramatic sources of fusion, particularly when the fusion target is not a group.

The narrowness of Whitehouse’s (2018) focus extends to his understanding of the cognitive mechanisms that foster fusion. He proposes that two ingredients are necessary for fusion to occur. The first ingredient—self-reflection—has received mixed support, with some studies lending support (Muzzulini, van Mulukom, Kapitány, & Whitehouse, 2021) and others not (Kapitány, Kavanagh, Buhrmester, Newson, & Whitehouse, 2019). The second ingredient—autobiographical “flashbulb” memories—are surely not the sole memorial cause of fusion. Consider semantic memory. The semantic knowledge that other group members share the same values (Swann, Buhrmester et al., 2014) or a specific sacred value (Sheikh et al., 2016) strengthens the links between fusion and willingness to self-sacrifice. More generally, a flashbulb memory of an emotionally significant event may be irrelevant when it comes to fusion with abstract values such as democracy or freedom. Instead, enculturation, reflection, or other semantic influences may be central.

One final thought: although dying for a group may be an important consequence of fusion, it is merely one of many. Whitehouse’s rather narrow focus on violence, terrorism, and similar extreme behaviors overlooks the fact that identity fusion may manifest itself in ways that do not involve aggression against others (e.g., pro-social behavior and doggedness
in pursuing pro-social goals). This is less a critique than an acknowledgment of a difference in goals. While Whitehouse (2018) proposes an evolution-tinged theory of extreme self-sacrifice, we propose a general theory of identity fusion that is intended to offer a relatively comprehensive account of its nature, causes, and consequences.

7. The role of fusion in understanding key contemporary issues

Identity fusion theory was originally developed to understand violent terrorism. Since then, it has been leveraged to explain humanity’s worst behaviors, from the schemes of corrupt politicians to the misbehavior of hooligans, to the mass murders of terrorists. We will now briefly explore the contributions of fusion theory to understanding several key contemporary issues: motivated reasoning in the political sphere, terrorism, football hooliganism, and the military. In addition, we consider some very recent work that focuses on some of the more positive consequences of fusion.

7.1 Motivated reasoning in the political sphere

While the relationship between fusion and intergroup violence has captured the limelight of many fusion researchers, fusion can also influence political behavior. One of the early studies in this tradition tested the prediction that identity fusion would be associated with even-handed responses to political outcomes, but that identification would be associated with self-enhancing responses (Buhrmester et al., 2012). The researchers focused on reactions to the 2008 presidential elections in the USA and Spain. They discovered that whereas respondents who were “fused” with their political party reveled in victory and mourned defeat, those who were identified with their party reveled in victory but distanced themselves from defeat. These findings supported fusion theory’s prediction that fusion would be linked to internalization of negative as well as positive outcomes and social identity theory’s prediction that identified participants would internalize victory only.

More recent research has tested the notion that fusion predicts motivated reasoning; that is, a willingness to distort the truth to benefit the ingroup and its interests. This is consistent with the tendency of strongly fused people to maximize ingroup advantage at the outgroup’s expense (e.g., Buhrmester, Newson, Vázquez, Hattori, & Whitehouse, 2018), and
could be viewed as another method of furthering the group’s agenda. However, we argue that fusion’s impact is more insidious than that. Strong fusion to a group, leader, or value has a profound effect on the way we interpret information, and tilts us towards a partisan truth.

At a fundamental level, fusion is related to attitude strength. In the case of fusion to specific ideas and values, this relationship is straight-forward: attitudes will be strong insofar as they are aligned with the personal self. For example, fusion to a cause is strongly correlated with attitude strength as well as moral convictions (Ashokkumar et al., 2020), with fusion representing a particularly strong predictor of willingness to fight and die for the relevant cause (Martel et al., 2021). Moreover, evidence that threatening the validity of the personal self strengthened the relationship between fusion and willingness to sacrifice for the group (Martel et al., 2021) suggests that the personal self is instrumental in motivating such compensatory behavior. As people become more certain in this alignment with the personal self, their endorsement of pro-group behavior increases (Paredes, Brñol, Petty, & Gómez, 2021). Here, fusion offers insight into the identity-related mechanisms that drive attitudes and moral convictions and suggests that moral convictions will be strong insofar as they are aligned with the personal self.

The effect of fusion on attitudes also becomes particularly relevant in a group context. For example, fusion may motivate the adoption of partisan versions of the truth. This follows the congruence principle: the more a person aligns a specific social self with the personal self, the more they will be inclined to adopt and promote a version of the truth endorsed by the ingroup. For example, fusion to country or religion predicts willingness to deny the group’s wrongdoing (Besta, Gómez, & Vázquez, 2014). Linguistic analysis also suggests that using fewer words indicating cognitive processing or questioning (e.g., think, unsure) is one of strongest linguistic predictors of fusion, with an index incorporating this measure predicting how long people remained active in an online political group (Ashokkumar & Pennebaker, 2022). Apparently, in this instance, fusion suppressed willingness to question the group narrative. Interestingly, the effect of fusion seemingly overwrites adherence to more fundamental morals. Despite evidence that conservatives generally care more about values that bind group members (i.e., loyalty, authority, and purity) than liberals, strongly fused liberals and conservatives display equally strong propensities towards ingroup binding (Talaifar & Swann, 2018). Together, this research shows that, in the case of strongly fused individuals, a devotion to the group and their version of the truth trumps objective truth and other moral concerns.
The adoption of a partisan truth among strongly fused people can also manifest itself in censorship and, ultimately, political violence. For example, fusion to values, such as gun or abortion rights, have been found to positively predict a willingness to censor online comments that opposed [versus supported] the participant’s position (Ashokkumar et al., 2020). Likewise, strongly fused people are willing to use unethical methods, such as evidence tampering, to hide damaging information about the ingroup (Ashokkumar, Galaif, & Swann, 2019) and, at its most extreme, persecute (Kunst et al., 2019) and commit acts of violence (Mason, Winter, Schmeer, & Berrington, 2021) against political opponents. Although this research fits into a broader context of fusion-based violence and pro-group behavior, the motivated reasoning and lack of a shared intergroup understanding of truth associated with political violence may make it especially difficult to combat.

An interesting example of fusion being strategically used for political ends can be seen in the case of President Zelenskyy of Ukraine. In the ongoing war against Russia, Zelenskyy has attempted to recruit supporters not only by appealing to their sympathy for Ukraine but also their support for freedom and democracy (Gómez et al., 2023). This broad appeal has proven extremely effective as of this writing. Alas, fusion may also be leveraged for nefarious political ends. Take former US president Donald Trump, for example. Fusion with Trump has been shown to predict a willingness among Republicans to violently persecute Muslims, immigrants, Iranians, and political opponents, particularly if induced to feel that the United States is under threat (e.g., a belief that Americans are losing their jobs, suffering from a drug epidemic; Kunst et al., 2019). After leaving office and being indicted for multiple crimes, Trump warned his supporters: “They’re not coming after me. They’re coming after you”. Such admonitions illustrate how threat-laden language can be directed at strongly fused followers to motivate allegiance and even dangerous partisan actions.

7.2 Fusion and terrorism

One such dangerous partisan action involves acts of international and domestic terrorism. Although identity fusion was originally developed to explain terrorism—specifically suicide bombings—it was only recently that fusion researchers have begun studying terrorism in earnest. Several terrorism researchers (e.g., Atran, 2016; Möller-Leimkühler, 2018) have recognized the contribution of fusion to such activities and others have incorporated fusion into models of terrorist risk assessments (González-Álvarez et al., 2022; Monahan, 2017). Analyses of radicalization have also included elements of
fusion theory. These analyses focus on three populations: (a) persons at risk of radicalization and terrorist proselytizers, (b) current terrorists, and (c) former terrorists. We will discuss each population in turn.

To understand the role of fusion among those most at risk of engaging with terrorism, early studies sampled populations that were suspected to harbor terroristic sympathies. For example, one research team conducted fieldwork in two neighborhoods in Tetouan and Casablanca, the home of most of the plotters of the 2004 Madrid train bombings. They discovered that individuals who were strongly fused with their friends and regarded Sharia to be a sacred value were particularly willing to use violence, go to jail, or even die to implement sharia in their country (Sheikh et al., 2016). Likewise, just as fusion with Islam has been positively associated with support for violence among Muslim Indonesians (Wibisono, Louis, & Jetten, 2022), fusion to Judaism has predicted support for violent retribution by Israelis against Palestinians (Fredman et al., 2017). Moreover, fusion to country among Pakistani supporters of the Kashmiri cause was positively associated with willingness to fight and die for Pakistan (Pretus et al., 2019). Furthermore, this willingness correlated positively with activity in brain areas associated with emotion regulation and negatively with activity in areas related to cost-benefit calculations. This pattern has been interpreted as evidence that these prospective ‘devoted actors’ are motivated by an emotional commitment to the cause rather than a calculated analysis. More generally, these findings point to a clear role of fusion among terrorist sympathizers, with stronger fusion to the target group (e.g., religion, country) motivating greater endorsement of violence to protect the group and advance its agenda.

The role of identity fusion with groups, values, and leaders has also been applied in research with people who are already radicalized; that is, actual terrorists in prisons. A study of imprisoned Islamic jihadists, for example, revealed that admiration for Islamist activist groups was positively associated with fusion with religion and costly sacrifices for religion (e.g., moving to a prison further away from the family; Gómez, Bélanger et al., 2021). Jihadists in prison also displayed particularly strong levels of fusion with Islam and Muslims, as well as a willingness to make costly sacrifices for both (Gómez, Atran et al., 2022). Moreover, comparison of estimated fusion levels prior to (versus after) incarceration revealed that the perception of fusion with the group and value was more stable for jihadists than for other groups of criminals, such as members of street gangs or organized-crime gangs.
Considered together, the foregoing findings suggest that fusion with a particular interpretation of religion is an especially strong predictor of violence among terrorists. In fact, fusion with religion and admiration towards terrorist groups have both been included as risk factors in an instrument designed to assess the risk of violent jihadist radicalization in prisons—the DRAVY-3. Scores on these measures have also been used to successfully discriminate between Jihadist and non-Jihadist Muslim inmates (González-Álvarez, et al., 2022). Further research demonstrated that the relationship between fusion with their most cherished group and value (Muslims and Islam for jihadists, and ‘the gang’ and honor for gangsters) and willingness to make costly sacrifices for the group and for the value were mediated by perceived physical and spiritual formidability of the group (Gómez et al., 2023). Fusion thus appears to play a role in empowering terrorists to commit acts of violence. These findings have been conceptually replicated in groups other than Jihadists. For example, linguistic analyses of a broad range of terrorist manifestos revealed that fusion-related language was characteristic of endorsement of violent extremism (Ebner, Kavanagh, & Whitehouse, 2023). Fusion has also been implicated in narrative accounts regarding long-term participation in violent extremism among former members of Northern Irish paramilitary groups (Ferguson & McAuley, 2020). In short, there is considerable evidence that fusion motivates and sustains involvement in terroristic activities.

Other research on incarcerated terrorists has revealed that perceptions of injustice and discrimination positively increased fusion with religion, which in turn predicted costly sacrifices for religion. By contrast, fusion with alternative groups, such as family, reduced perceived discrimination and feelings of injustice and its corresponding effects on sacrifices for religion, while fusion with the judicial system has a negative relationship with endorsement of sacrifices for religion (Gómez, Atran et al., 2022). This research suggests that former terrorists should be treated fairly and with respect while being processed by the judicial system. In addition, connections with alternative groups and values (i.e., replacement; see the “Defusion” section above) should be fostered if deradicalization is to occur and be sustained.

Of particular interest here is research on former violent extremists who had participated in deradicalization programs. One advantage of this group is methodological: because they have been encouraged to reflect on their involvement in terrorist groups, they are more accustomed to thinking and talking about the factors that influenced their participation. For example, one study examined ex-members of the Liberation Tigers of Tamil Eelam,
an ethno-nationalist separatist group from Sri Lanka, and ex-Islamist fighters (e.g., former members of ISIS, Al Qaeda). The remarks of these participants revealed that although they mentioned both collective ties (cognitions, emotions, and values strongly linked to group membership) and relational ties (close connections with significant others), they cited relational ties more frequently (Gómez, Martínez et al., 2021). Likewise, other research revealed that severing relational ties may be a crucial aspect of sustained defusion from a terrorist group following rehabilitation (Webber et al., 2018). In short, consistent with identity fusion theory, research on former terrorists underscores the tremendous impact of forming—and severing—relational ties to the group across all stages of radicalization.

7.3 Fusion and football hooliganism

This is another area to which fusion theory has been applied with great success. Hooliganism refers to the organized violence that occurs between the “die-hard” supporters of rival football teams. It results in hundreds of deaths each year, especially in Africa and South America. Fusion among football fans is shaped by the highly emotional experience of watching one’s team play (Newson, Buhrmester, & Whitehouse, 2016; Newson et al., 2020; Whitehouse et al., 2017), and is associated with a life-long loyalty to the team and hatred of rivals (Newson et al., 2016). This loyalty often manifests in violent ways. Relative to regular fans, fusion is significantly higher among Brazilian football hooligans (e.g., the violent torcidas organizadas or ‘super fans’). In fact, fusion is a better predictor of willingness to fight and die for the group than social maladjustment (Newson et al., 2018; Newson, White, & Whitehouse, 2022). This link between fusion and violence among football fans has been replicated in multiple studies (Bortolini et al., 2018; Knijnik & Newson, 2020). This linkage is exacerbated by arousal-elevating factors (Swann, Gómez et al., 2010) such as cocaine usage (Newson, 2021).

In addition to helping to illuminate the psychological underpinnings of football hooliganism, identity fusion theory may also offer some creative solutions to violence. The Twinning Project is one such example. It aims to reduce recidivism by fusing prisoners to local football clubs, thereby redirecting the strong bonds implicit in football fandom toward a rehabilitative goal (Newson & Whitehouse, 2020).

7.4 Fusion and the military

Fusion theory has great potential for application in a military context. An early finding was that soldiers fighting in the 2011 revolution in Libya
found that frontline combatants were as strongly fused to their battalion as to their families, while those who provided logistical support were more fused with their families (Whitehouse et al., 2014). Recent fieldwork conducted in northern Iraq with fighters against the Islamic State including Peshmerga (Kurdish Regional Government forces), Iraqi army Kurds, and Arab Sunni militia, revealed that all combatants were fused with the group they were fighting for (Gómez et al., 2017). Researchers have also reported that fusion with the military predicted pro-military attitudes and behaviors among current and former soldiers (Hart & Lancaster, 2017).

Much of the evidence linking fusion to military activity could reflect either a tendency for strongly fused soldiers to volunteer for frontline assignments or for exposure to dysphoric combat experiences to increase fusion. The former interpretation is supported by evidence that fusion is associated with willingness to fight in multiple contexts, including Cadets of US Air Force (Gómez, Vázquez et al., 2023). On the other hand, both interpretations seem to be supported by recent analysis of archival data from US soldiers in the Second World War. This work has indicated that soldiers bonded through past combat experiences report higher levels of fusion with their company, and that this is related to more positive appraisals about future combat experiences (Klein & Peitz, 2023).

Such findings suggest that fusion theory could clarify a host of issues within military psychology. Applying the theory to this domain could enrich fusion theory as well, as this is one of the very few domains in which one could argue that increasing violent action is desirable. An important caveat here is that even in a military context the promotion of indiscriminate violence is counterproductive, so future research would benefit from understanding how fusion can be harnessed to produce violence that is targeted, ethical, and compliant with international law.

7.5 Fusion and pro-social outcomes

From the very first article (Swann et al., 2009), identity fusion publications have emphasized the extreme behaviors that strongly fused persons enact when they are threatened. Nevertheless, there have been instances in which researchers measured more benign behaviors such as donating money to needy in-group members and increasing the speed of a fusion-related avatar (Swann, Gómez et al., 2010). Similarly, researchers have reported that fusion with one’s university predicted retention up to one
year later (Talaifar et al., 2020). Still others have documented linkages between identity fusion and physical well-being (Tunçgenç, van Mulukom, & Newson, 2023).

In recent years, theorists and researchers have taken this work on the light side of fusion a step further by developing a more expansive analysis of the beneficial effects of identity fusion. Drawing on both the fusion and attachment literature, Klein and Bastian (2023) proposed the fusion-secure base hypothesis, which turns on the notion that strongly fused people are especially likely to consider their ingroup a reliable and supportive safe haven or “secure base”. Imbued with the synergistic strength inherent in fusion, it argues that strongly fused people can overcome intergroup anxiety and cooperatively interact with outgroup members to benefit the self and the group.

Key tenants of fusion theory support this contention and point to the potency of fusion as a secure base. For example, the relational ties principle highlights the powerful, affectionate bonds between strongly fused members of a group, marked by feelings of visceral responsibility for one another (Chinchilla, Vázquez, & Gómez, 2022). Likewise, the agentic-personal-self principle indicates that strongly fused people remain agentic and sufficiently empowered to venture outside the group. This is further cemented by the identity synergy principle: the group will not leave them, because they are the group. These functions sum to create a powerful feeling of safety and security, which the attachment literature has consistently reported is strongly associated with a willingness to engage and interact with outgroup members (see Tropp, 2021). Correspondingly, several researchers have already reported that fusion—to a greater degree than identification (Klein, Greenaway, & Bastian, 2024)—predicts willingness to interact positively with outgroups (Gómez, Vázquez, & Atran, 2023; Klein, Greenaway, & Bastian, 2024; Newson, White, & Whitehouse, 2022; Vázquez, Gómez, López-Rodríguez, & Swann, 2023).

Central to fusion’s effect on intergroup relations is the context in which interactions take place. For example, if the outgroup is perceived as non-threatening, then cooperation and trust is more likely, simply because this is more likely to result in mutually beneficial outcomes than warfare. If the outgroup is threatening, however, then warfare may well represent the most logical means of protecting the group and its interests, and fusion will predict the violent outcomes for which it is renowned. The role of threat as a moderator has been established both experimentally, wherein fusion predicted warmth unless the outgroup was primed as a threat (Vázquez, Dovidio, et al., 2023), and in field research in the volatile Bangsamoro region.
of the Philippines, wherein fusion predicted trusting intergroup relations if the outgroup was well regarded (Klein, Bastian, Odjidja et al., 2023).

Based partly on the results of economic games experiments that suggest highly agentic people are quicker to shift from cooperative strategies to exploitative strategies when the pay structure of a game favors this strategy (Boone et al., 2002), we propose that strongly fused people are similarly empowered to agentically take control of their environment and adopt a strategy that most benefits the self and the ingroup.

The fusion-secure base hypothesis is significant due to both its incorporation of attachment theory constructs and its emphasis on the pro-social nature of fusion’s influence on intergroup relations. In contrast to the past emphasis on fusion’s association with violence, the secure base hypothesis suggests that fusion’s effect on intergroup relations is contextually determined, with the decision to cooperate or fight based on an assessment of what will most benefit the self and the ingroup. While that will often be intergroup cooperation, as the attachment literature emphasizes (Tropp, 2021), the oft-dark history of humankind suggests fusion can also predict intergroup exploitation and warfare if that serves the interests of the ingroup.

8. Summary and conclusion

The revised theory of identity fusion that we present here is at once both more streamlined and comprehensive than its predecessor. As can be seen in the first row of Table 1, we streamlined the theory by culling the traditional fusion principles to place sole emphasis on the identity synergy principle, which we regard as the most central to understanding fusion. The other changes to the theory were designed to both broaden and deepen our understanding of identity fusion. Row 2 of Table 1, for example, focuses on refinements centered around creating a comprehensive model of identity fusion that works for a variety of targets. That is, whereas original identity fusion theory (OIFT) emphasized groups as the principle target of fusion, comprehensive identity fusion theory (CIFT) includes other abstractions such as values, ideologies, individual persons, brands and creatures. We thereby formally liberate fusion theory from its original focus on groups and outline a coherent approach to understanding fusion to non-group abstractions. Among other things, this development paves the way for related innovations such as the existence of “fusion clusters”.

Comprehensive identity fusion theory (CIFT)
The increased focus on the targets of fusion also led to our discussion of properties that help predict when fusion to a target is particularly likely to occur. As shown in row 3 of Table 1, we proposed three such properties, namely clarity, consistency, and congruence. This new approach helps iron out theoretical inconsistencies (e.g., acknowledging that relational ties are only relevant to specific forms of fusion) while advancing the theory toward a streamlined account of fusion and the broad range of targets to which it can be applied.

Related to this, we built upon the original theory’s emphasis on dysphoric experiences as proximal causes of fusion by adding several additional causes: self-verification, emotionally compelling experiences, and shared essence between the self and fusion target (see row 4 of Table 1). Finally, we added additional causes of defusion. As shown in row 5 of Table 1, we identify three general causes of defusion: reduction, removal, and replacement of the fusion target.

To a large degree, CIFT is the beneficiary of the remarkable conceptual and empirical developments in the field since the initial identity fusion publication (Swann et al., 2009). What started as a theory largely concerned with a niche problem—that is, providing a group-based understanding of the motivations of Jihadist suicide bombers—expanded into a comprehensive account of the synergistic unions that humans form with multiple targets. In fact, the theory has moved beyond a fixation on violence and groups, and now incorporates fusion to a range of targets to explain both hostile (Atran, 2016) and cooperative (Klein & Bastian, 2023) behaviors. Our understanding of identity fusion itself has also progressed, with multiple explanations of its developmental origins (e.g., Gomez et al., under review; Reese & Whitehouse, 2021), an anthropological account postulating its potential evolutionary origins (Whitehouse, 2018), and a growing number of studies illuminating the psychological processes that encourage strongly fused people to engage in extreme behavior (e.g., Swann, Buhrmester et al., 2014; Swann, Gómez et al., 2010). We have written this chapter with the goal of continuing fusion theory’s refinement and expansion, specifically by taking stock of the fusion literature, clarifying theoretical ambiguities, developing new fusion targets and properties, and outlining practical applications.

As fusion theory matures, it should become increasingly able to offer answers to questions of practical significance. This chapter highlighted key developments in this regard. For example, fusion theory has been usefully applied to understanding terrorism, with research on the process of
defusion offering a framework for deradicalization strategies (i.e., reduction, removal, and replacement). More generally, fusion theory has proven to be a useful framework for understanding and combatting diverse violent phenomena, ranging from football hooliganism to political violence. Recent research has also highlighted the ‘light side’ of fusion, such as its capacity to promote intergroup relations. Although still in a nascent stage, this work represents a significant development from the theory’s violent origins and may represent a more positive area in which identity fusion can be applied.

In closing, we are encouraged by the large body of research conducted on identity fusion over the past 15 years and are excited to see what the next 15 years hold. We hope that the practical applications of the theory are seized upon by researchers and practitioners alike, particularly in deradicalization and violence reduction contexts, but perhaps also for harnessing fusion to achieve prosocial outcomes. This chapter is also unlikely to be the ‘final say’ on the principles or properties of fusion theory; as fusion continues to lift itself from the shadow of related constructs, further theoretical ambiguities will surely be identified and resolved.

Table 1 Contrasting qualities of the original identity fusion theory (OIFT) and comprehensive identity fusion theory (CIFT).

<table>
<thead>
<tr>
<th>OIFT</th>
<th>CIFT</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Principles</strong></td>
<td></td>
</tr>
<tr>
<td>Identity synergy, agentic-personal-self,</td>
<td>Identity synergy</td>
</tr>
<tr>
<td>relational ties, irrevocability</td>
<td></td>
</tr>
<tr>
<td><strong>Targets of fusion</strong></td>
<td></td>
</tr>
<tr>
<td>Groups</td>
<td>Abstractions including groups, individuals,</td>
</tr>
<tr>
<td></td>
<td>values, ideologies, brands, or creatures.</td>
</tr>
<tr>
<td><strong>Properties of targets that attract</strong></td>
<td></td>
</tr>
<tr>
<td>potential fusers</td>
<td>Clarity, consistency, congruence</td>
</tr>
<tr>
<td><strong>Proximal causes of fusion</strong></td>
<td></td>
</tr>
<tr>
<td>Dysphoric experiences</td>
<td>Self-verification, shared essence, emotion</td>
</tr>
<tr>
<td></td>
<td>ally compelling experiences</td>
</tr>
<tr>
<td><strong>Causes of defusion</strong></td>
<td></td>
</tr>
<tr>
<td>Removal of group</td>
<td>Reduction, removal, replacement of fusion</td>
</tr>
<tr>
<td></td>
<td>target</td>
</tr>
</tbody>
</table>
Nevertheless, if the past is any indication of future progress, we are confident that fusion theory will continue to grow, and we hope this chapter will help guide the next stage of its development.

References


